CSN 300 LAB BASED PROJECT
Department of Computer Science and Engineering
May 3, 2019. INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

APPLICATIONS OF MACHINE LEARNING IN MEDICAL IMAGING

O. Joseph Jeshurun, P. Ajay Kumar, S. Heerendra Kumar, Dr. Balasubramanian Raman

MOTIVATION AND PROBLEM STATEMENT

Machine learning has recently gained considerable attention because of promising results for a wide range of radiology applications. Radiomics is a field of medical image study that aims to extract large amount of quantitative features from medical images called radiomics features. Here we use Machine learning techniques along with Radiomics features to classify Glioma grading and also measure the Overall Survival of a brain tumor patient. Glioma constitutes 80% of malignant primary brain tumors and is usually classified as HGG and LGG. The LGG tumors are less aggressive, with slower growth rate as compared to HGG, and are responsive to therapy. Overall Survival is one way to see how a new treatment works and it also gives useful information for personalized therapy.

METHODOLOGY

• Radiomics enterprise can be divided into five distinct processes

Figure: Processes in Radiomics

• Dataset of BraTS 2018 challenge is used. This MRI data was collected from various institutions and acquired with different protocols, magnetic strength and MRI scanners

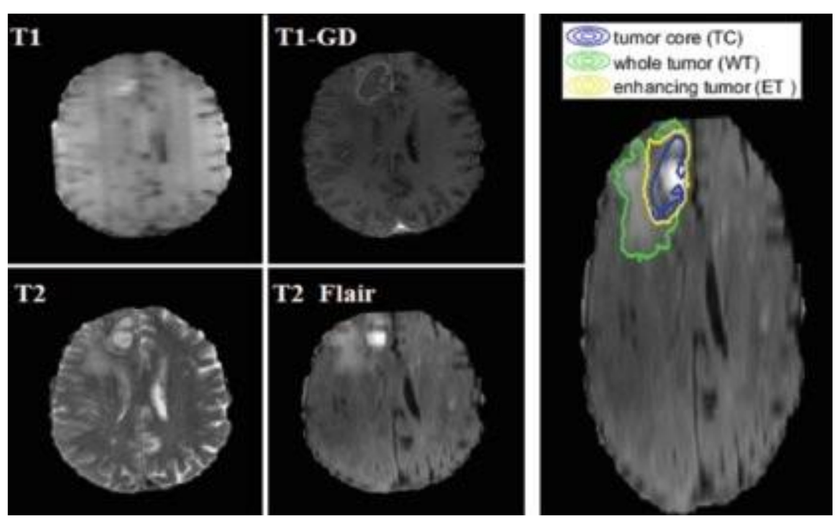

Figure: Segmented Training Dataset Images

• Image segmentation for Region of Interest(ROI) in this dataset is done manually by experts and there are 3 ROIs
• Radiomics Features are extracted from the 3 ROIs for each patient and this data is merged with the corresponding clinical data
• Data Pre-processing included elimination of categorical features and normalisation of features with real values
• ROI type 1 – Non enhancing tumour and necrotic region, ROI type 2 - Tumor region with enhancement added to region 1 and ROI type 3 – Area of Edema added to region 2

WORKFLOW AND DISCUSSION

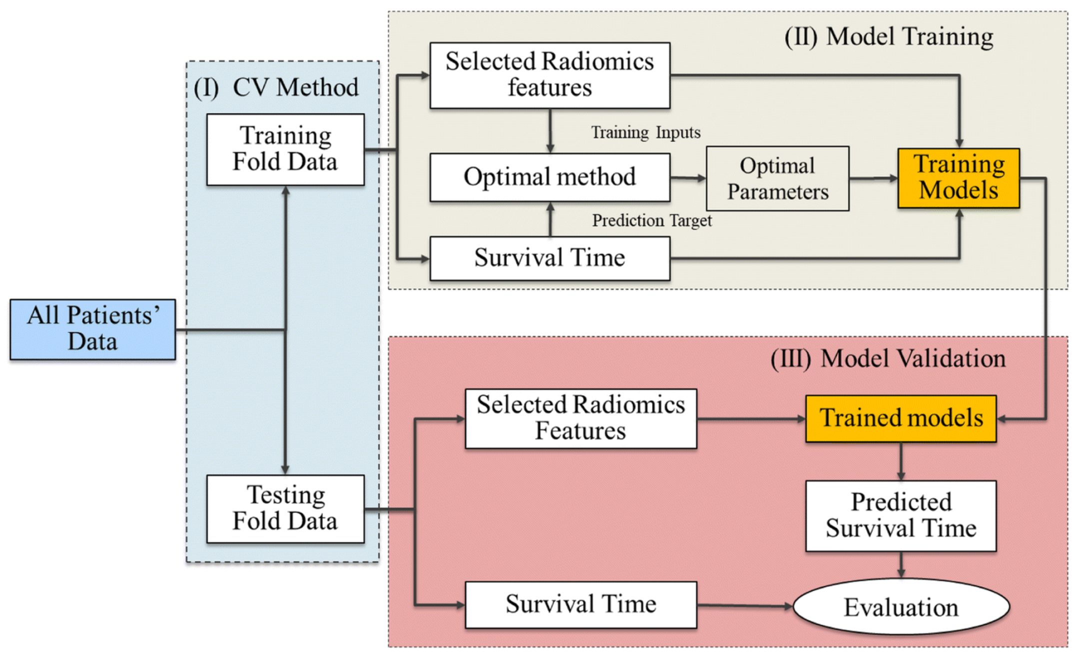

Figure: Stages in training ML model

• All the features may not influence the output variable on same level. So, features are ranked based on their influence on the output variable using minimum redundancy and maximum relevance algorithm
• SVM classifier was 5 fold cross validated for classification of Glioma grading on set of top 5, 10, 15, 25, 50 and top 100 features
• SVR model was cross validated for predicting the overall survival of a HGG Brain Tumour patient

Crossfold	Median Square Error	Mean Square Error	Explained Variance
5	154.2501461	123305.6802	0.0258623177
10	150.4155899	119170.234329	0.056009198628

Crossfold	Median Square Error	Mean Square Error	Explained Variance
5	186.9	116867.538947	0.13483178629
10	176.955	139769.542596	0.264398553649

Figure: Results for SVR and Random Forests Regressor

SUMMARY AND RESULTS

•Prediction of survival without more clinical data and treatment information is challenging
•As the number of cases for OS prediction are less there is a need to develop an efficient feature selection algorithm which will select potential features for accurate OS prediction
• The number of LGG patients is less than HGG patients so there may be a bias in the prediction. This is a challenge we faced

Number of Features	5	10	15	25	50	100
Mean Accuracy	0.8986	0.9046	0.8927	0.8868	0.8869	0.8748

Figure: 5 Fold Cross Validation Results

References:

1. Baid, Ujjwal, et al. "Deep Learning Radiomics Algorithm for Gliomas (DRAG) Model: A Novel Approach Using 3D UNET Based Deep Convolutional Neural Network for Predicting Survival in Gliomas." International MICCAI Brainlesion Workshop. Springer, Cham, 2018.
2. Kumar, Virendra, et al. "Radiomics: the process and the challenges." Magnetic resonance imaging 30.9 (2012): 1234-1248.
3. Cho, H. H., Lee, S. H., Kim, J., & Park, H. (2018). Classification of the glioma grading using radiomics analysis. PeerJ, 6, e5982

ACKNOWLEDGEMENT

We would like to sincerely thank Mr. Ankur Gupta and Mr. Rahul Kumar for guiding us through the project.