Automated Grading Of Gliomas using Deep Learning in Digital Pathalogy

- A modular approach with ensemble of ConvNets

Introduction

- Gliomas are the most common primary malignant brain tumors.
- Based on the morphological features of the tumor in histopathological slides they are divided into :
  - Lower Grade Glioma (Grade I, Grade II and Grade III)
  - Glioblastoma Multiforme (Grade IV)

Dataset

- Whole Slide Tissue Images (WSI) from The Cancer Genome Atlas (TCGA)
- Data included two types of brain cancer :
  - Glioblastoma Multiforme (GBM)
  - Lower Grade Glioma (LGG) { Grade I , Grade II and Grade III }

Preprocessing

- Tiling :
  - Whole Slide Tissue Images (WSI) are partitioned into tiles
  - Select tiles that contains tissue at least 90% of title area.
  - Tissue in tiles are distinguished by hysteresis thresholding on grayscale and 8-bit depth complemented image
  - Advances in information about a whole tumor and computational efficiency by parallel processing
- Segmentation :
  - Nuclei are segmented using morphological top-hat filtering
  - The nuclei are segmented such that there is no loss of nuclei distribution map

CNNs

- Each convolution layer performs a discrete 2D convolution operation and applies the activation function
- Pooling layers reduce the size by preserving the information
- We use ReLU activation function over other to train the data much faster
- Final layer (fully connected), which sums up to a loss function, which is aimed to be minimized

Inputs to the Deep Learning Pipeline

- After the segmentation, they are further tiled to reduce in size to form samples called E-Microbiopsy Samples

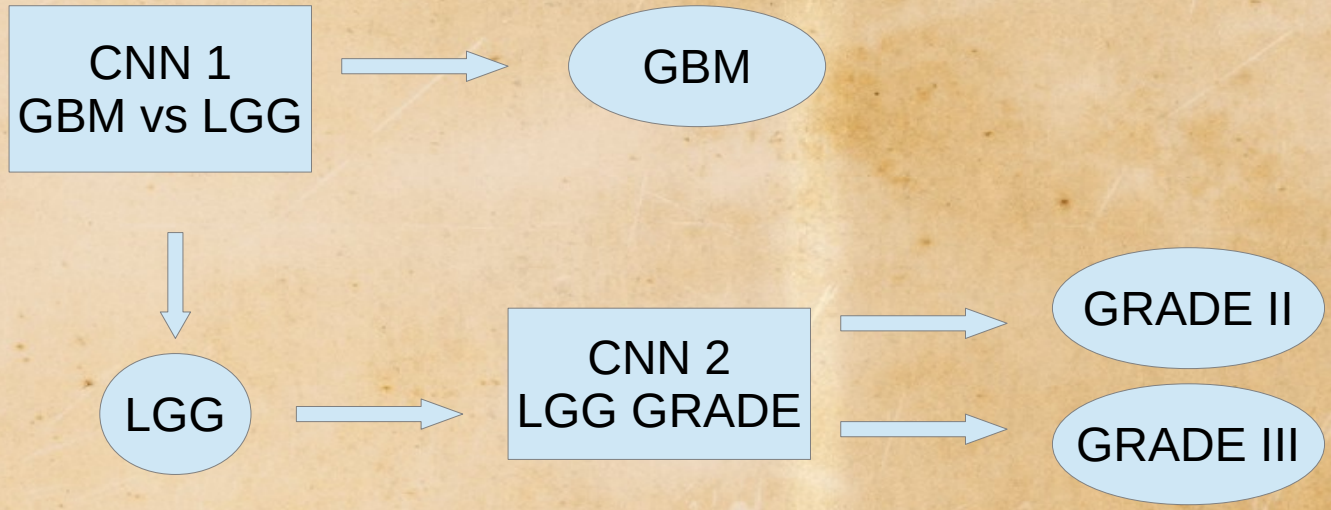
- These samples are inputs to the deep learning pipeline

Training CNNs

- Caffe is used for implementing the CNNs
- This trains the models by Stochastic Gradient Descent (SGD)
- Steps :
  - Forward Pass : inputs => ouputs
  - Backward Pass : Computes the gradients of the parameters

Model

- We have ensemble of two CNNs as shown :

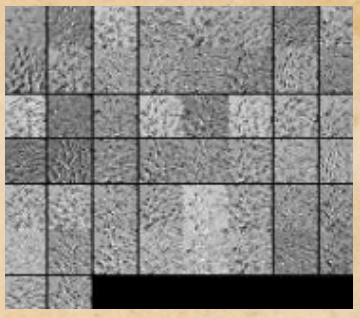


Data Split

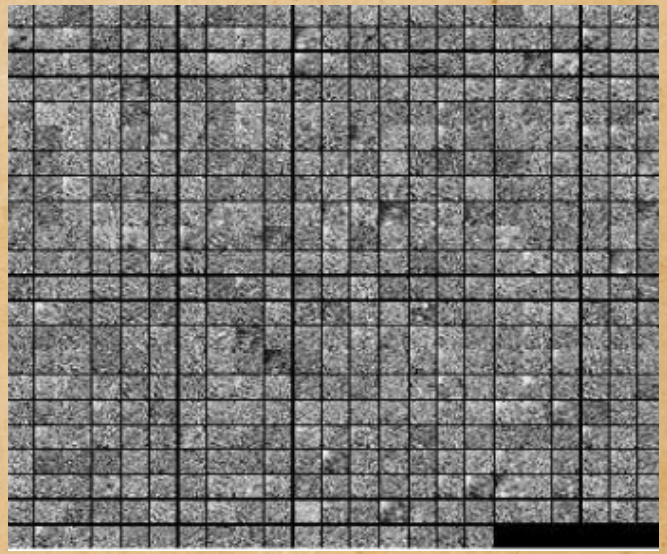
- For training the dataset was of 80% (train) – 20% (validation) split
- The training process is carried on till the training error is <= 2%, validation error <= 10%
- Regularization techniques as early stopping are used

Activation Maps of CNN1

Sample



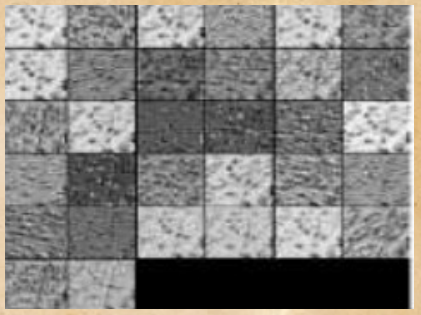
Layer 1



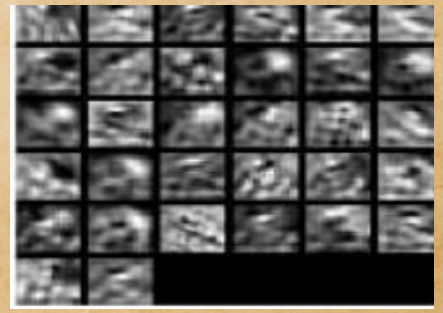
Layer 5

Activation Maps of CNN2

Sample



Layer 4



Layer 10

Evaluation Metric

- CNN1 (GBM VS LGG) gave a classification accuracy of 96%

	GBM	LGG
Precision	0.94	0.98
Sensitivity	0.98	0.94
Specificity	0.94	0.98

- CNN2 (Grade II VS Grade III) gave a classification accuracy of 71%

	LGG - GII	LGG - GIII
Precision	0.68	0.73
Sensitivity	0.64	0.77
Specificity	0.77	0.64