

# Multimodal Brain MRI Tumor Segmentation via Convolutional Neural Networks

## Introduction

- Gliomas are the most commonly occurring type of brain tumors and are potentially very dangerous [2, 1], with about 90% of gliomas belonging to a highly aggressive class of cancerous tumors known as glioblastomas [4].
- Multimodality magnetic resonance imaging is the primary method of screening and diagnosis for gliomas [4].
- Tumor regions currently are segmented manually by radiologists.
- Segmentation algorithms based on convolutional neural networks (CNNs) have been shown to be at least as effective as other automated segmentation methods.

## Architecture Overview

### Baseline Voxel-wise CNN

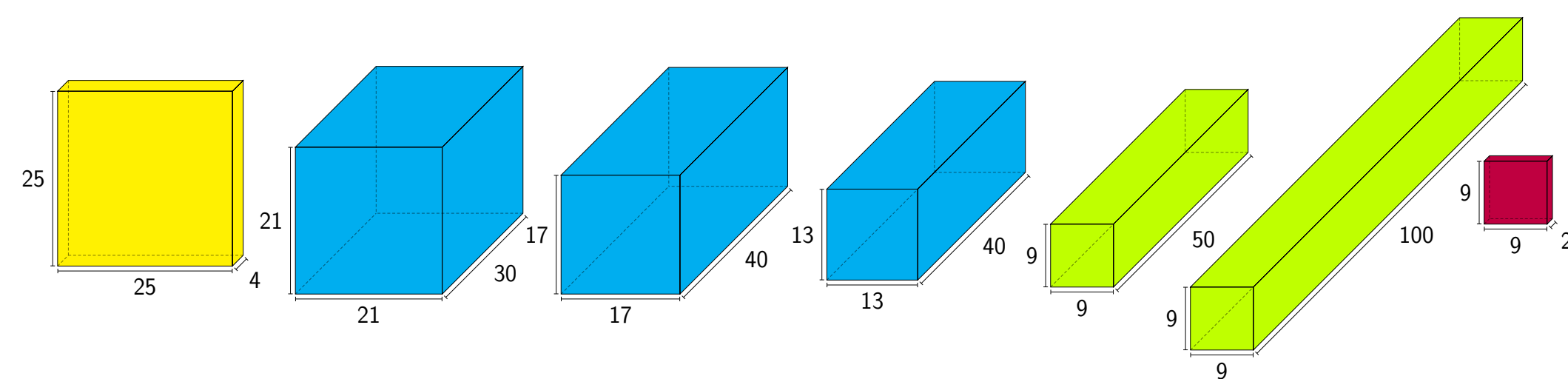


Figure 1: Baseline Architecture Diagram. Input is  $25^3 \times 4$  volume (single slice shown). Convolutional layers (cyan) followed by ReLU activation. Fully connected layers (green) implemented as  $1^3$ -kernel convolutions; first layer is followed by ReLU and dropout. Scores (red) for background and foreground.

### Patch-wise Fully-Convolutional Network Architecture

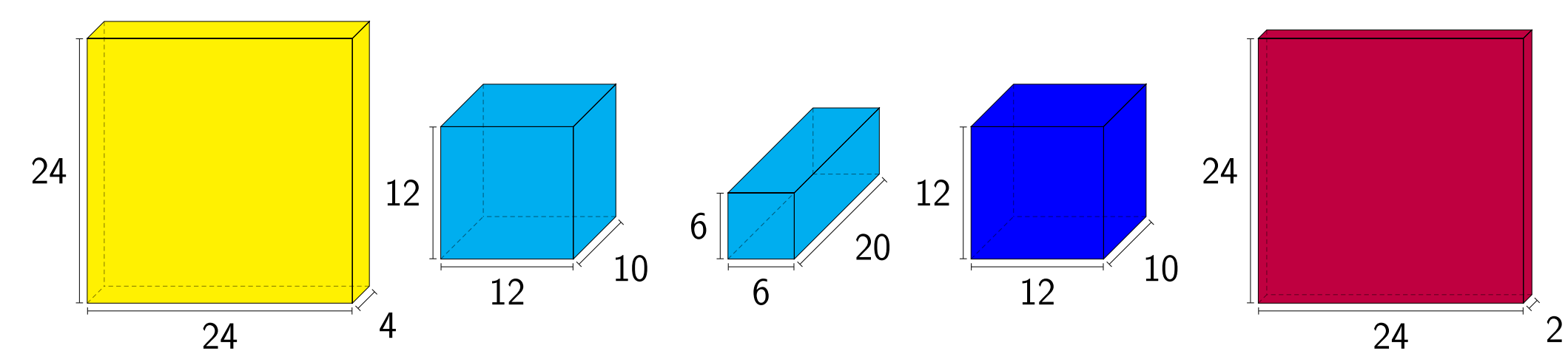
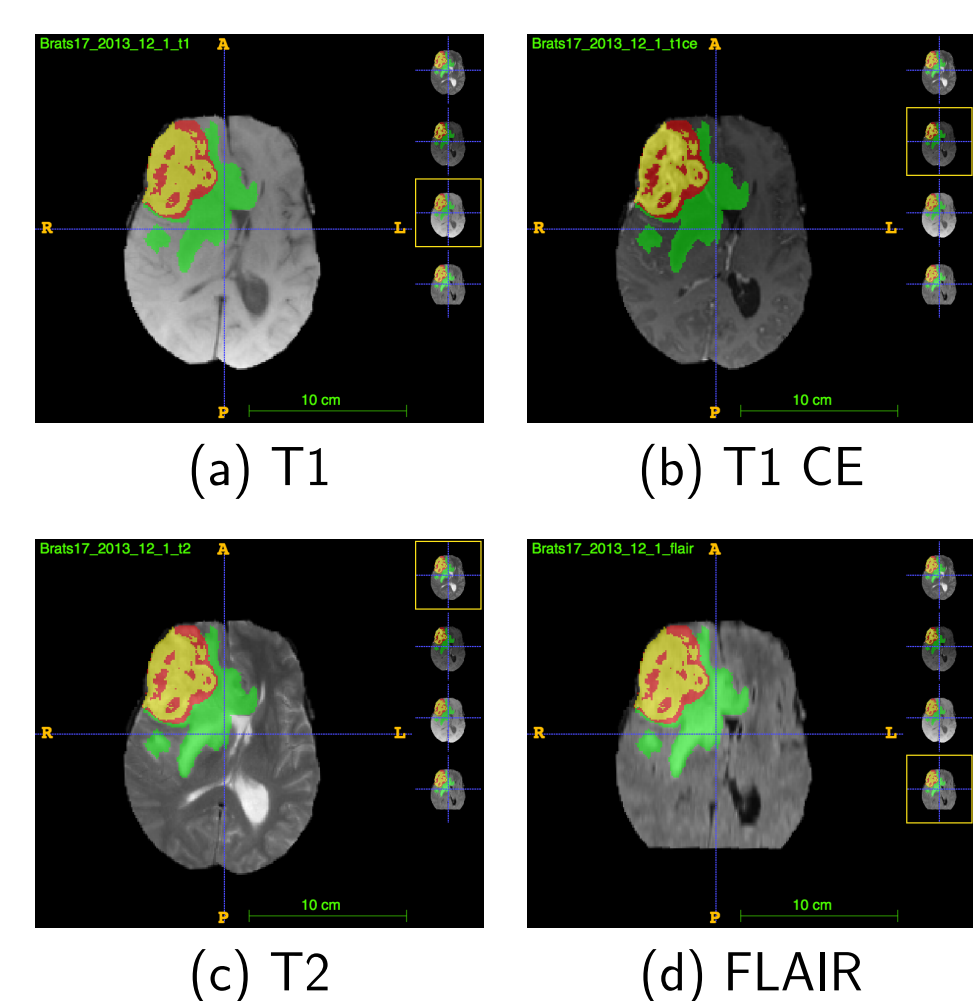


Figure 2: Patch-wise Fully-Convolutional Network (FCN) Architecture Diagram. Input same as baseline model. Convolutional layers (cyan) followed by ReLU activation, dropout, and  $2 \times 2$  max-pooling. Deconvolutional layers (blue) followed by ReLU and dropout. Scores (red) for background and foreground for entire patch.

## Dataset Overview



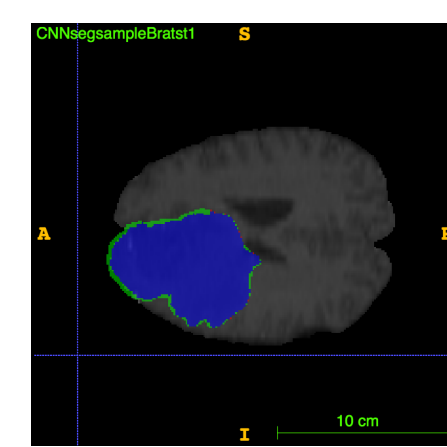
- MICCAI Brain Tumor Segmentation (BraTS) Challenge 2017 dataset from [2]
- Images segmented into edema (green), non-enhancing core (red), enhancing core (yellow), and necrotic core (dark red)
- Sample: 220 high-grade glioma, 75 low-grade glioma patients
- Images segmented by multiple radiologists
- All images  $240 \times 240 \times 155$  voxels, contain simultaneously acquired T1, T1 contrast enhanced, T2, and FLAIR images

## Segmentation Results

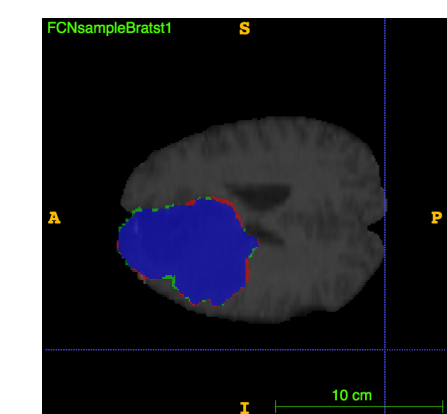
### Setup

- Goal: segment tumor from background
- Evaluated using dice score =  $\frac{2|Predicted Image \cap Ground Truth|}{|Predicted Image| + |Ground Truth|}$
- Tested using softmax cross entropy loss with  $L_2$  regularization

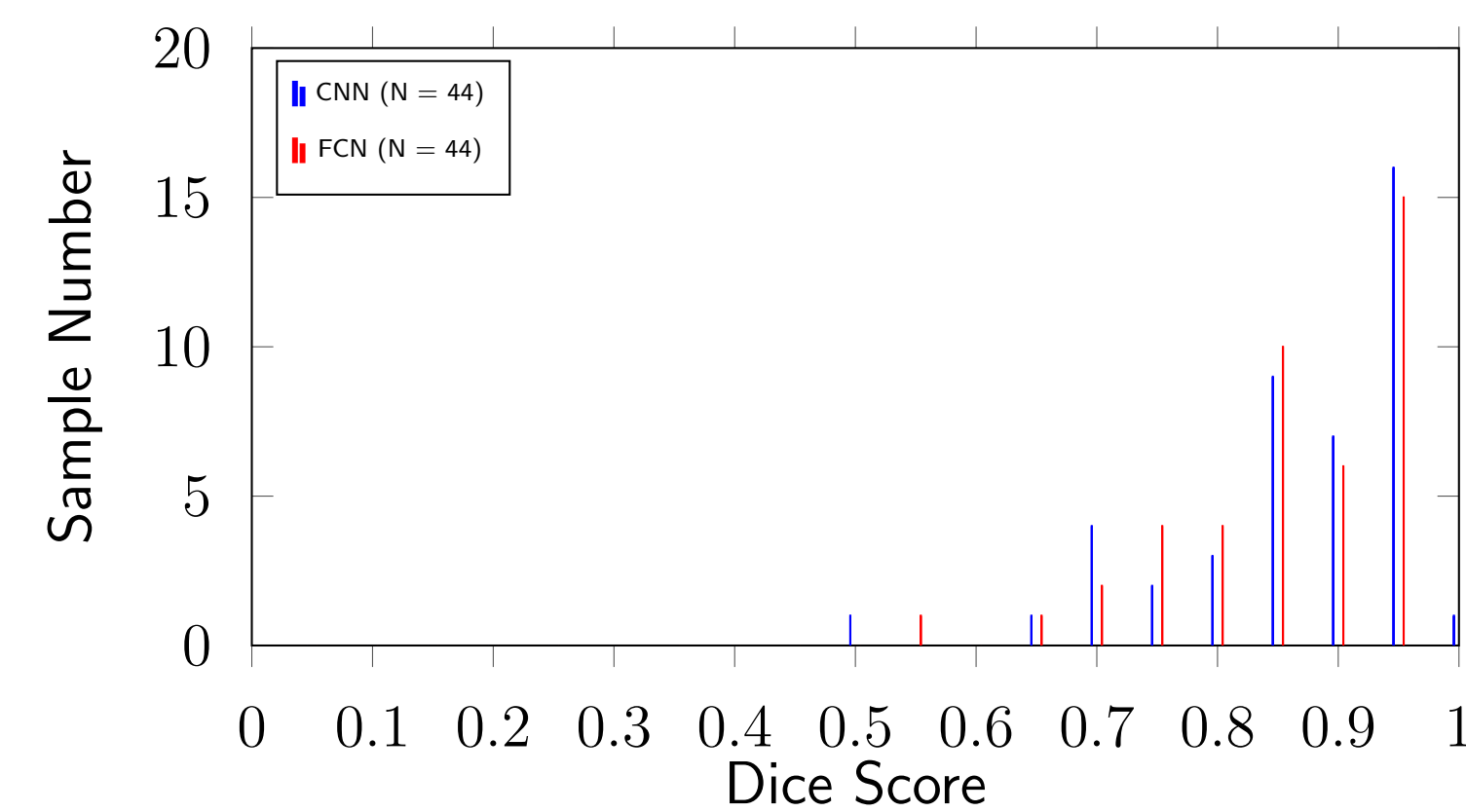
### Results



(a) CNN



(b) FCN



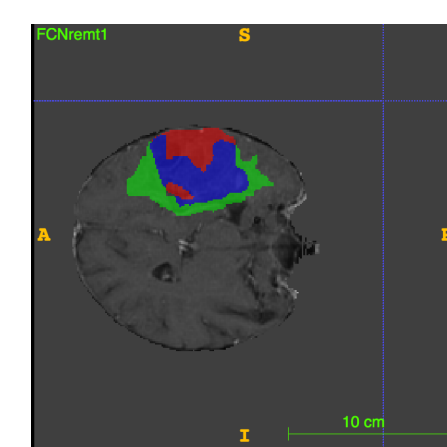
(c) Dice Score Histogram

Figure 3: (a-b) Examples of segmentation generated by algorithm. Correct pixels (blue), unidentified voxels (red), mis-identified voxels (green). (c) Histogram of dice scores across validation set. Voxel-wise approaches (mean dice 0.84) performs very strongly compared to benchmark (mean dice 0.89).

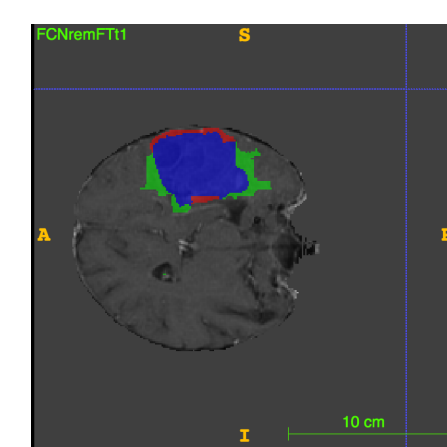
## Transfer Learning

- Segmented glioma for REMBRANDT data set (manual tumor segmentation for 130 patients) [3] with model pretrained on BraTS dataset
- Goal: segment tumor from background using different image data

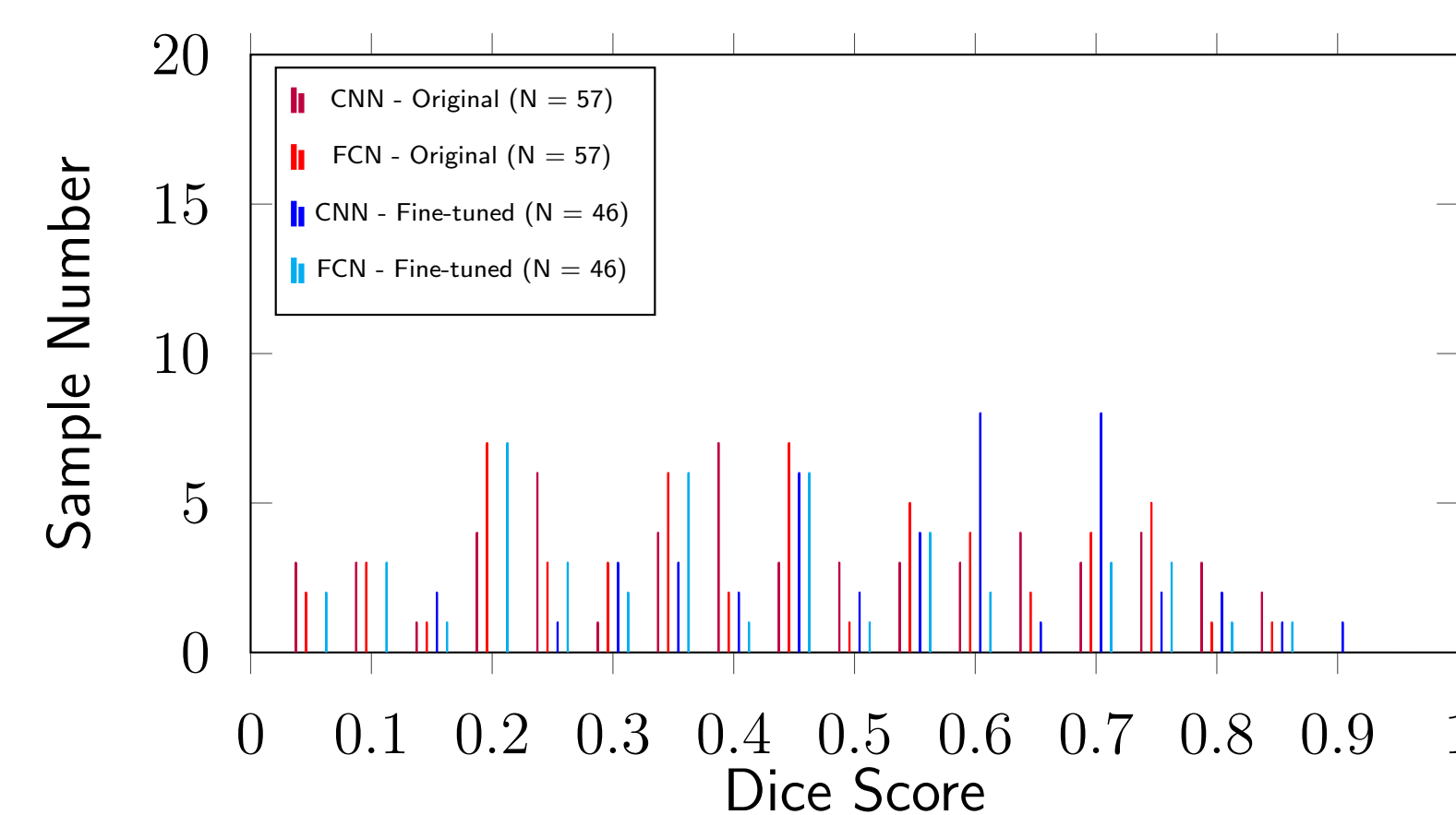
### Results



(a) FCN Original



(b) FCN Fine-tuned



(c) Dice Score Histogram

Figure 4: (a-b) Examples of segmentation generated by algorithm. Correct pixels (blue), unidentified voxels (red), mis-identified voxels (green). (c) Histogram of dice scores across validation set. Fine-tuning the model by adding batch-normalization increases accuracy. Overall, dice scores are inconsistent across the validation set.

## Future Work: Full-Image FCN

### Full-Image FCN Architecture

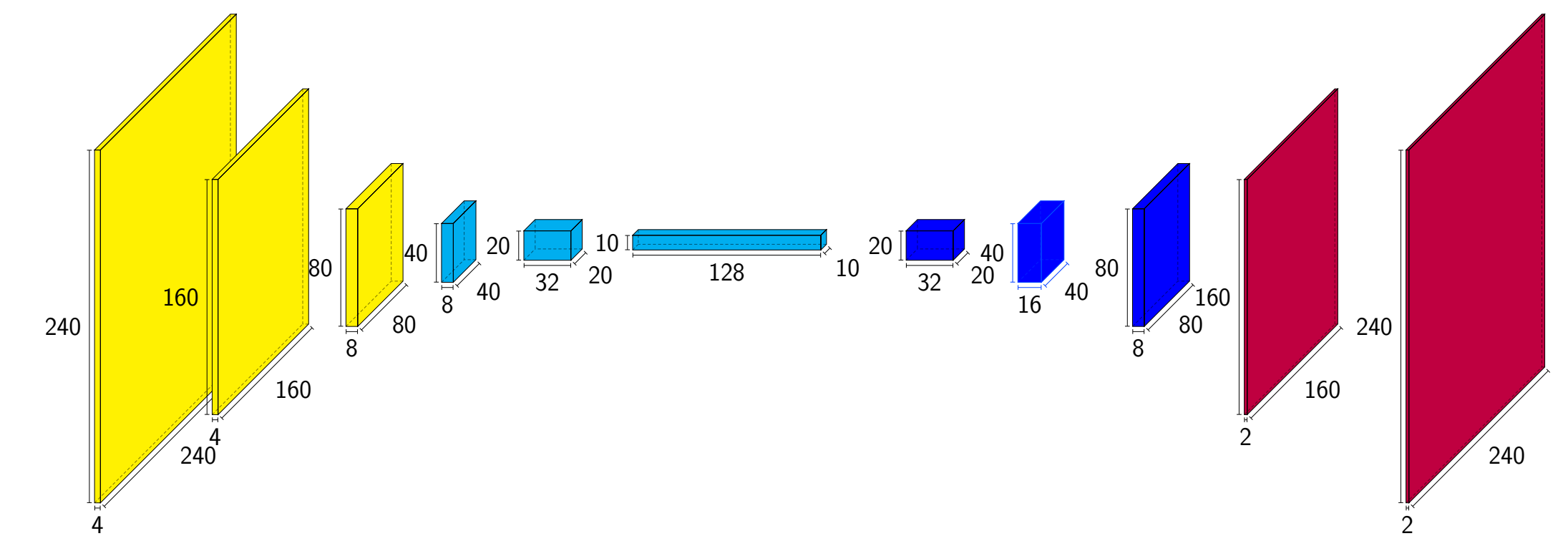


Figure 5: Full-Image FCN Architecture Diagram. Input is full image (single slice shown). Convolutional layers (cyan) followed by ReLU activation and  $2 \times 2$  max pooling. Deconvolutional layers followed by ReLU and batch normalization. Output (red) contains class scores for background and four tumor regions. Primary advantage of full-image FCN is significant speedup in throughput.

### Preliminary Results

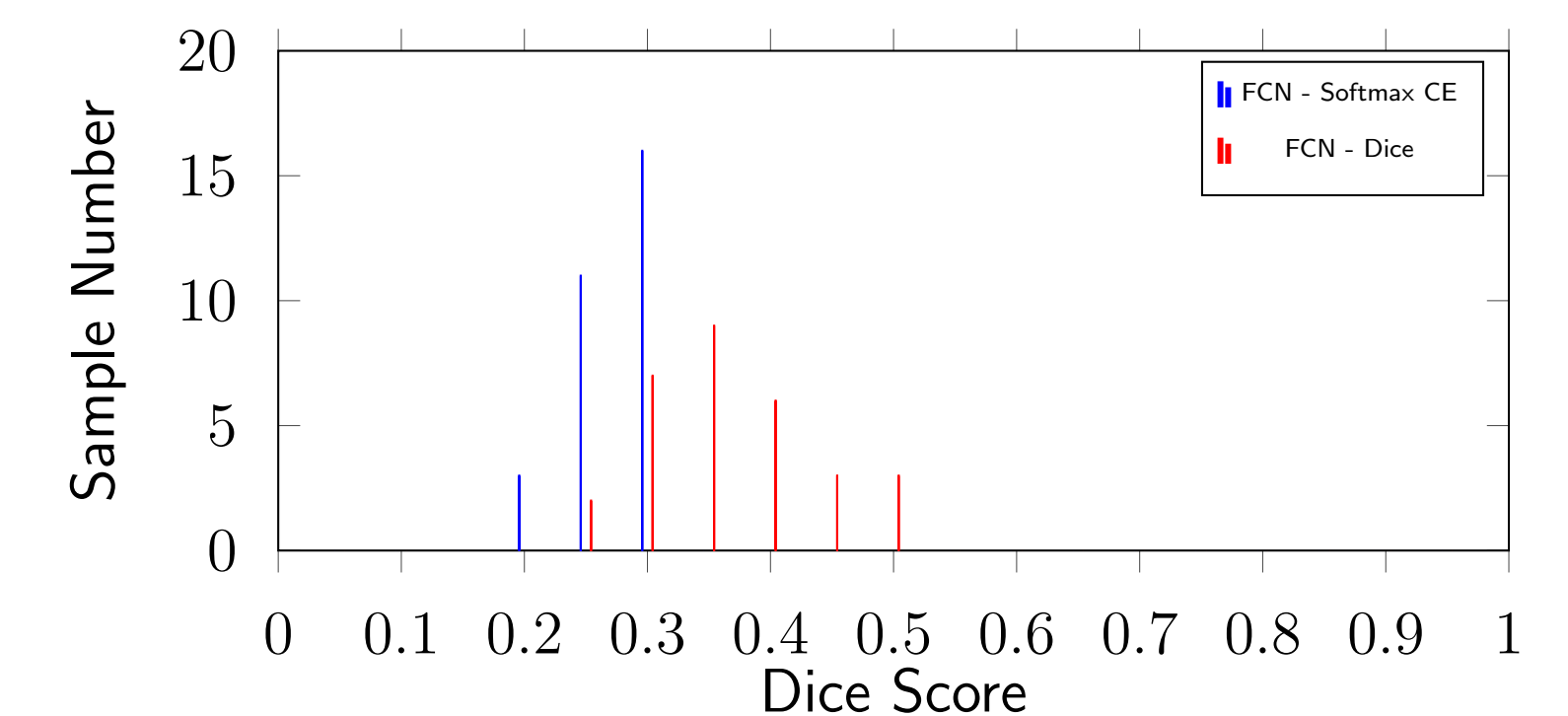


Figure 6: Histogram of dice scores for validation set samples. We tested using both cross-entropy loss and dice score loss. The Dice score loss model performs better, but the scores are poor compared to the voxel-wise architectures. This is most likely due to the model biasing very heavily towards the background class.

## Conclusion

- Voxel-wise approach obtains consistently high accuracy (compare with  $\sim 0.89$  dice score on previous BraTS benchmarks)
- Transfer learning accuracy is promising but accuracy inconsistent
- Much work to be done for full-image FCN approach

## Acknowledgements

Many thanks to Prof. Olivier Gevaert for his guidance on this project. Also thank you to the CS 231N teaching staff for their tireless support throughout the quarter.

### Citations

- [1] E. C. Holland. Progenitor cells and glioma formation. *Current Opinion in Neurology*, 14(6):683–688, 2001.
- [2] B. H. Menze et al. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging*, 34(10):1993–2024, 2015.
- [3] L. Scarpace, A. E. Flanders, R. Jain, T. Mikkelsen, and D. W. Andrews. Data from rembrandt, 2015.
- [4] K. Urbanska, J. Sokolowska, M. Szmidt, and P. Sysa. Glioblastoma multiforme - An overview. *Wspolczesna Onkologia*, 18(5):307–312, 2014.