

Project 4

Crumpler

Strategy

- Pre-processing
 - NMF
 - Other strategies
- Convolutional Neural Network (UNET)
- Grid Search
- Post-processing

Pre-Processing with NMF

- Dataset: m many images, each 512×512
- Flatten and stack to get one $m \times 512 \times 512$
- NMF_k returns two matrices:
 - W is $m \times k$
 - H is $k \times 512 \times 512$
- The H matrix may be used to directly estimate the ROIs
 - Thunder Extraction is particularly useful
- It also provides a “feature space” for the pixels
 - Each pixel now has k many features that describe it

Pre-Processing with Other Strategies

- **Average**

- Calculate the average value for each pixel across the time series
- Gives you one new image to use

- **Displacement**

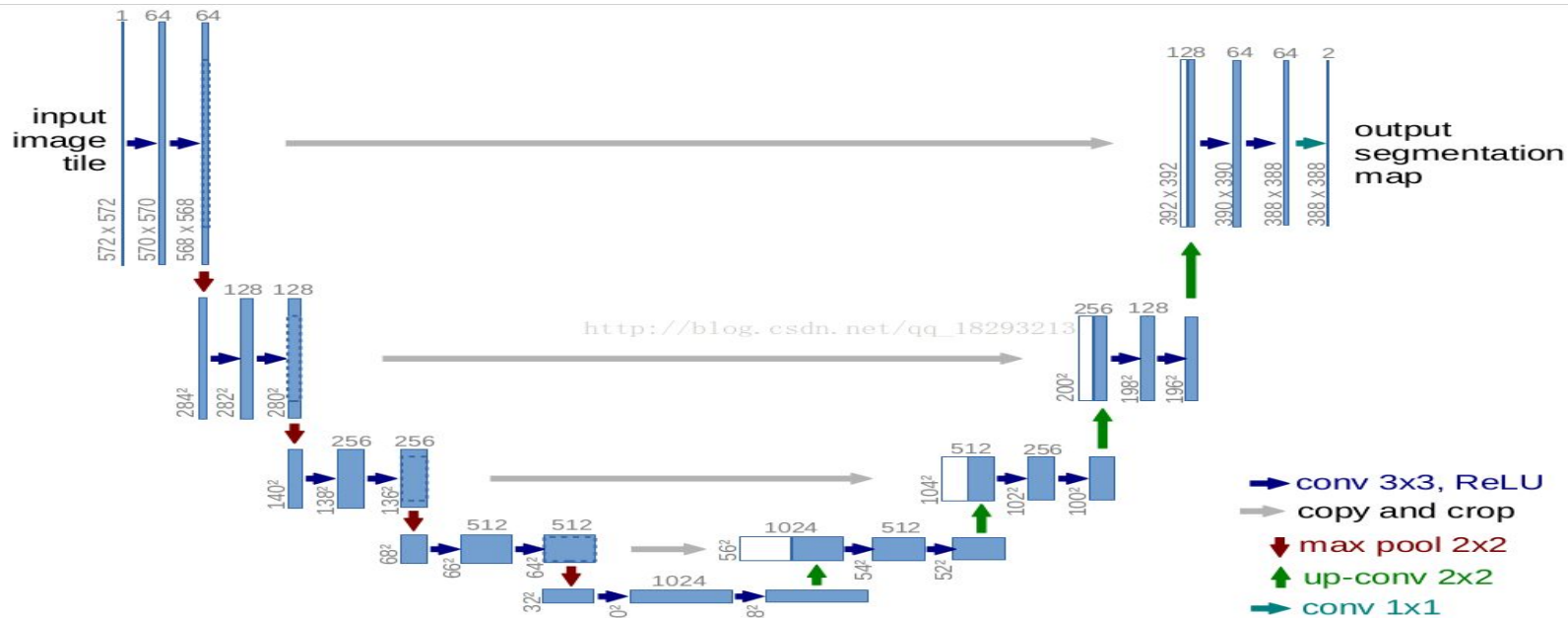
- Calculate the difference between the max and min value for each pixel across the time series
- Gives you another new image to use

- **Random**

- Snag random images across the time series to use as your training data

- The average, displacement, and random images for each time series may be used as input into a CNN for training data (probably best to use all...)

UNET Architecture(credits: <http://blog.csdn.net>)



Ensemble vs Individual

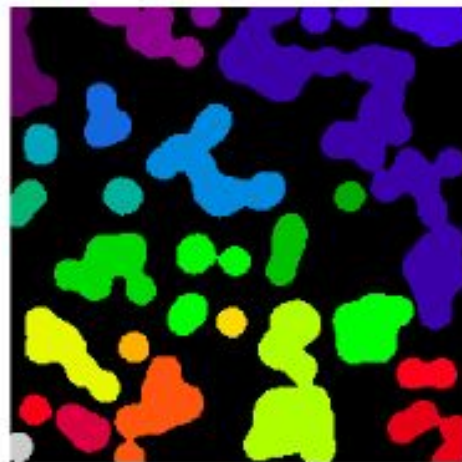
- Trained a separate UNET model on each cluster from NMF
- Combined output with simple voting to label pixels
- Also trained a single UNET model with all clusters stacked
- Individual had superior performance

Hyperparameter Tuning

- Model contains many parameters that affect training
 - learning rate, optimizer, batch size
 - weight of positives labels vs negatives
 - type / combination of preprocessing
- Configured training script to accept arguments for these
- Wrote bash script to run with different combinations of arguments
- Checkpointed each model and saved metrics to CSV
 - enables easy comparison of different models

Segmentation

- UNET outputs an image of 1s and 0s, representing neurons and non-neurons
- We need individual regions
- Scikit-image connected components algorithm
- Works well, but fails on overlapping regions



Limitations

TIME