



Brain Slices Microscopic Detection Using Simplified SSD with Cycle-GAN Data Augmentation

Weizhou Liu^{1,2}, Long Cheng^{1,2(✉)}, and Deyuan Meng³

¹ State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
long.cheng@ia.ac.cn

² School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China

³ School of Automation Science and Electrical Engineering, Beihang University, Beijing 100083, China

Abstract. Orderly automatic collection of brain slices on the silicon substrate is critical for understanding the working principle of the whole-brain neural network. Accurate and real-time brain slices detection with microscopic CCD is crucial for automatic collection of brain slices. To solve this task, an efficient simplified SSD detection model with Cycle-GAN data augmentation is presented in this paper. The proposed simplified SSD streamlines the detection network of the original SSD architecture, leading to a more rapid detection. Moreover, the proposed Cycle-GAN data augmentation method overcomes the limitation of training images. To verify the effectiveness of the proposed method, experiments are conducted with a self-made brain slices dataset. The experiment results suggest that, the proposed method has a good performance of rapidly detecting brain slices with only a small training dataset.

Keywords: Microscopic object detection · Deep learning
Data augmentation

1 Introduction

Three-dimensional (3D) representations of cerebral ultrastructure are essential for fully understanding the structure and working principle of the whole-brain neural network [1]. Currently, 3D imaging with the scanning electron microscopy provides the resolution necessary to reliably reconstruct all neuronal circuits contained within brain tissues [2].

For achieving high quality 3D imaging, brain tissues are sliced to brain slice ribbons, which need to be collected on silicon substrates. The slicing mechanism and the manually collection process are shown in Fig. 1, ribbons of brain slices floating on a knife boat are manually deposited on a silicon substrate. However,

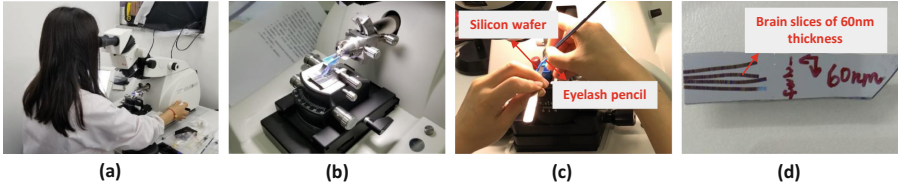


Fig. 1. Process for manually collection of brain slices. (a) shows the operating environment. (b) shows the slicing mechanism of Leica UC7 ultramicrotome. (c) shows the manually collection process. (d) shows the collected brain slices.

manually collection for huge amounts of brain slices requires the high experienced operation skills, and also consumes intensive work. Studies have been carried out for brain slices automatic collection, e.g., Schalek *et al.* developed an automated tape ultramicrotome (ATUM) which automatically collects thousands of serial sections on a plastic tape, but the charge accumulation of the sample caused by low conductivity of the plastic tape will influence the imaging quality [3].

In order to realize the automatic collection with silicon substrate, a simple automatic collection prototype was proposed, which is shown in Fig. 2. The collection device allows precise position and manipulation of a hydrophilic circular silicon wafer. The position and spin velocity of the silicon wafer could be adjusted by a displacement platform and a rotary motor, respectively. In the collection process, the silicon wafer was inclined dipped into a custom-built knife boat which has a baffle plate to ensure brain slices floating in the collection area, and as sectioning progressed the wafer was moved to the position of slices to be collected and slowly rotated to capture the nascent sections of the brain slice ribbons. This procedure needs an accurate and fast detection of brain slices in the microscopic live video.

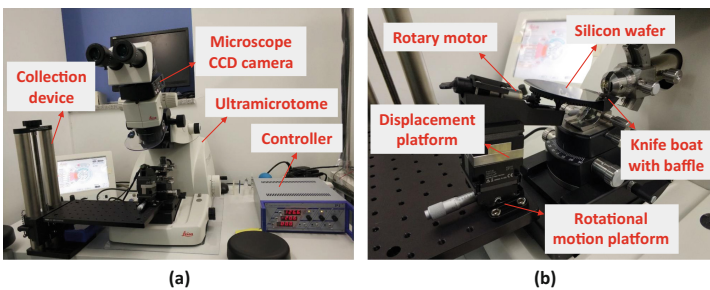


Fig. 2. The automatic collection device for brain slices. (a) shows the collection device and the ultramicrotome. (b) shows the detailed structure of the collection device.

This detection task is challenging due to factors like variations in brain slice shape, color and density. And the existence of reflection light of diamond knife and lighting condition changes in the background will also pose challenges to

detection. In addition, unlike natural image object detection task, the difficulty for obtaining large-scale dataset poses another challenge to train deep learning based detection models.

To address these challenges, a simplified SSD detection model was proposed for real-time brain slices detection task. In addition, Cycle-GAN data augmentation was proposed to improve the detection accuracy with limited training dataset. The proposed simplified SSD with Cycle-GAN data augmentation could achieve an accurate and real-time detection in the brain slices microscopic detection task.

2 Related Works

2.1 Data Augmentation

It is a common knowledge that a deep learning based algorithm would be more effective when accessing more training data. Previous studies have demonstrated the effectiveness of data augmentation through minor modifications to the available training data, such as image cropping, rotation, and mirroring [4].

In recent years, generative adversarial network (GAN) has been proposed as a powerful technique to perform unsupervised generation of desired image samples [5]. For improving image classifier performance, several data augmentation methods based on GAN have been proposed and have shown to be extremely good at augmenting training dataset [6,7].

2.2 Microscopic Object Detection

In recent years, deep learning based methods have achieved a great success in object detection. Unlike traditional handcrafted features, the features learned by deep learning based methods are more effective and general. A series of studies have been carried out for microscopic object detection with deep learning based methods. Mao *et al.* proposed a Convolutional Neural Network (CNN) with automatically learned features achieved better results than traditional methods [8]. Oscar *et al.* proposed a deep learning based method which was sufficient for visual detection of soil-transmitted helminths and schistosoma haematobium through a mobile, digital microscope [9].

For achieving better detection results, the state-of-the-art detection models have been applied to microscopy images. A series of methods based on Faster Region-based Convolutional Neural Network (Faster R-CNN) were used for microscopic cell detection, these studies reveal that using Faster R-CNN model to detect cells in microscopic image is very effective [10–12]. However, the two-stage detection models like Faster R-CNN based on region proposal still need expensive computations, and are difficult to use for real-time applications. To overcome this problem, end-to-end detection methods such as Single Shot Multi-box Detector (SSD) are much simple and faster than Faster R-CNN, which could be used for real-time applications [13]. Yi *et al.* proposed an efficient neural cell

detection method based on a light-weight SSD neural network [14], the detection speed is 10 FPS by testing on a single Nvidia K40 GPU. However, there is still room for improving the detection precision and speed with SSD model.

3 Data Augmentation

3.1 Data Augmentation with Traditional Techniques

Considering that the training image set is relatively small, data augmentation by some traditional methods was performed. In order to increase the robustness of the model for detecting various input object sizes and shapes, each training image was randomly sampled to patches which have minimum jaccard overlap with the ground truth bounding box of 0.1, 0.3, 0.5, 0.7, or 0.9. In addition, extra patches were randomly sampled from the input training image, and the overlapped part of the ground truth box will only be kept when the center of ground truth box was in the sampled patch. After the above-mentioned sample step, each sample patch was resized to 300×300 size, and some random flipping and photo-metric distorting were additionally performed on each sampled patch.

3.2 Data Augmentation with Cycle-GAN

In this paper, Cycle-GAN is used to augment the training data for improving microscopic object detection performance. Unlike traditional GAN, Cycle-GAN captures special characteristics of one image set and learns how these characteristics could be translated into other image sets [15].

For the microscopic images set augmentation task, Cycle-GAN is used to transfer images from training dataset into the artistic styles of Monet, Van Gogh, Cezanne, and Ukiyo-e. Figure 3 gives an augmentation example of a single training image. Through this augmentation process, additional data annotation is avoided due to Cycle-GAN kept the target locations in original images. The generated training images via Cycle-GAN have different color, density, lighting conditions compared to original training images, using these augmented images for training will improve the generalization ability of the detection network.

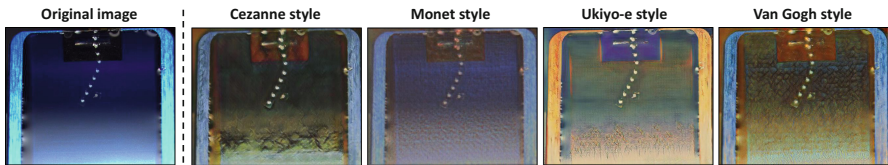


Fig. 3. An augmentation example of a single training image. The brain slices microscopic image in the training dataset was transferred into artistic styles of Monet, Van Gogh, Cezanne, and Ukiyo-e.

4 Simplified SSD Detection Model

4.1 Network Architecture

The overall pipeline of the proposed approach is shown in Fig. 4. Firstly, each input image is initially preprocessed into 300×300 input size. Then, the simplified SSD is formed by following the main architecture of the original SSD with some improvements, which aims to increase detection speed and accuracy in the brain slices microscopic detection task. The network architecture of the proposed simplified SSD model includes two main parts: a standard base network based on VGG-16 architecture (without any classification layers) and an auxiliary structure consists of some extra feature layers. The base network is used for feature extraction and the auxiliary structure is used to increase detection ability at multiple scales.

Compared to the original SSD model, the proposed simplified SSD model removes several extra feature layers on the top of the original SSD network. Each feature layer in the original SSD model was designed to detect objects within the specific range of scale, the large feature maps aim to detect small objects in the input image. The brain slices in the microscopic images are relatively small and without large scale change. For the brain slices detection task, three large feature layers (Conv4_3, Conv7, Conv8_2) which aim to detect small objects were used as the feature maps.

Similar to the original SSD, each feature map contains a specific number of default boxes, which are of certain shapes. Then object confidences and locations of these default boxes are predicted by two convolutional filters, i.e., localization filter and object filter. Finally, a fixed-size prediction of locations and scores for the presence of brain slices in those default boxes can be obtained, and a followed non-maximum suppression step will produce the final detection results.

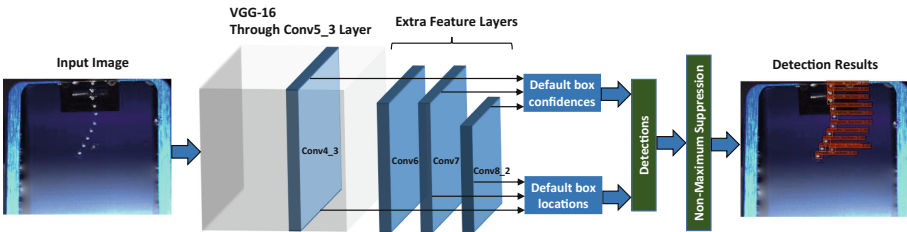


Fig. 4. The network architecture of the proposed simplified SSD.

4.2 Training

Unlike two-stage detection methods such as Faster R-CNN, SSD has a simple end-to-end training process. Next the training details are described in the following three aspects.

Transfer Learning. Several data augmentation methods have been used to solve the problem of the limited training images. However, it is not enough to train a large CNN model. In order to solve this problem, transfer learning were adopt to initialize the proposed simplified SSD model. In particular, the VGG-16 base network pre-trained on ImageNet classification dataset was transferred to the proposed simplified SSD model as the base network. This transfer learning process is an effective paradigm for helping the proposed detection model extract the features of brain slices when training data is scarce.

Matching Strategy. Each default box in feature maps needs to be matched with the corresponding ground truth bounding boxes. In this way, the positive examples and negative examples for training the proposed simplified SSD can be obtained. More specially, the jaccard overlap index of each default box with each ground truth box is calculated. For each ground truth box, the default box with the highest jaccard overlap index will be selected for matching. In addition, for each ground truth box, if it's jaccard overlap index with any default box is higher than a threshold (0.5), then they will be matched.

Training Objective. The proposed simplified SSD training objective contains two objective losses: the localization loss and the confidence loss. The localization loss is defined as a Smooth L1 loss, which can be written as:

$$L_{loc}(x, l, g) = \sum_{i \in Pos}^N \sum_{m \in \{cx, cy, w, h\}} x_{ij}^p \text{smooth}_{L_1}(l_i^m - \hat{g}_j^m), \quad (1)$$

$$\text{smooth}_{L_1}(z) = \begin{cases} 0.5z^2 & \text{if } |z| < 1 \\ |z| - 0.5 & \text{otherwise} \end{cases}, \quad (2)$$

$$\begin{aligned} \hat{g}_j^{cx} &= (g_j^{cx} - d_i^{cx})/d_i^w & \hat{g}_j^{cy} &= (g_j^{cy} - d_i^{cy})/d_i^h, \\ \hat{g}_j^w &= \log\left(\frac{g_j^w}{d_i^w}\right) & \hat{g}_j^h &= \log\left(\frac{g_j^h}{d_i^h}\right), \end{aligned} \quad (3)$$

where N is the total number of matched default boxes; $x_{ij}^p = \{1, 0\}$ is an indicator, $x_{ij}^p = 1$ only when the i -th default box and the j -th ground truth box of category p were matched; l and g represent the predicted box and the ground truth box, respectively; (cx, cy) represents the center offset of the default bounding box, and (w, h) represents the width offset and height offset of the default bounding box.

The confidence loss is a traditional softmax loss over multiple classes confidences(c), which is defined as:

$$L_{conf}(x, c) = - \sum_{i \in Pos}^N x_{ij}^p \log(\hat{c}_i^p) - \sum_{i \in Neg} \log(\hat{c}_i^0), \quad \text{where } \hat{c}_i^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)}. \quad (4)$$

The localization loss and confidence loss are combined to obtain the fully training objective:

$$L(x, l, g, c) = \frac{1}{N}(L_{loc}(x, l, g) + L_{conf}(x, c)). \quad (5)$$

5 Experiments

5.1 Data

In these experiments, rat hippocampus tissues are used for serial sectioning. Ribbons of rat hippocampus sections were cut at 30 nm section thickness using an ultramicrotome (Leica UC7, Wetzlar, Germany) equipped with a diamond knife angled at 35° (Ultra Diatome Knife, Biel, Switzerland). Then the CCD camera (Leica IC90E, Wetzlar, Germany) was adopted to collect several microscopy videos, from which 76 microscopic images which size of 1280×1024 were sampled. Among these images, 60 images were used for training, 16 images for testing.

The brain slices in these microscopic images were marked by bounding boxes as ground-truth annotations. Then data augmentation process was performed by Cycle-GAN. After the data augmentation process by Cycle-GAN, the training dataset contains 300 training images and 16 testing images. All the images and annotations were made into VOC format for convenient data reading by Caffe framework.

5.2 Results

In this subsection, the effectiveness of the proposed data augmentation method and simplified SSD are illustrated respectively. Firstly, the proposed Cycle-GAN data augmentation method is evaluated with three recent state-of-the-art detection algorithms: Faster R-CNN, SSD and YOLO-v3. Secondly, several comparison experiments were performed to illustrate the high performance of the proposed simplified SSD. The proposed data augmentation method and object detection model are trained and tested on a workstation with an Intel i7-6850K processor of 3.6 GHz and a single Nvidia Titan-xp GPU.

Cycle-GAN Data Augmentation. Faster R-CNN, SSD and YOLO-v3 are implemented on the brain slices microscopic dataset, respectively, and for comparison, the three state-of-the-art detection model are also implemented on the augmented dataset. For a fair comparison, VGG16 is chosen as the base network for all of the three detection models. The experiment results are given in Table 1. It can be seen that the Cycle-GAN data augmentation process could effectively improve the detection performance of all the three detection models.

Table 1. Effects of Cycle-GAN data augmentation on three state-of-the-art detection models

Method	mAP	
	Use Cycle-GAN augmentation?	
	Yes	No
SSD	0.950189	0.886907
YOLOv3	0.929375	0.869549
Faster R-CNN	0.909146	0.867713

Simplified SSD. Currently, SSD is an excellent detection model for its high speed detection performance while keep a high detection accuracy. It can be seen that SSD detection model is also very suitable for the brain slices detection task from Table 1. Since the ultimate objective is to detect brain slices with microscopic CCD in real time, the detection speed needs to be further accelerated by using the proposed simplified SSD.

The comparison experiments results are listed in Table 2. To compare with the original SSD detection model, the proposed simplified SSD has a better detection accuracy than the original SSD. Furthermore, the detection speed of the proposed simplified SSD has a distinct improvement than the original SSD detection model, and the GPU memory usage is reduced too. This characteristics enhanced the model portability of the proposed simplified SSD, specifically, it will be convenient to migrate the simplified SSD detection model to other graphics cards with lower computing ability.

Figure 5 demonstrates several examples of the final detection results. It can be seen that the proposed simplified SSD model is able to deal with the color change, background interference and shape deformation of the brain slices.

Table 2. Comparison of the proposed simplified SSD and the original SSD

Method	mAP		FPS	GPU memory
	Use Cycle-GAN augmentation?			
	Yes	No		
Original SSD	0.950189	0.886907	24	737 MB
Simplified SSD	0.965129	0.890029	29	689 MB

To illustrate the reason for selecting the three feature layers (Conv4.3, Conv7, Conv8.2) as the feature maps, layers of the original SSD are progressively removed for comparing the detection results. Table 3 shows that using Conv4.3, Conv7 and Conv8.2 as the feature maps can achieve the best detection results. From these experiments, some interesting trends can be observed. For example, it increases the detection performance if the small feature layers are progressively removed, the reason might be that the small feature maps aim to detect



Fig. 5. Visualization of the detection results using simplified SSD. The red boxes represent the final detection results. (Color figure online)

large objects which are not present in the microscopic images, so remove these small feature layers on the top of SSD network may reduce the interference of the useless large default boxes. But when only keeping Conv4_3 for prediction, although detection speed is greatly improved, the detection performance is the worst. In order to obtain a better detection accuracy and a comparable detection speed to the original SSD, Conv4_3, Conv7 and Conv8_2 are chosen as the feature maps for prediction.

Table 3. Comparison of using different output layers in original SSD

Prediction source layers from						mAP Use Cycle-GAN?		FPS
Conv4_3	Conv7	Conv8_2	Conv9_2	Conv10_2	Conv11_2	Yes	No	
✓	✓	✓	✓	✓	✓	0.950189	0.886907	24
✓	✓	✓	✓	✓		0.958645	0.890341	26
✓	✓	✓	✓			0.932281	0.890109	27
✓	✓	✓				0.965129	0.890029	29
✓	✓					0.963461	0.897847	33
✓					✓	0.900577	0.873513	42

6 Conclusions

A real-time detection model is proposed in this paper to detect brain slices in microscopic live video. To this end, in order to improve detection accuracy and speed with the relatively small training dataset, firstly, a data augmentation method based on Cycle-GAN is proposed, which significantly increases the detection accuracy. Secondly, a simplified SSD is proposed, which aims to increase the detection speed while reduce the GPU memory usage. The experiment results show that the proposed simplified SSD with Cycle-GAN data augmentation can overcome the problem of limited training dataset, and has a good ability to detect brain slices in microscopic images in real time.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China under Grants 61873268, 61633016, in part by the Research Fund for Young Top-Notch Talent of National Ten Thousand Talent Program, in part by the Beijing Municipal Natural Science Foundation under Grant 4162066.

References

1. Kubota, Y.: New developments in electron microscopy for serial image acquisition of neuronal profiles. *Microscopy* **64**, 27–36 (2015)
2. Mikula, S.: Progress towards Mammalian whole-brain cellular connectomics. *Front. Neuroanat.* **10**, 62–71 (2016)
3. Schalek, R., et al.: Development of high-throughput, high-resolution 3D reconstruction of large-volume biological tissue using automated tape collection ultramicrotomy and scanning electron microscopy. *Microsc. Microanal.* **17**, 966–967 (2011)
4. Perez, L., Wang, J.: The effectiveness of data augmentation in image classification using deep learning. arXiv preprint [arXiv:1712.04621](https://arxiv.org/abs/1712.04621) (2017)
5. Goodfellow, I., et al.: Generative adversarial nets. In: *Advances in Neural Information Processing Systems*, pp. 2672–2680 (2014)
6. Wei, L., Zhang, S., Gao, W., Tian, Q.: Person transfer GAN to bridge domain gap for person re-identification. arXiv preprint [arXiv:1711.08565](https://arxiv.org/abs/1711.08565) (2017)
7. Mariani, G., Scheidegger, F., Istrate, R., Bekas, C., Malossi, C.: BAGAN: data augmentation with balancing GAN. arXiv preprint [arXiv:1803.09655](https://arxiv.org/abs/1803.09655) (2018)
8. Mao, Y., Yin, Z., Schober, J.M.: Iteratively training classifiers for circulating tumor cell detection. In: *IEEE 12th International Symposium on Biomedical Imaging*, pp. 190–194 (2015)
9. Holmström, O., et al.: Point-of-care mobile digital microscopy and deep learning for the detection of soil-transmitted Helminths and *Schistosoma haematobium*. *Glob. Health. Action.* **10**, 1337325 (2017)
10. Yang, S., Fang, B., Tang, W., Wu, X., Qian, J., Yang, W.: Faster R-CNN based microscopic cell detection. In: *International Conference on Security, Pattern Analysis, and Cybernetics*, pp. 345–350 (2017)
11. Hung, J., Carpenter, A.: Applying faster R-CNN for object detection on malaria images. In: *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 808–813 (2017)
12. Zhang, J., Hu, H., Chen, S., Huang, Y., Guan, Q.: Cancer cells detection in phase-contrast microscopy images based on Faster R-CNN. In: *9th International Symposium on Computational Intelligence and Design*, pp. 363–367 (2016)
13. Liu, W., et al.: SSD: single shot multibox detector. In: *Proceedings of the European Conference on Computer Vision*, pp. 21–37 (2016)
14. Yi, J., Wu, P., Hoepfner, D.J., Metaxas, D.: Fast neural cell detection using light-weight SSD neural network. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 108–112 (2017)
15. Zhu, J.-Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. arXiv preprint [arXiv:1703.10593](https://arxiv.org/abs/1703.10593) (2017)