Journal of Machine Learning Research 24 (2024) 1-7 Submitted 7/24; Published 7/24

Panoptic Segmentation on Dendrites and Dendritic Spines

Sahil Sharma SAHIL.SHARMA@FAU.DE
Department of Artificial Intelligence in Biomedical Engineering
University of Erlangen-Nuremberg
91054 Erlangen, Germany

Editor: Sahil Sharma

Abstract

Dendritic spines are postsynaptic specializations thought to regulate the strength of synaptic transmission and play a critical role in neuronal plasticity, discussed in paper by Shen et al. (2008). They serve as storage sites for synaptic strength and help transmit electrical signals to the neuron's cell body. Accurate quantification of these spines is crucial for understanding their function and dynamics. Humans usually quantify light microscopy data in a painstaking and error-prone process. As stated in (Fernholz et al., 2024) paper, that the human-to-human variability is substantial (inter-rater reliability 82.2±6.4%), raising concerns about the reproducibility of experiments and the validity of using human-annotated 'ground truth' as an evaluation method for computational approaches of spine identification. To improve this process, a panoptic segmentation pipeline has been developed to automate the quantification of dendritic spines, which will help in understanding their structure.

Keywords: dendritic spines, panoptic segmentation, spines quantification, maskrcnn, fcn

1 Introduction

Dendritic spines are the seat of most excitatory synapses in the brain and are central to learning, memory, and activity-dependent plasticity (Fernholz et al., 2024). The entire DeepD3 open framework, including the fully segmented training data, a benchmark annotated by multiple experts, and the DeepD3 model zoo, is fully available. This addresses the lack of openly available datasets of dendritic spines while offering a ready-to-use, flexible, transparent, and reproducible spine quantification method (Fernholz et al., 2024).

Traditional methods for quantifying dendritic spines rely heavily on manual annotation by human experts, which is both time-consuming and prone to variability. To address the limitations of manual spine quantification, this work aims to develop a robust and automated approach using deep learning techniques. Specifically, the objective is to create a panoptic segmentation pipeline that enhances the accuracy and efficiency of dendritic spine quantification. By integrating both semantic and instance segmentation, this approach seeks to provide a more comprehensive and detailed analysis of dendritic spines, thereby reducing human error and variability.

2 Methodology

Below is the detailed explanation of the methods used in the creation of this project work.

2.1 Dataset

The dataset used in this project is taken from the DeepD3 project, which provides binary mask data for both dendrites and dendritic spines. The dataset is stored in .tiff format and includes 637 training samples and 67 validation samples. The binary mask of this dataset was used in DeepD3 as the ground truth mask for semantic segmentation of both dendrites and dendritic spines.

2.2 Preprocessing

The mask that we have for dendrites and dendritic spines is provided as binary masks. As we are performing instance segmentation, we need to preprocess the dendritic spine masks. In preprocessing, we need to convert these binary masks to instance masks, which can then be used as ground truth masks. For this step, the scikit-image library is used, specifically the 'label' method. This 'label' method identifies every region with connected components and returns an image where each connected region is assigned a unique integer value. This returned image will then be used as instance masks for the training process for instance segmentation.

2.3 Panoptic Segmentation Pipeline

As Kirillov et al. (2008) explained, panoptic segmentation unifies the typically distinct tasks of semantic segmentation (assigning a class label to each pixel) and instance segmentation (detecting and segmenting each object instance).

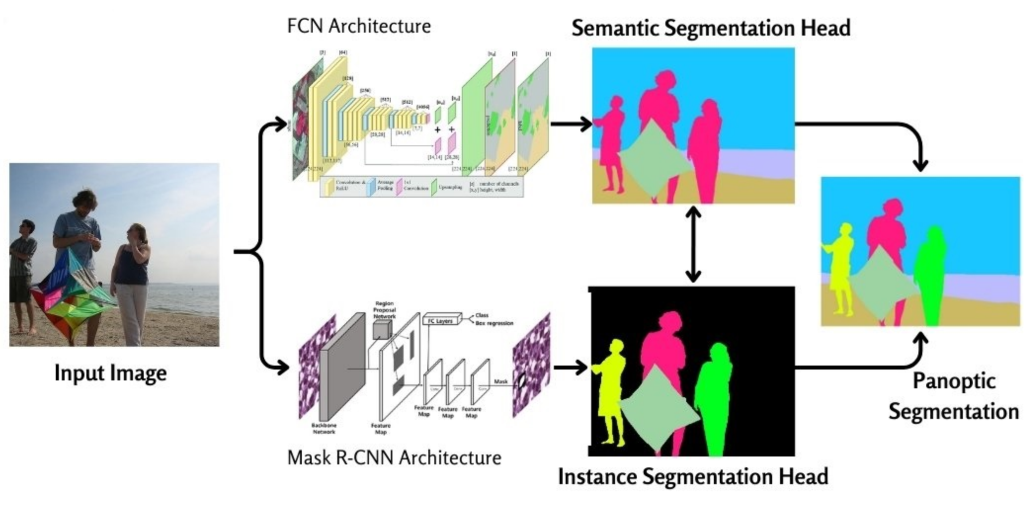

Figure 1: Panoptic Segmentation Pipeline

[1]Image available at https://viso.ai/deep-learning/panoptic-segmentation-a-basic-to-advance.

For this project, the panoptic segmentation pipeline is implemented using Mask R-CNN for instance segmentation of dendritic spines and FCN with a ResNet backbone for semantic segmentation of dendrites. Both models are trained individually, which allows for better evaluation of their performance. The Mask R-CNN focuses on detecting the boxes and predicting the masks of the dendritic spines, while the FCN provides a detailed segmentation mask of the dendrites.

The trained models are then combined to predict the panoptic mask. This is done by merging the predictions of both models during the final inference. The final panoptic mask provides a detailed representation of the dendritic structure in the input image.

3 Model Architecture

Below is the detailed explanation of both the instance and semantic models, along with the criteria used during training.

3.1 Mask R-CNN for Spines

Architecture Mask Region-based Convolutional Neural Network (Mask R-CNN) models are used for the instance segmentation task for spines. Instance segmentation deals with the correct detection of all objects in an image while also precisely segmenting each instance. It is the combination of object detection, object localization, and object classification. This type of segmentation gives a clear distinction between each object classified as similar instances. Mask R-CNN extends Faster R-CNN by adding a branch for predicting an object mask in parallel with the existing branch for bounding box recognition, mentioned in a paper on Mask R-CNN by He et al. (2017). The additional mask branch produces segmentation masks for each detected object.

Parameters The model was trained with a initial learning rate of 0.0001, a batch size of 2, and a learning rate decay schedule over 100 epochs. The model with the best validation loss is saved during training. The loss components include classification loss, bounding box loss, and mask loss. The Adam optimizer was used due to its adaptive learning rate capabilities.

3.2 FCN for Dendrites

Architecture Fully Convolutional Network (FCN) is used for semantic segmentation to segment dendrites from the background. Semantic segmentation involves classifying each pixel in an image into a predefined category. FCNs replace the fully connected layers in traditional CNNs with convolutional layers that produce dense predictions, mentioned in a paper on FCN by Long et al. (2015). These models are designed to handle pixel-level classification tasks. The network architecture includes encoder layers that downsample the input image and decoder layers that upsample it to produce pixel-wise predictions.

Parameters The model was trained with a learning rate of 0.001, a batch size of 4, and a learning rate decay schedule over 100 epochs. Binary cross-entropy loss with logits is used as the loss function. Similarly, the Adam optimizer is used in this case as well because of its adaptive learning rate capabilities.

4 Results

After training the model for 100 epochs, some evaluation metrics were used to evaluate the performance. The evaluation of these metrics is computed on the validation set. The evaluation metrics are given below:

Performance Metrics They include Precision, which calculates the proportion of true positive pixels among all pixels predicted as positive; Recall, which measures the proportion of true positive pixels among all actual positive pixels; and Intersection over Union (IoU), which computes the overlap between predicted masks and ground truth masks.

4.1 Mask R-CNN

After running the evaluation script on the validation set using Mask R-CNN, we obtained an IoU of around 0.27, Precision of 0.31, and Recall of 0.67, as shown in Table 1 below.

Table 1: Performance Metrics for Mask R-CNN

Metric	**Result**
IoU	0.2731
Precision	0.3109
Recall	0.6787

To further show the effectiveness of the Mask R-CNN, a qualitative evaluation is included. After running inference on this validation set, one sample of the predicted spine masks can be seen in Figure2 below. This representation provides a clearer understanding of the model's performance in practical scenarios. The masks and the boundary boxes of the predicted masks are colored in red.

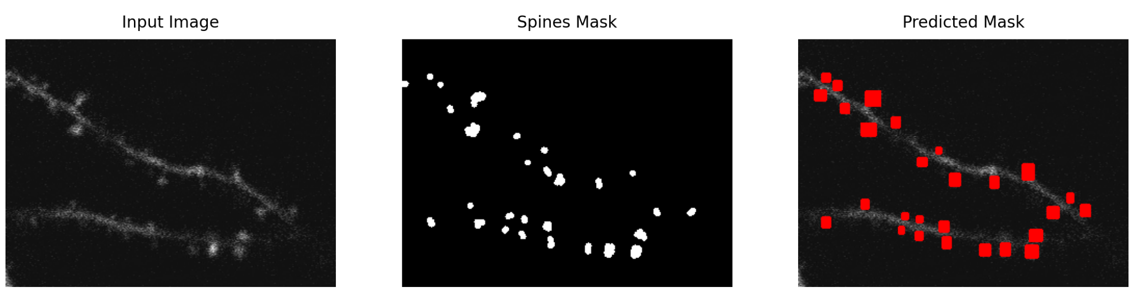

Figure 2: Predicted mask of spines using Mask R-CNN

4.2 Results from FCN

After running the evaluation script on the validation set using the Fully Convolutional Network, we obtained an IoU of around 0.4, Precision of 0.58, and Recall of 0.43, as shown in Table 2 below.

Table 2: Performance Metrics for FCN

Metric	**Result**
IoU	0.3996
Precision	0.5808
Recall	0.4350

Similarly, to further show the effectiveness of the Fully Convolutional Network, a qualitative evaluation is included. After running inference on this validation set, one sample of the predicted dendrite masks is plotted, which can be seen in Figure 3 below. This representation provides a clearer understanding of the model's performance in practical scenarios. The predicted mask of dendrites is filled with green for better visibility.

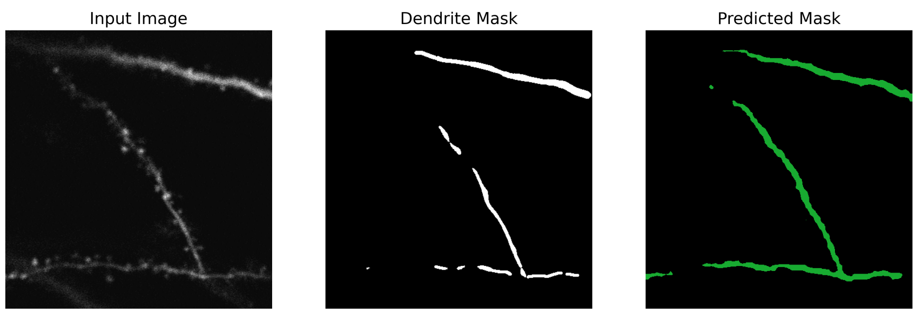

Figure 3: Predicted mask for dendrites using FCN

4.3 Panoptic Prediction

Now that we have both trained models, we can perform our last step of predicting the panoptic mask. To generate the panoptic prediction, we need to combine the predictions of both the instance and semantic models. We will get the instance mask prediction from the Mask R-CNN model and the semantic mask prediction from the Fully Convolutional Network (FCN). The process starts with Mask R-CNN by predicting the individual spines and producing the instance mask. These predicted spine instances will then be filtered based on their confidence scores. Similarly, for the same input image, FCN will predict a semantic mask that highlights the dendritic region. We then create a blank image, matching the size of the input image, which will be used to combine both masks. We apply the semantic mask with green color and overlay the instance masks with unique colors for each predicted spine mask. Instead of using the predicted boundary boxes, we will use contours for clarity. A threshold value is applied to show only the significant area, resulting in a detailed panoptic mask. A sample prediction from the panoptic inference can be seen in Figure 4.

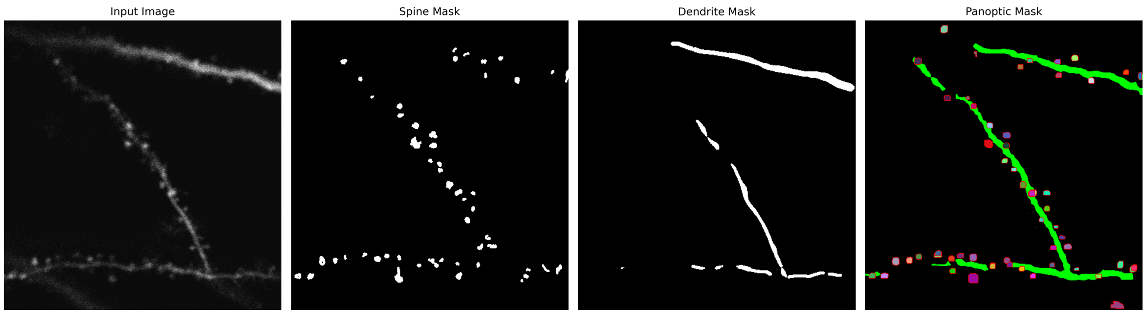

Figure 4: Panoptic Prediction

5 Discussion and Conclusion

This panoptic segmentation pipeline shows an effective integration of instance and semantic segmentation for a detailed representation of the input image on dendrites and dendritic spines. While the predictions we get are comparably sufficient, there are still areas for improvement. For instance, although the Mask R-CNN model effectively identifies the individual spines with clear boundaries, its Intersection over Union (IoU) and recall metrics indicate areas for improvement, particularly in handling closely spaced and overlapping spine instances. On the other hand, the Fully Convolutional Network (FCN) performs better in predicting the dendrite masks, as evidenced by its high precision and recall metrics. This approach surely helps in better understanding the input image of dendritic spines, but there are still challenges that need improvement. The instance segmentation can possibly be improved by using advanced architectures or by utilizing the latest instance detection models such as Detectron. For pipeline selection, FPN (Feature Pyramid Network) could be a better approach. Future work on this task should focus on enhancing instance segmentation performance by exploring new architectures or using additional data such as predefined instance masks. Expanding the dataset and trying new instance models could boost the robustness and results of the pipeline. Overall, this project provides an initial panoptic segmentation pipeline that can be used for the detection and quantification of dendritic spines. This pipeline can be further improved with the given suggestions to achieve better results in the future.

References

M. H. P. Fernholz, D. A. G. Nilo, T. Bonhoeffer, and A. M. Kist. Deepd3, an open framework for automated quantification of dendritic spines. *PLOS Computational Biology*, 2024.

K. He, G. Gkioxari, P. Dollár, and R. Girshick. Mask r-cnn. *International Journal of Computer Vision*, 2017.

A. Kirillov, K. He, R. Girshick, and et al. Panoptic segmentation. *International Journal of Computer Vision*, 2008.

J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. *International Journal of Computer Vision*, 2015.

H. Shen, S. R. Sesack, S. Toda, and P. W. Kalivas. Automated quantification of dendritic spine density and spine head diameter. *Brain Struct Funct*, 2008.