A Machine Learning Approach for Brain Tissue Recognition of Human Brain Slice Images

Jiawei Cui, cuijiaweiyes@gmail.com, KIT

Plamen Kiradjiev,kiradjiev@de.ibm.com, IBM

Johannes Kunze von Bischhoffshausen, kunzevb@de.ibm.com, KIT/IBM

Abstract

This work presents a Machine Learning (ML) approach for classifying areas of brain tissue in a stack of high resolution human brain slice images. Compared with standard image segmentations algorithms, this ML approach provides more reliable results by concentrating on pixel classification. The presented ML approach is fourfold. First, four feature extraction methods were developed to extract features as a basis for the classification procedure. Second, two feature selection approaches were developed and implemented in order to construct feature vectors. Third, Random Forest (RF), Neuronal Networks (NN), and a novel ensemble Meta classifier constructed by different multilayer perceptron (MLP) were implemented in our classifier construction procedure. Finally, a post-processing method based on graph cut algorithm was used to enforce a smoother classification result into coherent regions. This paper details the feature extraction part and illustrates its application conceiving initial results of a small subset of brain slices.

1 Introduction

1.1 Problem definition & research questions

The basis of the presented work is a stack of colored (RGB) images taken during histological sectioning of a tissue block from a post-mortem human brain. The tissue block was frozen at -80°C and cut into 843 slices, each 70 µm thick. Images were taken right after the extraction of each individual slice using a high-resolution industry-grade camera. Both, the histological processing and image collection has been carried out in the lab of Prof. Katrin Amunts at the Institute of Neuroscience and Medicine (INM1) at Research Center Juelich in Germany. At INM1, such “block-face” images are successfully reconstructed into consistent 3D volumes, and used as reference data for the significantly more difficult 3D reconstruction of the subsequent histological scans that typically suffer from nonlinear distortions and artifacts. Such 3D reconstructions play an important role in the EU FET flagship “Human Brain Project”.

Due to the lack of contrast between the sectioning plane and the adjoining underlying brain surface, standard image segmentations algorithms often produce unreliable results for this kind of data. Therefore, neuroscientists in Juelich have spent significant manual effort to segment the sectional planes of brain tissue manually.

Figure 1.1 Image ressource of slice 450

In order to reduce the manual effort for marking areas of brain tissue, a ML approach needs to be developed. Since the manually labeled data has been provided together with the images, INM1 is particularly able to investigate supervised learning techniques to solve the problem. Based on supervised learning techniques, INM1 should be able to segment future datasets based on a small sample of manually marked image sections as training data. Based on the trained model, INM1 should be able to automatically segment the remaining images by classifying pixels of slice image into two classes- one identifying a pixel as part of the sectioning plane, and the other associating it to the background (Figure 1.1 - Left image is the original brain slice image. On the is the “blockface” image, which provides label data.).

2 Methodology

This project is a collaborative project of KIT, IBM and INM1 at the Smart Data Innovation Lab (SDIL), which is an initiative intended for cutting-edge research in the area of data engineering. INM1 plays a role as data sponsor and propose requirements. KIT academics jointly conducted this research project with support of IBM-provided software and infrastructure (Figure 2.1).

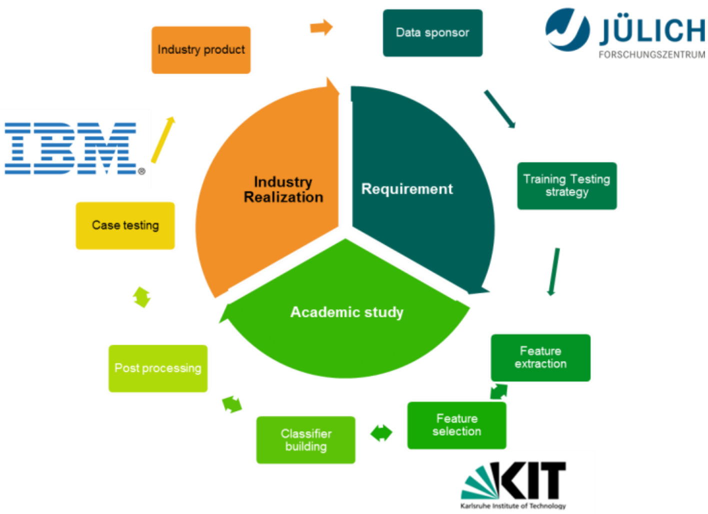

Figure 1.2 Consortium of smart brain project

2.1 Training Testing strategy

Strategy of training and testing laid the basis for each ML project. In order to improve the accuracy of models, classifiers learned by a certain slice are used for classification of neighboring slices. We define the distance between two training slice as scope N, number of training slice needed is 843/N. Training slices must be manually segmented, a large value of scope needs less slices for training, which reducing manual work obviously. In order to balance accuracy and manual work, scope distance is set as 20. Consequently only 5 percent of slices need to be manual segmented by INM1.

As an example, given training slices slice 450 and slice 470, testing scope of slice 450 is from slice 431 to slice 469. Our final generated images of slice 451 to slice 469 are created by classifiers, which trained by slice 450 and slice 470 indiviually.

2.2 Feature extraction

Each pixel is described by RGB values as a basis for feature extraction. We developed four feature extraction methods, which are presented in the following with an emphasis on the developed method based on LCH and linear regression.

First extraction method is filter based extraction method. We utilize three popular denoising filters to describe context information of pixel, i.e. Gaussian filter and two edge preserving filter Bilateral filter and Total Variation filter (C.Tomasi et al, 1998; Ivan W et al, 2010). Second method is the pre-segmentation extraction method. We append three additional image channels with help of Haematoxylin-Eosin-DAB color space (Ruifrok AC et al, 2001). Next feature selection method is used to emphasis edge pixel. Histograms of oriented gradients (HOG) and image entropy are chosen to implement this extraction method (Dalal, N et al, 2005; R.M. Haralick et al, 1985).

Last feature extraction method is based on LCH and linear regression. Main idea of this method is inspired by signal processing method. LCH generates red-green-blue signals of pixel by summarizing information of all pixels within the LCH window. Indeed, uniqueness of signals induces obvious difference between similar pixels, which meet the motivation of enhancing the difference between similar pixels thus improving descriptive ability of features (Figure 2.2).

For each pixel we use a LCH window to extract context information, i.e. the center of this window is the pixel we calculated for. All the pixels in this window serve for providing context information of this central pixel. Scale of LCH window is a critical parameter for the quality of features, since length of this window determines the scope of context information we calculated with. We set width (x’) and length (y’) of our LCH window as 151 with consideration of boundary limitations, i.e. for each pixel we utilize 151*151=22801 pixels to extract context information. In short, we translate feature extraction of pixel to feature extraction of signals under the assumption that features describe the signals efficiently also describe pixels efficiently.

Intuitively, the peaks of each signal can be regarded as features. We developed a peak finding algorithm (interpolate) and extract the maximal value and interpolate central of peaks as features. Instead of mean value of signal we extract media value for features with consideration of non-repeatability. With the help of Discrete Fourier Transformation we utilize the mean amplitude of a specific frequency bands as feature. In conclusion, with described approaches we extract 12 features for red-green -blue signals.

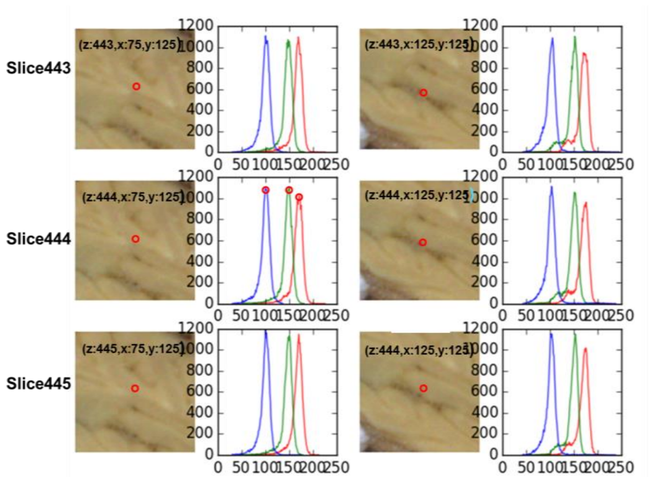

Figure 2.2 Red- green-blue signals of similar pixels

Features extracted from local two color space coordinate system are based on linear regression. Discriminating of linear regression line of each pixel make it plausible that features extracted from linear regression can be served as efficient features. A linear regression line can be described exactly with only two values, or rater slope and intercept. For each pixel we extract 2*3=6 features by implement this approach (Figure 2.2).

We regard the dimension of width of image, length of image and slice stack as x, y, z coordinate system. From the xy plane we extracted 18 features with window size 151*151. For xz and xy plane we extract also 18 features with 151*65 as window size respectively.

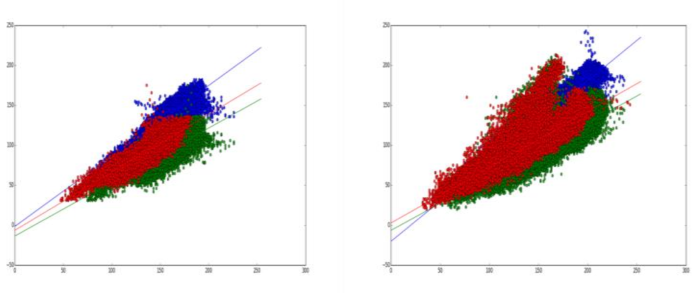

Figure 2.2 Linear regression of two similar pixels

Red-green-blue values are regarded as 3 basic features. We extracted another 81 features based on red-green-blue values. 22 features are extracted by filter based feature extraction algorithm, 3 features are extracted by pre-segmentation extraction method, 2 features are extracted by edge intensify extraction method, LCH provide most of the features, from 3 planes we extracted 54 features (Table 1).

Extraction method	Features number
basic	3
Filter based	22
Pre_segmentaion	3
Edge intensify	2
LCH	54
Total	84

Table 1 Overview of features

2.3 Feature selection

In consideration of loss original feature information widely used dimension reduction methods such as PCA, Autoencoder are not applied. Instead, two feature selection approaches are adopted. Feature filter method is implemented based on statistical test of the analysis of variance (ANOVA) with F-Score as ranking function (N. Elssied et al., 2014). In order to improve the performance of Filter feature selection method we also develop an Ensemble feature selection method, which implement a flexible forward selection strategy.

2.4 Classifier construction

Finding suitable classifiers are very important for testing result. We utilize IBM SPSS Modeler 17.1 to construct classifiers (IBM, 2015). Analytic server 2.1 of SPSS Modeler 17.1 supports Spark integration and provides the possibilities to integrate python and R. Hence, the feature extraction and selection algorithms can be also integrated into SPSS Modeler.

We choose two popular classification algorithms for classifier construction. Random Forrest (RF) is an ensemble learning method by constructing a multitude of decision trees during training time, which correct for decision trees' habit of overfitting (Ho. et al., 1995). Multi-layer perceptron (MLP) is chosen since our feature selection algorithm executes normalization for features. Normalization features proven to be efficient for MLP (Bishop et al., 1995). In order to avoid overfitting of MLP, we develop a method which implement ensemble method by changing topology and validation set of a series of MLPs.

2.5 Post processing method

Developing a method to post process our created image is helpful to improve quality of created images. We develop a post process algorithm for our created image with help of a graph cut algorithm and Gaussian smoothing algorithm (Egil Bae et al. 2009).

3 Preliminary results

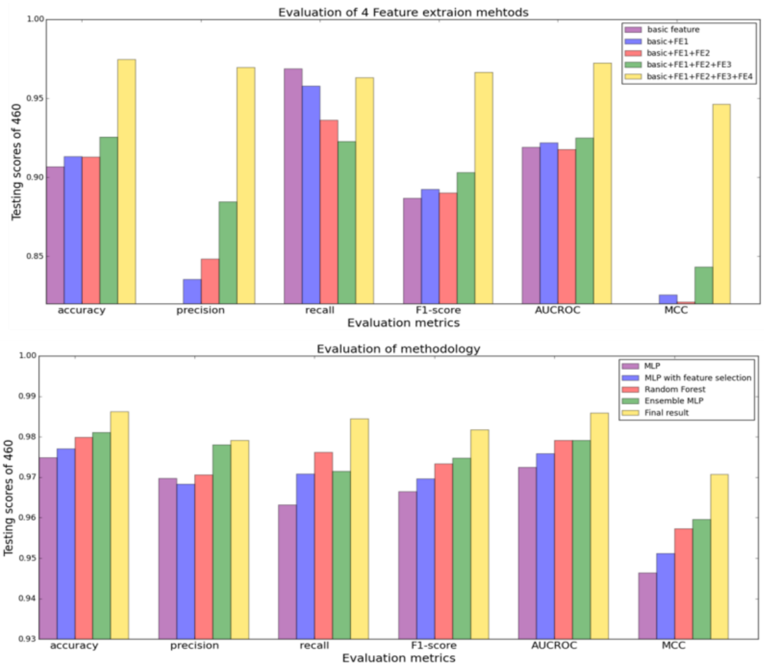

Figure 3.1 Evaluation of feature extraction methods

Preliminary evaluation results are conceived based on a set of 450 slices for training, 460 slices for testing, and the MLP classifier. These results illustrate the performance of those feature extraction methods. Another experiment denotes the final created image following described methodology. We utilize basic evaluation metrics to evaluate our created images, i.e. accuracy, precision, recall, F1-score, AUCROC, MCC.

Among different evaluation criterion features extract by LCH approach played an important role for the feature extraction procedure (Figure 3.1). FE1 denotes filter based method, FE2 denotes pre-segmentation method, FE3 denotes edge intensify method, FE4 denotes LCH based method.

4 Conclusion

This work presented a fourfold Ml approach for classifying areas of brain slice images. First, several feature extraction methods provide efficient features. Features extracted from LCH with help of signal processing and linear regression proven to be efficient that implicit enough contextual information of single pixel. In order to improve performance of classifier, we also utilize two feature selection methods to construct suitable feature vector. A classifier building module two popular classification algorithms RF and MLP is implemented. With purpose of handling of overfitting problem we develop another ensemble MLP classifier with help of IBM SPSS Modeler. Furthermore, post processing of result images is based on the implementation of graph cut algorithms. This provides a new way to combine different classified pixel created image, and proven to be an efficient outlier set reduction method. In conclusion, the initial results are promising and lay the basis for advancing this approach.

5 Acknowledgment

We would like to express our thanks to Dr. Timo Dickscheid and Dr. Stefan Koehnen for their suggestions and help during our research as well as reviewing of our paper. We are also thankful to Prof. Katrin Amunts, Prof. Morris Riedel and Christian Bodenstein for providing the source data, scientific material and allocating INM1 resources. We appreciate the responsiveness and support of the SDIL team, especially Erekle Magradze, Jan Erik Sundermann and Nico Schlitter.

References

Katrin Amunts, Claude Lepage, and Timo Dickscheid et al. “BigBrain - an ultra-high resolution 3D human brain model”. In: *Science* 340 (2013), pp. 1472-1475.

E. Davies and Royal Holloway. “Machine Vision: Theory, Algorithms and Practicali- ties ”. In: *Academic Press* (2012).

Navneet Dalal and Bill Triggs. “Histograms of oriented gradients for human detec- tion”. In: *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 1. IEEE. 2005, pp. 886–893.

C. Tomasi and R. Manduchi. “Bilateral Filtering for Gray and Color Images. ” In:*Proceedings of the1998 IEEE International Conference on Computer Vision (1998).*

Ivan W. Selesnick and Ilker Bayram. “Total Variation Filtering”. In: (2010).

Robert M Haralick and Linda G Shapiro. “Image segmentation techniques”. In: (1985), pp. 2–9.

AC Ruifrok and DA Johnston. “Quantification of histochemical staining by color deconvolution”. In: *Analytical and quantitative cytology and histology / the Inter- national Academy of Cytology [and] American Society of Cytology* 23.4 (Aug. 2001), pp. 291–299.

Massar and Melody L. “Local histograms for per-pixel classification”. In: (2012).

Grosan Crina, Abraham Ajith, and Chis Monica. “Swarm intelligence in data mining”.In: (2006).

Tin Kam Ho. “Random decision forests”. In: (1995), pp. 278–282.

Ji Liu, Ryohei Fujimaki, and Jieping Ye. “Forward-backward greedy algorithms for general convex smooth functions over a cardinality constraint”. In: *arXiv preprint arXiv:1401.0086* (2013).

Nadir Omer Fadl Elssied, Othman Ibrahim, and Ahmed Hamza Osman. “A novel feature selection based on one-way anova f-test for e-mail spam classification”. In: *Research Journal of Applied Sciences, Engineering and Technology* 7.3 (2014), pp. 625– 638.

IBM. “SPSS Modeler 17.1 Documentation”. In: (2015). url: http://www-01.ibm.com/ support/docview.wss?uid=showg27046116.

Christopher M Bishop. “Neural networks for pattern recognition”. In: (1995).

Egil Bae, Juan Shi, and Xue-Cheng Tai. “Graph cuts for curvature based image denoising”. In: *Image Processing, IEEE Transactions on* 20.5 (2011), pp. 1199–1210.