CSc 84020 Neural Networks and Deep Learning

Project 1, Part 1

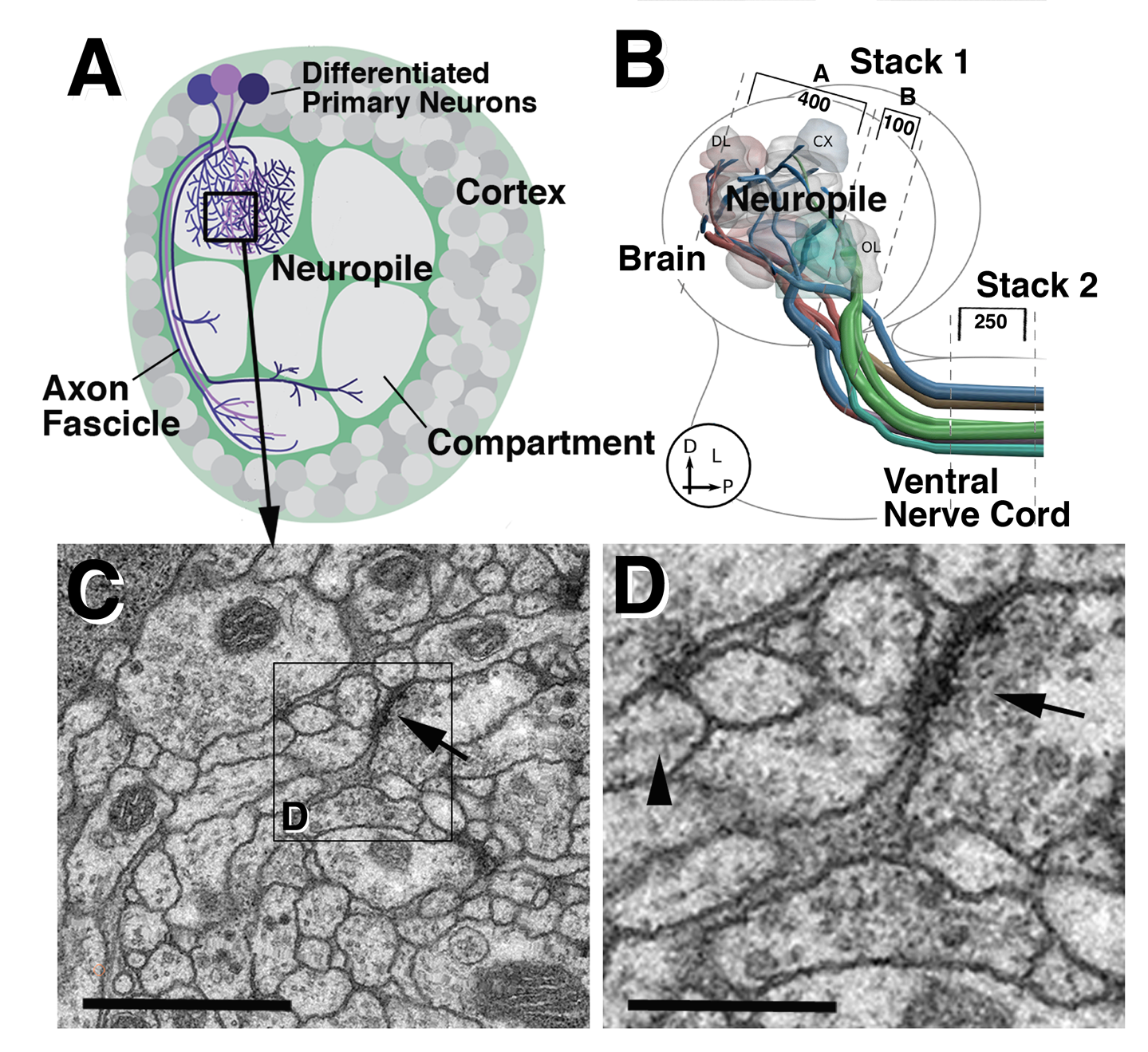
Andrea Ceres

**Project Description**

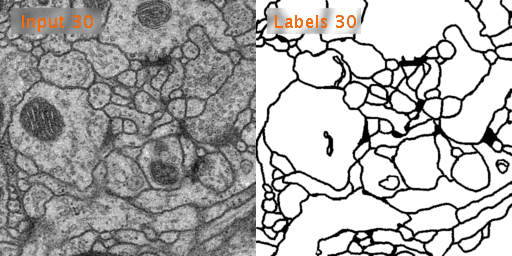
This project is an implementation of the original U-Net architecture first described in the 2015 International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI):

Ronneberger O., Fischer P., Brox T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4\_28

Unlike many other computer vision applications, biomedical image processing tasks are often limited by dataset size. The U-Net architecture introduced by Ronneberger et al. is a modification of the fully convolutional network with the expressed goals to better train using very few images and to yield more precise segmentations. In the first of two competitions entered by the authors, their model utilizing the U-Net architecture outperformed the prior leading model consisting of a sliding-window convolutional network on the ISBI challenge for segmentation of neuronal structures in electron microscopic (EM) stacks.



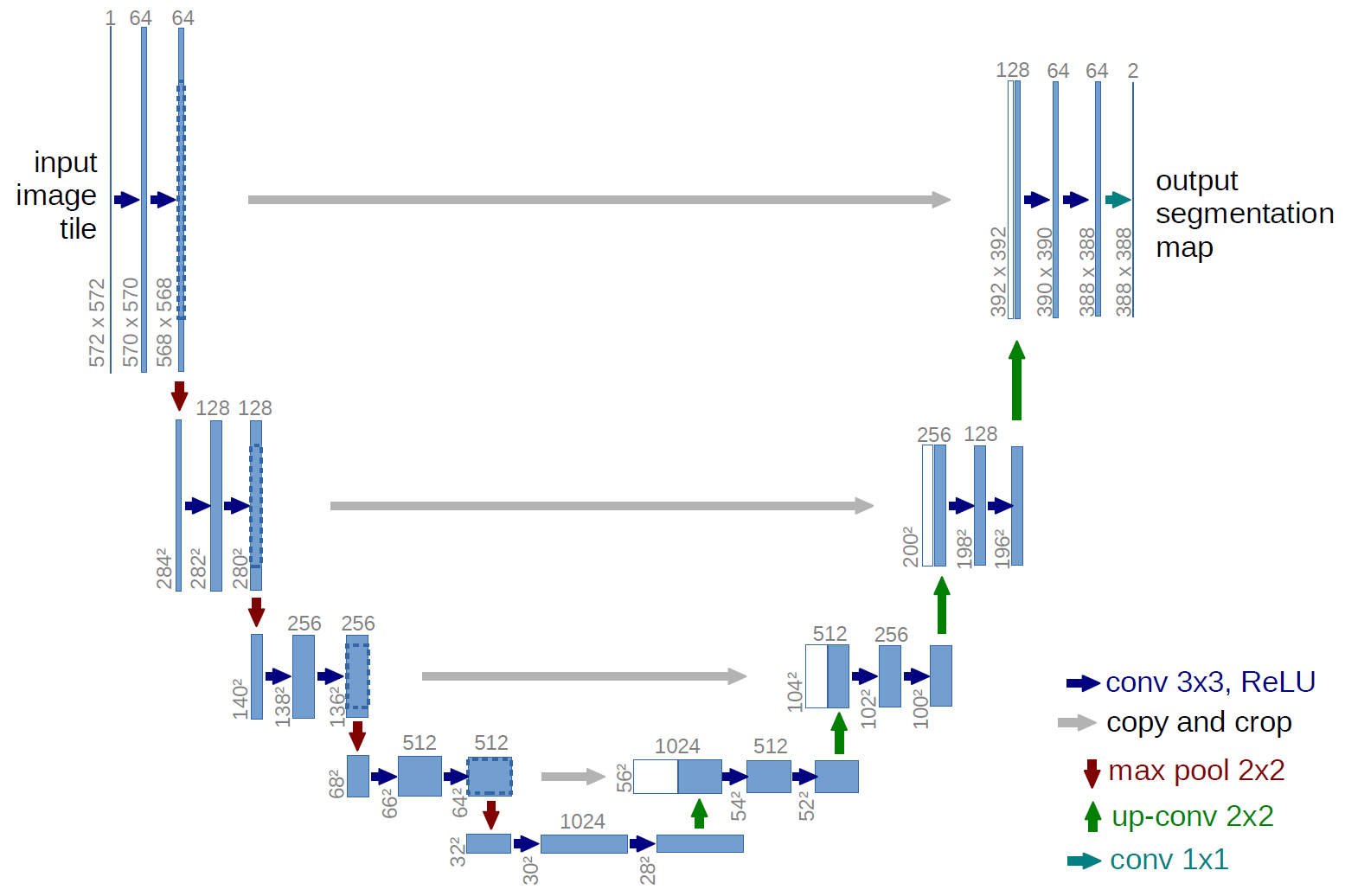
**Data**

The training data used for the first competition is taken from 30 serial section Transmission Electron Microscopy (ssTEM) images of *Drosophila* (fruit fly) first instar larva ventral nerve cord (VNC), illustrated in *Figure 1*. The corresponding labels reflect the segmentation of the membranes, whereby the white pixels represent the cell interiors and the black pixels delineate the membranes. An example is shown in *Figure 2*. Test data consists of another 30 images from the same *Drosophila* first instar larva VNC as the training data. Each image is 512x512 pixels with 3 RGB channels, although the images are, by and large, grayscale. This project utilizes the same datasets, but at 256x256 pixel resolution. This dataset and code are taken from the most popular TensorFlow implementation of U-Net on Papers with Code: https://github.com/zhixuhao/unet



**Methodology/Algorithm**

The U-Net architecture (*Figure 3*) is an extension of the fully convolutional network with 23 convolutional layers and no fully connected layers. A contracting network using pooling operators is followed by successive layers using upsampling operators, thereby increasing the resolution of the output. Localization is established by the concatenation of high resolution features from the contracting path with the expansive, upsampled path. The expansive path is approximately symmetric to the contracting path, giving a U-shape that lends itself to the naming of this architecture as U-Net. Given the limited training data available as is the case generally for many medical image datasets, data augmentation is performed via realistic elastic deformations, shifts, rotations, and gray value variations. Further implicit augmentation occurs from contracting path drop-out layers.



The contracting path consists of repeated application of two 3x3 unpadded convolutions, each followed by ReLU activation and downsampling by 2x2 max pooling operation with stride of 2 and doubling of the number of feature channels.

The expansive path consists of repeated upsampling of the feature map, each followed by a 2x2 up-convolution that halves the number of feature channels, a concatenation described above with the corresponding cropped feature map from the contracting path, and two 3x3 convolution and ReLU paired operations. During the upsampling, only the valid parts are used, and missing input data around the edges are extrapolated by mirroring. A 1x1 convolution is used at the last layer in order to map each 64-component feature vector to the desired number of classes. For this segmentation, the number of classes are 2 (white and black pixels).

Batch size is minimized to 1 to allow for larger input tiles. Stochastic gradient descent with a high momentum of 0.99 is implemented as the optimizer during training, so that each update is determined by a large number of previously seen training samples. Cross entropy is minimized in the loss function, and soft-max is used at the pixel level. Class frequencies are balanced with a pre-computed weight map that allows the network to learn the small separation borders between touching cells. These initial weights are drawn from a Gaussian distribution with a standard deviation of √2/N, where N is the number of incoming nodes of one neuron.

**Approach**

With code adapted from GitHub from the current most popular TensorFlow implementation of U-Net on Papers with Code, I attempt to recreate the U-Net architecture. The code does not follow the described topology precisely, but is a worthy approximation.

**Report**

Segmentation in the original paper is ascertained by thresholding the predicted probability map and judged by the computation of the warping error, the Rand error, and the pixel error. More exploration is needed to determine the relationship to the accuracy metric utilized in the TensorFlow architecture. Comparisons are difficult to make as a result.

**References**

[1] Ronneberger O., Fischer P., Brox T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham.

[2] Ignacio Arganda-Carreras, Srinivas C. Turaga, Daniel R. Berger, Dan Ciresan, Alessandro Giusti, Luca M. Gambardella, Jürgen Schmidhuber, Dmtry Laptev, Sarversh Dwivedi, Joachim M. Buhmann, Ting Liu, Mojtaba Seyedhosseini, Tolga Tasdizen, Lee Kamentsky, Radim Burget, Vaclav Uher, Xiao Tan, Chanming Sun, Tuan D. Pham, Eran Bas, Mustafa G. Uzunbas, Albert Cardona, Johannes Schindelin, and H. Sebastian Seung. Crowdsourcing the creation of image segmentation algorithms for connectomics. Frontiers in Neuroanatomy, vol. 9, no. 142, 2015.

[3] Albert Cardona, Stephan Saalfeld, Stephan Preibisch, Benjamin Schmid, Anchi Cheng, Jim Pulokas, Pavel Tomancak and Volker Hartenstein (10, 2010), "An Integrated Micro- and Macroarchitectural Analysis of the Drosophila Brain by Computer-Assisted Serial Section Electron Microscopy", PLoS Biol (Public Library of Science) 8 (10): e1000502, doi:10.1371/journal.pbio.1000502