601.461/661 Computer Vision

Segmentation of Subcortical Brain Structures with Convolutional Neural Networks

Li, Wang, Ye, Zhang

Introduction

Goal: Our project aims to perform 3D subcortical brain segmentations of open access MRI datasets through the use of a convolutional neural network (CNN).

Project significance:

- Speed, efficiency, accuracy over manual segmentations
- Aid in study of brain disorders

Individual contributions:

- Background research: Bryan/Judy/Jason/Will
- IBSR data acquisition and initial processing: Judy
- CNN-specific data pre-processing: Jason
- CNN implementation: Jason/Bryan
- F-CNN adaptation: Judy
- Performance evaluations: Will
- Report and presentation: Bryan/Judy/Jason/Will

Data Processing

NITRIC Internet Brain Segmentation Repository Dataset:

- 18 T1-weighted volumetric brain scans, health patients
- 14 males, 4 females
- Manual segmentations for 43 individual structures of principle grey and white matter
- Training set: IBSR patients 1-10
- Testing set: IBSR patients 11-18
- Normalized image resolution: 0.9375 x 1.5 x 0.9375 mm
- Consistent matrix size: 256 x 128 x 256

Final input data:

- File type conversion
- Bias field correction
- Image resolution normalization
- Skull-stripping

CNN Pre-Processing:

- Cropped image scans so that we could save on memory due to the fact that we had a limitation on the amount of RAM we had.
- Normalized our data such that the intensities were zero mean so that our neural net would train better
- In order just to get certain regions of the brain, we took the training data set and changed some of the labels of the ground truths such that all labels that were not relevant to our ROI we set to 0

CNN Architecture

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Eq 1: Parametric Rectified Linear Unit (PReLU)

$$f(x_i) = max(0, x_i) + a_i * min(0, x_i)$$

- $f(x_i)$: Output
- x_i: Input
- a_i: Coefficient that improves the adaptiveness of our PReLU for accuracy

Eq 2: Softmax function for voxel classification

$$\sigma(\mathbf{z})_\mathbf{j} = \frac{\mathbf{e}^{\mathbf{x_j}}}{\sum_{\mathbf{k=1}}^{\mathbf{K}} \mathbf{e}^{\mathbf{x_j}}}$$

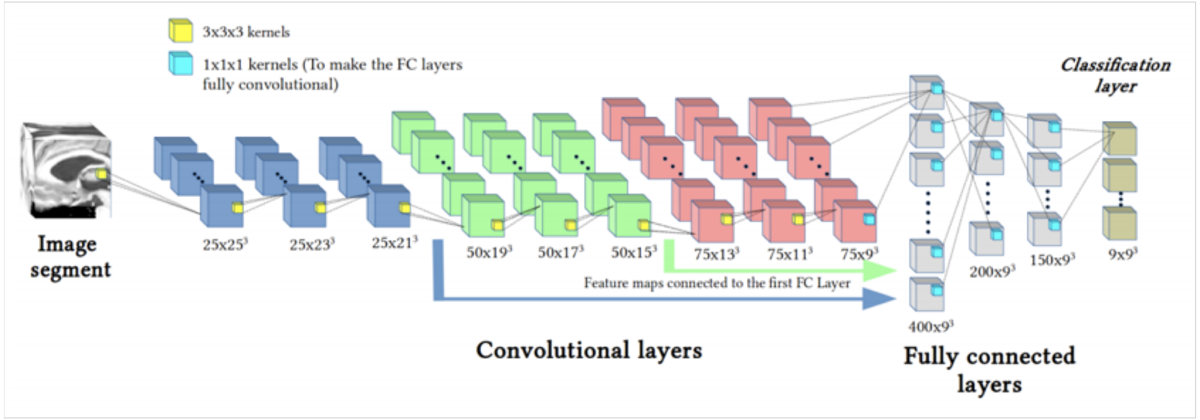


Fig. 1: Convolutional neural network proposed by J. Dolz

CNN Implementation

- First 9 layers project increasingly more complex features to each of the feature maps
- Repeat layers of the same filter size helps us attain a deeper architecture to learn more complex feature
- Fully Connected layers are fed the feature maps from the other convolution layers and allow us to learn a combination of these features to form the complex shape of the cortical regions
- Use softmax to classify into 7 classes(6 subcortical, 1 else)

Fig. 2: Subcortical segmentation produced by our CNN

Alternative F-CNN

Table I.
F-CNN architecture: Overview of layers and important parameters

Block	conv kernel	# filters	hole stride	pool kernel	pool stride	dropout
1	7×7	64	1	3×3	2	no
2	5×5	128	1	3×3	2	no
3	3×3	256	2	3×3	1	yes
4	3×3	512	2	3×3	1	yes
5	3×3	512	2	3×3	1	yes
6	4×4	1024	4	no pooling		yes
7	1×1	39	1	no pooling		no

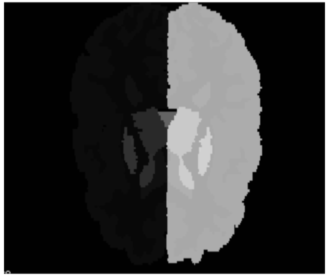
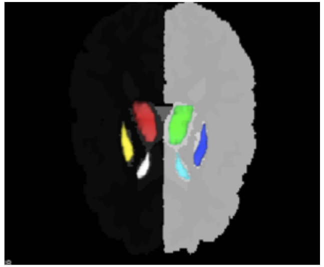




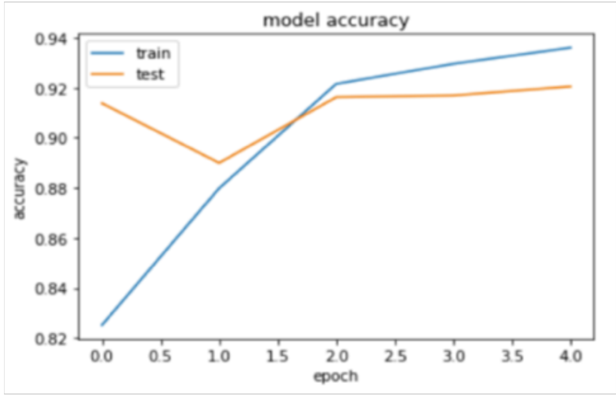
Fig. 3: Pre-processed, skull-stripped IBSR image data, patient 01, scan 138
Fig. 4: Manual segmentations of L./R. Caudate, L./R. Putamen, L./R. Thalamus
Fig. 5: F-CNN results overlayed on manual segmentations

Performance Evaluation

JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Table II.
CNN Optimized Training Parameters

ADAM Optimizer Parameter	Resulting Accuracy
0.1	40%
0.001	70%
0.0001	93%
0.00001	8%



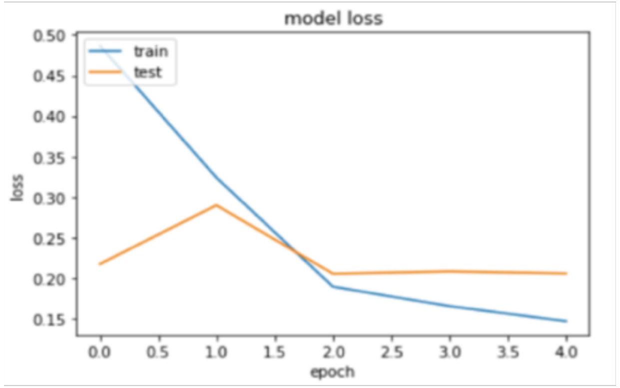


Fig. 6: Graphical representation of model accuracy vs. epoch number
Fig. 7: Graphical representation of model loss vs epoch number

Discussion

- Varied our CNN by changing kernel size
  - Larger meant more required memory for weights/more computation time for back propagation
  - Smaller meant drop in accuracy due to detection of smaller features only.
- CNN performed poorly when we experimented with training on all 43 individual structures
- Filter size provided limited feature detection (an area for improvement)
- Can improve network performance by increasing the number of convolutional layers to deepen network
  - While this would boost performance by learning more complex, non-linear behaviors, it will pose extensive training difficulty.

Conclusion:

- We were able to accurately segment subcortical structures in brain MRI scans, and is an effective/efficient approach for brain segmentation.
- Limitations include trouble with detection of more structures, system and memory limitations so we couldn’t define more structures.
- Hopefully with better infrastructure and deeper neural netowkr, segmentation is more accurate and better optimized.