Tumor segmentation from multimodal MRI using random forest with superpixel and tensor based feature extraction

H. N. Bharath[1,2], S. Colleman[3], D. M. Sima[1,2], S. Van Huffel[1,2]

[1] Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, KU Leuven, Leuven, Belgium.
[2] Imec, Leuven, Belgium.
[3] Department of Electrical Engineering (ESAT), KU Leuven, Leuven, Belgium.

Abstract. Identification and localization of brain tumor tissues plays an important role in diagnosis and treatment planning of gliomas. A fully automated superpixel wise two-stage tumor tissue segmentation algorithm using random forest is proposed in this paper. First stage is used to identify total tumor and the second stage to segment sub-regions. Features for random forest classifier are extracted by constructing a tensor from multimodal MRI data and applying multi-linear singular value decomposition. The proposed method is tested on BRATS 2017 validation and test dataset. The first stage model has a Dice score of 83% for the whole tumor on the validation dataset. The total model achieves a performance of 77%, 50% and 61% Dice scores for whole tumor, enhancing tumor and tumor core, respectively on the test dataset.

Keywords: Superpixel, Multilinear singular value decomposition, Random forest, MRI, tumor segmentation.

1 INTRODUCTION

Accurate characterisation and localization of tissue types plays a key role in brain tumor diagnosis and treatment planning. Neuro-imaging methods in particular magnetic resonance imaging (MRI) provide anatomical and pathophysiological information about brain tumors and aid in diagnosis, treatment planning and follow-up of patients. Manual segmentation of tumor tissue is a tedious and time consuming job, it also suffers from inter and intra-rater variability. An automated brain tumor segmentation algorithm will help to overcome those problems. However, automation of brain tumor tissue segmentation is a difficult problem and often fails when applied on MRI images from different centres/scanners.

Superpixels are gaining popularity in image segmentation algorithms, and have also been used in the context of brain tumor segmentation from MRI [1]. Performing superpixel-level image segmentation offers certain advantages over pixel-level segmentation like spatial smoothness, capturing image redundancy and reducing computational

complexity [1, 2]. Recently, tensor decompositions such as the canonical polyadic decomposition and the multilinear singular value decomposition (MLSVD) [3] have been used to extract features from high-dimensional data to use in classification algorithms [4]. MLSVD is the generalization of the matrix singular value decomposition, where the tensor is decomposed into a "all-orthogonal" core tensor multiplied by orthogonal factor matrices along each mode [5]. The factor matrices contain orthonormal bases for their respective modes. Multimodal MRI consisting of T2, T1, T1+contrast and FLAIR imaging after co-registration and re-sampling to the same resolution, can be naturally represented as a 3-D tensor. In this paper we develop a fully automatic tumor tissue segmentation algorithm using a random forest classifier, where both superpixel-level image segmentation and tensor decomposition methods are combined to extract features for the classifier.

2 METHOD

2.1 Preprocessing

First, each individual 3D image is scaled to the range: 0-1. Next, intensities are normalized by applying histogram equalization. A reference histogram is generated by selecting 10 random images from the training set and extracting a histogram from the combined image. Histogram equalization is applied separately to different modalities. Background is removed from each slice using Ostu's image threshold method [6].

2.2 Feature extraction

The MR images are divided into smaller patches which are better aligned with intensity edges, called superpixels [7]. The superpixels are generated from each slice from one of the modalities as shown in Fig.1. The tissue assignment is done on superpixel-level instead of individual pixel, which helps to reduce computational cost and improve spatial smoothness [1].

Different features extracted from each superpixel are explained below

- Feature1: Mean intensity values of all the 4 modalities and 6 difference images (e.g.: abs(T1-T2)) over each superpixel.
- Feature2: Entropy and standard deviation over each superpixel.
- Feature3: A 3-D tensor is constructed for each superpixel, where the frontal slices are the covariance matrix of pixel-level features and the third mode is the modality and the difference images of the modalities. Pixel-level features consist of mean, median, standard deviation and entropy over a 5×5 window. Features are extracted by applying a rank-2 truncated multilinear singular value decomposition (MLSVD) on the 3D Tensor as shown in Fig. 2.
- Feature4: A 4-D tensor is constructed for each superpixel where the first two modes are 5×5 image patches with the main voxel at the centre, third mode consists of image patches from all four modalities (T1, T2, T1C and FLAIR) and patches from six difference images of the modalities (abs(T1-T2), abs(T1-T1C), abs(T1-FLAIR), abs(T1C-T2), abs(T1C-FLAIR), abs(T2-FLAIR)) and the fourth mode consists of

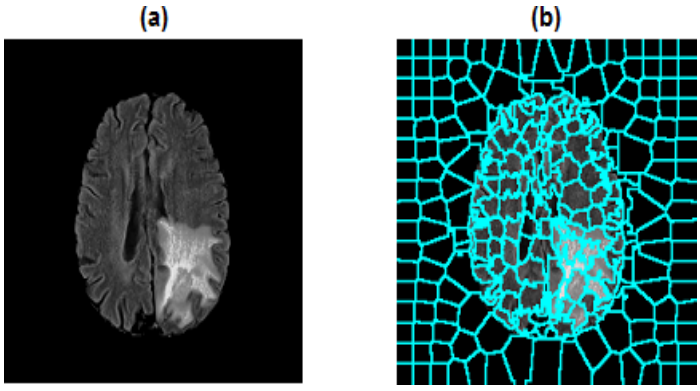


Fig. 1: (a) One slice of the FLAIR image. (b) Generated superpixels for the slice in (a).

voxels within the superpixel. Again MLSVD is used for feature extraction. Only the mode-1, mode-2, mode-3 and the core tensor are used as feature.

- Feature5: A 3-D tensor is constructed for each superpixel, where the first mode is the pixels from 5×5 image patches with the main voxel at the centre, the second mode is the modality and the difference images (e.g.: abs(T1-T2)) of the modalities and the third mode consists of the voxels within the superpixel. Again, features are extracted by applying rank-2 MLSVD. Only the mode-1, mode-2 and core tensor are used as feature.
- Feature6: For each superpixel a covariance matrix is estimated from the intensity values of all the modalities and the difference images. Covariance matrix plus two dominant eigenvectors and eigenvalues are used as features.
- Feature7: Local spectral histograms which are texture descriptors based on local distribution of filter responses [8] are estimated for all 4 modalities and 6 difference images. A 3D tensor is constructed for each superpixel, where the first mode includes local spectral histograms, second mode is the modality and the difference images of the modalities and the third mode consists of the voxels within the superpixel. Features are extracted by applying rank-2 MLSVD. Only the mode-1, mode-2 and core tensor are used as feature. The mean of the local spectral histograms over the superpixel are also used along with the MLSVD features.

2.3 Training and Tissue Segmentation

Tumor tissue segmentation was performed using a two-stage classifier. In the first stage a binary classification was performed on the superpixels to segment tumor and non-tumor regions. In the second-stage a multi-class classification was performed on the superpixels which are inside the estimated tumor region to segment enhancing tumor (ET), edema (ED), necrotic and non-enhancing tumor (NCR/NET) and healthy tissue. The two-stage operation is demonstrated in Fig. 3. For both stages a random forest classifier was used.

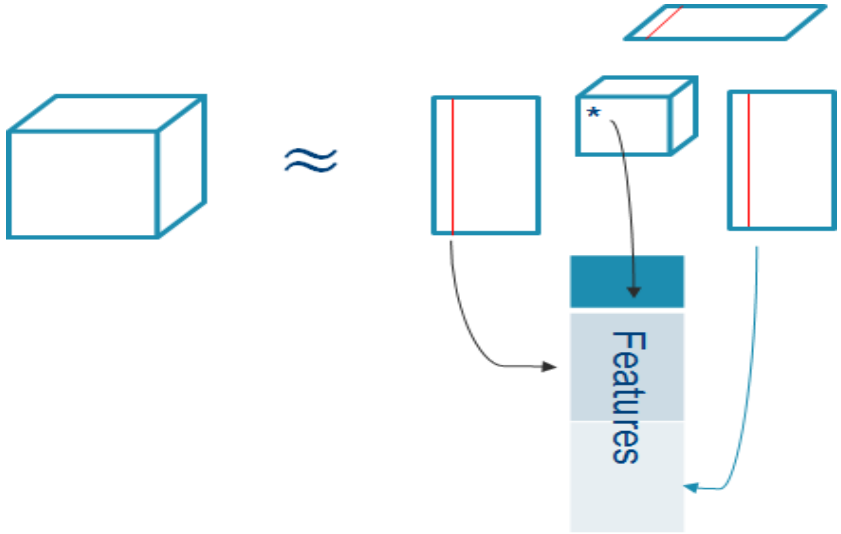


Fig. 2: Truncated multilinear singular value decomposition and feature extraction.

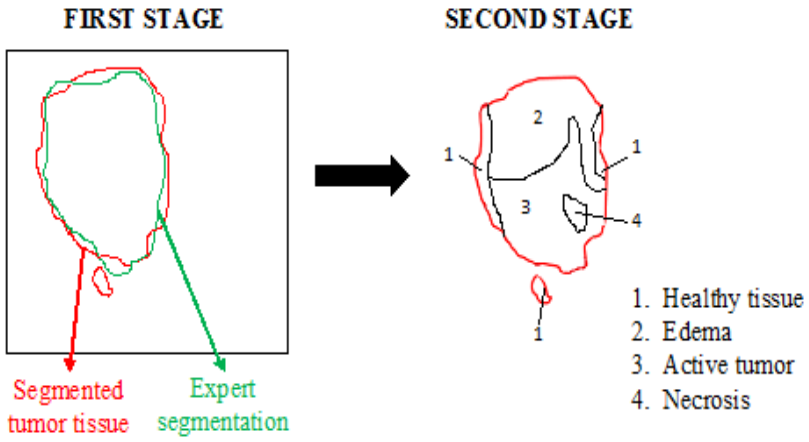


Fig. 3: Demonstration of whole tumor segmentation in first stage and sub-tissue segmentation in second stage.

First Stage: In the first stage, superpixels are obtained from FLAIR because the total tumor is brighter in this modality. The feature set for this stage consists of Feature1, Feature2, Feature3 and Feature4. The dataset is divided in three groups, and random forest classifiers with 100 trees are trained from each group. The prediction is result from majority voting of the classifiers learned from three data groups. For each model training is done iteratively, where a class balanced subset from the respective group is used for initial training. Next the trained model is tested on the remaining data from the respective group, the data that are classified wrongly are added to the initial subset and trained again with 100 tree random forest binary classifier. After the first stage classification at superpixel level, image filling and continuity-based denoising developed by [9] is performed on the whole tumor segmentation before going to the second stage.

Second Stage: In the second stage, superpixels are obtained from T1+contrast imaging modality because the enhancing tumor is brighter in this modality. Feature1, Feature5, Feature6 and Feature7 are used as features in this stage. Random forest classifiers with 250 trees are trained using a iterative method. Initially, 60 patients are randomly selected from the dataset for training and the trained model is tested on the remaining subset of the database. Next, the patients which resulted in low Dice scores are included in the training set and a new model is trained. This is continued until all the patients in the dataset are used. Initially, the training is started with balanced data. The list of all features with its corresponding dimension and the stage where they are used is shown Table 1.

Table 1: Features used in stage one and stage two along with there corresponding dimension.

	Feature1	Feature2	Feature3	Feature4	Feature5	Feature6	Feature7
Dimension	10	20	36	48	78	122	196
Stage One	✓	✓	✓	✓	X	X	X
Stage Two	✓	X	X	X	✓	✓	✓

3 Results and Discussion

3.1 First Stage: Whole tumor segmentation

A first stage model with three classifiers was trained using the BRATS 2017 training database [10–13] containing 210 HGG and 75 LGG patients. The performance of the trained model in segmenting the whole tumor is tested on the validation dataset of BRATS 2017 challenge. Average Dice score and sensitivity obtained from the trained first stage model over 46 HGG patients are shown in Table 2. The boxplots of the Dice scores and sensitivity are shown in Fig. 4.

Table 2: Mean, standard deviation, median 25 quantile and 75 quantile of Dice score and sensitivity for whole tumor (WT) over 46 patients using only first stage model.

	Dice WT	Sensitivity WT
Mean	0.8330	0.8574
Std	0.1186	0.1318
Median	0.8673	0.9024
25 quantile	0.8298	0.8114
75 quantile	0.9084	0.9415

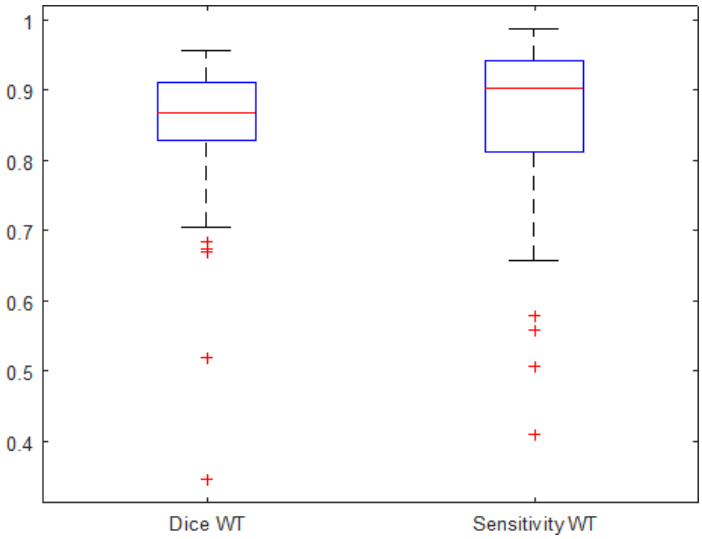


Fig. 4: Boxplots of Dice scores and sensitivity for whole tumor (WT) obtained from first stage model on BRATS 2017 validation dataset of 46 patients.

3.2 High Grade Glioma

The BRATS 2017 high grade glioma database [10–13] containing 210 patients is split into training set (70%) and test set (30%). A first stage model with a single classifier plus the second stage model is trained using only HGG and the trained model is tested on 63 HGG patients. Fig.5 shows the segmentation of tumor tissue for two different slices. We can observe from the figure that the enhancing tumor and edema region are segmented properly. However the NCR/NET region is not identified properly.

The boxplot of the Dice scores is shown in Fig. 6 and the average Dice score and sensitivity obtained from the trained model for 63 HGG patients are shown in Table 3. From the boxplot, we can observe that the algorithm performs well on most of the patients. however there are still some patients where the algorithm fails to segment properly.

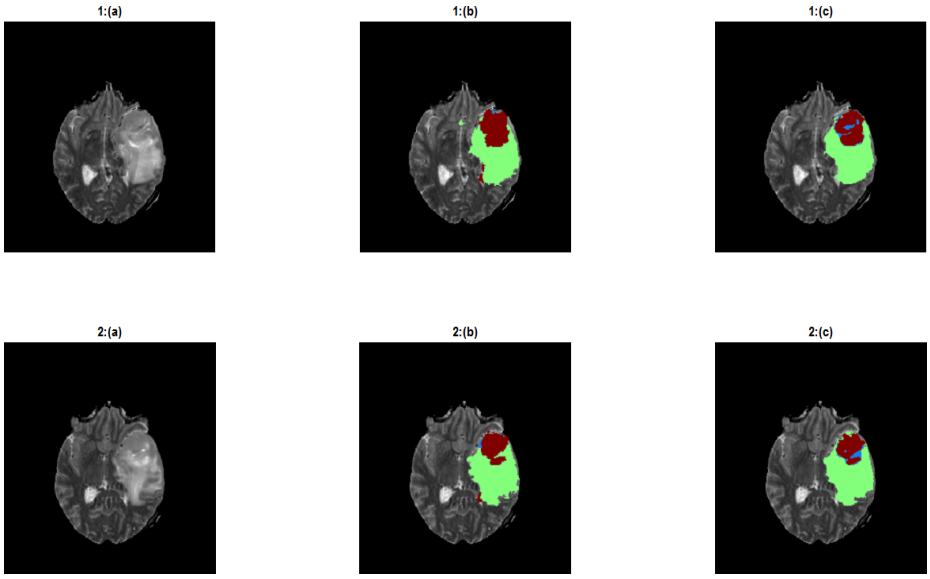


Fig. 5: Segmentation results on two slices 1-2. (a) T2 image of one slice, (b) Estimated segmentation (c) Expert segmentation. Green-Edema, Brown-enhancing tumor and Blue- Necrosis.

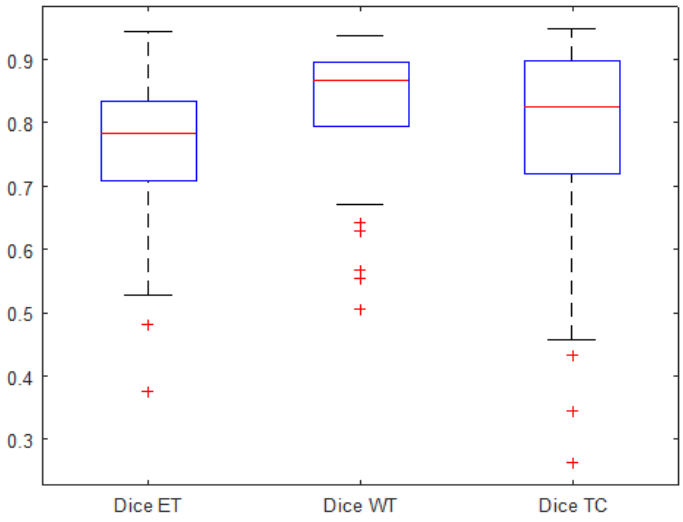


Fig. 6: Boxplots of Dice scores for enhancing tumor (ET), whole tumor (WT) and tumor core (TC) on BRATS 2017 training dataset of 63 patients.

Table 3: Mean, standard deviation, median 25 quantile and 75 quantile of Dice score and sensitivity for enhancing tumor (ET), whole tumor (WT) and tumor core (TC) over 63 HGG patients.

	Dice ET	Dice WT	Dice TC	Sensitivity ET	Sensitivity WT	Sensitivity TC
Mean	0.761	0.833	0.783	0.855	0.815	0.777
Std	0.106	0.096	0.147	0.126	0.090	0.191
Median	0.783	0.867	0.824	0.886	0.837	0.826
25 quantile	0.708	0.795	0.723	0.820	0.769	0.721
75 quantile	0.833	0.895	0.898	0.941	0.884	0.908

3.3 Validation and Test dataset results

The trained model consisting of both first and second stage is also tested on the BRATS 2017 validation dataset [10–13]. The results are shown in Fig. 7 and Table 4. The performance is worse when compared to only HGG case.

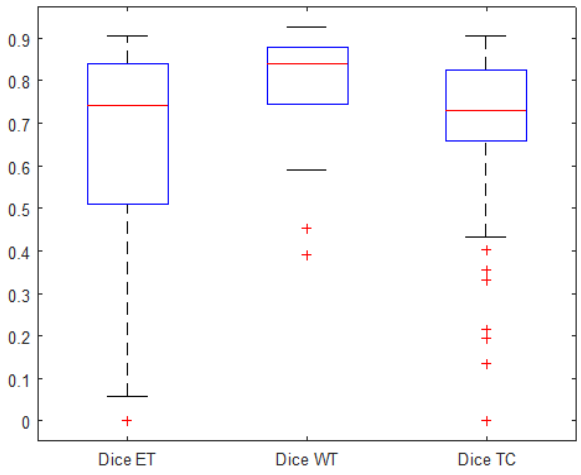


Fig. 7: Boxplots of Dice scores for enhancing tumor (ET), whole tumor (WT) and tumor core (TC) on BRATS 2017 validation dataset of 46 patients.

This algorithm does not identify necrotic and non-enhancing tumor (NCR/NET) tissue properly, it results in bad performance for LGG patients where NCR/NET tissue is larger than enhancing tumor in most cases. Also, when there is no enhancing tumor in patients, the algorithm identifies it falsely in some superpixels. This results in Dice score of zero for enhancing tumor, which can be seen from the boxplot in Fig. 7. Post-processing methods to remove such false positives may improve the average performance.

Table 4: Mean, standard deviation, median 25 quantile and 75 quantile of Dice score and sensitivity for enhancing tumor (ET), whole tumor (WT) and tumor core (TC) over 46 validation dataset.

	Dice ET	Dice WT	Dice TC	Sensitivity ET	Sensitivity WT	Sensitivity TC
Mean	0.6125	0.7928	0.6734	0.6971	0.8320	0.6763
Std	0.3013	0.1217	0.2215	0.2347	0.1221	0.2297
Median	0.7419	0.8410	0.7291	0.7683	0.8614	0.6921
25 quantile	0.5240	0.7441	0.6607	0.6365	0.8053	0.5961
75 quantile	0.8383	0.8788	0.8213	0.8468	0.9169	0.8482

Finally the complete algorithm is applied on BRATS 2017 test dataset consisting of 146 patients. The average results are shown in Table 5.

Table 5: Mean, standard deviation, median 25 quantile and 75 quantile of Dice score and Hausdorff95 for enhancing tumor (ET), whole tumor (WT) and tumor core (TC) over test dataset of 146 patients.

	Dice ET	Dice WT	Dice TC	Hausdorff95 ET	Hausdorff95 WT	Hausdorff95 TC
Mean	0.5032	0.7701	0.6105	71.27	31.33	38.90
Std	0.3054	0.1871	0.2954	132.94	30.61	84.13
Median	0.6376	0.8400	0.7416	6.56	14.56	9.56
25 quantile	0.2137	0.7202	0.4597	3.70	4.69	6.48
75 quantile	0.7447	0.8902	0.8358	39.46	54.40	29.73

4 Conclusion

In this paper, we developed a fully automated algorithm for brain tumor segmentation from multimodal MRI data. Superpixels and a tensor based feature extraction algorithm is proposed to be used with a two-stage random forest classifier for segmenting tumor tissue. The superpixels are restricted to 2-D slices because of the different resolution in the third dimension. In future work, the 2-D superpixels can be directly extended to 3D provided the difference in resolution is considered when constructing superpixels. The performance of the algorithm is comparable to the state-of-the-art methods when applied only to HGG patients. However, its performance deteriorates when tested on the BRATS 2017 validation and test database, which contains both low grade glioma (LGG) and HGG patients. The proposed superpixel method has good performance in segmenting the whole tumor using only the first stage model on the the BRATS 2017 validation set. However, this method does not perform well in segmenting the sub-regions, specifically the NCR/NET region. We assume that the features may not be discriminant enough to separate the NCR/NET region from others (mainly normal). In

future work, identification of the NCR/NET region can be improved by using texture based features like Gabor and applying feature selection for selecting dominant and removing redundant features. Therefore, the proposed method is more suited to segment the whole tumor and not good in identifying sub-regions, especially in LGG patients.

ACKNOWLEDGMENT

This research was supported by: Flemish Government FWO project G.0869.12N (Tumor imaging), G.0830.14N (Block term decompositions); IWT IM 135005; imec funds 2017; imec ICON project: ICON HBC.2016.0167, 'SeizeIT', #316679 and ERC Advanced Grant. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Advanced Grant: BIOTENSORS (n^{o} 339804). This paper reflects only the author's views and the Union is not liable for any use that may be made of the contained information.

References

1. Wu W, Chen AY, Zhao L, Corso JJ. "Brain tumor detection and segmentation in a CRF (conditional random fields) framework with pixel-pairwise affinity and superpixel-level features". International journal of computer assisted radiology and surgery, 9(2), 241-253 (2014).
2. Jia S, Zhang C. "Fast and robust image segmentation using an superpixel based fcm algorithm". Image Processing (ICIP), 2014 IEEE International Conference on. IEEE (2014).
3. N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis and C. Faloutsos, "Tensor Decomposition for Signal Processing and Machine Learning," in IEEE Transactions on Signal Processing, vol. 65, no. 13, pp. 3551-3582, July1, 1 2017.
4. Fargeas A, Albera L, Kachenoura A, Dréan G, Ospina JD, Coloigner J, Lafond C, Delobel JB, De Crevoisier R, Acosta O. "On feature extraction and classification in prostate cancer radiotherapy using tensor decompositions". Medical engineering & physics, 37(1), 126-131 (2015).
5. De Lathauwer L, De Moor B, Vandewalle J. "A multilinear singular value decomposition". SIAM journal on Matrix Analysis and Applications, 21(4), 1253-1278 (2000).
6. Otsu, N., "A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on Systems, Man, and Cybernetics, Vol. 9, No. 1, 1979, pp. 62-66.
7. Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S. "SLIC superpixels compared to state-of-the-art superpixel methods". IEEE transactions on pattern analysis and machine intelligence, 34(11), 2274-2282 (2012).
8. Liu, Xiuwen, and DeLiang Wang. "A spectral histogram model for texton modeling and texture discrimination" Vision Research 42, no. 23 (2002): 2617-2634.
9. Wu, Wei, Albert YC Chen, Liang Zhao, and Jason J. Corso. "Brain tumor detection and segmentation in a CRF (conditional random fields) framework with pixel-pairwise affinity and superpixel-level features." International journal of computer assisted radiology and surgery 9, no. 2 (2014): 241-253.
10. Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, Burren Y, Porz N, Slotboom J, Wiest R, Lanczi L, Gerstner E, Weber MA, Arbel T, Avants BB, Ayache N, Buendia P, Collins DL, Cordier N, Corso JJ, Criminisi A, Das T, Delingette H, Demiralp C, Durst CR, Dojat M, Doyle S, Festa J, Forbes F, Geremia E, Glocker B, Golland P, Guo

X, Hamamci A, Iftekharuddin KM, Jena R, John NM, Konukoglu E, Lashkari D, Mariz JA, Meier R, Pereira S, Precup D, Price SJ, Raviv TR, Reza SM, Ryan M, Sarikaya D, Schwartz L, Shin HC, Shotton J, Silva CA, Sousa N, Subbanna NK, Szekely G, Taylor TJ, Thomas OM, Tustison NJ, Unal G, Vasseur F, Wintermark M, Ye DH, Zhao L, Zhao B, Zikic D, Prastawa M, Reyes M, Van Leemput K. "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", IEEE Transactions on Medical Imaging 34(10), 1993-2024 (2015).

11. Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby JS, Freymann JB, Farahani K, Davatzikos C. "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features", Nature Scientific Data, (2017) [In Press].
12. Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby J, Freymann J, Farahani K, Davatzikos C. "Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-GBM collection", The Cancer Imaging Archive, 2017. DOI: 10.7937/K9/TCIA.2017.KLXWJJ1Q
13. Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby J, Freymann J, Farahani K, Davatzikos C. "Segmentation Labels and Radiomic Features for the Pre-operative Scans of the TCGA-LGG collection", The Cancer Imaging Archive, 2017. DOI: 10.7937/K9/TCIA.2017.GJQ7R0EF