

by Xuan Chen, Binh P. Nguyen,  
Chee-Kong Chui,  
and Sim-Heng Ong

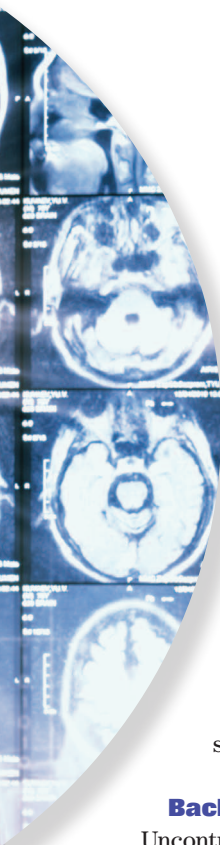
**A**dvances in the field over the years have made medical imaging an indispensable part of medicine. Today, the use of medical images is often critical for diagnosis and treatment planning. The efficient processing and analysis of the large imaging data sets that have accompanied the rise in popularity of medical imaging have presented significant challenges that have yet to be successfully overcome. Medical image analysis and volume visualization are topics that attract much attention from the image-processing community. The former extracts useful knowledge for specific purposes,



# Reworking Multilabel Brain Tumor Segmentation

An Automated Framework  
Using Structured Kernel Sparse Representation

Digital Object Identifier 10.1109/MSMC.2017.2664158  
Date of publication: 18 April 2017



such as tumor segmentation, while the latter creates vivid two-dimensional representations of three-dimensional volumetric data. Our research projects have focused on the sparse and compressible properties of signals in some representation systems with the general objective of fully utilizing these properties for the development of effective and efficient algorithms for medical image analysis and volume visualization. In this article, we introduce one of our research projects, which aims to address the significant but challenging task of multilabel brain tumor segmentation. In this work, we proposed a superpixel-based framework for this specific task using structured kernel sparse representation.

## Background

Uncontrollable cell proliferation in the brain is referred to as a *brain tumor*. Even though a brain tumor is not a common case, with a prevalence of less than 0.1% in Western countries, the high mortality it causes is eye-catching [1]. The goal of brain tumor segmentation is to separate tumor tissues from nontumor regions and partition them into spatially contiguous regions according to predefined criteria [2]. Magnetic resonance (MR) imaging is the most common technology in brain imaging due to its advantages in terms of safety, tissue contrast, and fewer artifacts compared to other modalities, such as computed tomography. This emphasizes the significance of developing effective and reproducible methods for brain tumor segmentation based on MR images. When dealing with large medical image data sets, semi-automated and automated approaches, which require very little human intervention, are preferable to manual segmentation, which is tedious and time-consuming.

Nevertheless, brain tumor segmentation is never an easy task. Brain tumors vary significantly in size, shape, and location from case to case. In addition, a common situation in MR images is that tumor and nontumor tissues overlap in their gray-level intensities. These present a huge challenge in semi-automated and fully automated brain tumor segmentation because the use of strong priors is impossible. Brain tumor segmentation is

considered a challenging topic, especially for the multilabel case.

Researchers have put much effort into the study of brain tumor segmentation for decades. Published approaches can be roughly categorized into either generative methods or discriminative methods [3]. Generative methods explicitly model the anatomy and statistics of brain tissues and usually use these models in the expectation-maximization algorithm [4], [5]. Generative methods generalize well on unseen images, but they suffer from difficulties in modeling the prior knowledge of brain tissues and elaborate nonrigid registration. Discriminative methods directly learn the features of task-relevant brain tissues from given training sets for the segmentation task [6], [7]. Although discriminative methods avoid the difficulties in modeling and registration, they are sensitive to the quantity and quality of the available training samples. The success of sparse coding and dictionary learning has been demonstrated in various vision problems, including image segmentation. In [8], we proposed a superpixel-based framework aimed at multilabel brain tumor segmentation using structured kernel sparse representation.

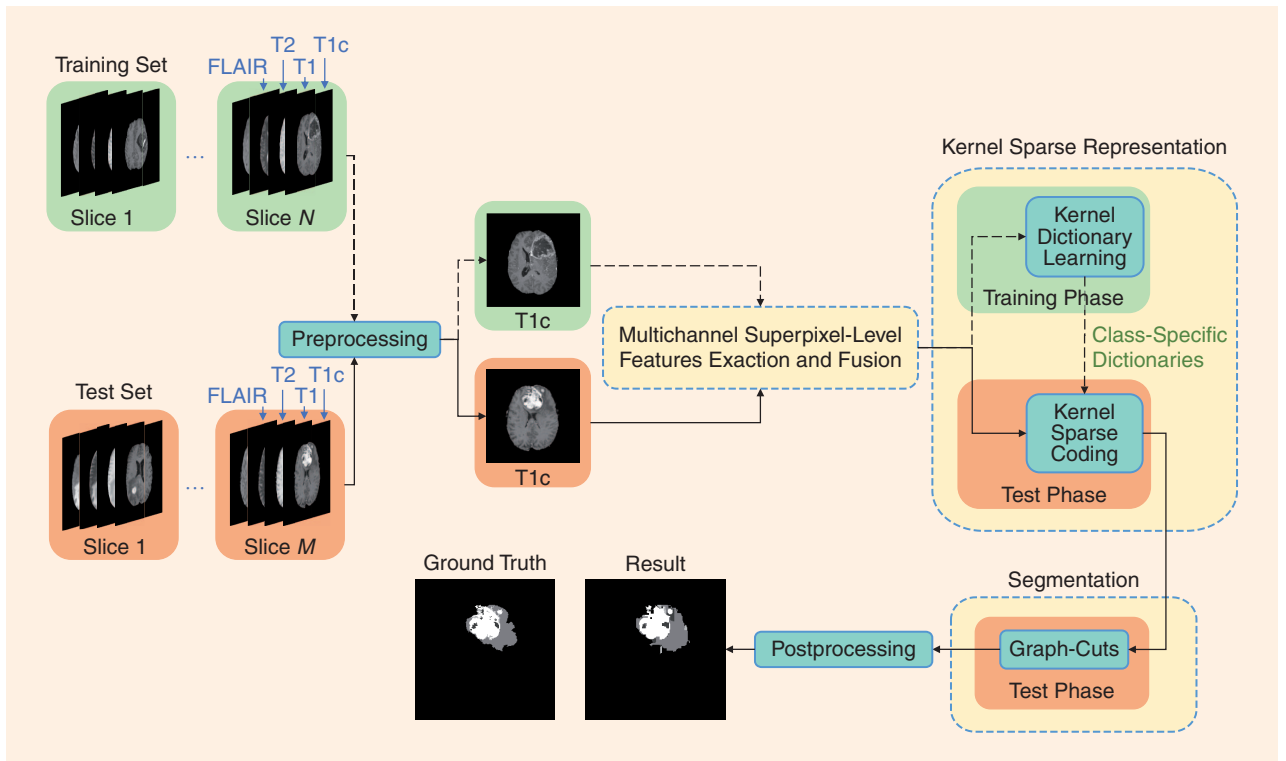
## The Proposed Framework

An overview of the proposed framework is shown in Figure 1. There are three main components: multichannel superpixel-level feature extraction and fusion, kernel sparse representation, and segmentation. In our approach, the superpixel, which refers to a group of pixels with similar perceptual meaning, is used as the basic processing unit for efficiency and compact representation. Sparse coding and dictionary learning, both part of kernel sparse representation, are implemented in a high-dimensional feature space  $\mathcal{F}$  with the help of the kernel trick, which simplifies the complicated, or even intractable, inner product calculation in  $\mathcal{F}$  to the calculation of known kernel functions.

MR imaging improves our knowledge of brain tumors by providing multichannel information [3], such as T1-weighted (T1), T2-weighted (T2), contrast-enhanced T1-weighted (T1c), and fluid-attenuated inversion recovery (FLAIR). Due to the fact that T1c images usually have higher spatial resolution

and present a clearer display of brain tumor structure compared to other channels, they are used as the reference image to generate superpixels. These superpixel regions are then applied to other channels for the

Today, the use of medical images is often critical for diagnosis and treatment planning.



**Figure 1.** An overview of the proposed automated framework for multilabel brain tumor segmentation. The dashed and solid arrows indicate the pipelines in the training phase and the test phase, respectively.

extraction of superpixel-level features, including multichannel histograms and spatial location. The extracted superpixel-level features are mapped to the feature space  $\mathcal{F}$  with a nonlinear transformation  $\Phi(\cdot)$ , where the nonlinear similarity between two samples are more discriminative than their linear similarity in the original space, and hence more discriminative sparse representations can be expected. As mentioned, the nonlinear similarity, which is computed by the inner product  $\Phi(\mathbf{x})^T \Phi(\mathbf{x}')$  in  $\mathcal{F}$ , given two samples  $\mathbf{x}$  and  $\mathbf{x}'$ , can be easily calculated with a predefined kernel function  $\mathcal{K}$ . The kernel trick also facilitates the fusion of multifeatures in such a way that it can be fulfilled by a simple entry-wise product. The fusion enables the dictionary-learning process to capture several features at the same time.

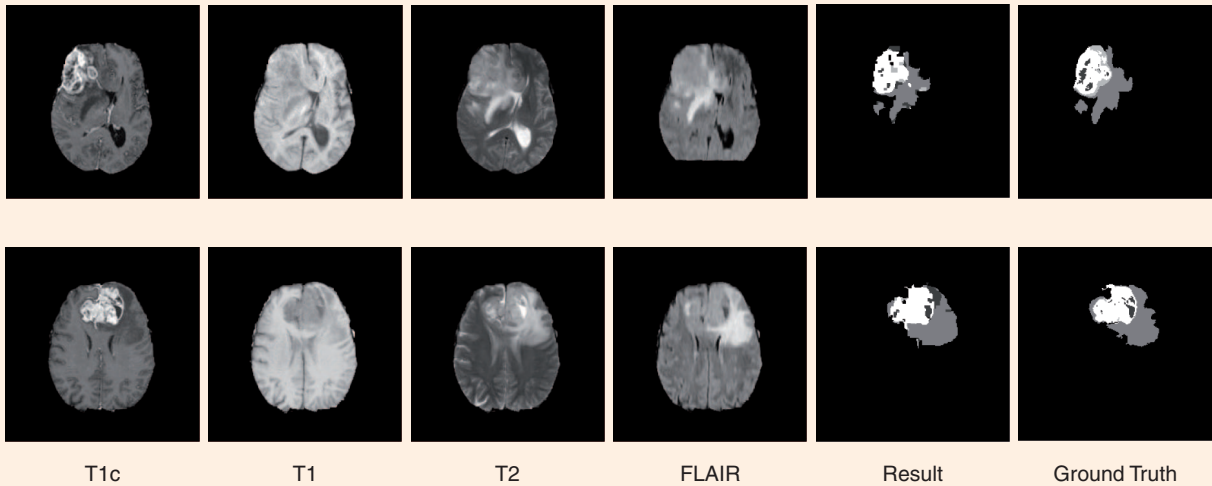
To adapt sparse coding and dictionary learning in feature space, their kernel extensions are used. Given a data matrix  $\mathbf{Y}$ , the objective function is expressed as:

$$(\hat{\mathbf{X}}, \hat{\mathbf{D}}) = \arg \min_{\mathbf{X}, \mathbf{D}} \|\Phi(\mathbf{Y}) - \Phi(\mathbf{D})\mathbf{X}\|_F^2 \text{ s.t. } \|\mathbf{x}_i\|_0 \leq T_0, \forall i,$$

**Magnetic resonance imaging is the most common technology in brain imaging due to its advantages in terms of safety, tissue contrast, and fewer artifacts compared to other modalities, such as computed tomography.**

where  $\mathbf{X}$  denotes the sparse code matrix;  $\mathbf{D}$ , the dictionary;  $\|\cdot\|_F$  the Frobenius norm;  $\|\cdot\|_0$  the  $\ell_0$  norm, and  $T_0$ , the sparsity level [9]. The sparsity level indicates the maximum number of nonzero entries in a sparse code  $\mathbf{x}_i$ . The optimization of this objective function gives a learned dictionary for the training samples in  $\mathcal{F}$  space and optimal sparse codes for samples with respect to the dictionary.

In the training phase, kernel-dictionary learning leads to class-specific dictionaries that are optimized for each task-relevant class, i.e., nontumor, necrotic core, edema, nonenhancing core, and enhancing core (labels 1 to 5, respectively) in our application. These dictionaries should be able to model their own classes well, even though they fail to well approximate the rest of the classes. In the test phase, these learned dictionaries are used in the generation of kernel sparse representations for test samples. If a test sample belongs to a specific class, the smallest approximation error can be achieved when the dictionary learned for this class is used in kernel sparse coding.



**Figure 2.** Two segmentation examples of the proposed framework, using T1c images, T1 images, T2 images, FLAIR images, segmentation results, and ground truths, respectively. The first row shows one slice of patient009, while the second row is a slice of patient015.

Kernel sparse representation does not consider the useful structural information of brain tissues. To address this problem, we incorporate this information by utilizing kernel sparse representations in the graph-cut method, which naturally takes structural information into consideration, to structurize the representations for segmentation. The energy function of the graph-cuts is formulated by a data term and a smoothness term. To make the graph-cuts more suitable for our application, the approximation errors given by kernel sparse coding are further utilized in the data term, while the smoothness term takes the multichannel four-way connectivity neighborhood into consideration. The segmentation result is thus more accurate, because a structured kernel sparse representation helps the decision making in segmentation, especially segmentation in boundaries.

## Result and Discussion

The proposed framework is evaluated on 20 high-grade gliomas cases in the training set of the multimodal brain tumor segmentation (BRATS) benchmark 2013 [3] with two-fold cross validation. The framework is implemented on MATLAB using a computer with an Intel processor (i7-3930K, 3.20 GHz) and 32 GB of random-access memory. We evaluate our framework in the following three regions:

- ◆ region 1: complete tumor (label 1+2+3+4)
- ◆ region 2: tumor core (label 1+3+4)
- ◆ region 3: enhancing core (label 4).

The Dice similarity coefficient, Jaccard index, and sensitivity are reported [3].

In Figure 2, two segmentation examples generated by the framework are shown. The qualitative results show

**Table 1. The evaluation of performance on BRATS 2013 training cases (high-grade gliomas).**

	Dice		Jaccard		Sensitivity	
	Mean	STD	Mean	STD	Mean	STD
Region 1	81.1	9.3	69.2	12.5	81.9	13.6
Region 2	62.9	17.6	48.0	17.3	69.3	22.4
Region 2*	65.3	14.2	50.0	15.2	72.4	18.2
Region 3	69.7	17.2	55.6	17.8	70.1	22.5
Region 3*	71.9	14.4	57.7	15.5	72.9	19.2

\*Denotes the scores are calculated excluding the result of patient0012. STD: standard deviation.

that our approach is able to identify where the tumor is and to differentiate different brain tumor tissues with relatively high accuracy. The averages and standard deviations of the Dice similarity coefficient, Jaccard index, and sensitivity that are achieved by our method is provided in Table 1. We report all scores twice in region 2 and region 3 due to the peculiar characteristics of one patient, which significantly worsened the overall performance. The tumor in this case consisted mainly of nonenhancing core and edema, which shared similar intensities in all channels. Therefore, a nonenhancing core is easily misclassified as edema, which causes low scores in the evaluation of region 2 and region 3. The performance of our framework is comparable to state-of-the-art discriminative approaches. In addition, we find that our framework generalizes well to unseen images, although smaller training sets are needed.



## Conclusion

In this work, we focused on developing an automated framework using structured kernel sparse representation to deal with the challenging multilabel brain tumor segmentation task. The sparse property of medical images was leveraged in our study, and results showed the effectiveness of our method. We plan to further research the power of sparsity in our future works for medical image analysis and volume visualization.

**In the training phase, kernel-dictionary learning leads to class-specific dictionaries that are optimized for each task-relevant class.**

## About the Authors

**Xuan Chen** (xuan.chen@u.nus.edu) earned her B.E. degree in communication engineering from Beijing University of Posts and Telecommunications, China, in 2014. She is currently pursuing her Ph.D. degree in the Department of Electrical and Computer Engineering at the National University of Singapore. Her research interests include medical image analysis, visualization, and machine learning.

**Binh P. Nguyen** (phubinh@ieee.org) earned his B.Eng. degree in information technology in 2002 and his M.Sc. degree in information processing and communications in 2004 from Hanoi University of Science and Technology, Vietnam, where he was also a lecturer in the School of Information and Communication Technology. He earned his Ph.D. degree in the Department of Electrical and Computer Engineering at the National University of Singapore. After obtaining his Ph.D. degree, he worked as a research fellow in the Department of Mechanical Engineering at the National University of Singapore and then as a research fellow in the Centre for Computational Biology, Duke-National University of Singapore Medical School, Singapore. Currently, he is a scientist at the Institute of High Performance Computing, Agency for Science, Technology and Research, Singapore. His research interests include medical image analysis and visualization, health data sciences and informatics, computational intelligence and interdisciplinary computing, multimedia analysis, and human-machine systems.

**Chee-Kong Chui** (mpecck@nus.edu.sg) earned his Ph.D. degree from the University of Tokyo, Japan. He is an associate professor with the Control and Mechatronics Division of the Department of Mechanical Engineering at the National University of Singapore, where his group focuses on medical devices, robotics, imaging, and simulations. His current focus on immersive media involves the provision of visual and haptic cues to assist humans in the training of hand-eye coordination and to augment hand-eye coordination in a mixed-reality environment. He is a Senior Member of the IEEE and was chairman of the IEEE Engineering in Medicine and Biology Society Singapore Chapter (2015–2016).

**Sim-Heng Ong** (eleongsh@nus.edu.sg) earned his B.E. degree from the University of Western Australia and his Ph.D. degree from the University of Sydney, Australia. He is an associate professor in the Department of Electrical and Computer Engineering at the National University of Singapore. His research interests are in signal and image processing. He has published extensively in international journals and conference proceedings.

## References

- [1] S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, "A survey of MRI-based medical image analysis for brain tumor studies," *Phys. Med. Biol.*, vol. 58, no. 13, pp. R97, July 2013.
- [2] N. Gordillo, E. Montseny, and P. Sobrevilla, "State of the art survey on MRI brain tumor segmentation," *Magn. Reson. Imaging*, vol. 31, no. 8, pp. 1426–1438, Oct. 2013.
- [3] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K.-V. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, L. Lanczi, E. Gerstner, M.-A. Weber, T. Arbel, B. B. Avants, N. Ayache, P. Buendia, D. L. Collins, N. Cordier, J. J. Corso, A. Criminisi, T. Das, H. Delingette, Ç. Demiralp, C. R. Durst, M. Dojat, S. Doyle, J. Festa, F. Forbes, E. Geremia, B. Glocker, P. Golland, X. Guo, A. Hamamci, K. M. Iftekharuddin, R. Jena, N. M. John, E. Konukoglu, D. Lashkari, J. A. Mariz, R. Meier, S. Pereira, D. Precup, S. J. Price, T. R. Raviv, S. M. S. Reza, M. Ryan, D. Sarikaya, L. Schwartz, H.-C. Shin, J. Shotton, C. A. Silva, N. Sousa, N. K. Subbanna, G. Szekely, T. J. Taylor, O. M. Thomas, N. J. Tustison, G. Unal, F. Vasseur, M. Wintermark, D. H. Ye, L. Zhao, B. Zhao, D. Zikic, M. Prastawa, M. Reyes, and K. van Leemput, "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015.
- [4] M. Prastawa, E. Bullitt, S. Ho, and G. Gerig, "A brain tumor segmentation framework based on outlier detection," *Med. Image Anal.*, vol. 8, no. 3, pp. 275–283, Sept. 2004.
- [5] B. H. Menze, K. van Leemput, D. Lashkari, M.-A. Weber, N. Ayache, and P. Golland, "A generative model for brain tumor segmentation in multi-modal images," in *Proc. Workshops Medical Image Computing and Computer Assisted Intervention Challenges*, Beijing, China, 2010, pp. 151–159.
- [6] E. Geremia, B. H. Menze, N. Ayache, "Spatial decision forests for glioma segmentation in multi-channel MR images," In the Workshop Proceedings of the Medical Image Computing and Computer Assisted Intervention Challenges (MICCAI), Nice, France, 2012, pp. 14–17.
- [7] J. J. Thiagarajan, K. N. Ramamurthy, D. Rajan, A. Spanias, A. Puri, D. Frakes, "Kernel sparse models for automated tumor segmentation," *Int. J. Artif. Intell. Tools*, vol. 23, no. 3, June 2014.
- [8] X. Chen, B. P. Nguyen, C.-K. Chui, and S.-H. Ong, "Automated brain tumor segmentation using kernel dictionary learning and superpixel-level features," in *Proc. IEEE Conf. Systems, Man, and Cybernetics*, Budapest, Hungary, 2016, pp. 2547–2552.
- [9] H. Nguyen, V. M. Patel, N. M. Nasrabadi, and R. Chellappa, "Kernel dictionary learning," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Kyoto, Japan, 2012, pp. 2021–2024.