HOCHSCHULE
MITTWEIDA
UNIVERSITY OF
APPLIED SCIENCES

MASTERARBEIT

Herr
Mirko Weber

Deep learning for dense reconstruction of neurons from electron microscopic images

2019

Faculty of **Applied Computer- and Biosciences**

MASTERARBEIT

Deep learning for dense reconstruction of neurons from electron microscopic images

Author:
Mirko Weber

Study Programme:
Molekularbiologie/Bioinformatik

Seminar Group:
MO17w1-M

First Referee:
Prof. Dr. rer. nat. Dirk Labudde

Second Referee:
Dr. rer. nat. Torsten Bullmann

Mittweida, August 2019

Bibliographic Information

Weber, Mirko: Deep learning for dense reconstruction of neurons from electron microscopic images, 69 pages, 35 figures, Software CD, Hochschule Mittweida, University of Applied Sciences, Faculty of Applied Computer- and Biosciences

Masterarbeit, 2019

Satz: LaTeX

Abstract

A relatively new research field of neurosciences, called *Connectomics*, aims to achieve a full understanding and mapping of neural circuits and fine neuronal structures of the nervous system in a variety of organisms. This detailed information will provide insight in how our brain is influenced by different genetic and psychiatric diseases, how memory traces are stored and ageing influences our brain structure. It is beyond question that new methods for data acquisition will produce large amounts of neuronal image data. This data will exceed the zetabyte range and is impossible to annotate manually for visualization and analysis. Nowadays, machine learning algorithms and especially deep convolutional neuronal networks are heavily used in medical imaging and computer vision, which brings the opportunity of designing fully automated pipelines for image analysis. This work presents a new automated workflow based on three major parts including image processing using consecutive deep convolutional networks, a pixel-grouping step called connected components and 3D visualization via neuroglancer to achieve a dense three dimensional reconstruction of neurons from EM image data.

I. Contents

Contents I

List of Figures II

List of Tables III

Nomenclature IV

Preface V

1 Introduction 1
- 1.1 Machine Learning in Biomedical Imaging . . . 1
- 1.2 Image acquisition by electron microscopy . . . 2
- 1.3 ISBI 2012/2013 Challenge - Image segmentation for connectomics . . . 3
- 1.4 Flood-filling networks for single neuron reconstruction . . . 4
- 1.5 Pix2Pix with conditional adversarial networks . . . 4
- 1.6 Goal of this thesis . . . 5

2 Methods 7
- 2.1 Dataset . . . 7
- 2.2 Dataset preparation . . . 8
- 2.3 Generation of cytoplasm, overlap and membrane images . . . 8
  - 2.3.1 Cytoplasm images . . . 8
  - 2.3.2 Membrane images . . . 9
  - 2.3.3 Overlap images . . . 10
- 2.4 Training of neuronal networks . . . 12
- 2.5 Prediction, reconstruction and 3D visualization . . . 13
- 2.6 Evaluation . . . 13
  - 2.6.1 Precision and recall . . . 13
  - 2.6.2 Adjusted RAND index . . . 14
  - 2.6.3 Split and merge errors . . . 14
  - 2.6.4 Shortest average distance . . . 15
- 2.7 Technical Information . . . 16

3 Results 17
- 3.1 Relabeling cytoplasm images with connected components . . . 17
- 3.2 Relabeling interleaved cytoplasm and overlap with connected components . . . 18
- 3.3 Neuronal network training approaches . . . 23
  - 3.3.1 Training of cytoplasm, overlap and membrane from EM images . . . 23
  - 3.3.2 Training of cytoplasm and overlap from predicted membrane . . . 25
  - 3.3.3 Training of overlap from predicted cytoplasm . . . 25
- 3.4 Different combinations to generate an interleaved cytoplasm/overlap image stack 26
- 3.5 Evaluation of neuronal network predictions . . . 27
  - 3.5.1 Reconstruction based on cytoplasm and overlap predicted from EM images 27

3.5.2 Reconstruction based on overlaps predicted from cytoplasm 34
3.5.3 Reconstruction based on cytoplasm and overlap predicted from membrane groundtruth 35
3.5.4 Reconstruction based on cytoplasm and overlap from predicted membrane 36
3.6 Proofreading the SNEMI3D dataset 38
3.7 Hyperparameter testing for membrane prediction 41

4 Discussion **45**
4.1 Evaluation of dense reconstruction 45
4.1.1 Metrics 45
4.1.2 Pre- and post-processing 46
4.1.3 Experiments 47
4.1.4 SNEMI3D annotations 48
4.1.5 NN architectures and hyperparameter tuning 49
4.2 Computational costs 49

5 Conclusion **51**

6 Future directions **53**

A Supplementary information **55**

B Poster **57**

C CD content **59**

Bibliography **61**

Glossary **67**

II. List of Figures

1.1 Overview of data acquisition methods 2
2.1 SNEMI3D dataset 7
2.2 SNEMI3D dataset preparation 8
2.3 Generation of cytoplasm images 9
2.4 Generation of membrane images 10
2.5 Generation of overlap images 11
2.6 Composition of the neuronal network architecture 12
2.7 Sparse matrix for calculating split and merge errors 15
2.8 Computation of the shortest average euclidean distance of two membrane labels . 16
3.1 Reconstruction of 3D objects from cytoplasm label only 18
3.2 Reconstruction of 3D objects from true cytoplasm and overlap labels 19
3.3 Split error induced by subtraction of membranes from thin objects 21
3.4 Split error already present in original dataset due to thin spine necks 22
3.5 Overview of different paths to train DCNN for automatic cytoplasm and overlap image generation 23
3.6 Training input, output and target of the cytoplasm training with EM input 24
3.7 Training input, output and target of the overlap training with EM input 24
3.8 Training input, output and target of the membrane training with EM input 24
3.9 Training input, output and target of the cytoplasm training with membrane input . . 25
3.10 Training input, output and target of the overlap training with membrane input 25
3.11 Training input, output and target of the overlap training with cytoplasm input 26
3.12 Experiment I and II visualized with neuroglancer 29
3.13 Original objects and experiment I visualized with neuroglancer 30
3.14 Evaluation results of experiment I after eroding overlapping regions in overlap images 31
3.15 Comparison of overlap images with erosion and skeletonization 32
3.16 Visual comparison between unchanged and eroded/skeletonized reconstruction of experiment I 33

3.17 Visual comparison between unchanged and eroded/skeletonized reconstruction of experiment III 34
3.18 Differences between EM and SNEMI3D derived membranes 39
3.19 Pixel-based assessment of membrane labels 39
3.20 Distance-based assessment of membrane labels 40
3.21 Comparison of SNEMI3D derived membranes and self annotated membranes . . . 41
3.22 Precision for different ResNet hyperparameter 42
3.23 Recall for different ResNet hyperparameter 43
3.24 Predicted membrane images of NNs with different loss functions 43
3.25 Comparison of generator loss during NN training for all three tested loss functions . 44
B.1 Poster - 4th Central German Meeting on Bioinformatics 57

III. List of Tables

2.1 Contingency table for membrane images 14
3.1 Evaluation metrics for relabeling 20
3.2 Overview of all used combinations to create a cytoplasm image stack with interleaved overlap images 26
3.3 Evaluation results for experiment I - III 29
3.4 Comparison of eroded objects and eroded objects with additionally skeletonization . 32
3.5 Evaluation results for experiment I - III with and without erosion and skeletonization . 33
3.6 Evaluation results for experiment IV - V 35
3.7 Evaluation results for experiment VI - VIII 36
3.8 Evaluation results for experiment IX - X 38
4.1 RAND scores of all competitors at ISBI with additional listing of experiment VIII . . . 48
4.2 Comparison of computational costs with FFN 50
A.1 Evaluation results for experiment I with altered overlap objects after training 55

IV. Nomenclature

ARI	Adjusted RAND index
ATUM-SEM	Automatic tape-collecting ultramicrotome scanning electron microscopy
CLAHE	Contrast limited adaptive histogram equalization
CNN	Convolutional neuronal network
DCNN	Deep convolutional neuronal network
FFN	Flood-filling network
FIB-SEM	Focused ion beam scanning electron microscopy
FOV	Field of view
GALA	Graph-based active learning of agglomeration
GFLOPS	Giga floating point operations per second
GPU	Graphics processing unit
GUI	Graphical user interface
ISBI	International Symposium on Biomedical Imaging
MAD	Mean absolute deviation
NN	Neuronal network
POM	Predicted object map
ReLU	Rectified linear unit
ResNet	Residual network
RGB	Red, green, blue additive color model
SBEM	Serial block-face scanning electron microscopy
SD	Standard deviation
SNEMI3D	3D Segmentation of neurites in electron microscopic images
ssTEM	Serial section transmission electron microscopy

V. Preface

Zunächst möchte ich Dr. rer. nat Torsten Bullmann meinen herzlichsten Dank aussprechen, der mir dieses Forschungsthema anvertraut und mir Ratschläge, Tipps und Denkanstöße in jeder vorliegenden Situation gegeben hat. Vielen Dank Torsten.

Außerdem möchte ich mich bei Prof. Dr. rer. nat. Dirk Labudde für die Einführungen und Grundlagen im Bereich Bioinformatik über den Zeitraum meines Bachelor- und Masterstudiums bedanken, die mir meinen zukünftigen Weg erleichtern werden.

Weiterer Dank gilt Christoph Leberecht, M.Sc., der mir im Bereich wissenschaftliches Arbeiten eine souveräne Grundlage geschaffen und mich mit dem Umgang mit Git und der Programmierung generell vertraut gemacht hat.

Darüber hinaus möchte ich auch einen Dank an die "*bioinformatics group Mittweida*" aussprechen. Mit Ihnen zu arbeiten, sich in Gruppenmeetings auszutauschen oder bei Grillabenden einfach mal abzuschalten, hat mir das Verfassen dieser Arbeit ungemein erleichtert.

Zum Schluss natürlich auch ein riesengroßes Dankeschön an meine Familie, die mich in dieser doch anstrengenden Zeit immer finanziell und mit guter Laune unterstützt hat.

1 Introduction

> We can build a much brighter future where humans are relieved of menial work using AI capabilities.
>
> *Andrew Ng*

1.1 Machine Learning in Biomedical Imaging

Deep Learning, a sub domain of machine learning, has surpassed several classic methods in speech recognition [Hinton et al., 2012], object detection [Szegedy et al., 2013] and other areas like genomics [Eraslan et al., 2019] and connectomics [Turaga et al., 2010] with ease and is currently a leading machine-learning tool in computer vision and imaging [Greenspan et al., 2016, LeCun et al., 2015]. The central aims in the research area of connectomics are dedicated to the mapping of neuronal structures and circuitry, which in contrast requires huge amounts of high resolution electron microscopic data [Lichtman et al., 2014]. Nonetheless, researchers hope that connectomics helps to solve important neurobiological questions like the physical correlate for memory traces, effects of age and environment on the brain as well as the pathogenesis of Alzheimer's disease.
In order to gain fast and precise 3D reconstructions of neurons, new approaches for automatic segmentation of objects from electron microscopy data by 2D convolutional neuronal networks (CNN) have been proposed [Arganda-Carreras et al., 2015]. Isolated structures like mitochondria are possible to trace through image stacks with a high z-distance that emerge through data acquisition via ATUM-SEM, ssTEM or SBEM [Urakubo et al., 2019]. However, neurons are separated through relatively thin membranes, which have to be predicted correctly for image segmentation. Therefore, many research groups designed their own custom segmentation and reconstruction pipelines of neurons, which mostly consists of membrane detection followed by watershed transformation and graph-based active learning of agglomeration (GALA) [Nunez-Iglesias et al., 2014] or connected components [Arbeláez et al., 2011, Zlateski and Seung, 2015]. Recently, 3D CNN such as flood-filling networks (FFN) by JANUSZEWSKI ET AL. achieved considerably higher accuracies than traditional 2D CNN methods [Januszewski et al., 2016]. FFNs utilise recurrent 3D CNNs to produce single neuron segments from raw image data. Unfortunately, the training of FFNs comes with high computational costs, which results in at least one week-long training with the use of several GPUs in parallel. Additionally, FFNs need huge subvolumes of annotated data. This annotation cannot be reused, because it is dependent on the imaging conditions as well as tissue and species.

1.2 Image acquisition by electron microscopy

The mapping of densely packed neuronal structures, such as dendrites, glial cells or synapses requires high resolution 3D imaging methods. In recent years, new image acquisition methods (for example ATUM-SEM) emerged and classical methods (TEM) were automated and scaled-up [Briggman and Bock, 2012]. Considering the resolution, almost all methods have one thing in common. They are able to produce high lateral resolutions, whereas the resolution of the z-direction, the spacing between two images, remains relatively low. Therefore, high anisotropic datasets like the SNEMI3D dataset [Arganda-Carreras et al., 2013] with a resolution of 3x3x29 nm are relatively difficult to annotate. Classic approaches for image segmentation include free annotation software like TrakEM2 [Cardona et al., 2012] and VAST-Lite [Berger et al., 2018] for outlining traces as well as Knossos [Helmstaedter et al., 2011] for centerline tracing. They facilitate the annotation by, e.g. multi-user access and the ability to quickly browse through EM data, but still require neuroanatomy experts. An overview of four current data acquisition methods, ssTEM [Harris et al., 2006], ATUM-SEM [Hayworth et al., 2006], SBEM [Denk and Horstmann, 2004] and FIB-SEM [Knott et al., 2008] with their corresponding resolution in x-,y- and z-axis is visualized in Figure 1.1. Whereas SEM uses backscattered electrons and can be used on serial sections as well as for blockface imaging, TEM uses transmitted electrons and yields higher resolution and contrast.

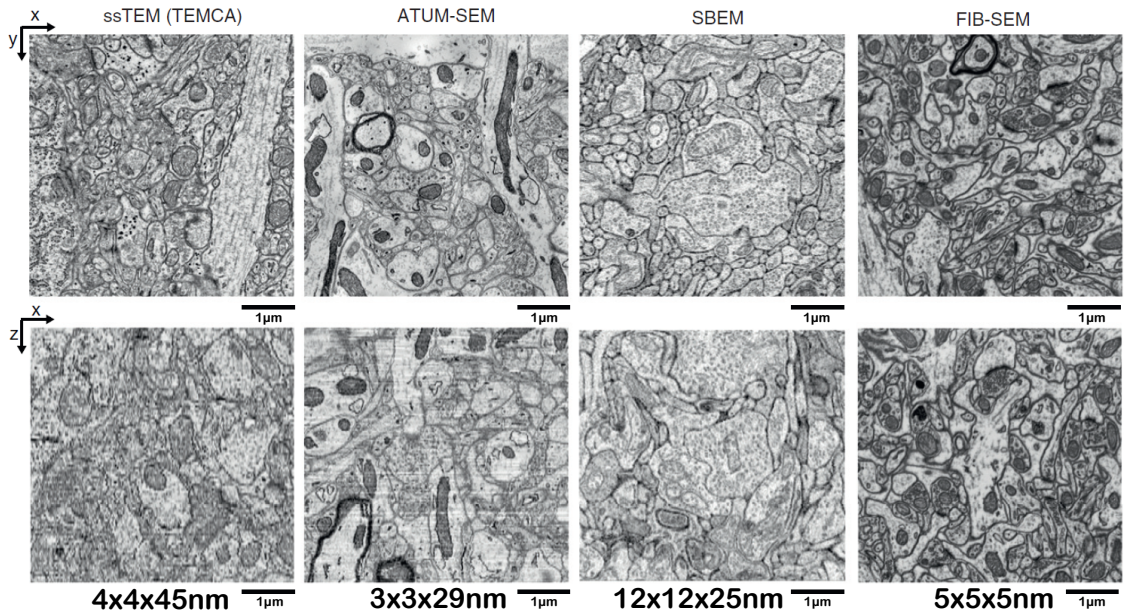


Figure 1.1: Overview of data acquisition methods (modified after [Briggman and Bock, 2012]). ssTEM illustrates mouse visual cortex, whereas SBEM shows the inner plexiform later of mouse retina. Both ATUM-SEM and FIB-SEM visualizes mouse cortex.

The SNEMI3D dataset used in this work was created with a ATUM-SEM and released in the context of the IEEE International Symposium on Biomedical Imaging [Hayworth et al., 2006]. The automatic tape-collecting ultramicrotome consists of a tape and con-

veyor belt, which retrieves thin brain sections cut by a diamond knife in a water basin. These thin brain sections were subsequently imaged by a scanning electron microscope, creating sufficient resolutions for detecting individual synaptic vesicles. It is also possible to generate a multi-scale dataset by re-imaging several sections at multiple resolutions. This approach provides additional context of surrounding regions [Kasthuri et al., 2015].

1.3 ISBI 2012/2013 Challenge - Image segmentation for connectomics

The first international challenge and approach to introduce many researcher and research groups to image segmentation for automatic reconstruction of neural circuits was originated by the IEEE International Symposium on Biomedical Imaging [Arganda-Carreras et al., 2012]. The importance of automatic segmentation is emphasized in a study from HELMSTAEDTER ET AL. in 2011. His team reconstructed 1000 neurons from a mouse retina spanning 20,000 h of human labor, which is barely enough to cover the smallest types of retinal neurons. Even this work used semi-automatic methods to reduce the required time tenfold and more [Arganda-Carreras et al., 2015]. With the development of faster data acquisition, the need of automatic methods is inevitable.
In the ISBI challenge, from over 100 registered groups, 17 different researchers submitted their results, of which 13 are ranked and presented in the article of ARGANDA-CARRERAS ET AL. in 2015 [Arganda-Carreras et al., 2015]. The training set included 30 consecutive images with 512x512 pixels and a resolution of 4x4x50 µm retrieved from a ssTEM of the *Drosophila* first instar larva ventral nerve cord [Cardona et al., 2010]. Most of the participants used deep convolutional network architectures for boundary detection [Ciresan et al., 2012]. In this case, boundary maps are in general equal to later introduced cytoplasm images used in the developed workflow. Neuronal objects were labeled as true, whereas all other structures received a false label, which are later used for segmentation with connected components.
One year later, a second challenge in the context of the ISBI 2013 was released. This challenge was specialised in 3D segmentation of neuronal structures in EM images, whereas the first introduced challenge 2012 was dedicated to 2D image segmentation. The corresponding website of the SNEMI3D challenge includes a leaderboard [Arganda-Carreras et al., 2013], but unfortunately there is no detailed information about used methods for each participant provided.

1.4 Flood-filling networks for single neuron reconstruction

First introduced in 2016, JANUSZWESKI ET AL. presented a new recurrent 3D CNN architecture for neuronal segmentation called flood-filling network, which placed first in the segmentation challenge of FIB-25 [Takemura et al., 2015] and second in the ISBI 2013 challenge [Arganda-Carreras et al., 2013].
The groundtruth dataset used in this work relied on the manual segmentation of 0.02% of a 96x98x114 µm subvolume of the zebra finch brain imaged with a serial block-face scanning electron microscope. These 131 million annotated voxels were used in 33 subvolumes of varying size for neuronal network training. The basic architecture of this CNN combines two input channels for training. The first one is the original 3D EM subvolume, whereas the second is a predicted object map (POM). This POM yields a probability for each pixel corresponding to the object or not. During training, several seed points were set to segment different objects. After each iteration, the networks weights are adjusted using per-voxel cross-entropy loss [Goodfellow et al., 2017, Januszewski et al., 2018].
Unfortunately, flood-filling networks bear some downsides, which are related with the training and prediction of objects. The training takes around one to two weeks with a custom server setup of 32 Nvidia K40 GPUs. For a single GPU computer, this training would last over two months. Second, objects can only be reconstructed one after another, which can take several minutes per object.
Nevertheless, flood-filling networks are considered as state-of-the-art algorithm for 3D image segmentation.

1.5 Pix2Pix with conditional adversarial networks

Many problems in computer vision and image processing include the translation of one image into a corresponding image [Isola et al., 2017]. Specialized CNNs are heavily used in image processing. Unfortunately, it is important to choose the best loss function for each present problem to achieve the desired results.
Therefore, in recent years, conditional generative adversarial networks (GANs) were heavily analysed and described [Denton et al., 2015, Goodfellow et al., 2014]. In general, GANs consists of two components: a generator and a discriminator. In this case the generator receives an input image and creates an output image, which is compared to an existing target image. After that, a loss function calculates the difference between prediction and target with a prior defined loss function and updates the generators weights. The discriminator receives the output image of the generator and guesses, whether this image matches target images or not. The underlying network architecture of a generator includes a typically encoder-decoder or U-Net. Each layer consists of convolution-BatchNorm-ReLu modules [Ronneberger et al., 2015], whereas the discriminator uses a convolutional *PatchGAN*. Conditional GAN implies the change of a

per-pixel classification to a structured loss, which penalizes joint configuration of the output [Isola et al., 2017].

1.6 Goal of this thesis

The major goal of this thesis was to combine machine learning and deterministic algorithms for a dense reconstruction of three dimensional objects such as neurons from a series of EM images. In particular, my workflow combined established 2D CNNs suited for image prediction with the classic connected components algorithm. Cytoplasm images represent cross-sections of putative three-dimensional objects in the image stack, whereas overlap images describe the individual connectivity of these cross-sections. Therefore an intermediate representation consisting of cross-sections and corresponding features for the final object prediction was used. There are several paths to go from EM images to mentioned features (see Figure 3.5).
Another goal was the development of several Python scripts that allow the user via command line to pre-process the groundtruth dataset, train neuronal networks, predict images from EM data, enable dense reconstruction of three dimensional objects with connected components and their interactive visualization with neuroglancer. These scripts together with instructions for installation and usage are provided as a GitHub repository at `https://github.com/mweber95/neuron3d`.

2 Methods

The following chapter is an extended version of the description found at the GitHub repository `https://github.com/mweber95/neuron3d`, which takes up image pre-processing in form of dataset preparation and membrane, cytoplasm and overlap image creation as well as used network architecture and evaluation metrics.

2.1 Dataset

The SNEMI3D dataset was released in the context of the IEEE International Symposium on Biomedical Imaging (San Francisco, CA, April 7-11th, 2013) [Arganda-Carreras et al., 2013]. It served as the basis for an open challenge concerned with the automatic segmentation of neurons and following 3D visualization. The data was obtained using a ATUM-SEM, producing consecutive image slices with a spacing of 29 nm and a resolution of 3 nm in x- and y-direction, generating highly anisotropic images.
The original SNEMI3D includes three different tif files with a pixel size of 1024x1024 and 100 consecutive images (cubic volume of around 3x3x3 µm), named *train-input*, *train-labels* and *test-input* illustrated in Figure 2.1. *Train-input* is representing original EM images, whereas *train-labels* is the associated labeled dataset. The *test-input* contains only the image data for the validation dataset for the challenge. In this work, only EM images and annotated object labels of the original training set were used. The annotation consist of exactly 400 manually annotated objects, spanning a variety of neuronal structures like axons, glial cells or dendrites. A complete description of the dataset together with the main biological findings and further information of the data acquisition was published 2015 by KASTHURI ET AL. [Kasthuri et al., 2015].

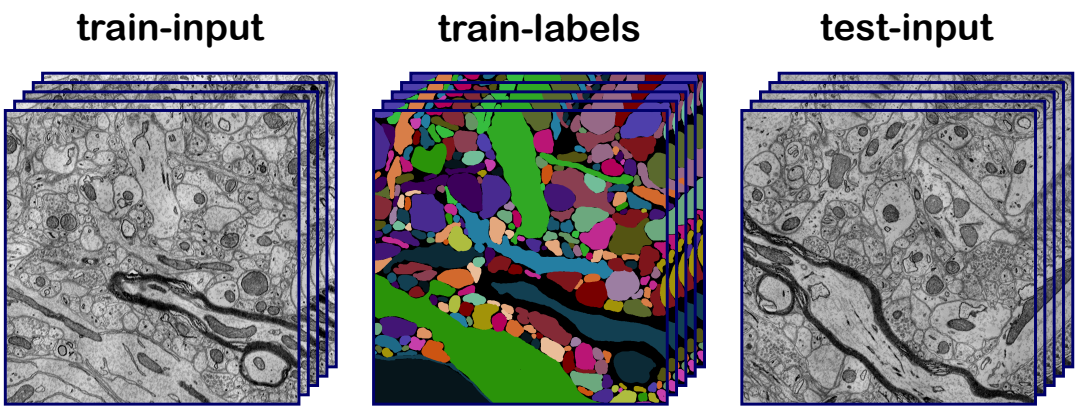


Figure 2.1: SNEMI3D dataset.

2.2 Dataset preparation

My workflow involves the training of multiple NN for predicting several specific images structures, such as membrane, cytoplasm and overlapping cross-sections of consecutive images. To increase the data basis available for training and allow independent training, prediction and validation, the *train-input* and *train-labels* image stack from the SNEMI3D dataset was split in four parts with equal size (512x512 pixels), which are enumerated from A to D (see Figure 2.2). The numeration of parts A-D was chosen by objects represented in the dataset. Part A contains a variety of differing objects, such as myelinated axons, dendrites and glial cells, which are mandatory for training and following prediction of these objects. In contrast, part D was only used once for prediction because of myelinated axons spanning the complete size of images were restricted to this sub volume and lead to false predictions when included in the training.

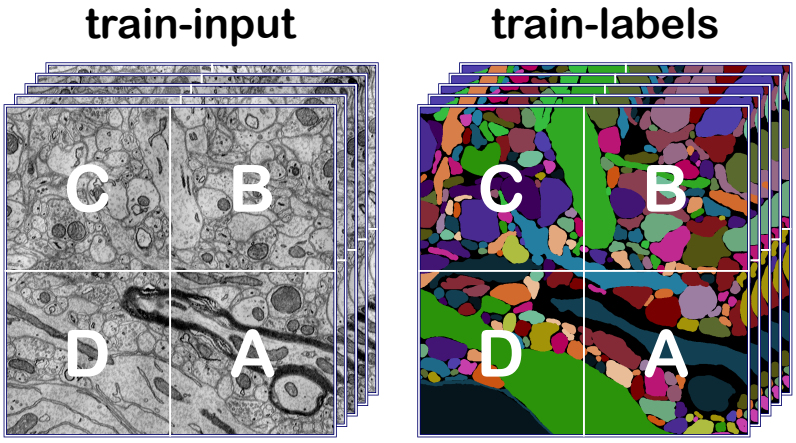


Figure 2.2: SNEMI3D dataset preparation.

2.3 Generation of cytoplasm, overlap and membrane images

The generation of cytoplasm, overlap and membrane labels are necessary for the subsequent training of 2D CNN in which the neuronal network translates raw images into target images. These three different labels were extracted from the original object index stack *train-labels* of the SNEMI3D challenge.

2.3.1 Cytoplasm images

Cytoplasm images represent the interior region of each individual object in the *train-labels* image stack without its boundary (membrane). The original *train-labels* image stack consists of long integers indexing each individual object. These integers were transformed to a binary image with white (true) pixels outlining cytoplasm and black (false) pixels for the remainder. The procedure is illustrated in Figure 2.3.

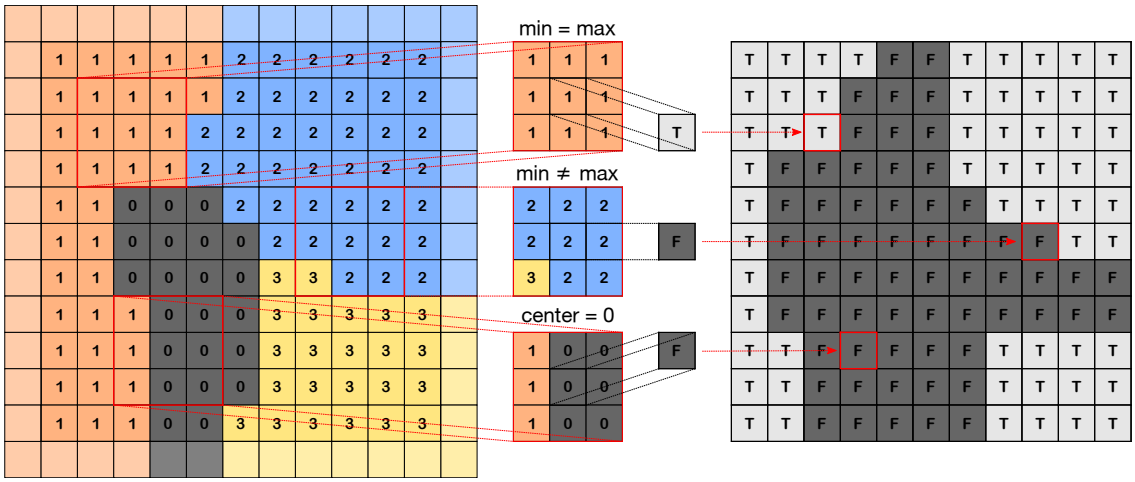


Figure 2.3: Generation of cytoplasm images.

First, the membrane was defined as one pixel wide boundary of each object. Furthermore, a 3x3 kernel traversing each image from top left to bottom right, was applied. This kernel checks, whether the minimum and maximum value of the kernel are equal or unequal. In the case of minimum equals maximum, the center pixel has the same index as all surrounding pixels and represents the interior of the object and therefore cytoplasm. If the values are not equal, the center pixel is next to another object or the background. This case does not represent the interior and is therefore set to not cytoplasm. If the center pixel has the index value zero it represents the background with no object and remains zero.

2.3.2 Membrane images

Membrane images were defined as the one pixel wide boundary of all individual labeled objects in the *train-labels* image stack. In later sections, they serve as intermediate step for the training of cytoplasm and overlap images.
The calculation of membrane images resembles the cytoplasm procedure. Nevertheless, there are some important differences between the cytoplasm and membrane image generation. If minimum and maximum value in the 3x3 kernel are the same, the corresponding center pixel is set to false. In contrast, if minimum and maximum are unequal, the center pixel is at the boundary of at least two objects including the background, thus representing a membrane label and is set to true. Center pixels that equal zero result in a false label. The membrane generation is illustrated in detail in Figure 2.4.

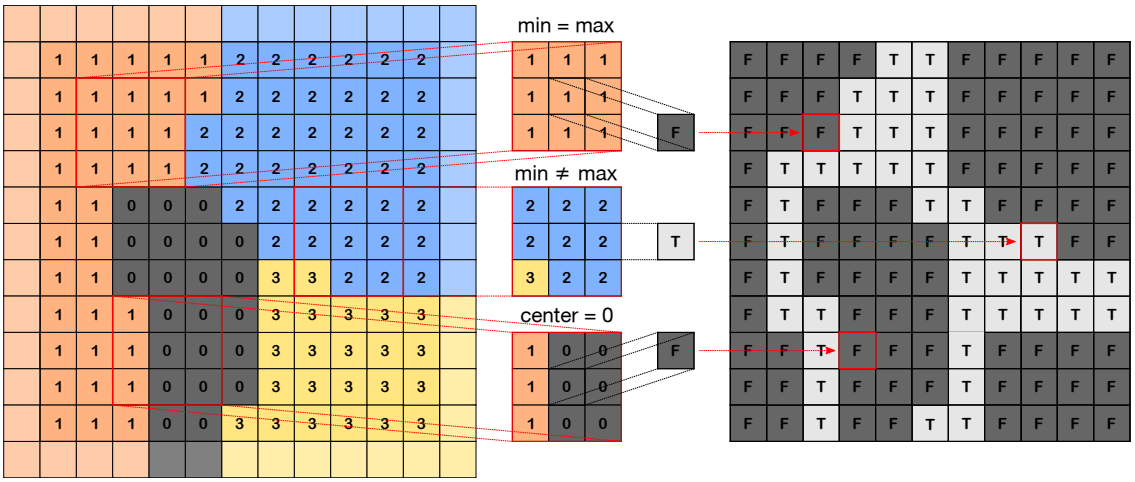


Figure 2.4: Generation of membrane images.

2.3.3 Overlap images

In comparison to the calculation of cytoplasm and membrane images, the generation of overlap images requires two consecutive images from the object index stack. All membrane pixels, the one pixel wide boundaries of each object were changed to a false label. This calculation resembles the procedure of generating cytoplasm images, with the exception that all non membrane and non background pixels retained their corresponding label. After the subtraction of membranes from the original labels, the superimposed pixels in both images were compared. The pixel was set to true, if both images possess the same integer value and both values are greater than zero (thus not representing the inter cellular space). The procedure is illustrated in Figure 2.5. This calculation reduces two subsequent images to one image describing the overlap between two cross-sections of each individual object spanning both images, which is the crucial information for the relabeling of neuronal objects after NN training.

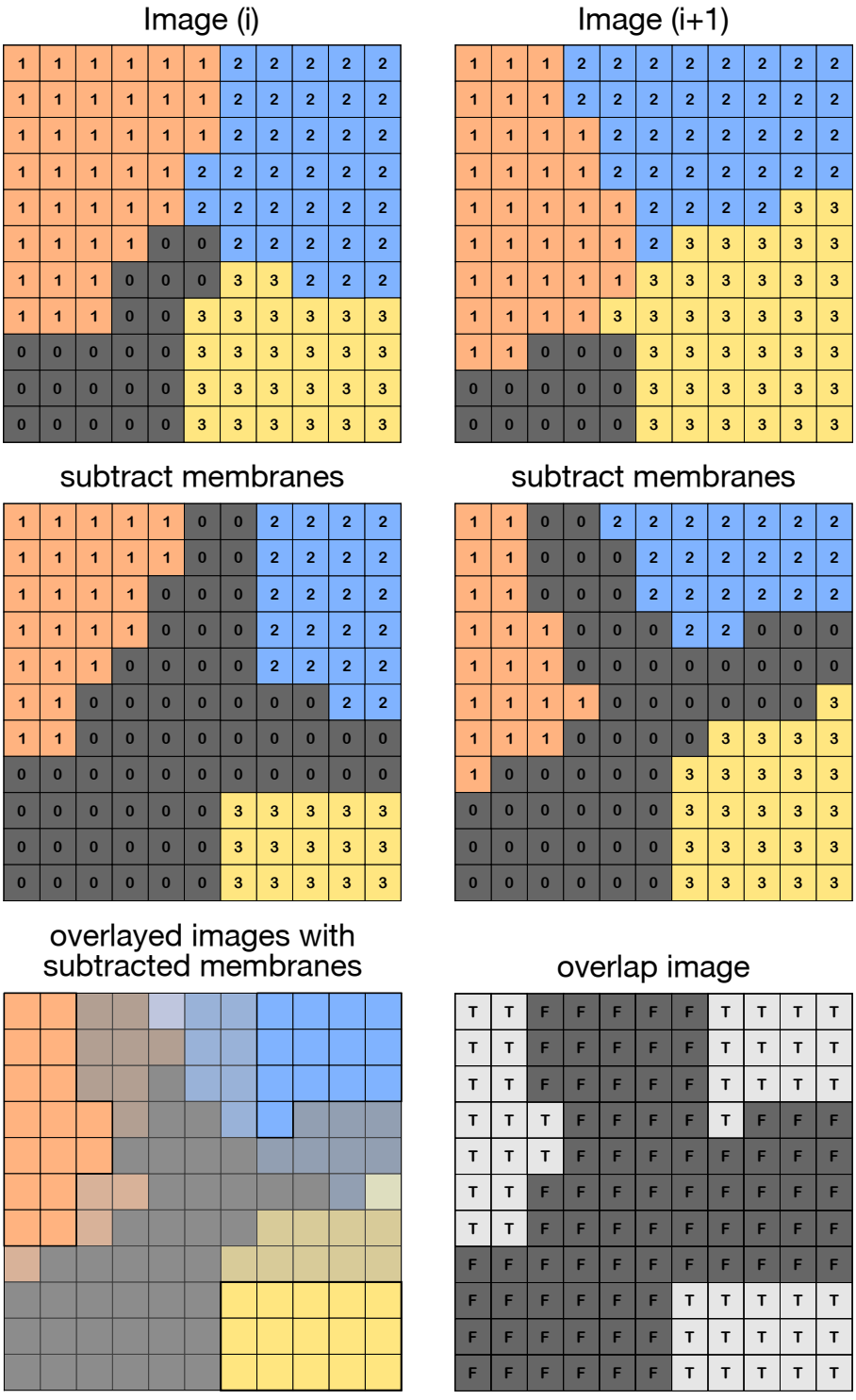


Figure 2.5: Generation of overlap images.

2.4 Training of neuronal networks

After the generation of cytoplasm, membrane and overlap images for all parts of the SNEMI3D dataset, the training of NN can be started. Therefore, the input images must be paired with the target images, corresponding to the images which should be learned to predict from the NN. For example, the training of a NN, which should predict membranes from EM images needs EM images as input and membrane images of the same section as target. EM images are available through the *train-input* dataset and membrane images were prior preprocessed and extracted from the object index (*train-labels*) image stack. A special case is dedicated to the training of overlap images. The input needs to provide two consecutive images (stored in the red and green color channel of a RGB image), because overlap images are representing the overlapping parts of two cross-sections.
Throughout the whole work, a residual network with prior encoder and subsequent decoder was used [Isola et al., 2017, Goodfellow et al., 2014]. An overview and abstract illustration of this structure is presented in Figure 2.6.

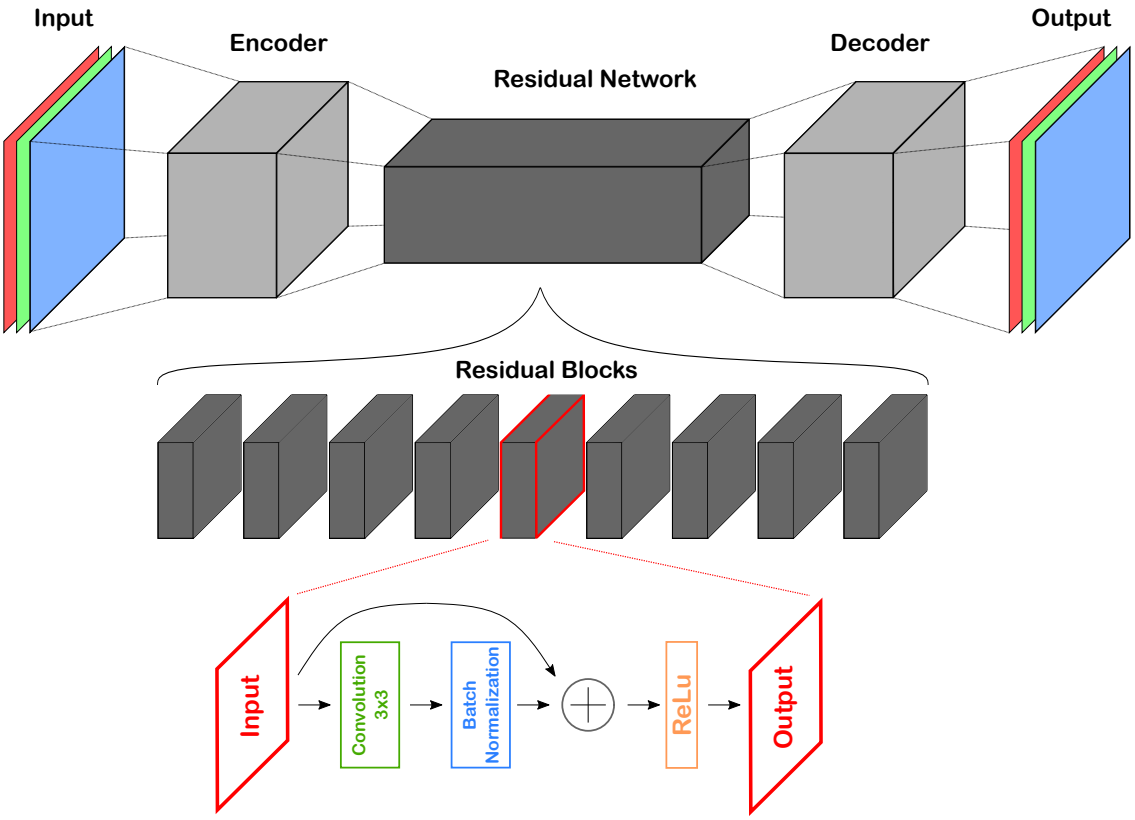


Figure 2.6: Composition of the neuronal network architecture.

The encoder was used to achieve a compressed representation of the input data and therefore a dimension reduction. Afterwards, the residual network, consisting of nine residual blocks, was used. One residual block is made up of a 3x3 convolution followed by a batch normalization. After this step, the normalized values and the original input

of the residual block are added up and further passed to a rectified linear unit function. This representation forms the output and the input of the subsequent residual block. In the end, the last output of the ninth residual block is decoded in the same manner as the encoder works to rebuild the output image.
Hyperparameters were set to 2000 epochs with a batch size of 100. The images in the batch were further augmented to avoid overfitting. The augmentation includes horizontal and vertical flipping and transposing. Additionally, 256x256 pixel random cutouts of the image were utilised and the GAN loss was set to zero to avoid blurry predictions.

2.5 Prediction, reconstruction and 3D visualization

After the training completed, the current state of the network is saved. With this checkpoint, the NN can predict outputs, if new input images are given.
The reconstruction of neuronal objects in the SNEMI3D dataset needs cytoplasm as well as overlap images. Furthermore, the reconstruction utilizes a pixel grouping step called connected components, that can be described as an algorithm, in which all vertices of a sub graph of an undirected graph are connected and labeled individually [Hopcroft and Tarjan, 1973]. The combination of cytoplasm and overlap images, that are arranged to an interleaved stack with overlap images between corresponding cytoplasm slices allows reliable reconstruction (see section 3.2).
The final 3D visualization was performed with neuroglancer. This WebGL-based viewer of volumetric data is able to visualize cross-sectional views of volumetric data, as well as 3-D meshes and line-segment based models, which also offers several visualization examples with python (see `https://github.com/google/neuroglancer`).

2.6 Evaluation

For the evaluation of predicted images and reconstruction of neuronal objects, precision and recall as well as adjusted RAND index, split and merge errors and the shortest average euclidean distance (for membrane labels) were chosen and further described in following subsections.

2.6.1 Precision and recall

Precision and recall served as two metrics to evaluate 2D predictions of the neuronal network. Both are statistical measurements of a binary classification, which is given in cytoplasm, membrane and overlap image due to true and false labels (see Table 2.1 for membrane images).
The precision, also called positive predicted value, describes the ratio of correct predicted true (membrane, cytoplasm, overlap) pixels in contrast to all true pixels in the im-

age. Furthermore, the recall, also called sensitivity or true positive rate, demonstrates the ratio of correct predicted false (background) pixels to all false pixels in the image. Both equations for the calculation of precision and recall are additionally illustrated in equation 2.1 and 2.2.

Table 2.1: Contingency table for membrane images.

	True membrane	**True background**
Predicted membrane	True positive (TP)	False positive (FP)
Predicted background	False negative (FN)	True negative (TN)

$$\text{Precision} = \frac{\text{TP}}{\text{FP + TP}} \tag{2.1}$$

$$\text{Recall} = \frac{\text{TP}}{\text{FN + TP}} \tag{2.2}$$

2.6.2 Adjusted RAND index

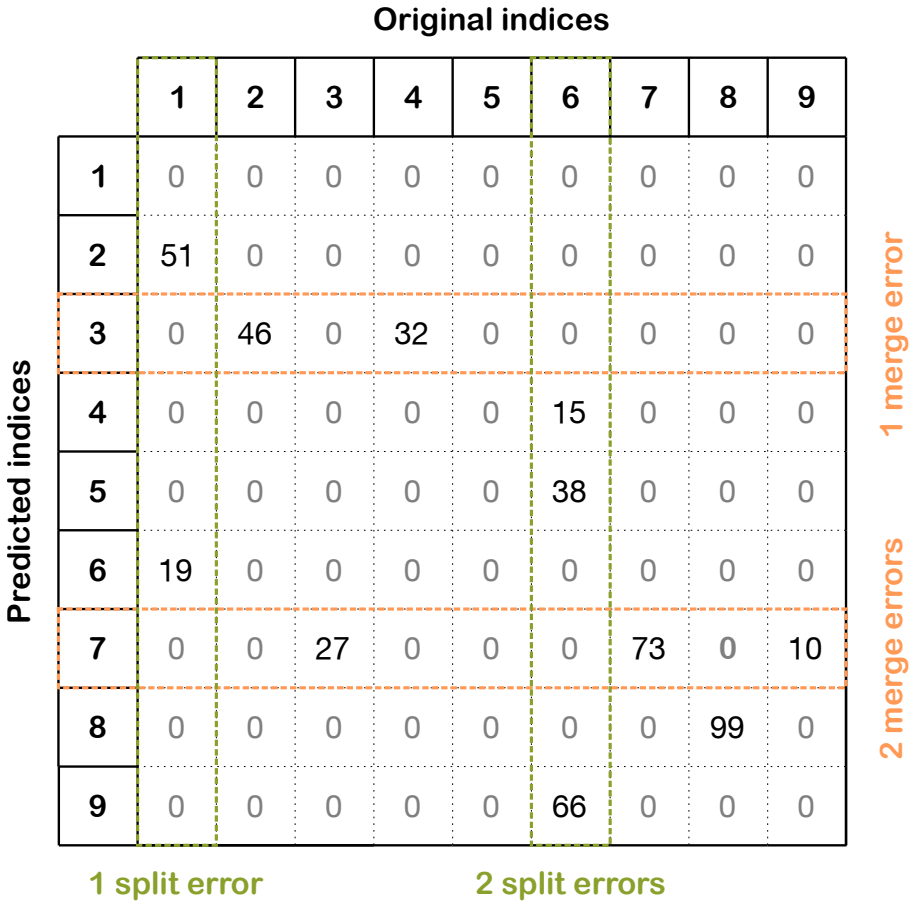
The RAND index is often used for the comparison of cluster or segmentation results [Rand, 1971]. In this work, the adjusted RAND index (corrected-for-chance version of the RAND index) was directly used from the scikit-learn library [Pedregosa et al., 2011], which is represented in equation 2.3 [Hubert and Arabie, 1985].

$$\text{adjusted RAND index} = \frac{\text{index - expected\_index}}{\text{max\_index - expected\_index}} \tag{2.3}$$

2.6.3 Split and merge errors

Split and merge errors are suitable evaluation metrics, which give an overview of the performance of a conducted segmentation. Unfortunately, there are multiple definitions of how split and merge errors are calculated. In my implementation, split errors represent the number of how often a original label is split into different predicted labels. In contrast, merge errors describe the amount of how often original labels are merged together into a single predicted label. For this, the overlap between true and predicted labels was represented as a matrix. Each element contains the number of pairs between a true and predicted label. An example is presented in Figure 2.7. The matrix representation gives the opportunity to simply count all rows or columns with more than one entry and add the resulting number of merge and split errors over all rows and

columns together. Because most elements of this matrix are zero (no overlap between most objects in the true and predicted labels), the implementation used a sparse matrix from the Scipy package [Jones et al., 2001].



	1	2	3	4	5	6	7	8	9
1	0	0	0	0	0	0	0	0	0
2	51	0	0	0	0	0	0	0	0
3	0	46	0	32	0	0	0	0	0
4	0	0	0	0	0	15	0	0	0
5	0	0	0	0	0	38	0	0	0
6	19	0	0	0	0	0	0	0	0
7	0	0	27	0	0	0	73	0	10
8	0	0	0	0	0	0	0	99	0
9	0	0	0	0	0	66	0	0	0

Figure 2.7: Sparse matrix for calculating split and merge errors.

2.6.4 Shortest average distance

For the comparison of self annotated and SNEMI3D derived membranes (see section 3.6), the shortest average euclidean distance for each image was measured. The computation is based on the function *cdist* of the Scipy package [Jones et al., 2001], which is able to calculate different distance metrics between two arrays. Therefore, the shortest euclidean distance from each self annotated membrane pixel to SNEMI3D derived membrane pixel was calculated, which is illustrated in Figure 2.8.

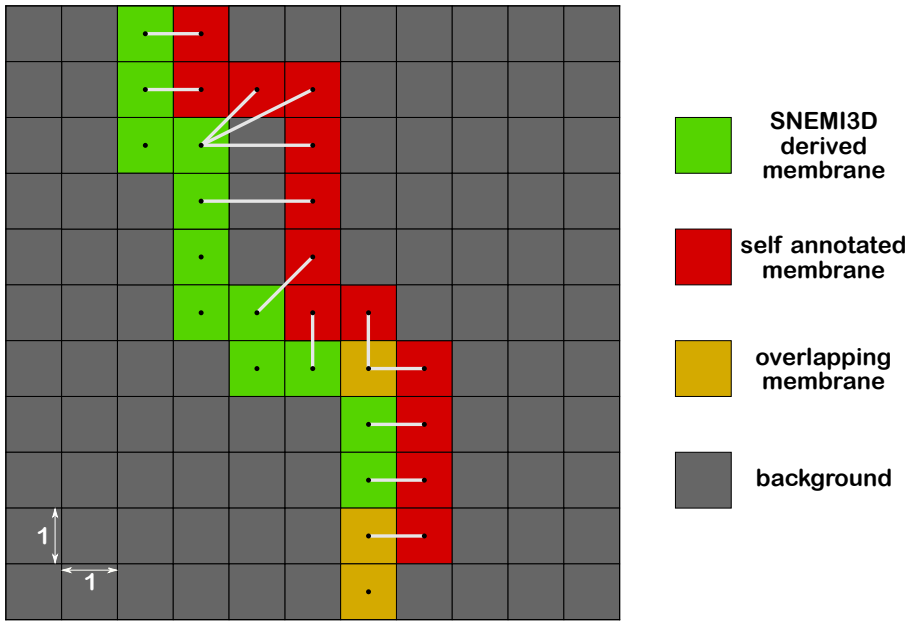


Figure 2.8: Computation of the shortest average euclidean distance of two membrane labels.

2.7 Technical Information

The training of the DCNN was carried out on a single Nvidia Tesla V100-SXM2-32Gb GPU. The whole server contains 32x Intel(R) Xeon(R) Gold 6132 CPU @ 2.60 GHz and 4x Nvidia Tesla V100-SXM2-32Gb and 126GB RAM. All other used scripts and computational operations, such as image pre-processing, training preparations and prediction of images were performed on a Intel(R) Core(TM) i7-6700 CPU @ 3.40 GHz with 16GB RAM.

3 Results

In this chapter, the results from initial testing with the relabeling algorithm "connected components" and 3D visualization via neuroglancer up to deep neuronal network algorithms are presented with the purpose of a fast and precise workflow for automatic segmentation and visualization of neurons from electron microscopic images.

3.1 Relabeling cytoplasm images with connected components

The first step was to find out which 2D Labels are necessary for reconstruction of 3D objects from a series of electron microscopic images. 2D labels provide spatial information about the cytoplasm of neurons, membranes or overlapping regions of two consecutive images. These labels served as input for the reconstruction of 3D objects with connected components and are predicted by machine learning algorithms.
A naive approach was to use cytoplasm label only. Therefore, an already existing and labeled image stack (SNEMI3D dataset) of mouse neocortex (see Figure 3.1 A left image) published by KASTHURI ET AL. [Kasthuri et al., 2015] consisting of one hundred consecutive images served as groundtruth. The cytoplasm images (see Figure 3.1 A middle image) derived from the object labels were then used to relabel the stack in order to test whether this information is sufficient to recreate the original segmentation by KASTHURI. The 3D visualization was realized with an open-source WebGL-based viewer of volumetric data called neuroglancer. Both image stacks (annotated and relabeled stack) were visualized and only separable objects were selected. The 3D visualization (see Figure 3.1 B-D) shows three nearly identical structures in both image stacks. Current state-of-the-art data acquisition methods generate highly anisotropic images with a high resolution in x- and y-direction (3 nm/pixel) and a low resolution in z-direction (29 nm/pixel). The green and red objects representing myelinated axons are only distinguishable through the relatively thick myelin that separates these neurons from surrounding objects. As shown in the right image of Figure 3.1 A, the cytoplasm of the remaining neurons are only separated by thin membranes. Therefore, the cytoplasm in consecutive image sections frequently overlaps with other neurons, which leads to a big connected structure illustrated in Figure 3.1 A-right image shown in blue.

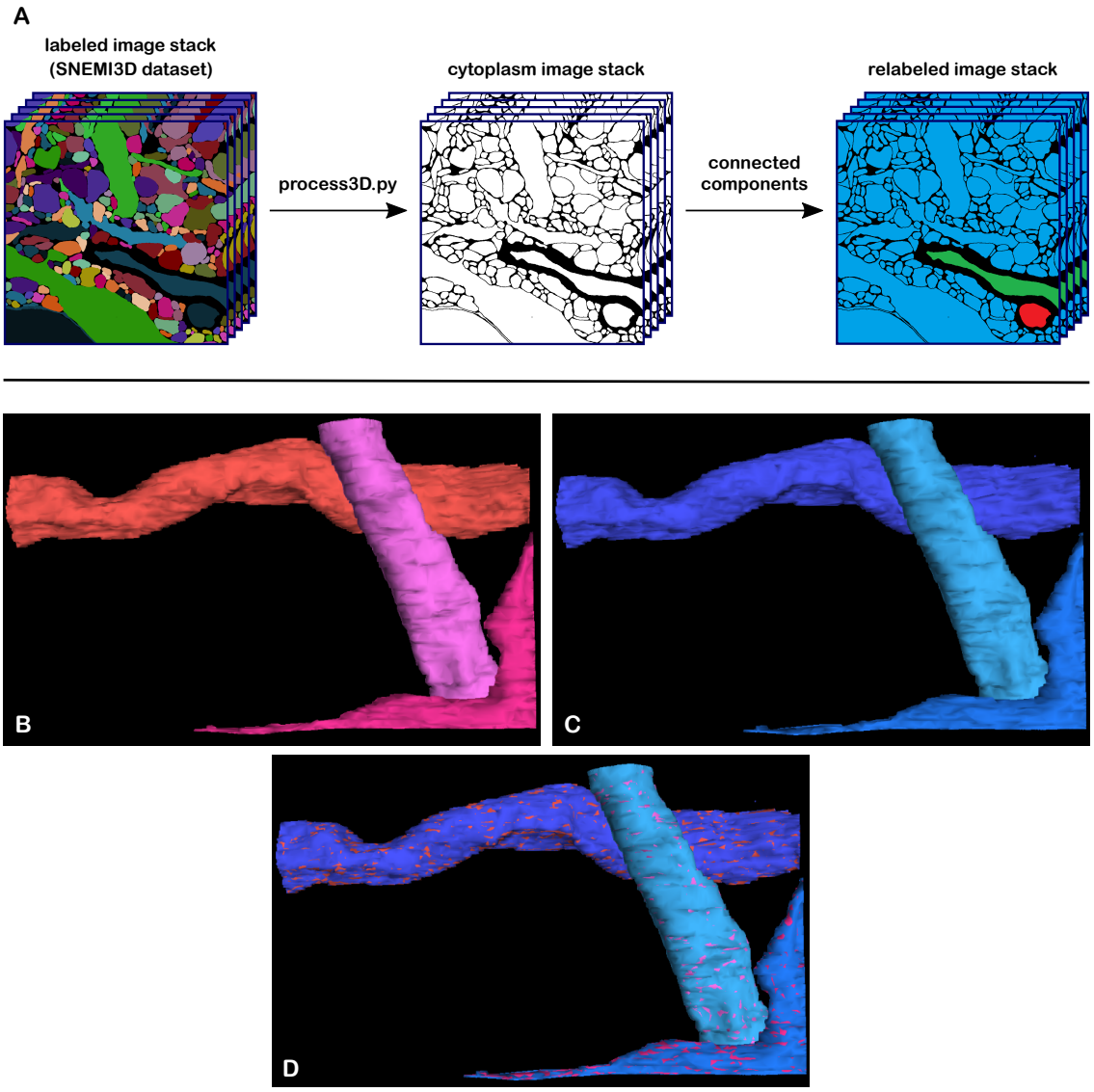


Figure 3.1: Reconstruction of 3D objects from cytoplasm label only. Labeled image stack of the SNEMI3D dataset (**A**-left) was processed with a custom script to extract only cytoplasm as true label and membranes, myelin and intercellular spaces as false label (**A**-middle). 3D objects were reconstructed from consecutive cytoplasm images with connected components (**A**-right). (**B**) Visualization of three distinguishable objects derived from cytoplasm relabeling. (**C**) Visualization of the ground-truth labeled image stack. (**D**) Reconstructed overlap with the original objects.

3.2 Relabeling interleaved cytoplasm and overlap with connected components

A much more sophisticated approach for the 3D reconstruction would include both cytoplasm label as well as the information about the overlap of cross-sections of objects in consecutive images (see Figure 3.2 A-middle-left). If the composition of the cytoplasm stack and overlap stack is changed to an interleaved image stack, in which cytoplasm and overlap images are alternately adjusted (overlap image between corresponding cy-

toplasm images), the segmentation of neurons resemble the already labeled SNEMI3D dataset (see Figure 3.2 A-middle-right and A-right).

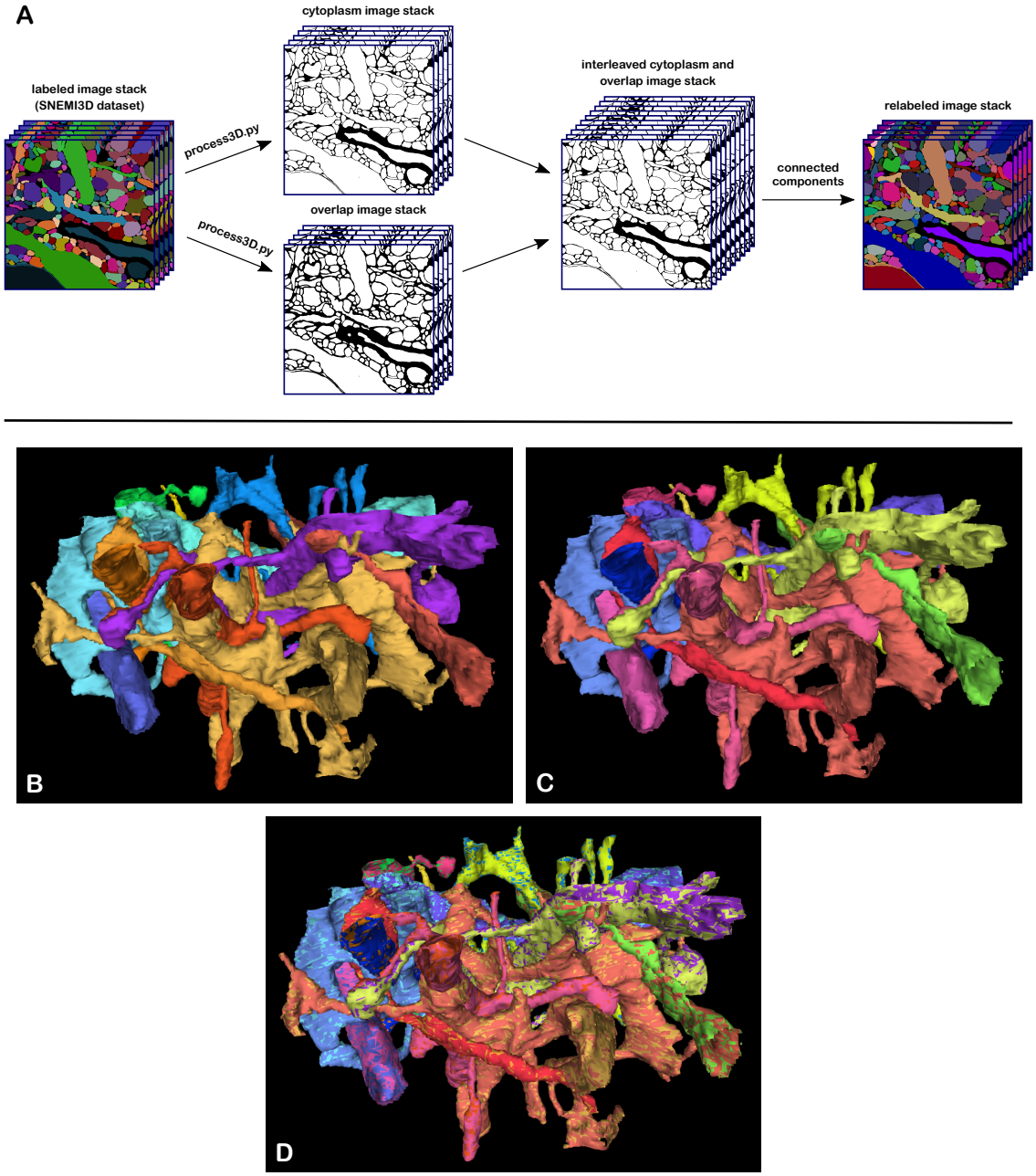


Figure 3.2: Reconstruction of 3D objects from true cytoplasm and overlap labels. Labeled image stack of the SNEMI3D dataset (**A**-left) was processed with a custom script to extract cytoplasm and overlap features. 3D objects resulting from interleaved cytoplasm and overlap images were then relabeled with connected components through the image stack (**A**-right). (**B**) Visualization of several neurons derived from interleaved cytoplasm with overlap images relabeling. (**C**) Visualization of the ground-truth labeled image stack by KASTHURI [Kasthuri et al., 2015]. (**D**) Reconstructed overlap with the original objects.

The dense reconstruction and segmentation of neurons from the interleaved cytoplasm with overlap image stack is superior in contrast to the relabeling of the cytoplasm stack.

The segmentation in Figure 3.2 (A-right) matches the original annotated data on the left nearly perfect. The 3D visualization of the originally annotated and interleaved image stack as well as the comparison of both stacks one above the other is illustrated in Figure 3.2 (D). Small differences are recognizable, which arise due to different annotation and processing, respectively. The original data consist of neuron labels for cytoplasm and membrane, whereas the processing and extracting of features only considered the cytoplasm without membrane. Therefore, the membranes of the SNEMI3D dataset were subtracted in order to receive better comparable evaluation values in further used methods.

To show the difference between the comparison of SNEMI3D data with and without subtracted membrane labels and relabeled data, precision, recall and adjusted RAND index (ARI) as well as split and merge errors were calculated as listed in Table 3.2. Furthermore, the original SNEMI3D dataset consisting of one hundred 1024x1024 pixel electron microscopic and labeled images had to be split in four 512x512 pixel parts each, enumerated as A, B, C and D (see section 2.2 and Table 3.2 Part: A-D) to have independent datasets for neuronal network training, testing and validation.

Table 3.1: Evaluation metrics for both relabeled image stacks consisting of only cytoplasm and interleaved cytoplasm with overlap images generated from the groundtruth with and without subtracted membranes as well as the results for all four split parts.

	adj. RAND Index	**Precision**	**Recall**	**Merges**	**Splits**
Relabeled cytoplasm images	0.0379	1.0	0.925	396	0
Relabeled cytoplasm and overlap images	0.717	1.0	0.925	0	8
Relabeled cytoplasm and overlap images with subtracted membranes from object labels in the groundtruth data	0.999	1.0	1.0	0	8
Part A	0.998	1.0	1.0	0	13
Part B	0.995	1.0	1.0	0	17
Part C	0.998	1.0	1.0	0	21
Part D	0.999	1.0	1.0	0	22

Table 3.2 represents the different evaluation metrics for three different approaches plus the split parts of the original data. For the relabeling of only cytoplasm images and of the interleaved cytoplasm and overlap image stack, ARI and recall are expected to be lower than 1.0 due to the subtraction of the membrane in cytoplasm labels. There-

fore a considerable high number of black pixels (false) exists that are originally white (true) in the groundtruth. In contrast to the relabel process of only cytoplasm images, the interleaved image stack consisting of cytoplasm and overlap images reaches zero merge errors. This indicates, that no objects are joined together and the addition of corresponding overlap images between two cytoplasm images offer a sufficient method to remove merge errors. However, this method still produces eight split errors. These errors show, that single objects in the groundtruth are separated in two or multiple objects in the relabeled image stack. For better comparison of the evaluation metrics in later sections of this thesis, membranes from the groundtruth data were subtracted and the metrics shown in Table 3.2 (third row) are chosen for the best possible outcome if the whole 1024x1024 pixel image stack is used. Part A, B, C and D serve as groundtruth for the in later sections described machine learning approach. A closer view reveals that split errors arise due to two different cases. The first case is illustrated in Figure 3.3. Through membrane subtraction from the groundtruth data, thin labels with at most three pixels in width are split into multiple single pixels, each with a different label (see Figure 3.3 B and C). Seven out of eight split errors listed in Table 3.2 (third row) belong to this category.

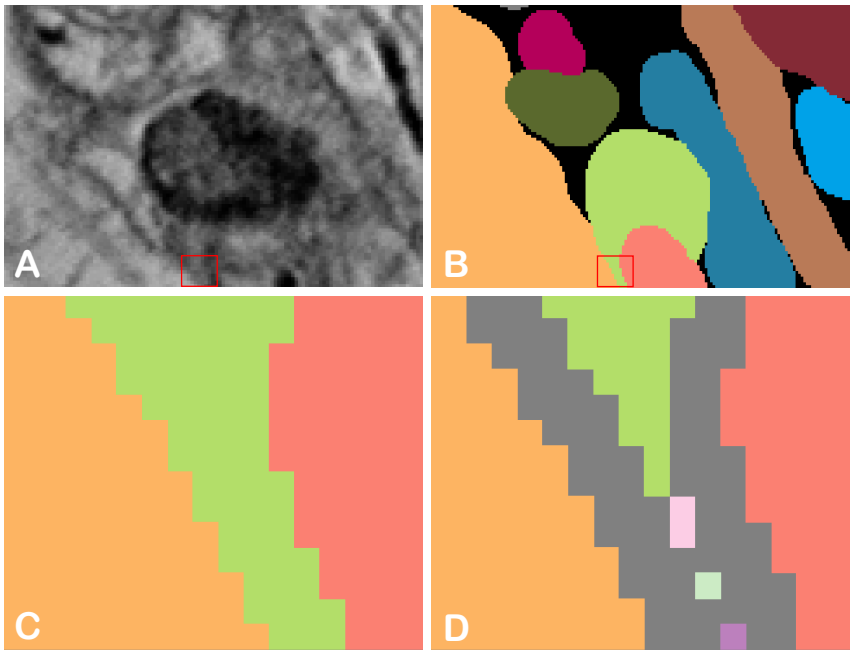


Figure 3.3: Split error induced by subtraction of membranes from thin objects. (**A**) shows the original electron microscopy image of the split error. (**B**) is the same part of the image with original published labels for this dataset. C and D are both zoomed in illustrations of the red rectangle shown in A and B. (**C**) illustrates the zoomed in part of B, whereas (**D**) shows the labeled data with subtraction of the membranes and induced split errors in pink, turquoise and purple.

The second case is correlated with the thickness of spine necks or neuronal structures and cutting angle in two consecutive images. This case occurred only once in the groundtruth dataset. The basic principle how this split error arises is illustrated in Figure 3.4 (D). The top and bottom of the cuboid represents two consecutive cytoplasm

images with a spacing of 29 nm. If a thin object is cut at a very small angle along its major axis, cytoplasm areas do not overlap in consecutive cutting frames resulting in no overlapping regions. Without this information the two cross-sections (Figure 3.4 brown circles) can not be associated with each other which leads to a split error seen in Figure 3.4 (B). However, human annotators can reattach isolated spine heads to their spine neck using the spatial relationship in the 3D visualization.

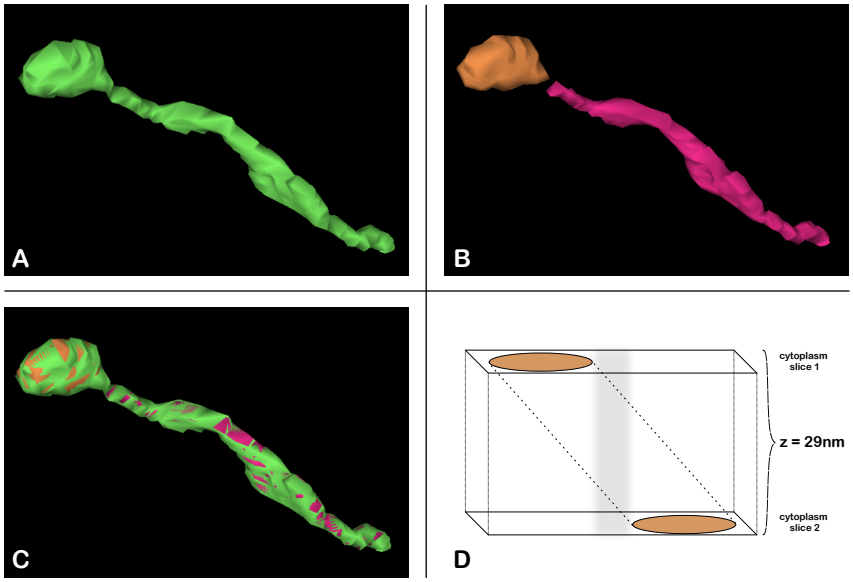


Figure 3.4: Split error already present in original dataset due to thin spine necks. (**A**) shows a single object visualized in neuroglancer. (**B**) shows the same object that is split into two different objects in the relabeled interleaved cytoplasm and overlap image stack. (**C**) illustrates both objects on top of each other. (**D**) represents two cytoplasm slices in the dataset with an object that is laying to collateral to the cytoplasm slices. This results in no overlapping information and therefore incorrectly tracing through the image stack.

3.3 Neuronal network training approaches

This section presents an overview of different training approaches to generate cytoplasm and overlap images, that are chosen to evaluate deep neuronal network algorithms for automatic segmentation and 3D reconstruction of neurons from electron microscopic images. As mentioned in section 3.2, cytoplasm and overlap images that are arranged to an interleaved image stack consisting of overlap images between corresponding cytoplasm images are capable of precise segmentation and 3D visualization. With this starting point, the following Figure 3.5 shows multiple paths to train deep convolutional neuronal networks (DCNN) for automatic generation of cytoplasm and overlap image stacks. In the next subsections, NN1-NN6 from Figure 3.5 are explained in detail.

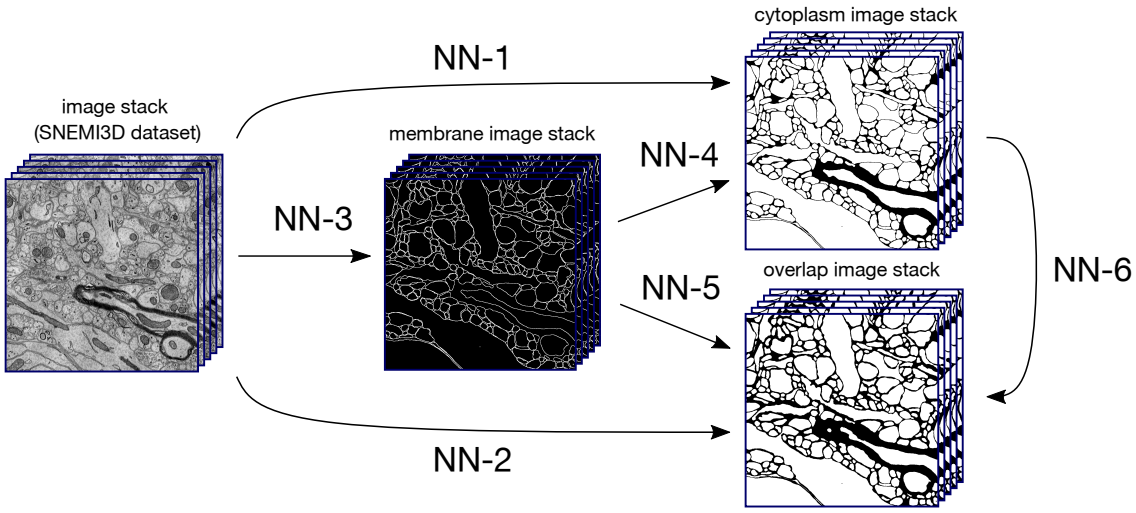


Figure 3.5: Overview of different paths to train DCNN for automatic cytoplasm and overlap image generation. One possibility is to calculate cytoplasm and overlap images directly from EM images (NN-1 and NN-2) . The second approach includes a training of membranes first (NN-3), that sets up the additional training to predict cytoplasm and overlap images out of membrane images (NN-4 and NN-5). A third path describes a training of cytoplasm images from EM images with subsequently training of predicted cytoplasm images to train a NN for overlap prediction (NN-6).

3.3.1 Training of cytoplasm, overlap and membrane from EM images

The first neuronal network training attempt was carried out for the generation of cytoplasm, overlap and membrane images from EM groundtruth (NN-1, NN-2, NN-3). In the following Figure 3.6, input, output and target images for the cytoplasm training are presented. Respectively, Figure 3.7 shows input, output and target images for the overlap training and Figure 3.8 input, output and target images for the membrane training. The network architecture behind this and following training section is an autoencoder consisting of an encoder and decoder with a ResNet (Residual Network) in between (for in

depth explanation of the NN architecture see section 2.4). All three neuronal networks are supervised learning approaches with different input and target images. For the cytoplasm and membrane training single EM images serve as input and cytoplasm and membrane images as target. In contrast, the overlap training requires two consecutive EM images stored in the red and green channel of an RGB image and overlap images as target image.

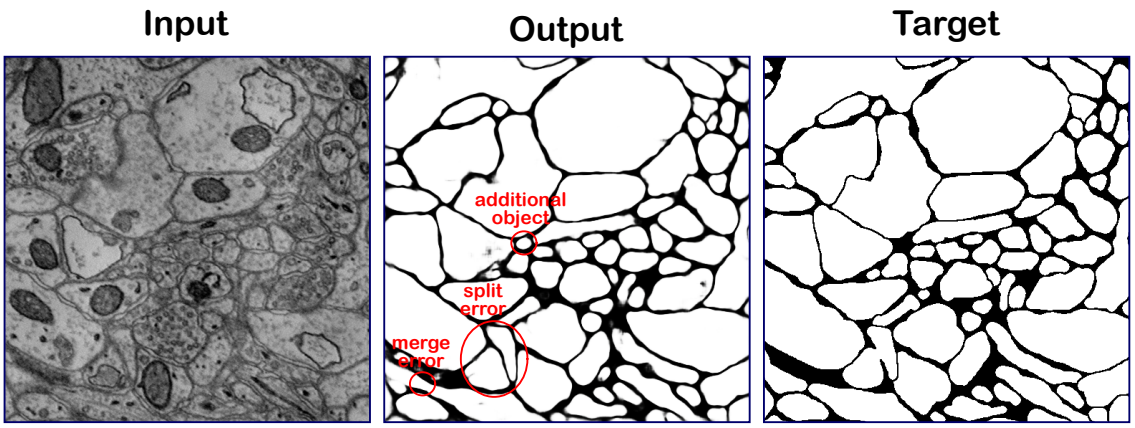


Figure 3.6: Presentation of one input, output and target image of the cytoplasm training.

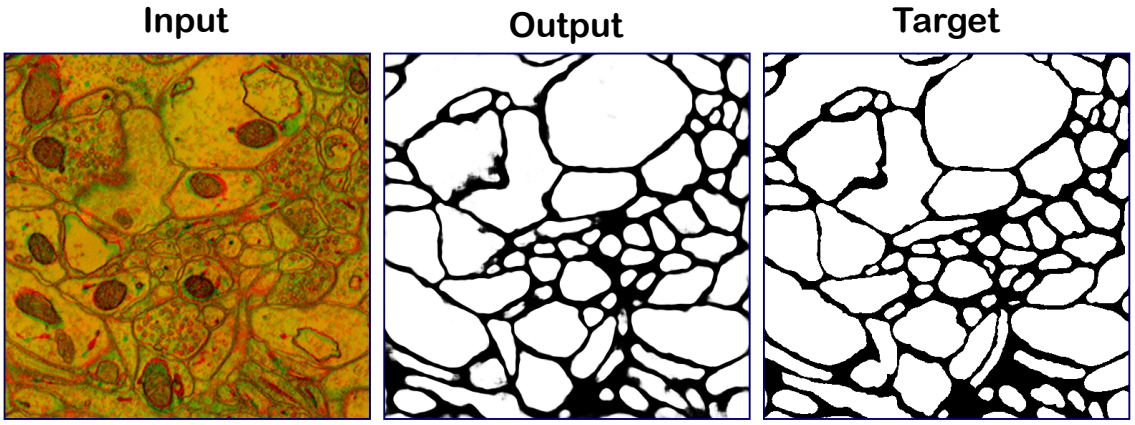


Figure 3.7: Presentation of one input, output and target image of the overlap training.

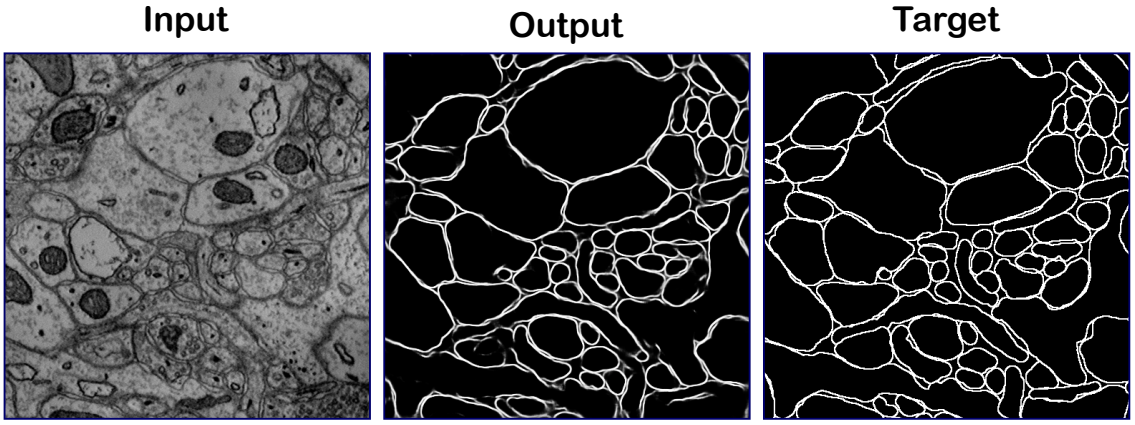


Figure 3.8: Presentation of one input, output and target) image of the membrane training.

3.3.2 Training of cytoplasm and overlap from predicted membrane

The second training approach includes cytoplasm and overlap training (target) with predicted membrane images from NN-3 as input (NN-4, NN-5). Cytoplasm training only requires a single membrane picture, whereas the overlap training needs information about two consecutive membrane images stored in the red and green color channel. An overview of input, output and target is illustrated in Figure 3.9 and 3.10.

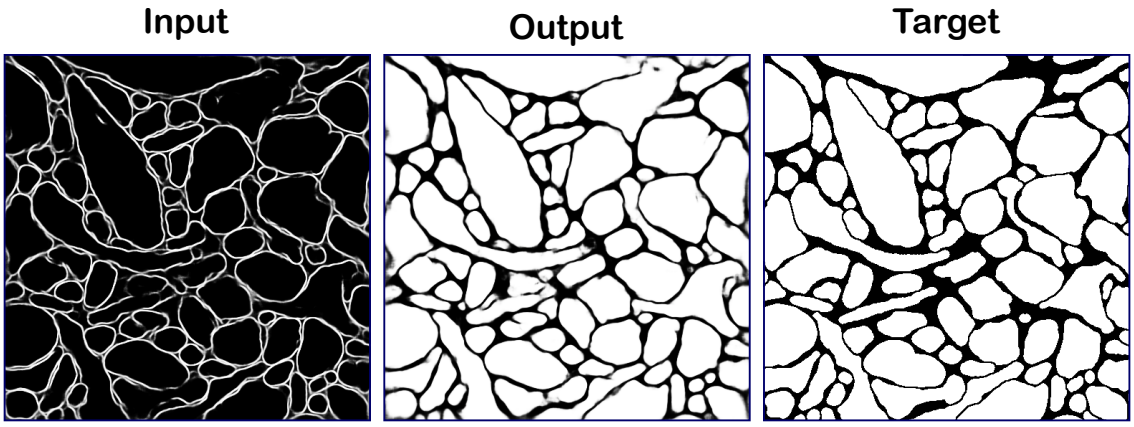


Figure 3.9: Presentation of one input, output and target image of the cytoplasm training with membrane input.

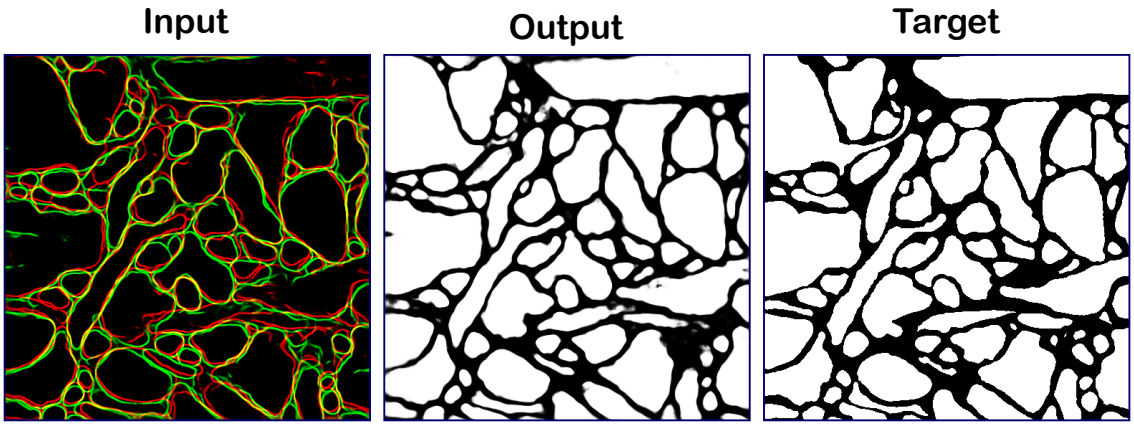


Figure 3.10: Presentation of one input, output and target image of the overlap training with membrane input.

3.3.3 Training of overlap from predicted cytoplasm

A third approach describes a training to detect overlap images from predicted cytoplasm images (NN-6). The input for this training are images with two consecutive cytoplasm images stored in the red and green color channel (see Figure 3.11).

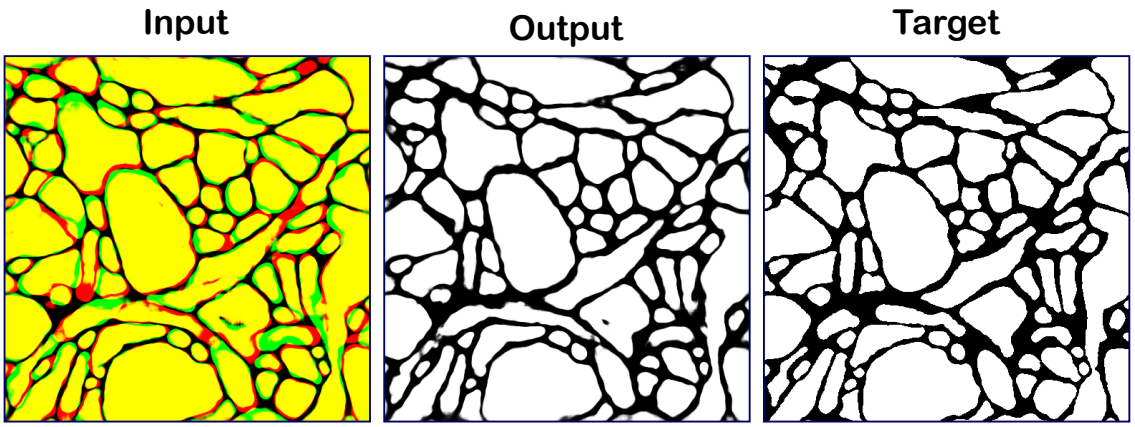


Figure 3.11: Presentation of one input, output and target image of the overlap training with cytoplasm input.

3.4 Different combinations to generate an interleaved cytoplasm/overlap image stack

The generation of a cytoplasm image stack with interleaved overlap images facilitates many different combinations. These images can be generated by label or through training of NN (see section 3.3 NN-1 - NN-6). In total, 10 different combinations to test reconstruction with the use of membrane, cytoplasm and overlap images were chosen. These combinations are all presented in Table 3.2. All listed experiments are endowed with an experiment number and are further used for validation and comparison with the original labels (Part B-D) from the dataset. The idea behind different combinations of predicted images and images generated from the labels is to gain insights which neuronal network introduces the largest errors when the reconstructed objects are compared with the original data.

Table 3.2: Overview of all used combinations to create a cytoplasm image stack with interleaved overlap images.

Exp.	Membrane images	Cytoplasm images	Overlap images
I	—	Groundtruth	Predicted from EM
II	—	Predicted from EM	Groundtruth
III	—	Predicted from EM	Predicted from EM
IV	—	Groundtruth	Predicted from cytoplasm
V	—	Predicted from EM	Predicted from cytoplasm
VI	Groundtruth	Groundtruth	Predicted from membrane

Exp.	Membrane images	Cytoplasm images	Overlap images
VII	Groundtruth	Predicted from membrane	Groundtruth
VIII	Groundtruth	Predicted from membrane	Predicted from membrane
IX	Predicted from EM	Predicted from membrane	Predicted from membrane
X	Predicted from EM	Predicted from membrane	Predicted from cytoplasm

3.5 Evaluation of neuronal network predictions

The evaluation metrics used for the comparison between reconstructed objects and the original objects consist of ARI, precision, recall as well as split and merge errors (see section 3.2). In the following subsections, all combinations listed in Table 3.2 marked with an Experiment number are analysed and compared with the associated part of the original data, respectively. Which part of the original label is associated with which relabeled label depends on how many NNs were used and how many parts are necessary for the testing of NN. All following subsections have their own workflows describing the procedure and use of training and testing images as well as images used for evaluation.

3.5.1 Reconstruction based on cytoplasm and overlap predicted from EM images

The first and most simple approach was to predict cytoplasm and/or overlap images directly from EM images. The different combinations tested include experiment I-III listed in Table 3.2. For example, the cytoplasm image stack with interleaved overlap images in experiment I consists of cytoplasm images generated from the original labels paired with predicted overlap images. The following formulas show the data acquisition of predicted overlap images.

$$\mathrm{EM}_{(A)} \xrightarrow{\text{Training}} \mathrm{Overlap}_{(A)} \tag{3.1}$$

$$\mathrm{EM}_{(B)} \xrightarrow{\text{Prediction}} \mathrm{Overlap}^{*}_{(B)} \tag{3.2}$$

Equation 3.1 describes the training of the NN-1. EM images from Part A serve as input,

whereas the overlap images of Part A generated from original labels provide the target. Equation 3.2 shows the prediction of overlap images with the trained parameters in equation 3.1. The prediction was performed on the EM images from Part B to generate predicted overlap images (predicted images marked with *). The penultimate step is to combine the predicted overlap images (part B) with the cytoplasm images from part B and change the arrangement to an interleaved image stack with overlap images between cytoplasm images for the relabeling with connected components. In the end, the overlap images need to be removed and the relabeled image stack of cytoplasm images is prepared for evaluation. The same procedure was used to generate an image stack consisting of predicted cytoplasm images and overlap images generated from original labels (experiment II). Equation 3.3 and 3.4 are similar to the explanation of equation 3.1 and 3.2 except that cytoplasm images are trained and predicted. Experiment III is the combination of both operations experiment I and experiment II. In this case, cytoplasm images and overlap images were trained and predicted and therefore used for relabeling. The results of experiment I - III are listed in Table 3.3.

$$\mathrm{EM}_{(A)} \xrightarrow{\text{Training}} \mathrm{Cytoplasm}_{(A)} \tag{3.3}$$

$$\mathrm{EM}_{(B)} \xrightarrow{\text{Prediction}} \mathrm{Cytoplasm}^{*}_{(B)} \tag{3.4}$$

The comparison between Part B (original label with subtracted membranes) and experiment I shows noticeable differences in their evaluation metrics. Precision and Recall are both 1.0, because the cytoplasm image stack from the original label was used. Thus, this result was predictable. However, ARI is extremely low and split and merge errors are comparatively high. In comparison, experiment II has clearly less split errors and about the same merge errors. The ARI is lower than experiment I probably due more merge errors and thus more objects labeled as the same object, which should be separated. Since only cytoplasm images are trained and predicted, it is possible to evaluate precision and recall. Both values are around 0.95, which means that every 20th white pixel is predicted black and every 20th black pixel is predicted white. Therefore, objects are in average one twentieth smaller than the groundtruth label.
To get a better concept about these evaluation metrics, Figure 3.12 shows the result of experiment I and II visualized with neuroglancer.

Table 3.3: Evaluation results for experiment I - III.

	adj. RAND Index	Precision	Recall	Merges	Splits
Part B	0.995	1.0	1.0	0	17
Experiment I	0.04	1.0	1.0	149	117
Experiment II	0.005	0.95	0.942	154	23
Experiment III	0.0066	0.95	0.942	164	154

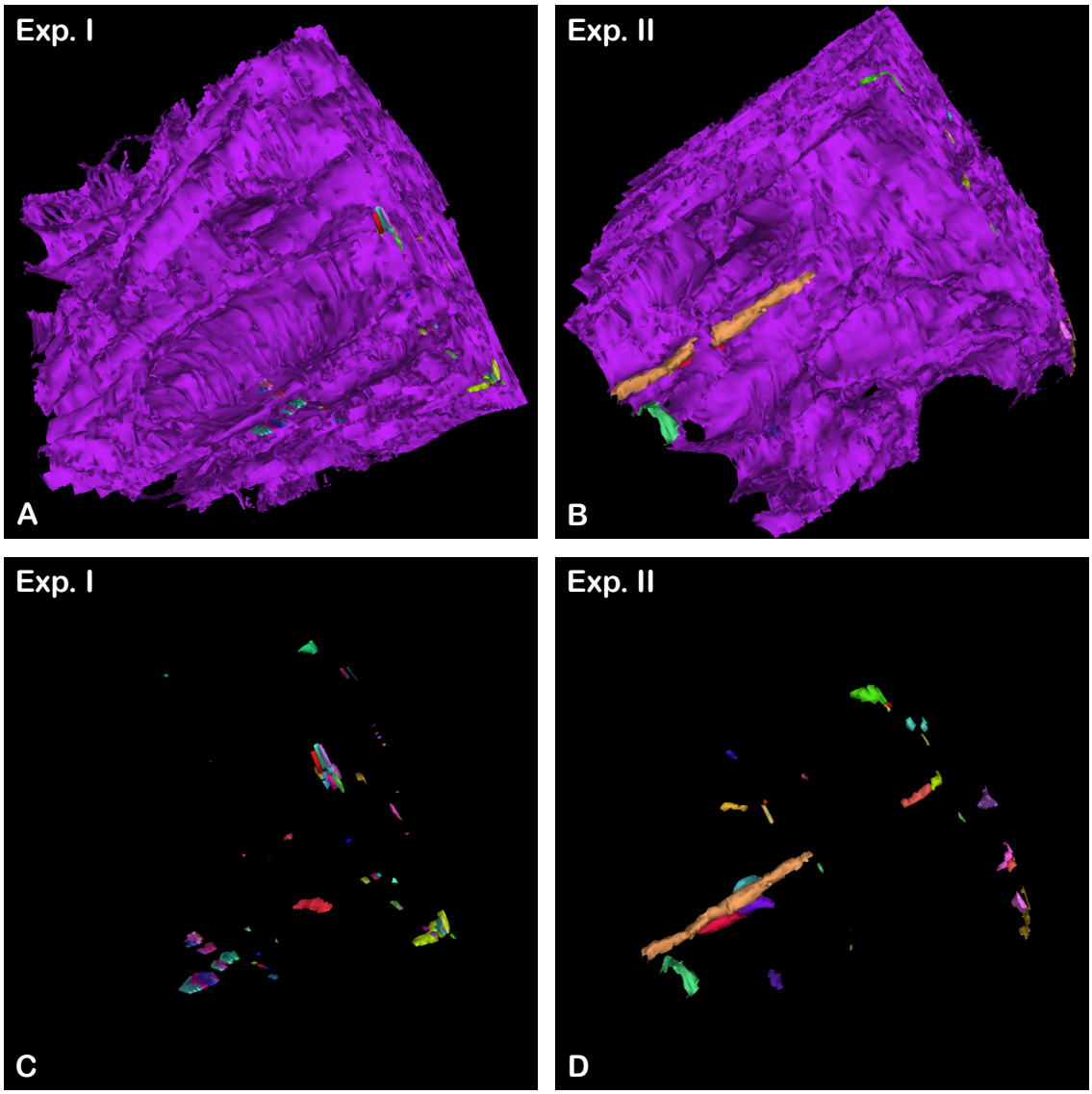


Figure 3.12: Experiment I and II visualized with neuroglancer. **(A)** and **(B)** shows all predicted objects from experiment I and II, whereas **(C)** and **(D)** presents single objects, that are not merged together.

Figure 3.12 (A) and (B) illustrates all predicted objects. The purple object seen in both images represents the merge error, that connects almost all bigger structures to one object. In contrast, (C) and (D) visualizes all objects except the connected purple structure. These objects are more or less pixel cluster that do not represent objects in the original label (see comparison of Figure 3.12 (C) and Figure 3.13 (A)). One exception is the orange object resembling an axon in Figure 3.12 (D).

Another illustration of the reconstructed merged object paired with a selection of single objects from the original data is shown in Figure 3.13 (B). This picture indicates, that original structures from the groundtruth label data are identifiable in experiment I as merged single structure. Therefore, post-processing of neuronal network predictions were taken into consideration to improve the evaluation results. The modification of predicted cytoplasm images is insufficient, since recall and precision are meant for evaluating the NN predictions. For this reason, overlap images are more suitable to use for modification, because they serve only to connect objects from cytoplasm slices for relabling. The slightest alteration of predicted objects from NN are capable of connecting single objects in cytoplasm images to massive connective structures. For this reason, overlapping regions were eroded in different magnitude. With the reduction of overlapping regions, it is possible to avoid merged objects. In contrast, split errors are more likely to arise due to missing overlap information of two consecutive cytoplasm images. To test this assumption, overlap images predicted by experiment I are binarized and eroded before relabeling. The evaluation results are visualized in Figure 3.14.

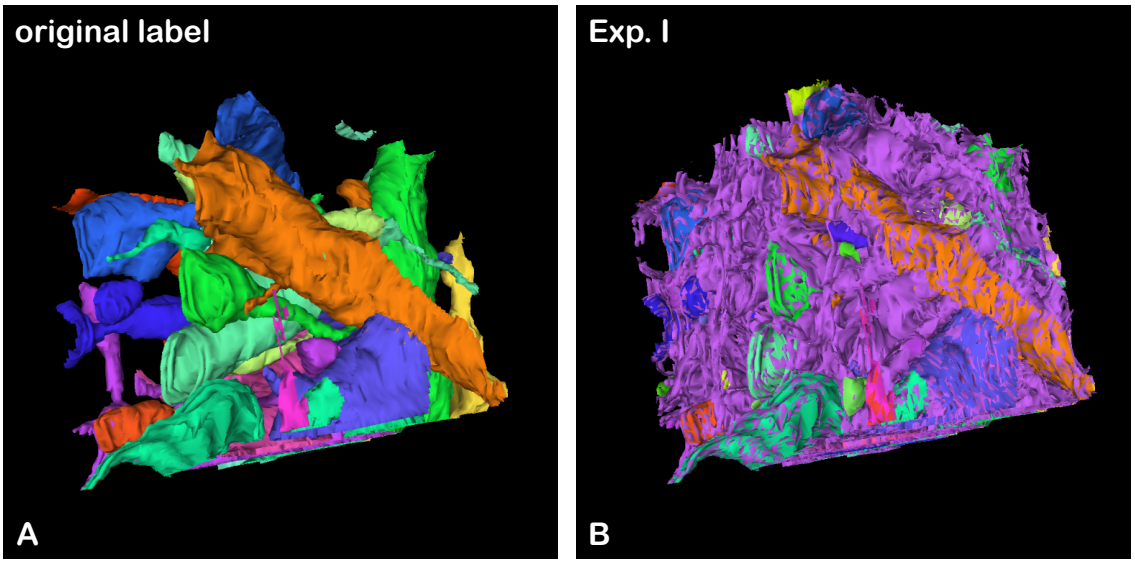


Figure 3.13: Original objects and experiment I visualized with neuroglancer. **(A)** shows a selection of different objects of Part B from the original label and **(B)** presents reconstructed objects from experiment I with original objects from (A).

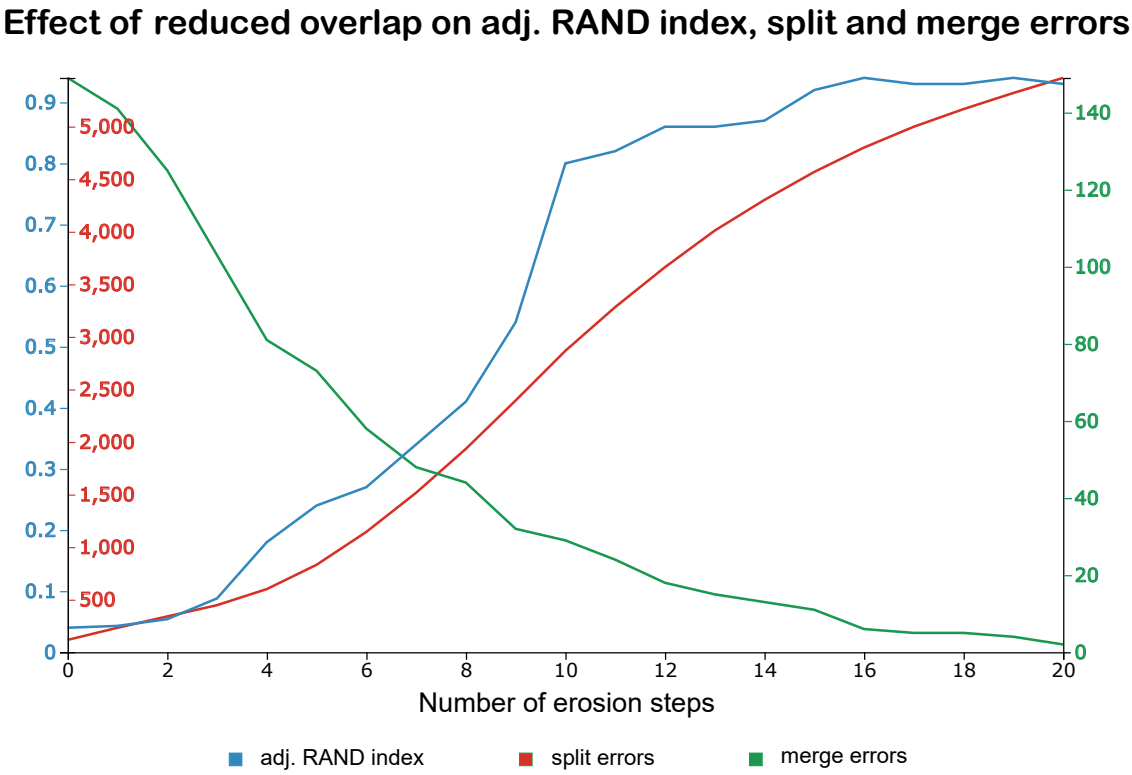


Figure 3.14: Evaluation results of experiment I after eroding overlapping regions in overlap images. Overlapping objects in overlap images were eroded up to twenty times (see appendix A.1 for exact values).

Already after two erosion steps merge errors start to decrease, whereas ARI remains the same and split errors are rising slowly. Additional erosion steps have impact on all three evaluation metrics. With sinking merge errors, ARI starts to rise because large objects are now correctly assigned to the objects in the original data. However, relatively small overlapping regions from the groundtruth data are going to vanish, thus objects that are connected in the original data are labeled in each cytoplasm slice individually. The reason for this phenomena are missing overlap information to assign objects in consecutive cytoplasm images to the same object. I decided to use tenfold erosion for further predicted overlap images to drastically increase correctly predicted objects.
In addition to the erosion of objects, it is also possible to skeletonize given overlapping objects in overlap images. The skeletonization of overlap images means to reduce the border of overlapping regions under the condition that connectivity of objects remain. The reason to use skeletonization after several erosion steps is to remove artefacts, that may occur after binarizing predicted images. Figure 3.15 illustrates overlap images after erosion and erosion paired with skeletonization. With this additionally step, merge errors were reduced and ARI was increased. Split errors remained almost the same. Table 3.4 shows the evaluation metrics of selected erosion steps compared to the skeletonized images with same amount of prior erosion.
With these findings, all experiments with predicted overlap images received post-processing steps in the form of tenfold erosion followed by skeletonization to improve object prediction (in later sections referred as post-processing).

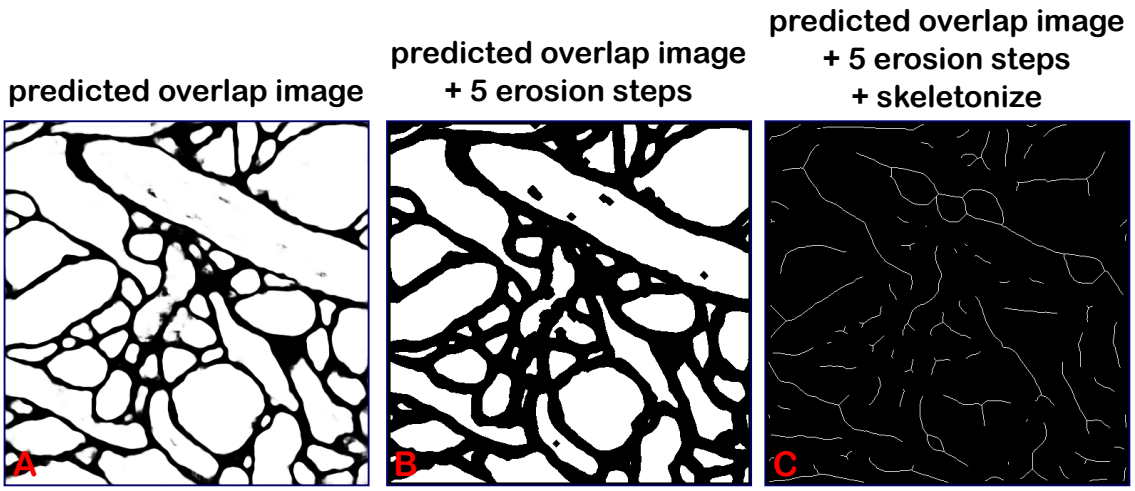


Figure 3.15: Comparison of overlap images with erosion and skeletonization. **(A)** shows an exemplary overlap image. **(B)** illustrates an overlap image with five erosion steps and **(C)** with additional skeletonizing.

Table 3.4: Comparison of eroded objects and eroded objects with additionally skeletonization for selected erosion steps.

	adj. RAND Index	Merges	Splits
Part B	0.995	0	17
Experiment I	0.04	149	117
1 erosion step	0.043	141	232
1 erosion step + skeletonize	0.088	112	200
5 erosion steps	0.24	73	831
5 erosion steps + skeletonize	0.42	55	851
10 erosion steps	0.80	29	2867
10 erosion steps + skeletonize	0.87	19	2872
15 erosion steps	0.92	11	4568
15 erosion steps + skeletonize	0.95	5	4570
20 erosion steps	0.93	2	5463
20 erosion steps + skeletonize	0.93	2	5464

Figure 3.16 and 3.17 illustrate experiment I and III, with and without post-processing in x-y-, x-z- and y-z-axis. The 3D reconstruction based on the prediction of NN-1 in experiment I can be improved by post-processing. This indicates the usefulness of the alteration of overlap images. Furthermore, the comparison between experiment I and III with post-processed overlap images shows the importance of exact cytoplasm images. Therefore, even small errors or displaced objects in predicted cytoplasm images are capable of merging objects back together due to overlapping regions reaching over multiple deviant cytoplasm objects. Table 3.5 presents the evaluation results of experiment I and III with post-processed overlap images.

Table 3.5: Evaluation results for experiment I - III with and without erosion and skeletonization.

	adj. RAND Index	Precision	Recall	Merges	Splits
Part B	0.995	1.0	1.0	0	17
Experiment I without post-processing	0.04	1.0	1.0	149	117
Experiment I with post-processing	0.87	1.0	1.0	19	2872
Experiment III without post-processing	0.0066	0.95	0.942	164	154
Experiment III with post-processing	0.036	0.95	0.942	314	2957

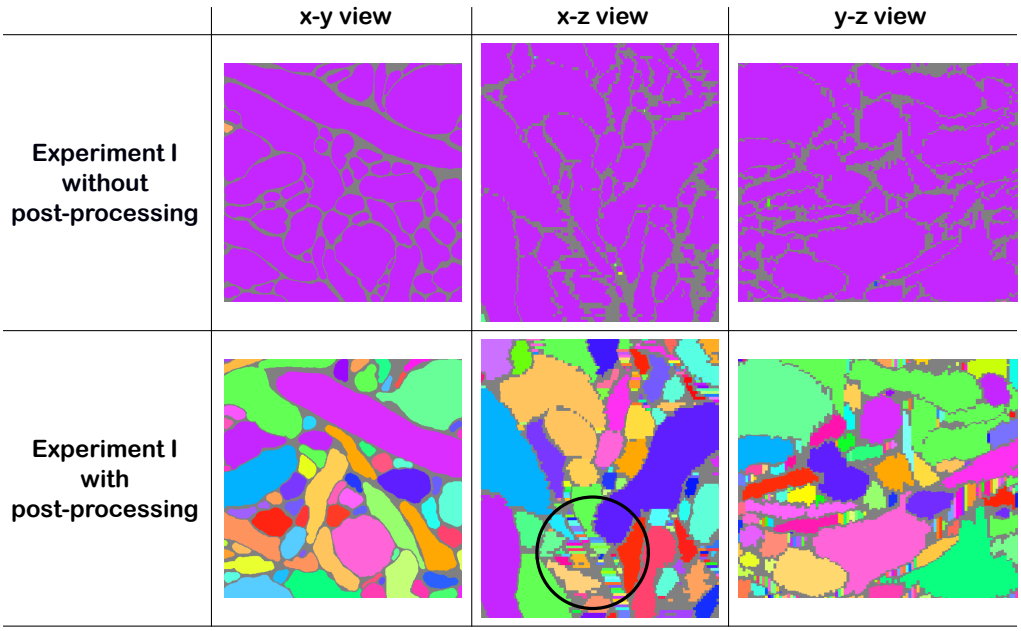


Figure 3.16: Visual comparison between unchanged and eroded/skeletonized reconstruction of experiment I. The black circle indicates the increased split errors arising from altering overlap images (each image receives individual label for the same object).

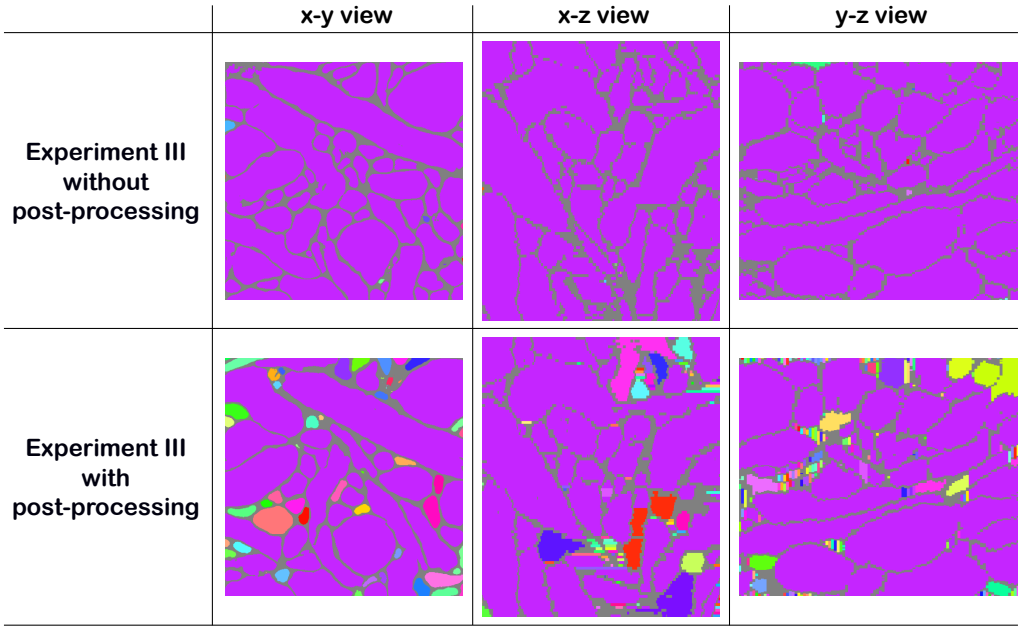


Figure 3.17: Visual comparison between unchanged and eroded/skeletonized reconstruction of experiment III.

3.5.2 Reconstruction based on overlaps predicted from cytoplasm

As shown in Figure 3.5, overlap images can also be predicted from cytoplasm images (NN-6). Therefore, experiment IV and V were carried out to test the reconstruction based on overlap images predicted from groundtruth and predicted cytoplasm images. The following equations 3.5 and 3.6 provide information on which image stack was used for training and prediction. Experiment IV consists of cytoplasm images generated from the groundtruth labels with predicted overlap images from cytoplasm images.

$$\text{Cytoplasm}_{(A)} \xrightarrow{\text{Training}} \text{Overlap}_{(A)} \tag{3.5}$$

$$\text{Cytoplasm}_{(C)} \xrightarrow{\text{Prediction}} \text{Overlap}^{*}_{(C)} \tag{3.6}$$

In contrast, experiment V used cytoplasm images predicted from EM images by NN-1 symbolized in equation 3.7 and 3.8. For 3D reconstruction, predicted cytoplasm images were then combined with overlap images predicted by NN-6 (see equation 3.9 and 3.10). The obtained evaluation results of experiment IV and V are both presented in Table 3.6.

$$\text{EM}_{(A)} \xrightarrow{\text{Training}} \text{Cytoplasm}_{(A)} \tag{3.7}$$

$$\text{EM}_{(B)+(C)} \xrightarrow{\text{Prediction}} \text{Cytoplasm}^{*}_{(B)+(C)} \tag{3.8}$$

$$\text{Cytoplasm}^{*}_{(B)} \xrightarrow{\text{Training}} \text{Overlap}_{(B)} \tag{3.9}$$

$$\text{Cytoplasm}^{*}_{(C)} \xrightarrow{\text{Prediction}} \text{Overlap}^{*}_{(C)} \tag{3.10}$$

Table 3.6: Evaluation results for experiment IV - V

	adj. RAND Index	**Precision**	**Recall**	**Merges**	**Splits**
Part C	0.998	1.0	1.0	0	21
Experiment IV without post-processing	0.14	1.0	1.0	94	27
Experiment IV	0.97	1.0	1.0	9	3154
Experiment V	0.023	0.937	0.937	313	2946

Experiment IV was carried out with and without post-processing. The results in Table 3.6 represent prior presumption of decreasing merge and rising split errors and ARI introduced due to post-processing. In comparison, the 3D reconstruction of experiment V consists of two subsequent NN training steps. Even after post processing, the ARI remains low and merge errors increase, which indicates that the NN training of already predicted cytoplasm images is unsuitable for overlap prediction and following 3D reconstruction.

3.5.3 Reconstruction based on cytoplasm and overlap predicted from membrane groundtruth

Previous studies often used membrane prediction as a first step in reconstructions and several very good membrane predictors have been described [Arbeláez et al., 2011, Hwang and Liu, 2015]. Watershedding and learned agglomeration were then used to reconstruct neurons. Therefore, this sections explores whether 3D reconstruction can be obtained using the framework from membranes only. With the assumption of correct membranes, membranes from the groundtruth were used in experiment VI - VIII. In experiment IX and X (see next subsection) an additional NN (NN-3) is used to predict the membranes from EM images.

The training and prediction processes of experiment VI are presented in equation 3.11 and 3.12 and experiment VII in equation 3.13 and 3.14. Experiment VIII consists of the training and prediction steps from Experiment VI + VII.

$$\text{Membrane}_{(A)} \xrightarrow{\text{Training}} \text{Overlap}_{(A)} \tag{3.11}$$

$$\text{Membrane}_{(B)} \xrightarrow{\text{Prediction}} \text{Overlap}^{*}_{(B)} \tag{3.12}$$

$$\text{Membrane}_{(A)} \xrightarrow{\text{Training}} \text{Cytoplasm}_{(A)} \tag{3.13}$$

$$\text{Membrane}_{(B)} \xrightarrow{\text{Prediction}} \text{Cytoplasm}^{*}_{(B)} \tag{3.14}$$

The evaluation results of experiment VI - VIII are presented in Table 3.7. Since experiment VI and VIII include an overlap prediction, post-processing was applied after prediction. In comparison to previous experiments, experiment VI - VIII are performing considerably better. Even with the prediction of overlap and cytoplasm images, merge errors, ARI, precision and recall are showing exceptional results.

Table 3.7: Evaluation results for experiment VI - VIII

	adj. RAND Index	**Precision**	**Recall**	**Merges**	**Splits**
Part B	0.995	1.0	1.0	0	17
Experiment VI	0.96	1.0	1.0	5	2824
Experiment VII	0.99	0.99	0.99	0	21
Experiment VIII	0.96	0.99	0.99	5	2836

3.5.4 Reconstruction based on cytoplasm and overlap from predicted membrane

The final results of experiment IX and X are based on a training to predict membrane images first. This step of membrane training and prediction is shown in equation 3.15 and 3.16. Predicted membrane images serve as input for the following training steps to

train cytoplasm and overlap images visualized in following equations: 3.17, 3.18, 3.19, 3.20.

$$\text{EM}_{(A)} \xrightarrow{\text{Training}} \text{Membrane}_{(A)} \tag{3.15}$$

$$\text{EM}_{(B)+(C)+(D)} \xrightarrow{\text{Prediction}} \text{Membrane}^{*}_{(B)+(C)+(D)} \tag{3.16}$$

$$\text{Membrane}^{*}_{(B)} \xrightarrow{\text{Training}} \text{Cytoplasm}_{(B)} \tag{3.17}$$

$$\text{Membrane}^{*}_{(C)+(D)} \xrightarrow{\text{Prediction}} \text{Cytoplasm}^{*}_{(C)+(D)} \tag{3.18}$$

$$\text{Membrane}^{*}_{(B)} \xrightarrow{\text{Training}} \text{Overlap}_{(B)} \tag{3.19}$$

$$\text{Membrane}^{*}_{(C)} \xrightarrow{\text{Prediction}} \text{Overlap}^{*}_{(C)} \tag{3.20}$$

The prediction of cytoplasm images for the image stacks Part C and D (see equation 3.18) were essential for a further used training and prediction implemented in experiment X. In addition to experiment IX, experiment X involves a training overlap images from predicted cytoplasm images generated from predicted membrane images. Therefore, equations 3.21 and 3.22 are appended in experiment X.

$$\text{Cytoplasm}^{*}_{(C)} \xrightarrow{\text{Training}} \text{Overlap}_{(C)} \tag{3.21}$$

$$\text{Cytoplasm}^{*}_{(D)} \xrightarrow{\text{Prediction}} \text{Overlap}^{*}_{(D)} \tag{3.22}$$

Table 3.8 shows the last evaluation results for experiment IX and experiment X with post-processing of predicted overlap images. Due to many training and prediction steps, experiment IX and X are expected to yield the worst results in regards to ARI, precision and recall. Unexpectedly, splits and merge errors seem to be decreasing, which should raise the ARI. However, the ARI is extremely low, which also indicates huge merged

objects. The explanation for this phenomena are probably predicted cytoplasm images from Part D. Most neurons are already connected (due to prior training and prediction), which indicates that larger connected objects in cytoplasm images can not produce more merge errors as available individual cytoplasm objects. Furthermore, overlap images are trained from these predicted cytoplasm images. Thus, the remaining neurons are connected through overlap images resulting in less split errors.

Table 3.8: Evaluation results for experiment IX - X

	adj. RAND Index	**Precision**	**Recall**	**Merges**	**Splits**
Part C	0.998	1.0	1.0	0	21
Part D	0.999	1.0	1.0	0	22
Experiment IX	0.01	0.926	0.935	462	3053
Experiment X	0.000018	0.87	0.93	144	125

3.6 Proofreading the SNEMI3D dataset

A closer observation of the SNEMI3D dataset reveals some inaccuracies, that might have affected cytoplasm, overlap and membrane image prediction by NN. Figure 3.18 shows electron microscopic images from the dataset with overlayed original annotated membranes (extracted from the labeled data).
There are substantial portions examples of the annotated membrane labels that are not overlapping with the membranes visible in electron microscopic images (red arrows pointing on regions with differences). Additionally, intracellular structures are labeled as "membrane" (see red circle), which is surprising because these structures belong to 3D objects in the SNEMI3D dataset. All of these mentioned inaccuracies may affect NN training and lead to a poor prediction of cytoplasm, overlap and membrane images. Therefore, a representative part of the original EM images was manually annotated for membrane training and prediction. Two different approaches for the comparison of the difference between self annotated and original annotated membranes were used.
The first was to calculate the precision and recall of the membranes created from the object boundaries in the SNEMI3D dataset. The manual re-annotated membranes were used as the groundtruth. However, precision and recall are pixel-wise measures that depend on the thickness of the membrane label. Thicker membranes would result in a larger overlap than thinner membranes. The pixel-based assessment (see Figure 3.19) shows that almost half of the membranes in the SNEMI3D do not belong to actual membranes (precision=0.55, mean+/- SD for n=10 images) and almost two-third of membranes are not labeled at all (recall=0.43, mean+/- SD for n=10 images).

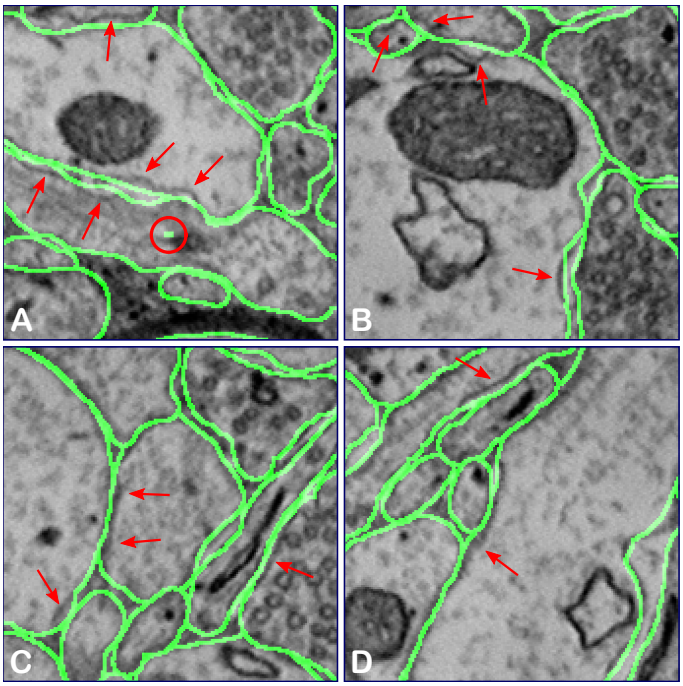


Figure 3.18: Differences between EM and SNEMI3D derived membranes. **(A)** - **(D)** shows exemplary differences in membranes marked with red arrows of EM images. Additional, the red circle in **(A)** indicates unannotated pixels appearing midst objects.

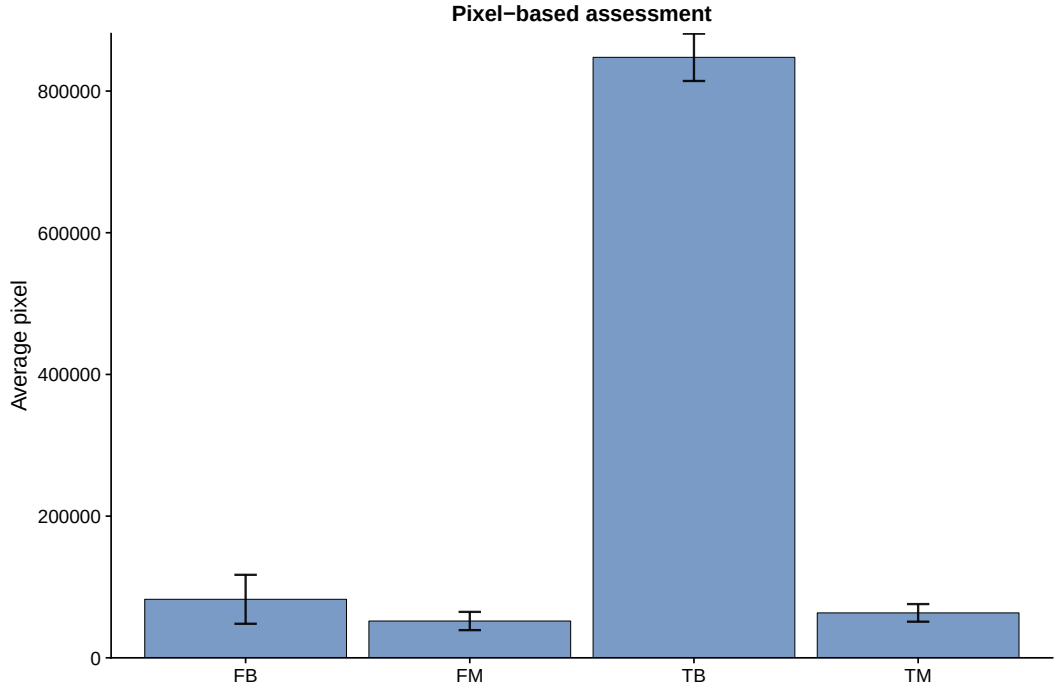


Figure 3.19: Pixel-based assessment of membrane labels. Bar graph showing the mean and standard deviation for n=10 manually annotated images.
FB = False background, FM = False membrane,
TB = True background, TM = True membrane.

Therefore, the second approach was to calculate the directed distance from the object boundaries in the SNEMI3D dataset and the midline of the manual annotated membranes. I calculated the shortest euclidean distance of the manually annotated membrane labels to each pixels in the SNEMI3D derived labels. By skeletonizing the membrane label first, a measurement independent from the membrane thickness was obtained. Note, that this measurement is related to the Hausdorff distance, which specifies the maximal distance of two subsets and is often used in image analysis [Taha and Hanbury, 2015]. However, this maximum gives a single value for all pixels and is also prone to outliers. Figure 3.20 shows the distance-based assessment of the SNEMI3D derived membrane labels. Almost 50 % of the SNEMI3D derived membranes are more than one pixel away from the manually annotated, true membranes (1 +/- 1.51, median +/- MAD, 10 images), which can be decisive for correct membrane prediction.

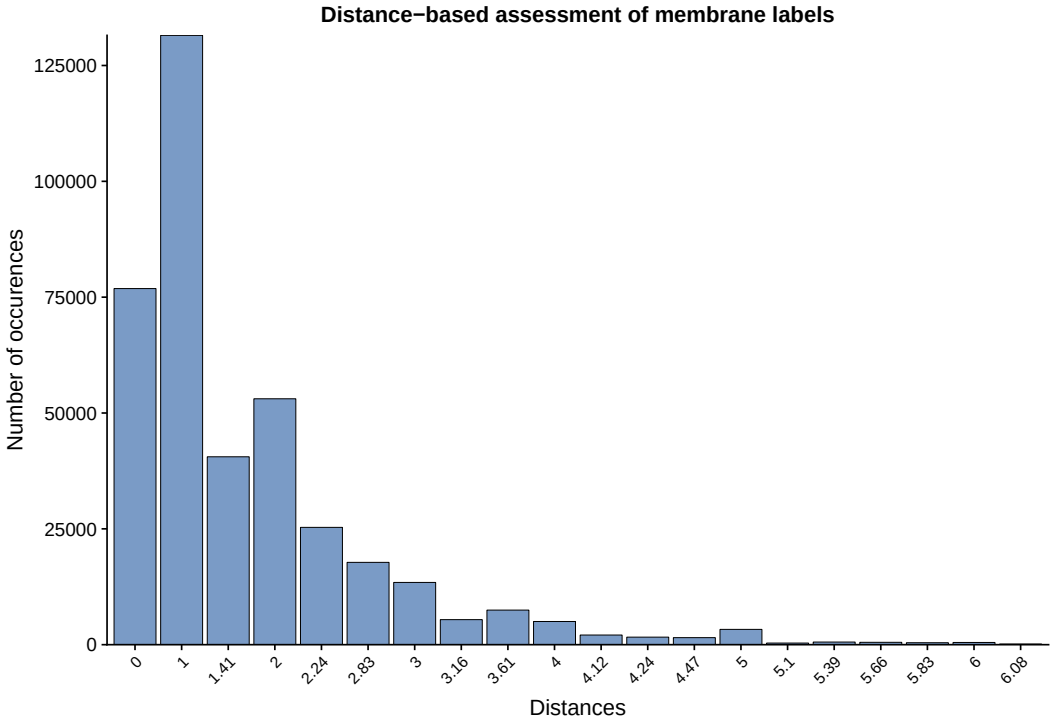


Figure 3.20: Distance-based assessment of membrane labels.

Although NN training can be based on data with noisy labels, I expect that a specific bias on membrane labels could have a great impact on the membrane training and prediction. Therefore, a comparison of the predicted membranes with the accuracy of the membranes derived from the SNEMI3D dataset (see Figure 3.21) was performed. It appears that small tunnels of extracellular space between neighbouring neurons were not accurately represented in the SNEMI3D dataset. This results in membrane labels that are somewhere in between the true membranes of the two neurons. Furthermore, Figure 3.21 (C) and (D) enables the elaboration of potential prediction fails correlated to not corresponding labels in (A). The presumption of failing membrane prediction and false negatives (referring to occurring gaps in membranes, see (D) may be associated with discrepancies of SNEMI3D derived and manual annotations. Unfortunately, there

is no real pattern visible that is able to describe missing membranes in the prediction. From a mere visual point of view, the training of NN-3 with manually annotated membranes resulted in a slightly better prediction compared with training on SNEMI3D derived membranes. Therefore, the following subsection is dedicated for hyperparameter tuning using only the manually annotated membrane labels.

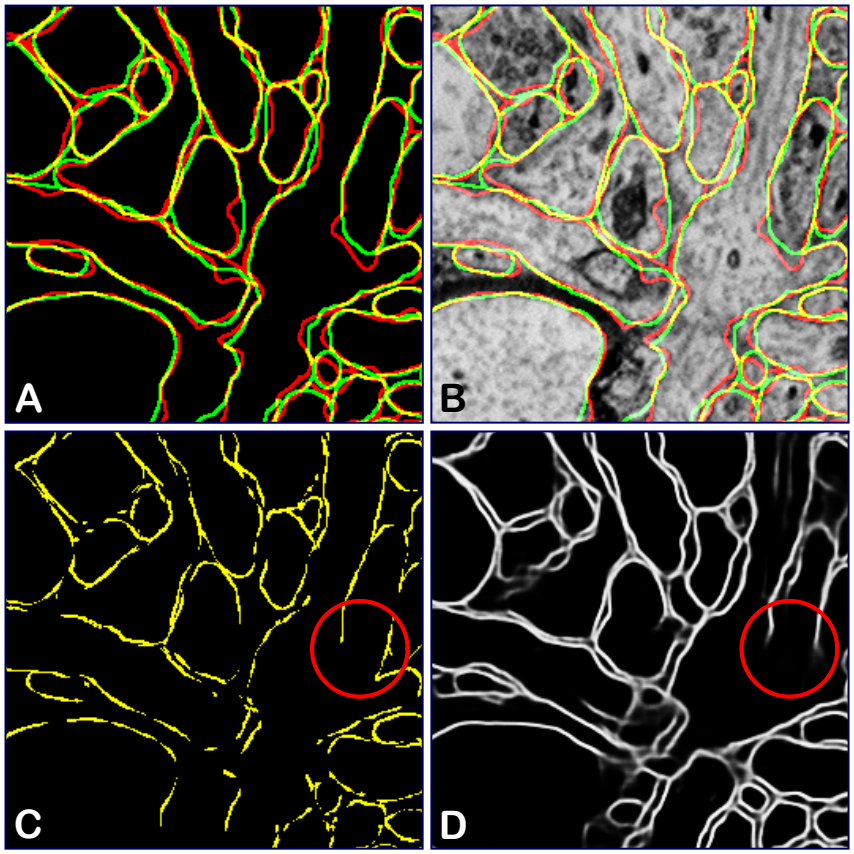


Figure 3.21: Comparison of SNEMI3D derived membranes and self annotated membranes of one section from the dataset. **(A)** shows SNEMI3D derived membranes (green) and self annotated membranes (red). **(B)** displays the associated EM image. **(C)** illustrates only the overlapping parts (yellow) of **(A)**. **(D)** shows the prediction of the NN of the same region. The red circles illustrate the dependence of correct annotated membranes for prediction.

3.7 Hyperparameter testing for membrane prediction

Previous experiments (experiment VI - VII) have shown excellent evaluation results for trained cytoplasm and overlap images from membrane groundtruth. Therefore, if it is possible to improve the membrane prediction from EM groundtruth data, neurons could be reconstructed from EM data. To test the influence of the NN architecture, different hyperparameter were varied for the training of membrane from EM images. Changed hyperparameter include the number of epochs, three different loss functions, the amount of residual blocks in the ResNet and the initial filters in the first convolutional layer. The number of epochs was tested for 20,000 and 100,000 epochs. The three different loss

functions are the mean squared error, logistic loss and dice coefficient, whereas the combination of residual blocks and filters in the first convolutional layer was set to 160 and 64 or 320 and 32 respectively. After the training, precision and recall of predicted and self annotated membrane images were calculated and illustrated in Figure 3.22 and 3.23. As shown in Figure 3.22, residual blocks, filter in the first convolutional layer and the number of epochs used for training have little impact on the precision. In contrast, the loss function (see Figure 3.25) seems to have a larger effect on the precision, which indicates, that many false positives (black pixel which are predicted white) are predicted. A closer look reveals, that predicted membrane images with the dice loss functions are mostly random white and black pixels with round white artifacts (see Figure 3.24). However, logistic and square loss function are able to recreate membrane images with gaps (as visualized in Figure 3.8).

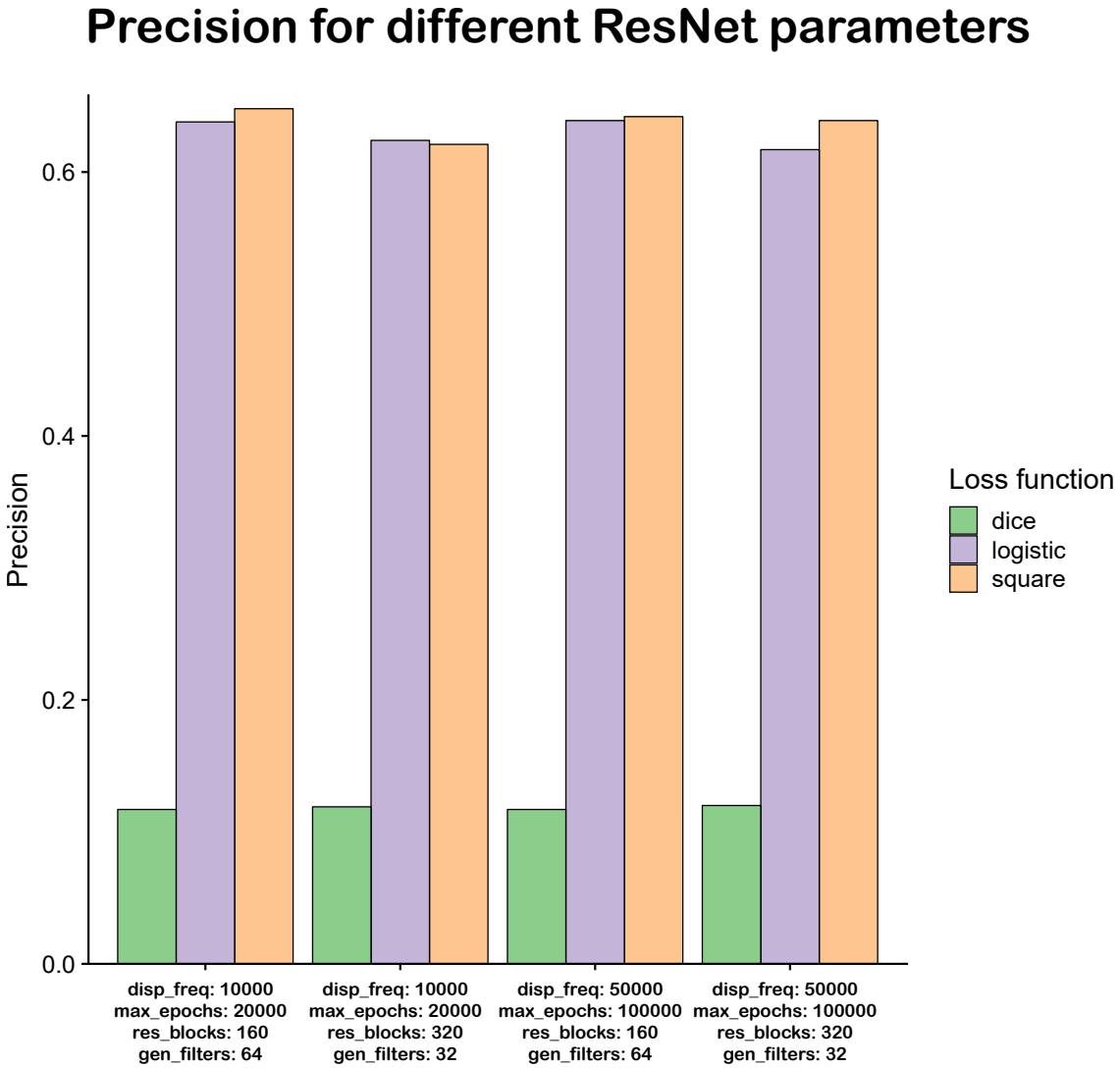


Figure 3.22: Presentation of precision values calculated from predicted membrane images with different ResNet hyperparameter.

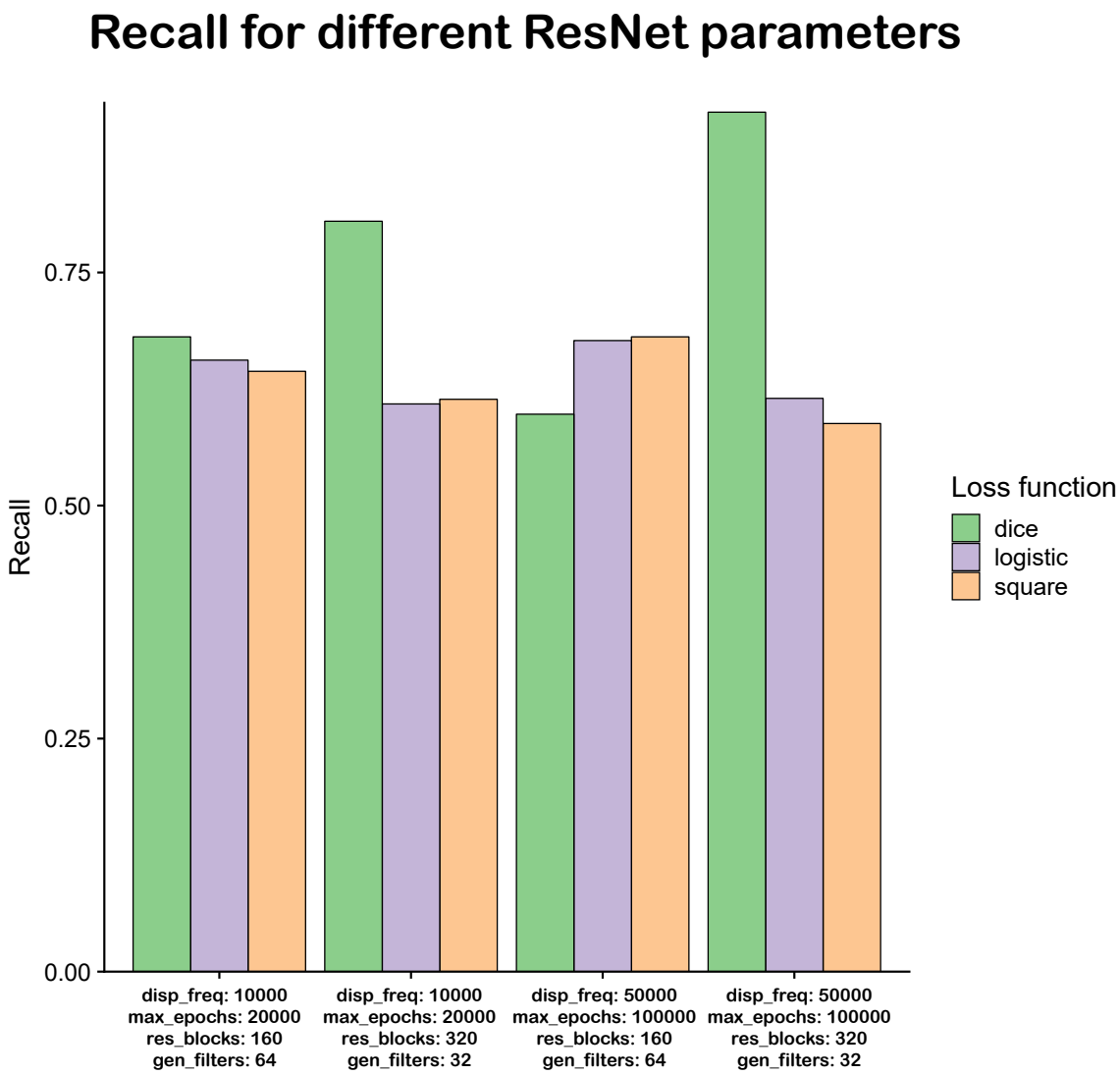


Figure 3.23: Presentation of recall values calculated from predicted membrane images with different ResNet hyperparameter.

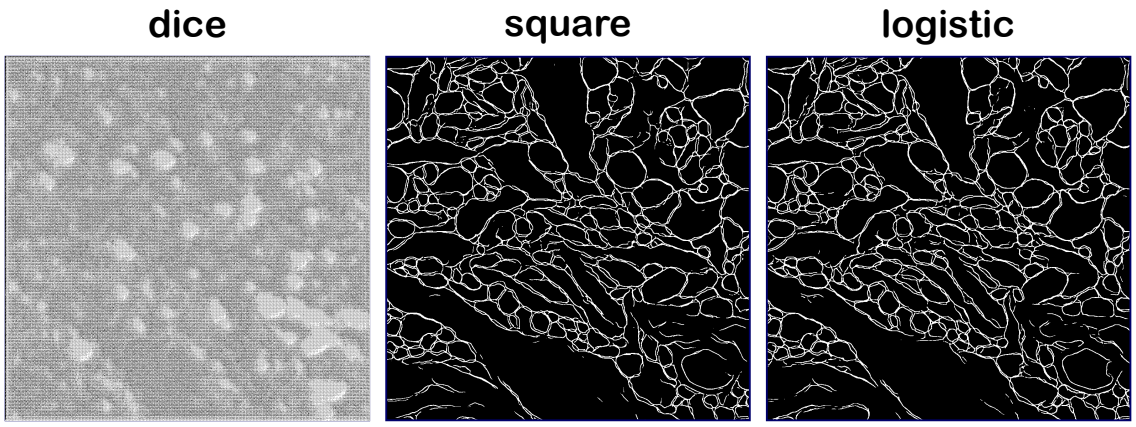


Figure 3.24: Predicted membrane images of NNs with different loss functions. Dice, square and logistic output images used for evaluating precision and recall against self annotated membrane images.

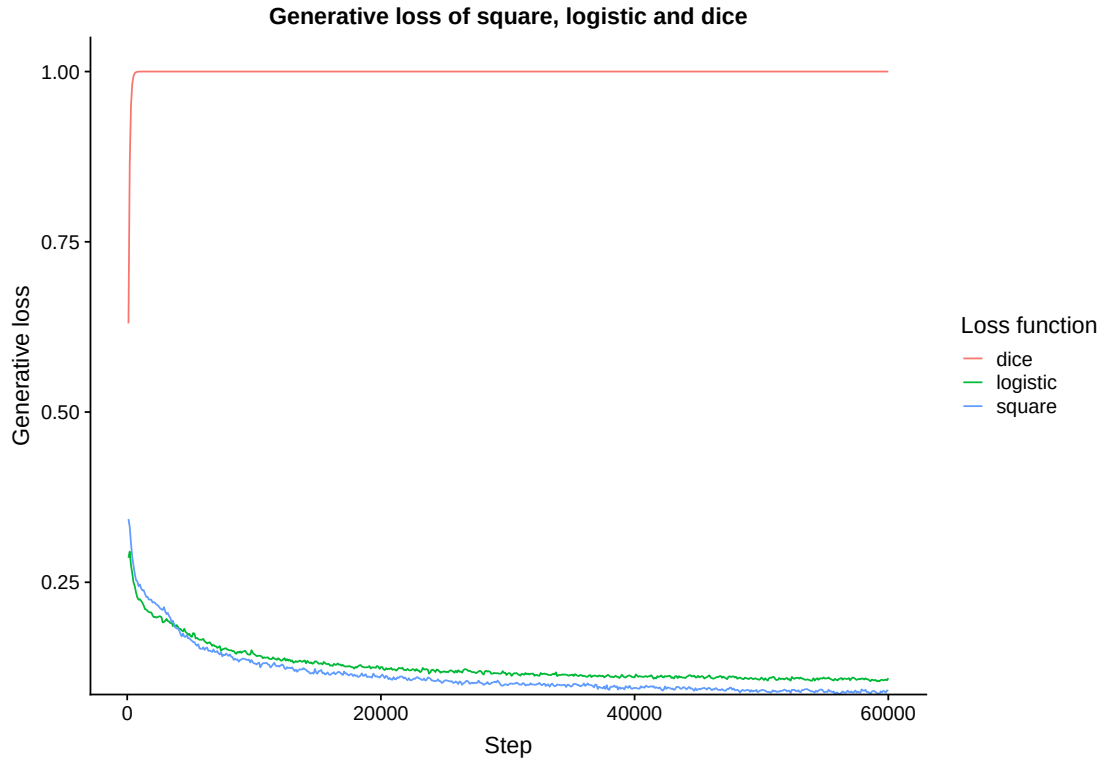


Figure 3.25: Comparison of generator loss during NN training for all three tested loss functions. Hyperparameter: 20.000 epochs, 160 residual blocks, 64 filter in the first convolutional layer

4 Discussion

I developed a new workflow for dense reconstruction of three dimensional objects from electron microscopic (EM) image stacks.
Recently, high throughput electron microscopic imaging has been used to capture the nano-scale structure of neuronal connections [Briggman and Bock, 2012]. However, the large number of high resolution images precludes manual and semi-automatic reconstruction [Lichtman et al., 2014]. Therefore, full automated segmentation and classification gained much interest [Arganda-Carreras et al., 2015, Januszewski et al., 2017].
My approach combines deep learning with deterministic algorithms in order to obtain dense, three dimensional reconstructions from serial section EM images stacks. I was able to solve the so called *contour correspondence problem* in serial sectioning [Herbert and Jones, 2001] by using deep learning to predict the overlap between consecutive objects. By restoring all objects in parallel, a faster reconstruction than the best 3D CNN algorithm reconstructing isolated objects with flood-filling networks [Januszewski et al., 2016] was achieved. However, detailed analysis shows, that the accuracy of the proposed workflow critically depends on the correct prediction of membranes, which in turn depends on the correct annotation in the training set.

4.1 Evaluation of dense reconstruction

First, I showed that the combination of cytoplasm and overlap information extracted from the annotated data with subsequent use of connected components is able to solve the *contour correspondence problem* in serial sectioning [Herbert and Jones, 2001]. Thus, I fully reconstructed objects of the high anisotropic SNEMI3D dataset (see Table 3.2). Proceeding with this knowledge, multiple ways of predicting cytoplasm and overlap images with NN were tested (see Figure 3.5). For example, it is possible to traverse the workflow through NN-1 and NN-2 or predict membrane images first with NN-3 followed by NN-4 and NN-5. Therefore, several experiments were performed successively in order to find the best prediction of cytoplasm and overlap images.

4.1.1 Metrics

For the evaluation of my experiments, precision and recall for cytoplasm images, adjusted RAND index (ARI) plus merge and split errors were calculated. Precision and recall served as metrics to examine how accurate the NN predicts cytoplasm images. Furthermore, ARI is the most used evaluation metric for segmentation purposes [Hubert and Arabie, 1985], whereas merge and split errors give more intuitive insight of reconstructed objects.

However, it is important to note how merge and split errors are calculated. The definition of these metrics specifies whether two objects in the relabeled data are merged together to one object despite being two single objects in the groundtruth data (merge error) or one object in the groundtruth data is assigned to two different objects in the predicted image stack (split error). If more than two objects are involved, I counted with how many objects the predicted label is merged or to how many objects the groundtruth label is split, respectively. Additionally, objects with a size of less than 10 voxels were discarded because they frequently arise due to thresholding grayscale images. All in all, these metrics emphasizes the correct prediction of object boundaries. In comparison, JANUSZEWSKI ET AL. are utilizing a different approach for evaluating FFN predictions, in which they defined a skeleton edge accuracy metric, which is based on skeletonized objects. This metric depends especially on correct, merged, split and ommited edges [Januszewski et al., 2016].

4.1.2 Pre- and post-processing

Pre- and post-processing allows a wide variety of operations to improve original EM images from the SNEMI3D dataset or obtained results after NN training. Many biomedical imaging methods require pre-processing image filtering methods due to poorer noise-to-signal ratios, low spatial resolutions, artefacts and low contrast between distinct anatomical structures [Behrenbruch et al., 2004]. The groundtruth dataset (SNEMI3D) used in this work also possesses varying brightness in consecutive images and artefacts in form of missing annotated pixel midst objects. For this reason, further trained NN used EM data with an applied local contrast enhancement filter *CLAHE* [Arici et al., 2006].
Furthermore, post-processing describes the improvement and optimization of images retrieved after prediction. Commonly used morphological operations include erosion, dilation, skeletonization, opening, closing, etc. [Busch and Eberle, 1995]. The application of tenfold erosion and following skeletonization to received overlap images of NN-2 and NN-4 (see Figure 3.5) led to an extreme improvement of the ARI illustrated in Figure 3.14 and reconstructible objects resembling the groundtruth objects. Unfortunately, the number of erosion steps used for all conducted experiments also led to extreme split error rates, because small overlapping regions in the dataset were shrank to extinction. Therefore, it could be better to establish an exception for the erosion, that overlapping regions can not be smaller than one pixel in total.
Another important issue includes the preparation of predicted cytoplasm and overlap images for connected components. For each predicted image, the NN returns a probability between 0 and 1 for each pixel, whether it is part of the background or cytoplasm and overlap respectively. Since the output is a grayscale image, ranging from 0 (black = background) to 255 (white = cytoplasm/overlap), and connected components, precision and recall require binary images for relabeling and evaluation, all pixels were thresholded and set to zero for pixel values below 127 ($P(Label) < 0.5$) and one for pixel values higher or equal 127 ($P(Label) >= 0.5$), in which label refers to cytoplasm,

overlap or membrane images. Thresholding images with different probabilities for true and false label can affect all evaluation metrics, thus requiring more investigation for all implemented experiments.

4.1.3 Experiments

The first conducted experiments (I, II, III) on the training and prediction of cytoplasm and overlap from EM images were performed without any post-processing. Experiment III, which included NN-1 and NN-2 for predicting both cytoplasm and overlap images turned out as the simplest and on the same time worst reconstruction of neuronal objects, resulting in one big merged object. Experiment I and II were either using NN-1 or NN-2 for cytoplasm or overlap prediction. Therefore, non NN predicted image stacks were generated through the groundtruth data. Despite producing less merge and split errors than experiment III, both experiments were unsuitable for dense reconstruction (see Table 3.3).
The next methodology included additional post-processing to predicted overlap images in experiments I and III, which improved the ARI of experiment I significantly, resulting in distinguishable objects. However, post -processing induced massive split errors and was not beneficial, if cytoplasm and overlap are trained and used for reconstruction (see Table 3.5). Experiments IV and V examined whether the training of overlap images from cytoplasm images is applicable. Evaluation results indicated extremely good reconstruction for trained overlap images from true cytoplasm labels (see experiment IV), illustrated in Table 3.6. If predicted cytoplasm serves as input for the training of overlap images, objects can not be separately reconstructed and are merged to one massive structure again (see experiment V).
Due to the very good prediction of overlap images from groundtruth cytoplasm in experiment IV, the prediction of cytoplasm and overlap from membrane images derived from the groundtruth dataset (see experiment VIII) was tested. Membrane images are object edges with one pixel width, representing the boundary of each object. The training and subsequent prediction of cytoplasm (NN-4) and overlap (NN-5) showed almost perfect segmentation, resulting in nearly perfect 3D reconstructed objects (see Table 3.7). A wide variety of approaches that include boundary detection have been already proposed [Arganda-Carreras et al., 2015] and presented in Table 4.1. My approach ranks right under the *"human 1 versus consensus"* and above all other competitors as well as *"human 2 versus consensus"*. Therefore, I extended this approach and trained an additional NN (NN-3) to predict membrane images from EM and use the predicted membrane images for further training of cytoplasm and overlap (see experiment IX). However, the quality of membrane predictions did not allow for correct prediction of cytoplasm and overlap, even if the consecutive networks were trained on predicted membranes (see experiment X). Therefore, membrane prediction is the crucial step in the proposed workflow for a complete 3D reconstruction of neurons from EM images.

Table 4.1: RAND scores of all competitors at ISBI with additional listing of experiment VIII calculated with the evaluation metric used by ARGANDA-CARRERAS ET AL. (table modified after [Arganda-Carreras et al., 2015])

Method	**RAND score**
human 1 vs. consensus	0.997
Experiment VIII (see section 3.5.3)	0.996
human 2 vs. consensus	0.971
IDSIA	0.944
BlackEagles	0.929
MLL-ETH	0.927
SCI	0.915
CellProfiler	0.904
Harvard	0.892
CoMPLEX	0.877
UCL	0.86
TSC+PP	0.843
IMMI	0.826
CLP	0.809
Freiburg	0.8
NIST	0.73

4.1.4 SNEMI3D annotations

As mentioned in the previous subsection, the cytoplasm and overlap prediction from SNEMI3D derived membrane groundtruth revealed almost perfect 3D visualization results. However, predicted membranes from EM images show conspicuous gaps midst membrane and are insufficient for further training. Furthermore, the original dataset was examined in more detail, which revealed inaccuracies in the annotated objects leading to deviations in derived membranes of the groundtruth labels (see Figure 3.21).
Therefore, I annotated a small portion of the dataset myself, trained NN-3 and received better visual predictions of membranes. Nevertheless, predicted membranes are still

unsatisfactory and contain a lot of unpredicted regions, which may distort cytoplasm and overlap prediction (see section 3.6 for in-depth information).

4.1.5 NN architectures and hyperparameter tuning

An additional step towards better membrane prediction includes the test of different NN architectures and loss functions for training, which might be able to overcome the used ResNet architecture with mean squared loss. URAKUBO ET AL. tested different loss functions for neuron, synapses and mitochondria prediction in EM images [Urakubo et al., 2019]. The three best working included mean squared error, logistic loss and dice coefficient. Therefore, logistic loss and dice coefficient were added to the testing regime with additional changes of hyperparameter including epochs, network depth and filter in the first convolutional layer (see Figure 3.25). The best training curves achieved mean squared loss and logistic loss, whereas the dice coefficient led to distorted predictions. Further testing should be dedicated to different network architectures such as U-Net [Ronneberger et al., 2015], Highway-Net [Srivastava et al., 2015] and DenseNet [Huang et al., 2016], which are heavily applied in bioimaging.

4.2 Computational costs

Another important aspect of this work was to minimize the computational cost needed for 3D reconstruction of neurons from EM images, including NN training and prediction, pre- and post-processing as well as reconstruction and relabeling. NN training yields the major percentage of computational costs by far. The NN training of FFN by JANUSZEWSKI ET AL. proceeded about seven days with a distributed setup of 32 NVIDIA Tesla K40 GPUs with 4290 GFLOPS per GPU [Januszewski et al., 2017]. My setup consists of one NVIDIA Tesla V100 with 15000 GFLOPS and an Intel Core i7-6700 3,4GHz CPU. The training of one of my NN with only one GPU takes around 1h (for 2000 epochs and 100 batch size) with subsequent prediction on CPU of less than 1 min. However, the FFN training on the best computational GPU currently available (NVIDIA Tesla V100 as of August 2019) would last at least two to three months. In addition, the NN training of FFN can last even longer, depending on how many objects belong to the current field of view (FOV) ranging from one up to several objects (two months corresponding to estimated 2 objects per FOV) [Januszewski et al., 2017]. Although FFN ranked first on the FIB-SEM (Drosophila ventral cord) challenge and currently second on the SNEMI3D challenge (mouse cortex), it does not imply overwhelming predictions on other test sets including neuronal structures, which may need additional training.
The prediction and relabeling of my workflow (<1min) and the prediction of FFN (7 min) are barely different in relation to the NN training. Another point, that could be of interest is the annotation of groundtruth necessary for training. The labeling of objects in the SNEMI3D dataset was described as "many hours" by KNOWLES-BARLEY ET AL.,

whereas the annotation of groundtruth data for FFN networks takes around two week labor [Urakubo et al., 2019]. Due to significantly lower computational cost of my NN training, it is also possible to train deeper ResNets. In this case, the depth of the neuronal network with more than eight residual blocks has no substantial effect on the training as well as the increase of epochs. A complete overview of computational costs is presented in Table 4.2.

Table 4.2: Comparison of computational costs with FFN

	Annotation	**Training**	**Prediction**	**Reconstruction**
FFN	∼two weeks	1-2 weeks with 32x Tesla K40 (∼2 months with 1x Tesla V100)	∼7 min on GPU	—
my approach	∼many hours	∼3h with 1x Tesla V100	<1 min on CPU	∼1 min

5 Conclusion

The main objective of this work, the investigation of 3D reconstruction of neurons from EM images with a Python-based command line software tool, was successful. This software package includes multiple scripts, each representing a different step in the software pipeline, such as pre-processing, NN training, reconstruction and relabeling and 3D visualization.
Initial testing with the SNEMI3D dataset showed that an interleaved image stack consisting of cytoplasm images in combination with overlap images is able to solve the so called *contour correspondence problem* in serial sectioning [Herbert and Jones, 2001] and is sufficient for the relabeling of original objects with connected components. Furthermore, many ways were introduced and used for the generation of ideal cytoplasm and overlap images from EM images using CNNs.
The first approach included the direct training to predict cytoplasm and overlap from original EM images. Precision and recall for the predicted labels reached values of around 0.95, indicating excellent predictions. Nevertheless, small errors in the segmentation resulted in completely merged objects. The best results indicated by few merge and split errors and a high ARI, was achieved with tenfold erosion plus skeletonization as post-processing step.
Since the prediction of objects was still not satisfactory, training to predict cytoplasm and overlap from SNEMI3D derived membrane images was performed. The results showed that the training with clear boundaries (membranes) is sufficient enough for the prediction of cytoplasm and overlap, reaching an ARI of 0.96. Therefore, an additional training was initialised to predict membrane images from EM first, with consecutive training of cytoplasm and overlap images to achieve a full automated processing pipeline. The results showed a better prediction of objects, than the direct prediction of cytoplasm and overlap from EM images. Thus, the next step included the optimization of membranes predicted by EM images.
Some of the SNEMI3D derived membranes may not be ideal for membrane prediction. Due to that assumption, membranes of ten EM images were manually labeled and used for membrane prediction, which showed visual improvement, but not perfectly predicted membranes.
Furthermore, some hyperparameters including number of epochs, filter in the first convolutional layer and number of residual blocks with three different loss functions, e.g mean square loss, logistic loss and dice coefficient, were tested to improve predicted membrane images even further. Mean square loss resulted in the best prediction, which was already used for previous predictions. More than 2,000 epochs, different filter in the first convolutional layer and an increasing number of residual blocks of more than eight have not led to any improvements of membrane prediction.
Nevertheless, a major advantage over the state-of-the-art algorithm FFNs is the computational cost needed for training. The training of FFNs takes around one to two weeks

with a custom server setup of 32 paralleled Tesla K40 GPUs. In comparison, the training of one NN in my pipeline takes only one hour with a Tesla V100 GPU. Moreover, the prediction of cytoplasm and overlap images from already trained NNs up until reconstruction and 3D visualization of reconstructed neurons takes only around ten minutes. In consideration of the large number of high resolution images generating image data in zetabyte range, fast and precise predictions are mandatory for future methods.

6 Future directions

The image pre- and post-processing pipeline, NN training and 3D reconstruction of neurons from EM images offers several ways for further improvement, that should be implemented, improved and investigated.
The next experiments should focus on the improvement and testing of different NN architectures. Already reviewed and suitable architectures for bioimaging include U-Nets [Ronneberger et al., 2015], Highway-Nets [Srivastava et al., 2015] and DenseNets [Huang et al., 2016]. Additionally, hyperparameter tuning of the currently used ResNet [He et al., 2015] could enhance the evaluation results of cytoplasm and overlap images predicted from SNEMI3D derived membranes even further, enabling an improved prediction of finer neuronal structures like dendrite spine necks.
Nevertheless, the main priority of NN training should be dedicated to membrane prediction from EM images to allow for an accurate, dense 3D reconstruction of neurons. In reference to SNEMI3D derived membranes, which are not optimized solely for membrane prediction, the NN training with self annotated membranes generated a better prediction of membranes. Therefore, it would be useful to annotate even more membranes of SNEMI3D EM images to increase the number of individual pictures for training.
In addition, membranes represent only one-seventh of the whole image, indicating a class imbalance, which can lead to poor predictions of the membranes. This could be prevented by higher cost sensitivities for the membrane class or adjusted threshold in favor of the minority class. The method *Large Margin Local Embedding* described by HUANG ET AL. could also yield improvements for the membrane prediction [Huang et al., 2018, Dong et al., 2018].
Furthermore, the Python-based pipeline offers scripts for image pre-processing, NN training, segmentation and 3D visualization, which are executed in a specific order in the command line including different parameters. Although this pipeline was designed for biologists, and contains scripts with precisely documented functions, it can be quite difficult to use for someone who has limited programming knowledge. Therefore, further software development should be oriented on a graphical user interface, including all steps from reading and pre-processing the groundtruth data to final 3D visualization of neurons, which ultimately can be integrated in the UNI-EM software presented by URAKUBO ET AL. [Urakubo et al., 2019]. In addition, the extension of developed scripts reading different data formats could come in handy for future work.
Moreover, it would be interesting to apply this pipeline to other datasets, e.g. from different species and imaging modalities.

Appendix A: Supplementary information

Table A.1: Evaluation results for experiment I with altered overlap objects after training

	adj. RAND Index	Precision	Recall	Merges	Splits
Part B	0.995	1.0	1.0	0	17
Experiment I	0.04	1.0	1.0	149	117
edges decreased by 1 pixel	0.043	1.0	1.0	141	232
edges decreased by 2 pixel	0.054	1.0	1.0	125	340
edges decreased by 3 pixel	0.088	1.0	1.0	103	447
edges decreased by 4 pixel	0.18	1.0	1.0	81	600
edges decreased by 5 pixel	0.24	1.0	1.0	73	831
edges decreased by 6 pixel	0.27	1.0	1.0	58	1144
edges decreased by 7 pixel	0.34	1.0	1.0	48	1516
edges decreased by 8 pixel	0.41	1.0	1.0	44	1934
edges decreased by 9 pixel	0.54	1.0	1.0	32	2394
edges decreased by 10 pixel	0.80	1.0	1.0	29	2867
edges decreased by 11 pixel	0.82	1.0	1.0	24	3282
edges decreased by 12 pixel	0.86	1.0	1.0	18	3662
edges decreased by 13 pixel	0.86	1.0	1.0	15	4010
edges decreased by 14 pixel	0.87	1.0	1.0	13	4302

	adj. RAND Index	Precision	Recall	Merges	Splits
edges decreased by 15 pixel	0.92	1.0	1.0	11	4568
edges decreased by 16 pixel	0.94	1.0	1.0	6	4800
edges decreased by 17 pixel	0.93	1.0	1.0	5	4998
edges decreased by 18 pixel	0.93	1.0	1.0	5	5166
edges decreased by 19 pixel	0.94	1.0	1.0	4	5319
edges decreased by 20 pixel	0.93	1.0	1.0	2	5463

Appendix B: Poster

A poster (see Figure B.1) has been designed and presented during the Mittelerde Meeting 2019 (4th Central German Meeting on Bioinformatics) in Dresden at BIOTEC on June 13, 2019.

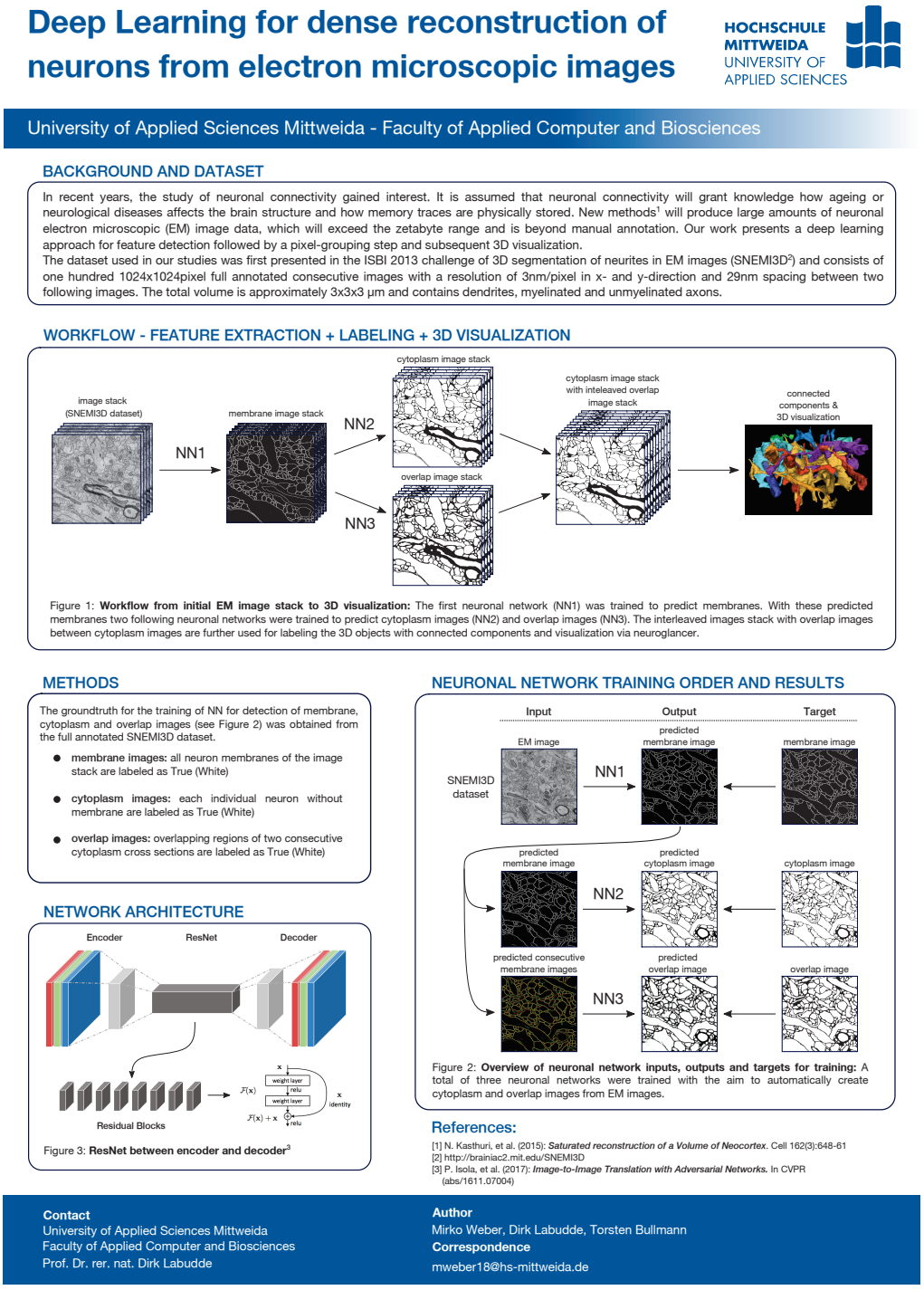


Figure B.1: Poster - 4th Central German Meeting on Bioinformatics

Appendix C: CD content

The enclosed CD contains the following content:

- thesis in pdf format
- poster in pdf format
- tensorflow graph of the used neuronal network as png

Bibliography

[Arbeláez et al., 2011] Arbeláez, P., Maire, M., Fowlkes, C., and Malik, J. (2011). Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5):898–916.

[Arganda-Carreras et al., 2012] Arganda-Carreras, I., Seung, S., Cardona, A., and Schindelin, J. (2012). ISBI Challenge: Segmentation of neuronal structures in EM stacks. `http://brainiac2.mit.edu/isbi_challenge/home`. [Online; accessed 12-August-2019].

[Arganda-Carreras et al., 2013] Arganda-Carreras, I., Seung, S., Vishwanathan, A., and Berger, D. R. (2013). SNEMI3D: 3D Segmentation of neurites in EM images. `http://brainiac2.mit.edu/SNEMI3D/`. [Online; accessed 09-August-2019].

[Arganda-Carreras et al., 2015] Arganda-Carreras, I., Turaga, S. C., Berger, D. R., Cireşan, D., Giusti, A., Gambardella, L. M., Schmidhuber, J., Laptev, D., Dwivedi, S., Buhmann, J. M., Liu, T., Seyedhosseini, M., Tasdizen, T., Kamentsky, L., Burget, R., Uher, V., Tan, X., Sun, C., Pham, T. D., Bas, E., Uzunbas, M. G., Cardona, A., Schindelin, J., and Seung, H. S. (2015). Crowdsourcing the creation of image segmentation algorithms for connectomics. *Frontiers in Neuroanatomy*, 9.

[Arici et al., 2006] Arici, T., Dikbas, S., and Altunbasak, Y. (2006). Local contrast enhancement using 2-dimensional recursive filters. In *2006 IEEE Workshop on Multimedia Signal Processing*. IEEE.

[Behrenbruch et al., 2004] Behrenbruch, C. P., Petroudi, S., Bond, S., Declerck, J. D., Leong, F. J., and Brady, J. M. (2004). Image filtering techniques for medical image post-processing: an overview. *The British Journal of Radiology*, 77(suppl_2):S126–S132.

[Berger et al., 2018] Berger, D. R., Seung, H. S., and Lichtman, J. W. (2018). VAST (volume annotation and segmentation tool): Efficient manual and semi-automatic labeling of large 3d image stacks. *Frontiers in Neural Circuits*, 12.

[Briggman and Bock, 2012] Briggman, K. L. and Bock, D. D. (2012). Volume electron microscopy for neuronal circuit reconstruction. *Current Opinion in Neurobiology*, 22(1):154–161.

[Busch and Eberle, 1995] Busch, C. and Eberle, M. (1995). Morphological operations for color-coded images. *Computer Graphics Forum*, 14(3):193–204.

[Cardona et al., 2010] Cardona, A., Saalfeld, S., Preibisch, S., Schmid, B., Cheng, A., Pulokas, J., Tomancak, P., and Hartenstein, V. (2010). An integrated micro- and macroarchitectural analysis of the drosophila brain by computer-assisted serial section electron microscopy. *PLoS Biology*, 8(10):e1000502.

[Cardona et al., 2012] Cardona, A., Saalfeld, S., Schindelin, J., Arganda-Carreras, I., Preibisch, S., Longair, M., Tomancak, P., Hartenstein, V., and Douglas, R. J. (2012). TrakEM2 software for neural circuit reconstruction. *PLoS ONE*, 7(6):e38011.

[Ciresan et al., 2012] Ciresan, D., Giusti, A., Gambardella, L. M., and Schmidhuber, J. (2012). Deep neural networks segment neuronal membranes in electron microscopy images. In Pereira, F., Burges, C. J. C., Bottou, L., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems 25*, pages 2843–2851. Curran Associates, Inc.

[Denk and Horstmann, 2004] Denk, W. and Horstmann, H. (2004). Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLoS Biology*, 2(11):e329.

[Denton et al., 2015] Denton, E., Chintala, S., Szlam, A., and Fergus, R. (2015). Deep generative image models using a laplacian pyramid of adversarial networks.

[Dong et al., 2018] Dong, Q., Gong, S., and Zhu, X. (2018). Imbalanced deep learning by minority class incremental rectification.

[Eraslan et al., 2019] Eraslan, G., Avsec, Ž., Gagneur, J., and Theis, F. J. (2019). Deep learning: new computational modelling techniques for genomics. *Nature Reviews Genetics*, 20(7):389–403.

[Goodfellow et al., 2017] Goodfellow, I., Bengio, Y., and Courville, A. (2017). *Deep Learning*. The MIT Press.

[Goodfellow et al., 2014] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial networks.

[Greenspan et al., 2016] Greenspan, H., van Ginneken, B., and Summers, R. M. (2016). Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Transactions on Medical Imaging*, 35(5):1153–1159.

[Harris et al., 2006] Harris, K. M., Perry, E., Bourne, J., Feinberg, M., Ostroff, L., and Hurlburt, J. (2006). Uniform serial sectioning for transmission electron microscopy. *Journal of Neuroscience*, 26(47):12101–12103.

[Hayworth et al., 2006] Hayworth, K., Kasthuri, N., Schalek, R., and Lichtman, J. (2006). Automating the collection of ultrathin serial sections for large volume TEM reconstructions. *Microscopy and Microanalysis*, 12(S02):86–87.

[He et al., 2015] He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition.

[Helmstaedter et al., 2011] Helmstaedter, M., Briggman, K. L., and Denk, W. (2011). High-accuracy neurite reconstruction for high-throughput neuroanatomy. *Nature Neuroscience*, 14(8):1081–1088.

[Herbert and Jones, 2001] Herbert, M. J. and Jones, C. B. (2001). Contour correspondence for serial section reconstruction: complex scenarios in palaeontology. *Computers & Geosciences*, 27(4):427–440.

[Hinton et al., 2012] Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A.-r., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Kingsbury, B., and Sainath, T. (2012). Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal Processing Magazine*, 29:82–97.

[Hopcroft and Tarjan, 1973] Hopcroft, J. and Tarjan, R. (1973). Algorithm 447: efficient algorithms for graph manipulation. *Communications of the ACM*, 16(6):372–378.

[Huang et al., 2018] Huang, C., Li, Y., Loy, C. C., and Tang, X. (2018). Deep imbalanced learning for face recognition and attribute prediction.

[Huang et al., 2016] Huang, G., Liu, Z., van der Maaten, L., and Weinberger, K. Q. (2016). Densely connected convolutional networks.

[Hubert and Arabie, 1985] Hubert, L. and Arabie, P. (1985). Comparing partitions. *Journal of Classification*, 2(1):193–218.

[Hwang and Liu, 2015] Hwang, J.-J. and Liu, T.-L. (2015). Pixel-wise deep learning for contour detection.

[Isola et al., 2017] Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE.

[Januszewski et al., 2017] Januszewski, M., Kornfeld, J., Li, P. H., Pope, A., Blakely, T., Lindsey, L., Maitin-Shepard, J., Tyka, M., Denk, W., and Jain, V. (2017). High-precision automated reconstruction of neurons with flood-filling networks.

[Januszewski et al., 2018] Januszewski, M., Kornfeld, J., Li, P. H., Pope, A., Blakely,

T., Lindsey, L., Maitin-Shepard, J., Tyka, M., Denk, W., and Jain, V. (2018). High-precision automated reconstruction of neurons with flood-filling networks. *Nature Methods*, 15(8):605–610.

[Januszewski et al., 2016] Januszewski, M., Maitin-Shepard, J., Li, P., Kornfeld, J., Denk, W., and Jain, V. (2016). Flood-filling networks. *CoRR*, abs/1611.00421.

[Jones et al., 2001] Jones, E., Oliphant, T., Peterson, P., et al. (2001). SciPy: Open source scientific tools for Python. [Online; accessed 20-August-2019].

[Kasthuri et al., 2015] Kasthuri, N., Hayworth, K. J., Berger, D. R., Schalek, R. L., Conchello, J. A., Knowles-Barley, S., Lee, D., Vázquez-Reina, A., Kaynig, V., Jones, T. R., Roberts, M., Morgan, J. L., Tapia, J. C., Seung, H. S., Roncal, W. G., Vogelstein, J. T., Burns, R., Sussman, D. L., Priebe, C. E., Pfister, H., and Lichtman, J. W. (2015). Saturated reconstruction of a volume of neocortex. *Cell*, 162(3):648–661.

[Knott et al., 2008] Knott, G., Marchman, H., Wall, D., and Lich, B. (2008). Serial section scanning electron microscopy of adult brain tissue using focused ion beam milling. *Journal of Neuroscience*, 28(12):2959–2964.

[LeCun et al., 2015] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444.

[Lichtman et al., 2014] Lichtman, J. W., Pfister, H., and Shavit, N. (2014). The big data challenges of connectomics. *Nature Neuroscience*, 17(11):1448–1454.

[Nunez-Iglesias et al., 2014] Nunez-Iglesias, J., Kennedy, R., Plaza, S. M., Chakraborty, A., and Katz, W. T. (2014). Graph-based active learning of agglomeration (GALA): a python library to segment 2d and 3d neuroimages. *Frontiers in Neuroinformatics*, 8.

[Pedregosa et al., 2011] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

[Rand, 1971] Rand, W. M. (1971). Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, 66(336):846–850.

[Ronneberger et al., 2015] Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation.

[Srivastava et al., 2015] Srivastava, R. K., Greff, K., and Schmidhuber, J. (2015). Highway networks.

[Szegedy et al., 2013] Szegedy, C., Toshev, A., and Erhan, D. (2013). Deep neural networks for object detection. In *NIPS*.

[Takemura et al., 2015] Takemura, S., Xu, C. S., Lu, Z., Rivlin, P. K., Parag, T., Olbris, D. J., Plaza, S., Zhao, T., Katz, W. T., Umayam, L., Weaver, C., Hess, H. F., Horne, J. A., Nunez-Iglesias, J., Aniceto, R., Chang, L.-A., Lauchie, S., Nasca, A., Ogundeyi, O., Sigmund, C., Takemura, S., Tran, J., Langille, C., Lacheur, K. L., McLin, S., Shinomiya, A., Chklovskii, D. B., Meinertzhagen, I. A., and Scheffer, L. K. (2015). Synaptic circuits and their variations within different columns in the visual system ofDrosophila. *Proceedings of the National Academy of Sciences*, 112(44):13711–13716.

[Turaga et al., 2010] Turaga, S. C., Murray, J. F., Jain, V., Roth, F., Helmstaedter, M., Briggman, K., Denk, W., and Seung, H. S. (2010). Convolutional networks can learn to generate affinity graphs for image segmentation. *Neural Computation*, 22(2):511–538.

[Urakubo et al., 2019] Urakubo, H., Bullmann, T., Kubota, Y., Oba, S., and Ishii, S. (2019). UNI-EM: An environment for deep neural network-based automated segmentation of neuronal electron microscopic images.

[Zlateski and Seung, 2015] Zlateski, A. and Seung, H. S. (2015). Image segmentation by size-dependent single linkage clustering of a watershed basin graph. *CoRR*, abs/1505.00249.

Glossary

adjusted RAND index Corrected-for-chance version of the RAND index.

ATUM-SEM Automatic tape-collecting ultramicrotome scanning electron microscope producing image resolutions of 3x3x29 nm.

connected components A sub graph of an undirected graph, in which all vertices are connected and labeled individually.

Connectomics A research area of neurosciences, which is dedicated to the mapping and visualization of all neuronal structures within an organism.

convolutional neuronal network By YANN LECUN established form of a NN consisting of convolutional and pooling layer especially suited for processing image- and audio data.

cytoplasm images Binary images representing all objects without the outmost pixel border of each object as true and the remainder as false.

deep convolutional neuronal network CNNs with a high number of convolutional and pooling layers.

DenseNet DenseNets use the same architecture as ResNets with adding skip connections from every previous layer to the actual layer.

FIB-SEM Focused ion beam scanning electron microscope producing maximal image resolutions of 5x5x5 nm.

Flood-filling networks State-of-the-art neuron segmentation algorithm with the use of 3D CNNs.

FLOPS Measurement for computational performance of a computer. Often used for GPU comparison.

Git A free version control system for non-linear software development.

HighwayNet The HighwayNet is based upon a ResNet with a additional learning parameter, which decides whether skip connections should be used or not.

Hyperparameter Hyperparameters in NN are parameters, which need to be defined before training (such as epochs, batch size, loss function, etc.).

membrane images Binary images representing the outmost pixel border of each individual object as true and the remainder as false.

merge errors Two or more individual labels in the groundtruth dataset, that are combined to one label in the predicted dataset.

Neuroglancer A WebGL-based viewer for volumetric data.

neuronal network A NN consists of a variety of calculations to learn relationships in a given set of data inspired by the composition of the human brain.

overlap images Binary images representing the cross-sections of all objects in two consecutive images as true and the remainder with false.

precision A performance measurement of binary classification. The precision describes the fraction of relevant instances among retrieved instances.

RAND index Measurement often used in data clustering to measure the similarity of two cluster.

recall A performance measurement of binary classification. The recall describes the fraction of relevant instances that have been retrieved over the total amount of relevant instances.

ResNet Residual networks are NNs containing skip connections between convolutional layer with rectified linear units activation functions allowing to train deep NN with less computational cost.

sBEM Serial block-face scanning electron microscope producing image resolutions of 4x4x45 nm.

split errors One label in the groundtruth dataset, which is divided into two or more individual labels in the predicted dataset.

ssTEM Serial section transmission electron microscope producing image resolutions of 12x12x25 nm.

U-Net U-Net is a CNN especially for bioimage segmentation. It consist of an encoder and decoder, whereas the encoder consists of a typical CNN and the decoder combines spatial information and features through up-convolutions.

Erklärung

Hiermit erkläre ich, dass ich meine Arbeit selbstständig verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und die Arbeit noch nicht anderweitig für Prüfungszwecke vorgelegt habe.

Stellen, die wörtlich oder sinngemäß aus Quellen entnommen wurden, sind als solche kenntlich gemacht.

Mittweida, im August 2019