CLASS-CONTROLLING COPY–PASTE AUGMENTATION FOR NUCLEAR SEGMENTATION

Heeyoung Ahn[1] *Yiyu Hong*[1] *Kyoung-Mee Kim*[2,3,4]

[1]Department of R&D Center, Arontier Co., Ltd, Seoul, Republic of Korea
[2]The Samsung Advanced Institute for Health Sciences & Technology (SAIHST), Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea
[3]Department of Pathology and Translational Genomics, Samsung Medical Center, Sungkyunkwan University School of Medicine, #81, Irwon-ro, Gangnam-Gu, Seoul 06351, Korea
[4]Center of Companion Diagnostics, Samsung Medical Center, Seoul, Republic of Korea.

ABSTRACT

Building segmentation models that can deal with rare and small nuclear objects in hematoxylin and eosin (H&E) stained pathologic images is a challenging task in digital pathology. Applying image augmentation can help alleviate this challenge. Hence, we propose new class-controlling copy–paste augmentation using a prepared nuclear objects set. Several image augmentations have been developed in computer vision for improving model performance; however, most of them are general-purpose methods and have not been designed for a specific domain. Our proposed method is appropriate for the pathology domain and provide strong regularization to make the model robust. In addition, it has the advantage of alleviating class imbalance problem, which is very common in histology datasets for nuclear segmentation. In our cross-validation experiments on a multi-tissue histology dataset, our method improves PQ and mPQ+ from 64.31 to 64.52 and 52.3 to 52.9, respectively.

***Index Terms*—**Nuclear segmentation, digital pathology, copy–paste, deep learning, image augmentation

1. INTRODUCTION

With the rapid development of deep learning in computer vision, the potential of automated algorithms that can assist doctors has drawn the attention of researchers for the past few years. Especially in digital pathology, where pathologists perform visual assessment using whole-slide-images acquired from histology slides, it has been expected that by leveraging the power of deep learning, diagnostic costs, which are inevitably incurred, could be reduced. Hence, nuclear segmentation and classification algorithms in hematoxylin and eosin (H&E) stained tissue images have been developed over the past few years. However, it is usually difficult to acquire sufficient annotated datasets that are of good quality in this field because of the expensive and laborious annotation work. In addition, most pathologic tissue images have nuclear class imbalance issues because of their diverse appearance. Therefore, it is essential to develop a method that can exploit limited data to improve data efficiency. Here, we focus on image augmentation for improving data efficiency. Our proposed augmentation is a variant of that developed in [1], which has some limitations. One of the limitations of the technique in [1] is that it makes unnatural images in which the pasted objects could lie on existing objects. This could damage the core information of small nuclear objects and has little impact. Hence, we paste nuclear objects to empty regions in training images, where they will not overlap with existing nuclear objects and the essential nuclear characteristics can be preserved. Another limitation is that objects can be copy–pasted between two images only. To overcome this restriction and provide more regularization to the model, we generate the whole nuclear set, which includes whole nuclear objects from training data; hence, we can efficiently extract and paste diverse objects from random images to random images. With these features, combining the proposed copy–paste augmentation, with base augmentation provided gains of + 0.21 PQ, 0.6 mPQ in 5-fold cross-validation setting.

The main contributions of our work were as follows:
(1) Proposal of a new augmentation technique that is readily applicable and appropriate for nuclear segmentation in the digital pathology domain.
(2) Advantage of alleviating nuclear class imbalance problem that inevitably occurs in histology datasets during nuclear segmentation.

2. METHOD

Fig. 1 illustrates the overall process of our proposed augmentation. In contrast to the method in [1] that copy–pastes objects within two randomly selected images, we first prepared a pre-defined nuclear set, which included all nuclear objects solely from training images. Then, striding a 15-pixel size window, which is selected considering various size of nuclei, we calculated candidate box coordinate $B^t = (x_1^t, y_1^t, x_2^t, y_2^t)$ (t = 1, 2, …, T, T is the total number of boxes) in an image indicating the region that did not overlap with existing objects. Lastly, a nuclear subset was extracted from the nuclear set and pasted onto B^t individually. If the pasted object still overlapped with existing objects, it skipped.

Adjusting the number of boxes to use
To provide more regularization to the model, we set α from

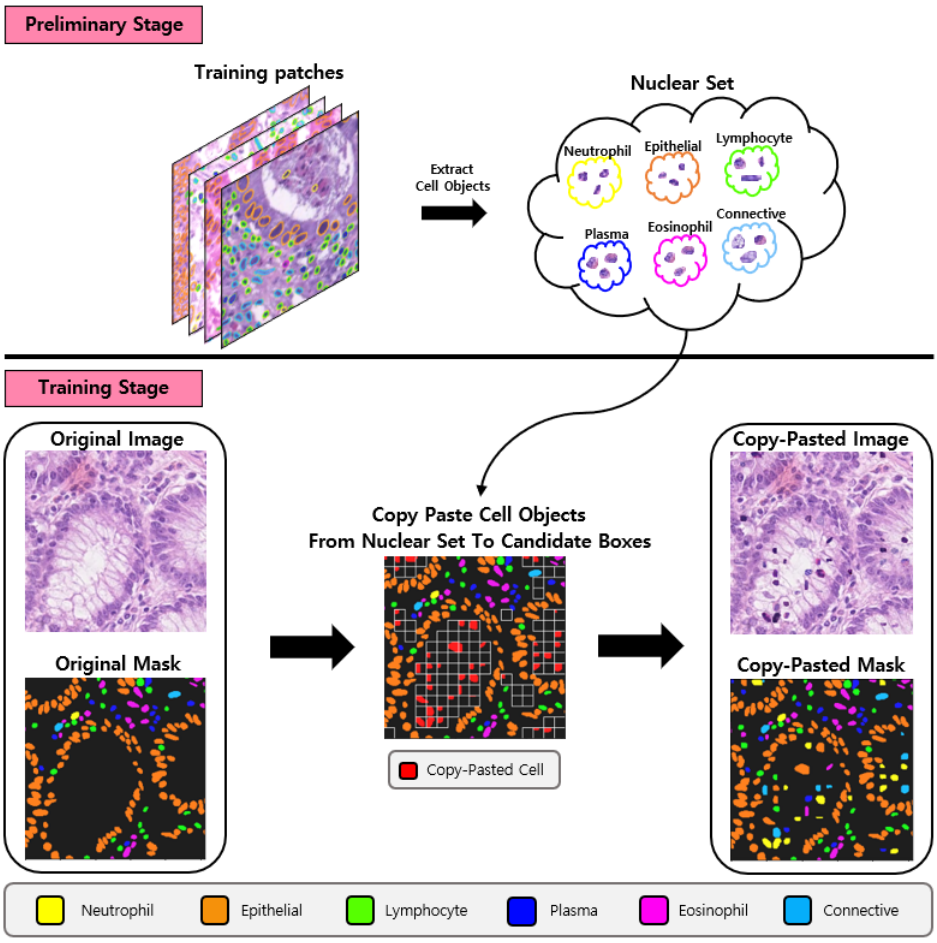

Fig. 1. Overall process for our proposed augmentation

uniform distribution (0, 1), which is an indicator for adjusting the number of candidate boxes that can be used. Then, random boxes were selected by the value of αT from total candidate boxes. For the selected *t* th candidate box in random boxes, we filled the box region with a nuclear object extracted from nuclear subset. In contrast to the technique in [1], which applied large- or small-scale jittering to pasted objects, we applied horizontal and vertical flip. αT was converted to an integer by rounding computation if its data was of floating type.

Assigning different copy–paste weight for each class

Through this method that calculate the total number of candidate boxes, copy–paste weight for each class can be assigned. That is, by multiplying the number of selected candidate boxes by the weight of each class we set, we assigned the ratio of how many objects of specific class to paste. Focus could be set on specific classes that had fewer objects than others. For example, in our case, to focus more on specific classes with fewer objects (i.e., Neutrophil, Eosinophil), we assigned the weight of 0.3, 0.1, 0.1, 0.1, 0.3, 0.1 to Neutrophil, Epithelial, Lymphocyte, Plasma, Eosinophil, Connective, respectively.

Copy–paste level considering metadata

For further application of our method, we designed two different copy–paste levels to consider metadata inherent in histology image dataset: source-level and random-level. In source-level copy–paste, considering there are several image sources (i.e. GlaS , CRAG , CoNSeP , DigestPath, Pannuke), which refers to the organization from where the histology images were taken, nuclear objects from images with specific source should be pasted onto images with same source. In other words, nuclear objects should be copy–pasted within different images with the same source; hence, this is referred to as source-level copy–paste. In contrast to this, in random-level copy–paste, nuclear objects from images with any source could be pasted onto images with random source, without considering image sources.

3. EXPERIMENTS AND RESULTS

3.1 Training Pipeline

To validate our method, we adopted the HoVer-Net [3] as the nuclear segmentation pipeline. HoVer-Net [3] predicts the horizontal and vertical distances between nuclear object and

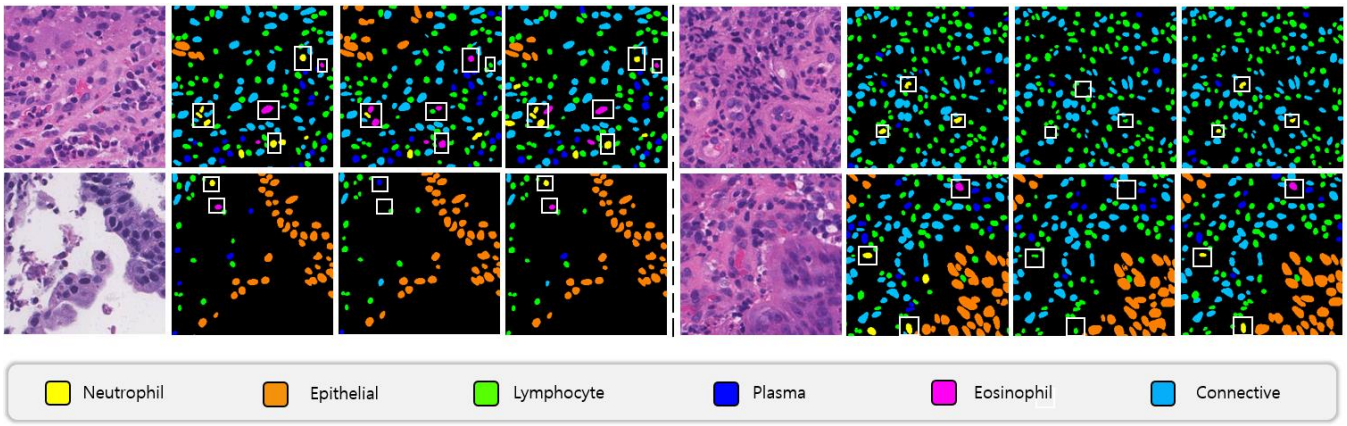

Fig. 2. Visual comparison for each method. White rectangles show improvements of our method for classes with fewer objects (i.e., yellow for Neutrophil, purple for Eosinophil). From left to right, original image, original mask, prediction mask of Unet++ with base augmentation, prediction mask of Unet++ with base augmentation combined with proposed copy-paste (source-level)

their centers of mass, which are leveraged to segment adjacent nuclei, precisely in subsequent post-processing. Only the model architecture is replaced with Unet++ [5] architecture, whose encoder is EfficientNet-b7 [6] pretrained on ImageNet [7], because it yielded slightly better performance than HoVer-Net architecture, whose encoder is pretrained on ImageNet [7] in our empirical experiments.

3.2 Dataset

We evaluated the proposed method on CoNIC nuclear segmentation dataset [2], which is provided in CoNIC Challenge 2022 [4]. In particular, CoNIC dataset [2] comprises various sizes of 238 H&E stained images at 20x objective magnification (0.5 microns/pixel) with a minimum height and width of 337 and 500 and a maximum of 2175 and 2145, respectively. For training, we strided the window of size 256 x 256 with 256 intervals on each image and prepared 4981 patch data.

3.3 Implementation Details

For base augmentation, we used colorjitter, flip, rotate, and blur. We used the Radam [8] optimizer, with an initial learning rate of 1e-3. The learning rate was downscaled from 1e-3 to 1e-5, using a step scheduler with 0.1 ratio every 100 epochs. We trained the model with a batch size of 8 for 300 epochs. All experiments were conducted under a five-fold cross-validation setting for clear verification.

3.4 Experimental Results

For evaluation, two different metrics were used for the challenge: panoptic quality (PQ) and multi-class panoptic quality (mPQ). PQ is defined as the multiplication of detection quality (DQ) and segmentation quality (SQ). DQ is defined as F1-score to evaluate DQ for nuclei segmentation. If the intersection of union (IoU) between a single predicted object and its ground-truth object was greater than 0.5, it was considered as true positive (TP), else false positive. SQ is defined as the summation of IoU for all objects that are considered as TP divided by TP in an image. mPQ is defined as the average of PQ_t for each type t. All experiments were conducted under the five-fold cross-validation settings and all scores were the average of all folds. Table 1 summarizes the

Table 1. Comparative results for each method

Architecture	Augmentation	PQ	mPQ
HoVer-Net	Base	64.3 ±0.5	51.52±0.47
Unet++	Base	64.31±0.58	52.30±0.51
Unet++	Base + copy–paste (random-level)	64.51±0.6	52.83±0.75
Unet++	Base + copy–paste (source-level)	**64.52**±0.44	**52.90**±0.54

Table 2. Comparative results for each method in terms of classes

Classes	Method	PQ
1. Neutrophil	Unet++	27.14
	+ Copy–Paste (random-level)	28.37
	+ Copy–Paste (source-level)	**28.64**
2. Epithelial	Unet++	61.53
	+ Copy–Paste (random-level)	**61.76**
	+ Copy–Paste (source-level)	61.65
3. Lymphocyte	Unet++	71.04
	+ Copy–Paste (random-level)	71.04
	+ Copy–Paste (source-level)	**71.05**
4. Plasma	Unet++	54.72
	+ Copy–Paste (random-level)	54.57
	+ Copy–Paste (source-level)	**54.58**
5. Eosinophil	Unet++	38.75
	+ Copy–Paste (random-level)	40.27
	+ Copy–Paste (source-level)	**40.54**
6. Connective	Unet++	60.68
	+ Copy–Paste (random-level)	61.02
	+ Copy–Paste (source-level)	**61.04**

Table 3. Comparative results according to α

Random Level	Source Level	α	PQ	mPQ
✓		Soft	64.48	52.71
	✓	Soft	64.48	52.9
✓		Hard	64.51	52.83
	✓	Hard	**64.52**	**52.9**

Table 4. Comparative results according to application of overlapping in copy–paste.

Random Level	Source Level	Overlap	PQ	mPQ
✓		✓	64.35	52.41
	✓	✓	64.37	52.45
✓			64.51	52.83
	✓		**64.52**	**52.9**

comparative results for each method. Under the same base augmentation setting, using Unet++ [5] yields slightly better performance of 0.82 mPQ than using HoVer-Net [3]. Compared with Unet++ with base augmentation, combining proposed random-level copy–paste augmentation with base augmentation provided gains of 0.2 PQ, 0.53 mPQ. Combining the proposed source-level copy–paste augmentation with base augmentation provided gains of 0.21 PQ, 0.6 mPQ, which was slightly better performance than random-level. Table 2 summarizes the comparative results for each method in terms of classes. Combining source-level copy–paste augmentation with baseline provided gains of 1.5, 0.12, 0.01, 1.79, 0.36 PQ for Neutrophil, Epithelial, Lymphocyte, Eosinophil, and Connective, respectively, except for slight decrease of 0.15 PQ for Plasma. Note that classes that were assigned high copy–paste weight (i.e., Neutrophil, Eosinophil) had significant improvement compared with other classes, which demonstrated that the proposed method had an effect of alleviating class imbalance, by focusing more on insufficient classes. In the same vein, Fig. 2 shows the visualization comparison between base augmentation and the proposed copy–paste augmentation with base augmentation, demonstrating SQ improvements for insufficient classes. Table 3 summarizes the comparative results for each copy–paste augmentation according to α value. For the convenience of comparison, we set just two categories for α value: Soft and Hard. Soft means α is from uniform distribution (0, 0.5), which results in copy–pasting small number of objects. Hard means α is from uniform distribution (0, 1), which results in copy–pasting small and large number of objects simultaneously. As summarized in Table 3, there was no significant difference in mPQ for each copy–paste level between Soft and Hard, but there is slight improvement (+0.3) in PQ for each copy–paste level when applying Hard. Table 4 summarizes the comparative results for each copy–paste augmentation according to the application of overlapping when copy–pasting. In other words, assuming that the number of objects to copy–paste per class was the same, the comparative results between pasting the objects to the candidate boxes (i.e., not overlapping with existing objects) or to a random area (i.e., overlapping with existing objects) are demonstrated. Pasting objects to a random area instead of candidate boxes had little effect.

4. CONCLUSIONS

Maximizing data efficiency in data-hungry pathologic field and solving the class imbalance problem, which is common in histology images, are essential. This paper presented a novel class-controlling copy–paste augmentation for nuclear segmentation. The proposed method can control the degree of how many objects to paste by calculating candidate boxes in an image, and therefore, exhibits an advantage of alleviating the class imbalance problem in histology images for nuclear segmentation. In nuclear segmentation for histology images, the proposed method can be effective additive, along with other image augmentations.

5. ACKNOWLEDGEMENTS

This study was supported by a grant from the Korea Health Technology R&D Project via the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (Grant Numbers: HI21C1137).

6. REFERENCES

[1] G. Ghiasi, Y. Cui, A. Srinivas, R. Qian, T.-Y. Lin, E.D. Cubuk, Q.V. Le, B. Zoph, "Simple Copy-Paste is a Strong Data Augmentation Method for Instance Segmentation," in *CVPR*, 2021.

[2] G. Simon, et al., "Lizard: A Large-Scale Dataset for Colonic Nuclear Instance Segmentation and Classification," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021.

[3] S. Graham, Q.D. Vu, S.E.A. Raza, A. Azam, Y.W. Tsang, J.T. Kwak, and N. Rajpoot, "Hover-net: Simultaneous Segmentation and Classification of Nuclei in Multi-Tissue Histology Images," *Medical Image Analysis*, vol. 58, pp. 101563, 2019.

[4] S. Graham, et al., "CoNIC: Colon Nuclei Identification and Counting Challenge 2022." arXiv preprint arXiv:2111.14485 (2021).

[5] Z. Zhou, M.M.R. Siddiquee, N. Tajbakhsh, J. Liang, "UNet++: A Nested U-Net Architecture for Medical Image Segmentation," in *CVPR*, 2018.

[6] M. Tan, Q.V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *ICML*, 2019.

[7] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., "Imagenet Large Scale Visual Recognition Challenge," in *IJCV*, 2015.

[8] L. Liu, H. Jiang, P. He, W. Chen, X. Liu, J. Gao, and J. Han, "On the Variance of the Adaptive Learning Rate and Beyond," in *ICLR*, 2019.