

Ischemic Stroke Lesion Segmentation

www.isles-challenge.org

Proceedings
5th October 2015
Munich, Germany



UNIVERSITÄT ZU LÜBECK



Technische Universität München

u^b

UNIVERSITÄT
BERN

MICCAI15
MUNICH

Preface

Stroke is the second most frequent cause of death and a major cause of disability in industrial countries. In patients who survive, stroke is generally associated with high socioeconomic costs due to persistent disability. Its most frequent manifestation is the ischemic stroke, whose diagnosis often involves the acquisition of brain magnetic resonance (MR) scans to assess the stroke lesion's presence, location, extent, evolution and other factors. An automated method to locate, segment and quantify the lesion area would support clinicians and researchers alike, rendering their findings more robust and reproducible.

New methods for stroke segmentation are regularly proposed. But, more often than desirable, it is difficult to compare their fitness, as the reported results are obtained on private datasets. Challenges aim to overcome these shortcomings by providing (1) a public dataset that reflects the diversity of the problem and (2) a platform for a fair and direct comparison of methods with suitable evaluation measures. Thus, the scientific progress is promoted.

With ISLES, we provide such a challenge covering ischemic stroke lesion segmentation in multi-spectral MRI data. The task is backed by a well established clinical and research motivation and a large number of already existing methods. Each team may participate in either one or both of two sub-tasks:

SISS Automatic segmentation of ischemic stroke lesion volumes from multi-spectral MRI sequences acquired in the sub-acute stroke development stage.

SPES Automatic segmentation of acute ischemic stroke lesion volumes from multi-spectral MRI sequences for stroke outcome prediction.

The participants downloaded a set of training cases with associated expert segmentations of the stroke lesions to train and evaluate their approach, then submitted a short paper describing their method. After reviewing by the organizers, a total of 20 articles were accepted and compiled into this volume. At the day of the challenge, each teams' results as obtained on an independent test set of cases will be revealed and a ranking of methods established.

For the final ranking and more information, visit WWW.ISLES-CHALLENGE.ORG.

Oskar Maier, Universität zu Lübeck

Mauricio Reyes, University of Bern

Björn Menze, TU Munich

Organizers

Oskar Maier, Universität zu Lübeck, Germany

Mauricio Reyes, University of Bern, Switzerland

Björn Menze, TU Munich, Germany

Sponsoring Institutions

Institute of Medical Informatics, Universität zu Lübeck, Germany

Institute for Surgical Technology & Biomechanics, University of Bern, Switzerland

Computer Science, TU Munich, Germany

Table of contents

SISS (p. 01)

- | | | |
|-------------------|--|------|
| ISLES-P-01 | Segmentation of Stroke Lesions in Multi-spectral MR Images Using Bias Correction Embedded FCM and Three Phase Level Set
<i>Chaolu Feng, Dazhe Zhao and Min Huang</i> | p.03 |
| ISLES-P-02 | A Novel Framework for Sub-acute Stroke Lesion Segmentation Based on Random Forest
<i>Liang Chen, Paul Bentley and Daniel Rueckert</i> | p.09 |
| ISLES-P-03 | Multi-Scale 3D Convolutional Neural Networks for Lesion Segmentation in Brain MRI
<i>Konstantinos Kamnitsas, Liang Chen, Christian Ledig, Daniel Rueckert and Ben Glocker</i> | p.13 |
| ISLES-P-04 | Random Forests with Selected Features for Stroke Lesion Segmentation
<i>Oskar Maier, Matthias Wilms and Heinz Handels</i> | p.17 |
| ISLES-P-05 | Ischemic Stroke Lesion Segmentation Using Local Gradient and Texture Features
<i>Syed MS Reza, Linmin Pei and Khan M Iftekharuddin</i> | p.23 |
| ISLES-P-06 | ISLES Challenge 2015: A Voxel-wise, Cascaded Classification Approach to Stroke Lesion Segmentation [also in SPES]
<i>David Robben, Daan Christiaens, Janaki Raman Rangarajan, Jaap Gelderblom, Philip Joris, Frederik Maes and Paul Suetens</i> | p.27 |
| ISLES-P-07 | ISLES (SISS) Challenge 2015: Segmentation of Stroke Lesions Using Spatial Normalization, Random Forest Classification and Contextual Clustering
<i>Hanna-Leena Halme, Antti Korvenoja and Eero Salli</i> | p.31 |
| ISLES-P-08 | Stroke Lesion Segmentation of 3D Brain MRI Using Multiple Random Forests and 3D Registration
<i>Ching-Wei Wang and Jia-Hong Lee</i> | p.35 |

Table of contents

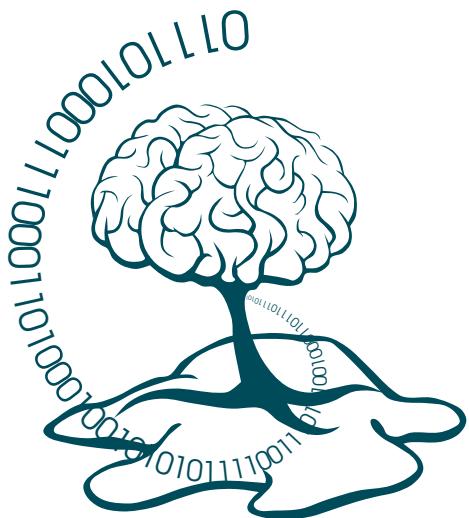
SISS (p. 01) cont.

- ISLES-P-09** Input Data Adaptive Learning (IDAL) for Sub-acute Ischemic Stroke Lesion Segmentation p.39
Michael Goetz, Christian Weber and Klaus Maier-Hein
- ISLES-P-10** Distributed Deep Learning Framework for Large-Scale 3D Medical Image Segmentation [also in SPES] p.43
Tassilo Klein, Nematollah K Batmanghelich and William M Wells III
- ISLES-P-11** Automatic Ischemic Stroke Lesion Segmentation in Multi-Spectral MRI Images Using Random forests Classifier p.47
Qaiser Mahmood and Abdul Basit
- ISLES-P-12** ISLES Challenge 2015: Automated Model-Based Segmentation of Ischemic Stroke in MR Images [also in SPES] p.51
Tom Haeck, Frederik Maes and Paul Suetens
- ISLES-P-13** A Convolutional Neural Network Approach to Brain Lesion Segmentation [also in SPES] p.55
Francis Dutil, Mohammad Havaei, Chris Pal, Hugo Larochelle and Pierre-Marc Jodoin
- ISLES-P-14** Hierarchical Segmentation of Normal and Lesional Structures Combining an Ensemble of Probabilistic Local Classifiers and Regional Random Forest Classification p.61
Andrew Jesson and Tal Arbel
- ISLES-P-15** A Vascular Territory Prior for Bayesian Sub-Acute Ischemic Stroke Lesion Segmentation p.67
Senan Doyle, Florence Forbes, and Michel Dojat
- ISLES-P-16** Segmentation of Ischemic Stroke Lesions Using Dictionary Learning p.73
Jean Huang, Lawrence Staib, and James Duncan
- ISLES-P-20** Prediction of Ischemic Lesions using Local Image Properties and Random Forests p.79
John Muschelli

Table of contents

SPES (p. 83)

- ISLES-P-17** Segmenting the Ischemic Penumbra: A Spatial Random Forest Approach with Automatic Threshold Finding p.85
Richard McKinley, Levin Häni, Roland Wiest and Mauricio Reyes
- ISLES-P-18** Lesion Segmentation of the Penumbra in Acute Stroke in the MICCAI 2015 ISLES Challenge p.89
Elias Kellner, Karl Egger, Maddalena Strumia, Valerij G Kiselev, Horst Urbach and Marco Reisert
- ISLES-P-19** Random Forests for Acute Stroke Penumbra Estimation p.93
Oskar Maier, Matthias Wilms and Heinz Handels
- ISLES-P-06** ISLES Challenge 2015: A Voxel-wise, Cascaded Classification Approach to Stroke Lesion Segmentation [also in SISS] p.27
David Robben, Daan Christiaens, Janaki Raman Rangarajan, Jaap Gelderblom, Philip Joris, Frederik Maes and Paul Suetens
- ISLES-P-12** ISLES Challenge 2015: Automated Model-Based Segmentation of Ischemic Stroke in MR Images [also in SISS] p.51
Tom Haeck, Frederik Maes, and Paul Suetens
- ISLES-P-13** A Convolutional Neural Network Approach to Brain Lesion Segmentation [also in SPES] p.55
Francis Dutil, Mohammad Havaei, Chris Pal, Hugo Larochelle and Pierre-Marc Jodoin
- ISLES-P-10** Distributed Deep Learning Framework for Large-Scale 3D Medical Image Segmentation [also in SISS] p.43
Tassilo Klein, Nematollah K Batmanghelich and William M Wells III



Sub-acute Ischemic Stroke Segmentation

ISLES 2015: SISS

Proceedings
5th October 2015
Munich, Germany



UNIVERSITÄT ZU LÜBECK



Technische Universität München



Segmentation of Stroke Lesions in Multi-spectral MR Images Using Bias Correction Embedded FCM and Three Phase Level Set

Chaolu Feng^{1,2 *}, Dazhe Zhao^{1,2}, and Min Huang^{1,3}

¹ College of Information Science and Engineering, Northeastern University, Shenyang, Liaoning, 110819, China,
fengchl@ise.neu.edu.cn

² Key Laboratory of Medical Image Computing of Ministry of Education, Northeastern University, Shenyang, Liaoning, 110819, China

³ State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang, Liaoning, 110819, China

Abstract. **Keywords:** lesion segmentation; bias correction; Fuzzy C-Means; level set

1 Introduction

Ischemic stroke is the third leading cause of death in industrialized countries [8]. Due to its excellent soft tissue contrast, magnetic resonance imaging (MRI) has become to be the modality of choice for clinical evaluation of ischemic stroke lesions [4]. As ischemic stroke lesions usually change over time and secondary and remote changes may occur, it is therefore necessary characterizing the tissue changes with different acquisition parameters to produce images of the same physical space in distinctive spectral signatures [1].

In clinical practice, Diffusion weighted images (DWI), T1-weighted (T1W), T2-weighted (T2W) and fluid attenuated inversion recovery (FLAIR) images are often acquired to monitor progression of strokes [7]. In acute stage, hyperintense signal observed on DWI provides important information about the anatomical location and extent of the infarcted territory. In more chronic phase, T2W and FLAIR images are normally used to delineate the final lesion volume. Chronic ischemic lesions appear as hyperintense regions in FLAIR with some heterogeneity within the lesion volume due to ongoing gliosis and demyelination [6].

Early and accurate diagnosis of brain lesion by multi-spectral magnetic resonance images is the key for implementing successful therapy and treatment planning [5]. However, the diagnosis is a very challenging task and can only be performed by professional neuro-radiologists. Lesion segmentation can improve this situation and help radiologists diagnose and make treatment plan. However, due to the variety of the possible shapes, locations, and intensity inhomogeneity, accurate segmentation is still a challenging task [2]. Manual segmentation can be

* Corresponding Author.

performed by trained radiologists, but it is a tedious and time consuming task and is non-reproducible [4].

In this paper, we propose a automatic ischemic stroke lesion segmentation algorithm in multi-spectral images (DWI, T1-w, T2-w, and FLAIR) using bias correction embedded FCM and three phase level set method. The rest of this paper is organized as follows.

2 Method

Before performing lesion segmentation, the input images of different modalities are first rigidly registered in the same coordinate system. Non-brain tissues are then removed from the images. Lesion segmentation is finally carried out with two major steps: 1) preliminary classification of normal brain tissues and lesions in multi-spectral MR images using an improved FCM with the capability of dealing with intensity inhomogeneities, and 2) boundary refinement of preliminary classification using a three phase level set designed for multiple spectral images. More details will be given in the following subsections.

2.1 Image Model

Given an observed MR brain image I defined on a continuous domain $\Omega \subset R^2$, its inhomogeneous intensities can be viewed as a product of the true image J and the bias field b , i.e,

$$I(\mathbf{x}) = b(\mathbf{x})J(\mathbf{x}) + n(\mathbf{x}) \quad (1)$$

where $\mathbf{x} \in \Omega$ and n is zero-mean additive noise. For multi-spectral MR images, we rewrite the above model into the following vector form:

$$\mathbf{I}(x) = \mathbf{b}(x) \cdot \mathbf{J}(x) + \mathbf{n}(x) \quad (2)$$

where $\mathbf{I}(x) = (I_1(x), I_2(x), \dots, I_L(x))$, $\mathbf{b}(x) = (b_1(x), b_2(x), \dots, b_L(x))$, $\mathbf{J}(x) = (J_1(x), J_2(x), \dots, J_L(x))$, $\mathbf{n}(x) = (n_1(x), n_2(x), \dots, n_L(x))$, \cdot is the multiplication operator of corresponding components of two vectors, and L is the total channel number.

2.2 Preliminary Segmentation of Lesions and Normal Tissues

For each image channel, true image characterizes an intrinsic physical property of human brain, which ideally takes a specific intensity for each type of tissue (CSF,WM,GM) and lesions and is therefore assumed to be piecewise constant. That is to say, the true image J_i of the i -th channel approximately takes distinct constant values c_{i1}, c_{i2}, \dots , and c_{iN} for $N - 1$ tissues and lesions in disjoint regions $\Omega_1, \Omega_2, \dots$, and Ω_N , i.e. $J_i(\mathbf{x}) \approx c_{ij}$ for $\mathbf{x} \in \Omega_j$. Then, in view of the image model in Eq. (2), we have

$$I_i(\mathbf{x}) \approx b_i(\mathbf{x})c_{ij} \quad \text{for } \mathbf{x} \in \Omega_j. \quad (3)$$

Therefore, intensities in the set $I_{ij} = \{I_i(\mathbf{x}) : \mathbf{x} \in \Omega_j\}$ form a cluster with the cluster centroid $m_{ij} \approx b_i(\mathbf{x})c_{ij}$. This clustering property indicates that intensities in the image domain Ω can be classified into N clusters with centroids $m_{i1} \approx b_i(\mathbf{x})c_{i1}$, $m_{i2} \approx b_i(\mathbf{x})c_{i2}$, ..., and $m_{iN} \approx b_i(\mathbf{x})c_{iN}$, respectively. To classify these intensities, we define

$$\mathcal{F}_i = \int_{\Omega} \sum_{j=1}^N \lambda_j \| I_i(\mathbf{x}) - b_i(\mathbf{x})c_{ij} \|^2 u_j^q(\mathbf{x}) d\mathbf{x}. \quad (4)$$

where λ_j is any real number that is not less than 1, $\lambda_1, \lambda_2, \dots, \lambda_N$ positive weighting coefficients for the N clusters, and $u_j(\mathbf{x})$ is the membership function that indicates whether pixel \mathbf{x} belongs to the j -th tissue or not.

For the multi-spectral MR images, we define

$$\mathcal{F} = \sum_{i=1}^L \gamma_i \mathcal{F}_i = \sum_{i=1}^L \gamma_i \int_{\Omega} \sum_{j=1}^N \lambda_j \| I_i(\mathbf{x}) - b_i(\mathbf{x})c_{ij} \|^2 u_j^q(\mathbf{x}) d\mathbf{x}. \quad (5)$$

where γ_i are positive weighting coefficient for the i -th spectral image. This objective function is minimized when high membership values are assigned to pixels, intensities of which are close to the centroid, and low membership values are assigned when the pixels are far from the centroids under the condition $\sum_{j=1}^N u_j(\mathbf{x}) = 1$ where $u_j(\mathbf{x}) \in [0, 1]$.

Energy minimization of \mathcal{F} can be achieved by alternately minimizing it with respect to each of its variables. For fixed $b_i(\mathbf{x})$ and $u_j(\mathbf{x})$, $i = 1, 2, \dots, L$, and $j = 1, 2, \dots, N$, we minimize \mathcal{F} with respect to c_{ij} by resolving $\frac{\partial \mathcal{F}}{\partial c_{ij}} = 0$. It is obvious that \mathcal{F} is minimized at $c_{ij} = \hat{c}_{ij}$, given by

$$\hat{c}_{ij} = \frac{\int_{\Omega} b_i(\mathbf{x}) I_i(\mathbf{x}) u_j^q(\mathbf{x})}{\int_{\Omega} b_i^2(\mathbf{x}) u_j^q(\mathbf{x})} \quad (6)$$

For fixed $u_j(\mathbf{x})$ and c_{ij} , $i = 1, 2, \dots, L$, and $j = 1, 2, \dots, N$, we minimize \mathcal{F} with respect to $b_i(\mathbf{x})$ by resolving $\frac{\partial \mathcal{F}}{\partial b_i(\mathbf{x})} = 0$. It can be shown that \mathcal{F} is minimized at $b_i(\mathbf{x}) = \hat{b}_i(\mathbf{x})$, given by

$$\hat{b}_i(\mathbf{x}) = \frac{I_i(\mathbf{x}) \sum_{j=1}^N \lambda_j c_{i,j} u_j^q(\mathbf{x})}{\sum_{j=1}^N \lambda_j c_{i,j}^2 u_j^q(\mathbf{x})} \quad (7)$$

For the case $q > 1$, minimization of \mathcal{F} with respect to $u_j(\mathbf{x})$ can be implemented by resolving the following Lagrangian equation:

$$\sum_{i=1}^L \gamma_i \int_{\Omega} \sum_{j=1}^N \lambda_j \| I_i(\mathbf{x}) - b_i(\mathbf{x})c_{ij} \|^2 u_j^q(\mathbf{x}) d\mathbf{x} - \lambda \left(\sum_{j=1}^N u_j(\mathbf{x}) - 1 \right) = 0 \quad (8)$$

where λ is the Lagrangian multiplier and $\sum_{j=1}^N u_j(\mathbf{x}) = 1$ is the extremum condition. For fixed $b_i(\mathbf{x})$ and c_{ij} , $i = 1, 2, \dots, L$, and $j = 1, 2, \dots, N$, we take

partial derivative of the above equation with respect $u_j(\mathbf{x})$, set the result to 0, and resolve the equations with the constraint that $\sum_{j=1}^N u_j(\mathbf{x}) = 1$. Then, it can be shown that \mathcal{F} is minimized at $u_j(\mathbf{x}) = \hat{u}_j(\mathbf{x})$, given by

$$\hat{u}_j(\mathbf{x}) = \frac{\left(\lambda_j \sum_{i=1}^L \gamma_i \| I_i(\mathbf{x}) - b_i(\mathbf{x})c_{i,j} \|^2 \right)^{\frac{1}{1-q}}}{\sum_{k=1}^N \left(\lambda_k \sum_{i=1}^N \gamma_i \| I_i(\mathbf{x}) - b_i(\mathbf{x})c_{i,k} \|^2 \right)^{\frac{1}{1-q}}} \quad (9)$$

Preliminary segmentation of CSF, WM, GM and stroke lesions is performed in this step in an iterative process.

2.3 Lesions Segmentation Using a Three Phase Level Set Method

To refine boundaries of the preliminary segmentation, we propose a three phase level set formulation as the second step of the proposed method in this subsection. The proposed level set formulation can be seen as an extension of the local intensity clustering (LIC) model with the capability of dealing with intensity inhomogeneities [3]. Preliminary segmentation results are used to initialize the level set function, such that the zero level contour of the initial level set function is near the true lesion boundaries.

Consider a relatively small circular neighborhood with a radius ρ centered at a given point $\mathbf{y} \in \Omega$, defined by $\mathcal{O}_{\mathbf{y}} \triangleq \{\mathbf{x} : |\mathbf{x} - \mathbf{y}| \leq \rho\}$. For each image channel, the bias field in the neighborhood can be ignored due to its slowly and smoothly varying property. Taking into account the constant intensity c_{ij} of the true image J in Ω_j , we obtain

$$b_i(\mathbf{x})J_i(\mathbf{x}) \approx b_i(\mathbf{y})c_{ij} \quad \text{for } \mathbf{x} \in \Omega_j \cap \mathcal{O}_{\mathbf{y}}. \quad (10)$$

This local intensity clustering property allows us to apply the standard K-means algorithm in the following continuous form to classify these local inhomogeneous intensities in the neighborhood $\mathcal{O}_{\mathbf{y}}$. Therefore, taking all the L channel images into account, we define

$$\mathcal{E}_y = \sum_{j=1}^N \lambda_j \int_{\mathcal{O}_y} \left(\sum_{i=1}^L \chi_i \| I_i(\mathbf{x}) - b_i(\mathbf{y})c_{ij} \|^2 \right) u_j(\mathbf{x}) d\mathbf{x} \quad (11)$$

where λ_j is the weighting coefficient used to control size of j -th tissue, χ_i is the weighting coefficient for the i -th channel, and u_j is the membership function of Ω_j . On account of the inherent property of the membership function u_j in representing the subregion Ω_j , \mathcal{E}_y can be rewritten as

$$\mathcal{E}_y = \sum_{j=1}^N \lambda_j \int_{\Omega_j} K_\sigma(\mathbf{x} - \mathbf{y}) \left(\sum_{i=1}^L \chi_i \| I_i(\mathbf{x}) - b_i(\mathbf{y})c_{ij} \|^2 \right) d\mathbf{x} \quad (12)$$

where K_σ is a nonnegative kernel function with the property $\int_{|\mathbf{u}| \leq \sigma} K_\sigma(\mathbf{u}) = 1$.

To ensure the partition $\{\Omega_j\}_{j=1}^N$ of the entire domain Ω to be the one such that \mathcal{E}_y is minimized for all y in Ω , we minimize the integral of \mathcal{E}_y with respect to y over the entire image domain Ω and define

$$\mathcal{E} = \int_{\Omega} \left(\sum_{j=1}^N \lambda_j \int_{\Omega_j} K_{\sigma}(\mathbf{x} - \mathbf{y}) \left(\sum_{i=1}^L \chi_i \| I_i(\mathbf{x}) - b_i(\mathbf{y}) c_{ij} \|^2 \right) d\mathbf{x} \right) d\mathbf{y}. \quad (13)$$

As our goal is to segment lesions, we consider the background and CSF as one region, GM and WM as the second region, and the lesions as the third region. We therefore use $M_1(\phi_1, \phi_2) = (1 - H(\phi_1))(1 - H(\phi_2))$, $M_1(\phi_1, \phi_2) = H(\phi_1)(1 - H(\phi_2))$, and $M_1(\phi_1, \phi_2) = H(\phi_2)$ to represent these regions and rewrite \mathcal{E} as

$$\mathcal{E} = \int_{\Omega} \left(\sum_{j=1}^N \lambda_j e_j(\mathbf{x}) M_j(\phi_1(\mathbf{x}), \phi_2(\mathbf{x})) \right) d\mathbf{x} \quad (14)$$

where

$$e_j(\mathbf{x}) = \int_{\Omega} K_{\sigma}(\mathbf{x} - \mathbf{y}) \left(\sum_{i=1}^L \chi_i \| I_i(\mathbf{x}) - b_i(\mathbf{y}) c_{ij} \|^2 \right) d\mathbf{y} \quad (15)$$

The energy \mathcal{E} defined above is used as the data term of the final energy functional of the proposed level set formulation, which defined by

$$\mathcal{F} = \mathcal{E} + \mathcal{P} + \mathcal{L}. \quad (16)$$

where \mathcal{P} and \mathcal{L} are the regularization term and arc length term defined below to maintain the regularity of the level set functions and smooth the 0-level set contours of the level set functions, respectively.

$$\mathcal{P} = \mu_1 \int \frac{1}{2} (|\nabla \phi_1(\mathbf{x})| - 1)^2 d\mathbf{x} + \mu_2 \int \frac{1}{2} (|\nabla \phi_2(\mathbf{x})| - 1)^2 d\mathbf{x} \quad (17)$$

$$\mathcal{L} = \nu_1 \int |\nabla H(\phi_1(\mathbf{x}))| d\mathbf{x} + \nu_2 \int |\nabla H(\phi_2(\mathbf{x}))| d\mathbf{x} \quad (18)$$

where μ_1, μ_2, ν_1 and ν_2 are weighting coefficients and H is the Heaviside function.

For fixed b and \mathbf{c} , we minimize the final energy functional \mathcal{F} using the standard gradient descent method and obtain

$$\frac{\partial \phi_1}{\partial t} = \delta(\phi_1)(1 - H(\phi_2))(\lambda_1 e_1 - \lambda_2 e_2) + \mu_1 \left(\nabla^2 \phi_1 - \operatorname{div} \left(\frac{\nabla \phi_1}{|\nabla \phi_1|} \right) \right) + \nu_1 \delta(\phi_1) \operatorname{div} \left(\frac{\nabla \phi_1}{|\nabla \phi_1|} \right). \quad (19)$$

For fixed ϕ and b , the optimal \mathbf{c} that minimizes the final energy functional \mathcal{F} is given by

$$c_{ij} = \frac{\int I_i(\mathbf{x}) M_j(\phi_1(\mathbf{x}), \phi_2(\mathbf{x}))(b_i * K_{\sigma})(\mathbf{x}) d\mathbf{x}}{\int M_j(\phi_1(\mathbf{x}), \phi_2(\mathbf{x}))(b_i^2 * K_{\sigma})(\mathbf{x}) d\mathbf{x}}, \quad i = 1, 2, \dots, L \quad \text{and} \quad j = 1, 2, \dots, N. \quad (20)$$

For fixed ϕ and \mathbf{c} , the optimal b that minimizes the final energy functional \mathcal{F} is given by

$$b_i = \frac{\left(I_i \sum_{j=1}^N c_{ij} M_j(\phi_1, \phi_2) \right) * K_\sigma}{\sum_{j=1}^N c_{ij}^2 M_j(\phi) * K_\sigma}, \quad i = 1, 2, \dots, L. \quad (21)$$

2.4 Implementation

The implementation of the proposed method can be straightforwardly expressed as follows.

- Step 1. Remove non-brain tissues from the images and register them in the same coordinate system.
- Step 2. Preliminary classification of normal brain tissues and lesions. Update each variables of energy function defined in Eq. (5) iteratively until convergence criterion has been reached or the iteration number exceeds a predetermined maximum number.
- Step 3. Initialize the level set functions using preliminary classification results and keep ϕ_2 fixed. Update each variables of energy functional defined in Eq. (16) iteratively until convergence criterion has been reached or the iteration number exceeds a predetermined maximum number.

References

1. Chyžhyk, D., Dacosta-Aguayo, R., Mataró, M., Graña, M.: An active learning approach for stroke lesion segmentation on multimodal MRI data. *Neurocomputing* 150, 26–36 (2015)
2. de Haan, B., Clas, P., Juenger, H., Wilke, M., Karnath, H.O.: Fast semi-automated lesion demarcation in stroke. *NeuroImage: Clinical* (2015)
3. Li, C., Huang, R., Ding, Z., Gatenby, J.C., Metaxas, D.N., Gore, J.C.: A level set method for image segmentation in the presence of intensity inhomogeneities with application to mri. *Image Processing, IEEE Transactions on* 20(7), 2007–2016 (2011)
4. Lladó, X., Oliver, A., Cabezas, M., Freixenet, J., Vilanova, J.C., Quiles, A., Valls, L., Ramió-Torrentà, L., Rovira, À.: Segmentation of multiple sclerosis lesions in brain MRI: a review of automated approaches. *Information Sciences* 186(1), 164–185 (2012)
5. Maier, O., Wilms, M., von der Gabeltz, J., Krämer, U.M., Münte, T.F., Handels, H.: Extra Tree forests for sub-acute ischemic stroke lesion segmentation in MR sequences. *Journal of neuroscience methods* 240, 89–100 (2015)
6. Mitra, J., Bourgeat, P., Fripp, J., Ghose, S., Rose, S., Salvado, O., Connelly, A., Campbell, B., Palmer, S., Sharma, G., et al.: Classification Forests and Markov Random Field to Segment Chronic Ischemic Infarcts from Multimodal MRI. In: *Multimodal Brain Image Analysis*, pp. 107–118. Springer (2013)
7. Mitra, J., Bourgeat, P., Fripp, J., Ghose, S., Rose, S., Salvado, O., Connelly, A., Campbell, B., Palmer, S., Sharma, G., et al.: Lesion segmentation from multimodal MRI using random forest following ischemic stroke. *NeuroImage* 98, 324–335 (2014)
8. Rekik, I., Allassonnière, S., Carpenter, T.K., Wardlaw, J.M.: Medical image analysis methods in MR/CT-imaged acute-subacute ischemic stroke lesion: Segmentation, prediction and insights into dynamic evolution simulation models. A critical appraisal. *NeuroImage: Clinical* 1(1), 164–178 (2012)

A novel framework for sub-acute stroke lesion segmentation based on random forest

Liang Chen^{1,2}, Paul Bentley², and Daniel Rueckert¹

¹ BioMedIA group, Department of Computing, Imperial College London, UK

² Division of Brain Sciences, Department of Medicine, Imperial College London, UK

Abstract. Neuroimaging in the context of stroke is becoming more and more important. Quantifying and characterizing stroke lesions is still an open challenge. In this paper, we propose a novel framework to solve this problem. The features we use are intensities of patches from multiscale multimodal magnetic resonance (MR) images. We have built random forest classifiers for different parts of the whole brain. A leave-one-out cross-validation result on SISS training data yields 0.55 in Dice score.

1 Introduction

Stroke is a cerebrovascular accident, in which part of the function of the brain is lost through a decrease of the blood supply [3]. It is the second major cause of death and it may lead to long-term disability [2]. Advanced neuroimaging techniques have been widely used in the diagnosis of stroke. It is normally recommended that patients should undergo either MR or computer tomography (CT) imaging [4]. Diffusion-weighted imaging (DWI), T2-fluid attenuated inversion recovery (FLAIR), T1-weighted imaging, and T2-weighted imaging should be included in the MR sequences, which are regarded as the gold standard in stroke treatment since they are able to show different types of lesions.

Based on MR images, quantifying lesions is important for assessing the progression of the disease and predicting the functional outcomes for patients. However, manual delineation of lesions is extremely time-consuming and the inter-expert consistency is not satisfactory. In this paper, we propose a novel framework for sub-acute stroke lesion segmentation based on the data from the ISLES challenge, MICCAI 2015.

2 Methods

The ISLES challenge released 28 cases for model training, each of which consists of T1-weighted, T2-weighted, diffusion-weighted, and FLAIR images and a corresponding manual delineation of the actual lesions. The overview of our framework is shown in the Figure 1. It consists of six steps.

In the first stage we normalize all images in terms of intensity. For each image X , we apply the formula $\frac{X-\mu}{s}$, where μ is the mean intensity of the tissue in

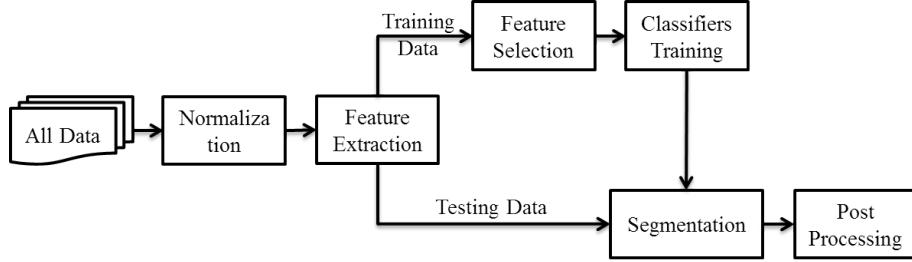


Fig. 1: Overview of the segmentation process.

X and s is its standard deviation. Notably, we exclude 5% outlier voxels with minimum and maximum intensities of the tissue, respectively.

Secondly, we extract features from all images. Intensities of multi-scale patches in each modality are extracted. Specifically, we blur all images at the lower resolutions using Gaussian kernels $\sigma = 1$ and $\sigma = 2$. 5×5 patches are extracted at each scale of each modality. Finally, all patches are converted to vectors and concatenated into a long vector of dimension 300. In a real clinical scenario, the acquired axial slices are typically thick and their thickness can vary significantly. Resampling them into thin slices leads to additional errors. Therefore we prefer pixels, rather than voxels as our features. We parcellate the whole brain into three parts (see Figure 2), including top, middle, and bottom. In the given dataset, there are 65, 40, and 49 slices in the bottom, middle, and top part, respectively. Patches are separated according to their locations and classifiers will be trained for each part individually. The main reason for this is that each part of the brain contains different anatomical structures. The top part contains relatively homogeneous structures. The middle part contains the ventricles and the bottom part contains complicated structures such as the cerebellum. Another reason is that strokes occur most frequently in the middle part of the brain because the main arteries are located there so that the numbers of lesion and normal patches are extremely unbalanced in the top and the bottom part.

In the third step, the data is divided into training and testing sets. In this work, as we will perform the leave-one-out cross-validation, one patient is left out for validation in each round. The remaining training patches will be selected to train classifiers. Since we have a limited number of subjects and not all images contain large lesions, there are significantly less lesion patches than normal ones. If we sample the same number of normal patches as the lesion ones globally, there will be many normal patterns that will be excluded and therefore the test performance will not be satisfactory. We propose to sample different numbers of normal patches for different parts of the brain. In the bottom part, we randomly select 5 times normal patches more than lesion ones since we would like to cover all kinds of normal patterns. For the middle and the top parts, the rates, where the number of normal patches versus lesion ones are 1.5 and 3, respectively.

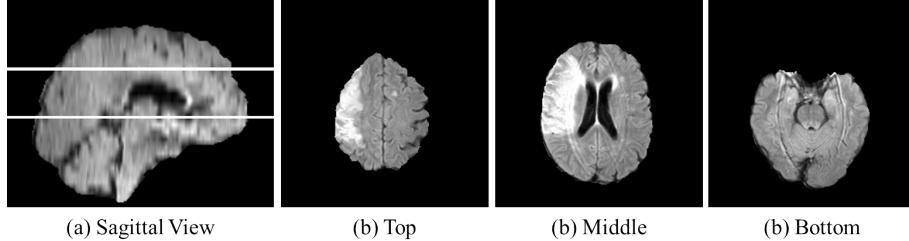


Fig. 2: Illustration of brain parcellation. This is a FLAIR example showing each part of the brain. The structure of the top part is relatively simple. The middle part has ventricles and more lesions. The bottom consists of more complex structures and less lesions.

Subsequently, we can train three classifiers based on the patches selected from three parts of the brain. In this paper, the standard random forests [1] are used as patch classifiers. In each forest, 100 trees are developed. Afterwards, the classifiers can be evaluated with the test data to distinguish how abnormal they are. The outputs of the classifiers are the probabilities that characterize the abnormality of the test patches.

Finally, we perform some post-processing operations. Considering that the lesions in the brain are typically continuous, we smooth the probabilities of the slice at the joint of bottom and middle part of the brain by averaging the probabilities of the neighbouring slices where the outputs given by the classifier of the bottom volume and the classifier for the middle volume have sharp difference occasionally. Based on the resulting probabilities, a threshold $\Theta = 0.6$ is applied to obtain the binary lesion map. For some patients with lacuna infarction, the lesion appearance on the FLAIR image used to be a hyperintense clot with a dark 'hole' inside, which can hardly be detected by the classifiers. Therefore we perform a morphological operation to fill up these 'holes'.

3 Experiments and Results

The method mentioned above is performed on the given training data and we achieve the results presented in Table 1. The leave-one-out cross-validation is used. It is obvious that the results are good if the subject have large lesions. The very small lesions shown in Case 26 and 27 can never been detected.

4 Discussion and Conclusion

We have presented a novel framework for sub-acute stroke lesion segmentation and we achieved an average Dice score of 0.55. In the future, we proposed to collect more data so that there are sufficient data for all kinds of lesions. As a result, different classifiers can be trained for different conditions, where the lesion sizes vary.

Table 1: The results on the SISS training data.

Case ID	ASSD	Dice	Hausdorff Distance	Precision	Recall
1	0.96	0.93	48.88	0.91	0.94
2	2.15	0.83	52.43	0.75	0.92
3	1.59	0.62	49.75	0.48	0.89
4	2.34	0.79	55.29	0.93	0.69
5	1.44	0.87	45.52	0.85	0.88
6	1.05	0.90	28.46	0.86	0.95
7	1.76	0.86	58.60	0.84	0.88
8	19.4	0.49	94.22	0.35	0.84
9	2.11	0.86	24.19	0.90	0.82
10	5.19	0.67	73.74	0.83	0.56
11	10.09	0.59	92.50	0.43	0.94
12	8.30	0.53	67.60	0.41	0.74
13	12.56	0.23	70.80	0.15	0.57
14	1.55	0.81	81.65	0.89	0.75
15	2.22	0.83	48.93	0.73	0.95
16	40.52	0.02	120.59	0.01	0.14
17	11.02	0.49	93.01	0.67	0.38
18	8.41	0.59	83.96	0.47	0.80
19	13.78	0.16	56.86	0.09	0.69
20	4.44	0.77	126.15	0.81	0.74
21	41.43	0.07	140.13	0.04	0.46
22	6.75	0.52	79.76	0.72	0.41
23	22.80	0.38	90.14	0.25	0.77
24	15.38	0.44	102.51	0.31	0.76
25	9.46	0.60	87.87	0.43	0.97
26	29.92	0	85.65	0	0
27	59.99	0	124.96	0	0
28	11.13	0.67	76.69	0.56	0.56
Average		12.42 \pm 14.29	0.55 \pm 0.29	77.17 \pm 28.60	0.52 \pm 0.32
				0.68 \pm 0.27	

References

- [1] Breiman, L.: Random forests. *Machine learning* 45(1), 5–32 (2001)
- [2] Donnan, G., Fisher, M., Macleod, M., Davis, S.M.: Stroke. *The Lancet* 371, 1612–1623 (2008)
- [3] Sims, N.R., Muyderman, H.: Mitochondria, oxidative metabolism and cell death in stroke. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* 1802(1), 80–91 (2010)
- [4] Wintermark, M., Albers, G.W., Alexandrov, A.V., Alger, J.R., Bammer, R., Baron, J.C., Davis, S., Demaerschalk, B.M., Derdeyn, C.P., Donnan, G.A., et al.: Acute stroke imaging research roadmap. *American Journal of Neuroradiology* 29(5), e23–e30 (2008)

Multi-Scale 3D Convolutional Neural Networks for Lesion Segmentation in Brain MRI

Konstantinos Kamnitsas*, Liang Chen, Christian Ledig, Daniel Rueckert, and Ben Glocker

Biomedical Image Analysis Group, Imperial College London, UK

Abstract. We present our 11-layers deep, double-pathway, 3D Convolutional Neural Network, developed for the segmentation of brain lesions. The developed system segments pathology voxel-wise after processing a corresponding multi-modal 3D patch at multiple scales. We demonstrate that it is possible to train such a deep and wide 3D CNN on a small dataset of 28 cases. Our network yields promising results on the task of segmenting ischemic stroke lesions, accomplishing a mean Dice of 64% (66% after postprocessing) on the ISLES 2015 training dataset, ranking among the top entries. Regardless its size, our network is capable of processing a 3D brain volume in 3 minutes, making it applicable to the automated analysis of larger study cohorts.

1 Introduction

The blockage of an artery during a stroke may disrupt the supply of oxygen and other required substances to brain regions, leading to neuronal death. If the blood flow is restored quickly enough, parts of the affected brain tissue may survive and gradually recover [1]. Automatic detection of the affected but salvageable penumbral tissue can accelerate the decision making and treatment of the patient in the acute clinical setting, increasing the likelihood of a more favourable outcome. Development of robust and accurate segmentation techniques could also facilitate the longitudinal monitoring and analysis of stroke lesions which evolve over time [2] and enable larger-scale studies that can further our understanding of the relations between tissue damage and functional deficits.

Following their success on challenging tasks in the field of Computer Vision [3,4,5], Convolutional Neural Networks (CNNs) have been subsequently applied successfully on a variety of biomedical segmentation problems. Most developed approaches relied on the adaptation of 2D CNNs for processing 3D volumes [6,7,8], with difficulties being reported when training of 3D CNNs was attempted. While these architectures of 2D networks might be successful in some problems, they are suboptimal in their use of available 3D information. The first pure 3D CNN reviewed was presented in [9], where it was employed for brain tumour

* The first author was funded by an Imperial College Phd Scholarship.
Email correspondence to: konstantinos.kamnitsas12@imperial.ac.uk

segmentation. Regardless of its small size, the system demonstrated excellent performance and formed the starting point for our approach.

2 Method

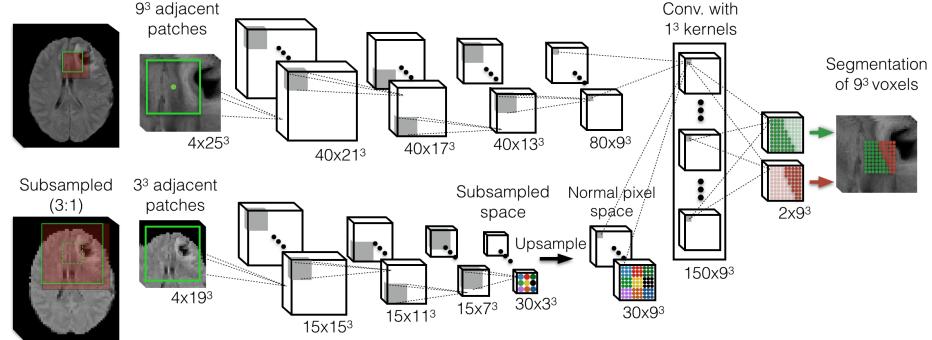


Fig. 1: The architecture of one of our earlier networks. Number and dimensions of the feature maps (FMs) are depicted in the format (*Number* × *Dimension*). Please consult the text for more details on our latest 11-layer network. Multi-modal input is not depicted, in order to avoid cluttering the figure.

2.1 Network Architecture

Although inherently classifiers, CNNs can tackle segmentation tasks by casting them to voxel-wise classification. The network processes a 3D patch around each voxel of an image. It is trained to predict whether the central voxel is pathology or normal brain tissue, depending on the content of the surrounding 3D patch. During training, the parameters of the kernels are optimized using gradient descent, with the target of minimizing the error between the predictions and the true labels.

One of the limitations in the above setting is that the segmentation of each voxel is performed solely by processing the contents of a small patch around it. It is intuitive that greater context is likely to lead to better results. However, a straight-forward increase in the size of the 3D input patch would prohibitively increase the memory requirement and computational burden. Our proposed solution is to perform parallel processing of the image at multiple scales. Our network architecture consists of two parallel convolutional pathways, where both have receptive fields of the same size. The input to the second pathway, however, is a patch extracted from a subsampled-version of an image, thus allowing it to efficiently process greater area around each voxel. This architectural design is presented in figure 1.

Another important feature of our architecture is its full convolutional nature, which allows its efficient application on larger parts of the image. By providing as input segments of an image larger than the receptive field of the final layer's neurons, the network can efficiently process the larger input and provide as output

predictions for multiple neighbouring voxels. Following [9,5], we also exploit this feature during training, constructing our training batches by extracting image segments of size larger than the network’s receptive field.

An earlier version of our system is depicted in figure 1. Our final network resulted from the replacement of each convolutional layer with a kernel of size 5^3 with two layers with 3^3 kernels and the addition of another layer before the final classification. The final 11-layers deep network exhibited significantly more accurate segmentation performance. The network is regularised using Dropout [10] at the 9th and 10th layers with a rate of 50%, on top of L1 (10^{-6}) and L2 (10^{-4}) regularisation. Initial learning rate was set at 0.01 and was gradually reduced during training, along with constant momentum equal to 0.6. ReLu activation functions [11] and batch-normalisation [4] were used for the acceleration of convergence that they provide. The training time required for convergence of the final system is roughly one day using a NVIDIA GTX Titan Black GPU. Segmentation of a 3D scan of a brain with four modalities requires 3 minutes.

2.2 Data Preprocessing, Augmentation and Postprocessing

The modalities of each patient were individually normalised to have zero mean and unary standard deviation, as preliminary experiments showed the networks to behave better on input in this intensity range. In order to regularise the network, we augment the dataset by reflecting the images with respect to the sagittal axes. This processing, along with the subsampling of the images, is performed in parallel with training, when the image segments for the next training iteration are extracted, thus effectively adds no computational time. The output of the network was postprocessed by our version of the CRF presented in [12], extended in order to be able to process 3D biomedical images.

3 Evaluation

The system was evaluated on the training dataset of the ISLES 2015 Challenge. Four modalities (Flair, DWI, T1, T2) were available for 28 cases of patients with ischemic stroke lesions. We performed 5-fold validation for this evaluation. Our system achieved a mean Dice coefficient of 66%, with our network alone achieving 64% and an additional increase of 2% achieved through post-processing using the CRF. In figure 2 we present our results for each case and the current results on the evaluation platform, where we rank among the top entries.

4 Conclusion and Discussion

Our architecture exhibits promising performance, with capabilities for delicate segmentations. Difficulties are observed in the segmentation of lesions of particularly small size. The separation of lesions into different categories, for instance according to their size, and their treatment by separate classifiers could simplify the task for each learner and help alleviating the problem.

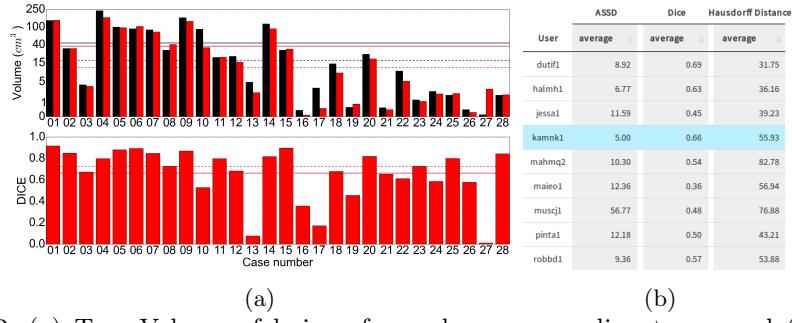


Fig. 2: (a) Top: Volume of lesions for each case according to manual (black) and our system’s segmentation (red). Bottom: DICE coefficient for our system’s segmentation for each case, along with mean (continuous) and median (dashed) values. (b) Online evaluation on training data, with our system ranking 1st and 2nd with respect to mean ASSD and Dice metrics at the time of writing.

References

- Rekik, I., Allinsonnière, S., Carpenter, T.K., Wardlaw, J.M.: Medical image analysis methods in MR/CT-imaged acute-subacute ischemic stroke lesion: Segmentation, prediction and insights into dynamic evolution simulation models. A critical appraisal. *NeuroImage. Clinical* **1**(1) (January 2012) 164–78
- Mitra, J., Bourgeat, P., Fripp, J., Ghose, S., Rose, S., Salvado, O., Connelly, A., Campbell, B., Palmer, S., Sharma, G., Christensen, S., Carey, L.: Lesion segmentation from multimodal MRI using random forest following ischemic stroke. *NeuroImage* **98** (September 2014) 324–35
- Krizhevsky, A., Sutskever, I., Hinton, G.: Imagenet classification with deep convolutional neural networks. *Advances in neural ...* (2012) 1–9
- Szegedy, C., Ioffe, S.: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *arXiv:1502.03167* (February 2015)
- Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. *arXiv preprint arXiv:1411.4038* (2014)
- Ciresan, D., Giusti, A.: Deep neural networks segment neuronal membranes in electron microscopy images. *Advances in neural ...* (2012) 1–9
- Prasoon, A., Petersen, K., Igel, C., Lauze, F.: Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network. *...Image Computing and ...* (09) (2013) 246–253
- Roth, H., Lu, L., Seff, A., Cherry, K.: A new 2.5 D representation for lymph node detection using random sets of deep convolutional neural network observations. In: *... Image Computing and* (2014)
- Urban, G., Bendszus, M., Hamprecht, F., Kleesiek, J.: Multi-modal brain tumor segmentation using deep convolutional neural networks. in proc of BRATS-MICCAI (2014)
- Hinton, G., Srivastava, N.: Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580* (2012) 1–18
- Nair, V., Hinton, G.: Rectified linear units improve restricted boltzmann machines. *Proceedings of the 27th International ...* (3) (2010)
- Krähenbühl, P., Koltun, V.: Efficient inference in fully connected crfs with gaussian edge potentials. *arXiv:1210.5644* (2012) 1–9

Random forests with selected features for stroke lesion segmentation

Oskar Maier^{1,2}, Matthias Wilms¹, and Heinz Handels¹

¹ Institute of Medical Informatics, Universität zu Lübeck

² Graduate School for Computing in Medicine and Life Sciences, Universität zu Lübeck

maier@imi.uni-luebeck.de

Abstract. From clinical practise as well as research methods rises the need for accurate, reproducible and reliable segmentation of ischemic stroke lesions from brain MR scans. This article details a contribution to the Sub-acute Ischemic Stroke Lesion Segmentation (SISS) sub-task of the Ischemic Stroke Lesion Segmentations Challenge (ISLES), organized in conjunction with the MICCAI 2015. The proposed method bases on previous works, which showed the approach to handle various stroke appearances well and to be applicable to potentially flawed data acquired in clinical routine. The method is described in detail and all chosen parameter values are disclosed. Preliminary results on the training data places the approach among the highest ranking contributions.

Keywords: ischemic stroke, lesion segmentation, magnetic resonance imaging, brain MRI, random forest, RDF

1 Introduction

Ischemic stroke is caused by an obstruction of the blood supply to the brain and the subsequent death of brain tissue. Its diagnosis often involves the acquisition of brain magnetic resonance (MR) scans to assess the strokes presence, location, extent, evolution and other factors. An automated method to locate, segment and quantify the lesion area could support the clinicians and render their findings more robust and reproducible. Another demand for automatic stroke lesion segmentation comes from neuroscientists, who employ a research method termed lesion to symptom mapping, which is used to detect correlations between brain areas and cognitive functions by means of negative samples [4]. A number of methods for stroke lesion segmentation have been proposed over the years [2, 8, 12, 3, 13], but none proofed satisfactory to date [11]. The ISLES 2015 challenge offers the first platform for researchers to compare their methods directly and fairly. Our contribution has been previously published [7] and showed good results. It is based on carefully selected features extracted from the MR sequences and used to train a random forest (RF).

2 Method

The challenge’s training data consists of multi-spectral (T1, T2, Flair, DWI) scans of 28 patients displaying sub-acute ischemic stroke. A highly diverse range of stroke types is supplied, ranging from large, single-hemisphere MCA to small, embolic cerebellum strokes. Other complexities are non-stroke white matter lesions, midline shifts, ventricular enhancement and the presence of haemorrhages. For training, the manual segmentations of a single expert rater has been provided. For the testing data, two distinct ground truth will be supplied. More details on the data can be found on www.isles-challenge.org.

2.1 Pre-processing

The image data is provided with a 1 mm isotropic resolution, already co-registered and skull-stripped. Nevertheless, the training cases of the challenge display high intensity differences, a normal occurrence for MRI, where intensity ranges are not standardized. With a learning based intensity standardization method implemented in MedPy [6] and based on [9] we harmonize each sequences intensity profile after a prior bias-correction step with CMTK [5].

2.2 Forest classifier

We employ the RF classifier implemented in [10], which is similar to the propositions made by [1]. The classification of brain lesions in MRI is a complex task with high levels of noise [7], hence a sufficiently large number of trees must be trained.

2.3 Features

The primary distinction criteria for identifying pathological tissue of stroke lesions is the MR intensity in the different sequences. The bulk of our voxel-wise features therefore bases on the intensity values.

intensity First feature is the voxel’s intensity value.

gaussian Due to the often low signal-to-noise ratio in MR scans and intensity inhomogeneities of the tissue types, we furthermore regard each voxel’s value after a smoothing of the volume with a 3D Gaussian kernel at three sizes: $\sigma = 3, 5, 7\text{ mm}$.

hemispheric difference While stroke can affect both hemispheres, it usually does not display symmetric properties. Therefore, we extract the hemispheric difference (in intensities) after a Gaussian smoothing of $\sigma = 1, 3, 5\text{ mm}$ to account for noise. Here, the central line of the sagittal view is taken as sufficiently close approximation of the sagittal midline.

local histogram Another employed feature is the local histogram, as proposed in [7], which provides information about the intensity distribution in a small neighbourhood around each voxel. The neighbourhoods considered were $R = 5^3, 10^3, 15^3 \text{ mm}$, the histogram was fixed to 11 bins.

center distance Finally, we extract the distance to the image center (assumed here to coincide roughly with the brain's center of mass) in mm as final feature. Note that this is not intensity based, but rather discloses each voxel's rough location inside the brain.

All features are extracted from each of the MR sequence, hence in total we obtain 163 values per multi-spectral voxel. Note that all of these features are implemented in MedPy [6].

2.4 Post-processing

After thresholding the a-posteriori class probability maps for a crisp segmentation, all unconnected components with a size smaller than 1000 ml are removed under the assumption that they represent outliers. In all remaining binary components possibly existing holes are closed and a binary dilation of size 1 mm applied to compensate for the methods tendency to under-segment the stroke lesions slightly.

3 Experiments

3.1 Training choices and parameter values

For training our RF, we sample 1,000,000 voxels randomly from all training cases. The ratios between classes in each case are kept intact (i.e. stroke class samples will be highly under-represented). A total of 50 trees are trained for the forest. As split criteria the Gini impurity is employed, a maximum of $\sqrt{163}$ features is considered at each node. No growth restrictions are imposed. The a-posteriori class probabilities produced by the forest are thresholded at a value of 0.4, to counter a slight under-segmentation.

3.2 Preliminary results

Online evaluation is provided with the Dice's coefficient (DC), the average symmetric surface distance (ASSD) and the Hausdorff distance (HD) as quality metrics. Using a leave-one-out evaluation scheme, we have obtained the scores presented in Tab. 1.

Table 1. Mean evaluation results and standard deviation on 28 training cases. See the text for details on the abbreviations employed.

	DC	ASSD	HD
28 cases	0.58 ± 0.29	7.91 ± 13.09	34 ± 29

4 Discussion and conclusion

The favourable placement of our proposed method among the contributions confirms the suitability of our approach for stroke lesion segmentation as has already been observed in [7].

By employing RFs, we have a powerful classifier at our hand that is robust against uninformative features, generalized well and produces good results for a wide range of parameters. Mixing widely used with specially designed features, we can successfully learn to discriminate between ischemic stroke lesion and other, not only healthy, tissue.

On the downside, RFs suffer from the same drawbacks as all other machine learning based methods: The training set must be carefully chosen and types of cases not present in the training data can not be processed.

References

1. Criminisi, A., Shotton, J., Konukoglu, E.: Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning. *Foundations and Trends® in Computer Graphics and Vision* 7(2–3), 81–227 (2012)
2. de Haan, B., Clas, P., Juenger, H., Wilke, M., Karnath, H.O.: Fast semi-automated lesion demarcation in stroke. *NeuroImage: Clinical* pp. 1–1 (2015), <http://www.sciencedirect.com/science/article/pii/S2213158215001199>, available online 17 July 2015
3. Hevia-Montiel, N., Jimenez-Alaniz, J., Medina-Banuelos, V., Yanez-Suarez, O., Rosso, C., Samson, Y., Baillet, S.: Robust nonparametric segmentation of infarct lesion from diffusion-weighted MR images. In: *IEEE EMBS 2007*. pp. 2102–2105 (2007)
4. Krämer, U.M., Solbakk, A.K., Funderud, I., Løvstad, M., Endestad, T., Knight, R.T.: The role of the lateral prefrontal cortex in inhibitory motor control. *Cortex* 49(3), 837–849 (2013)
5. Likar, B., Viergever, M.A., Pernus, F.: Retrospective correction of MR intensity inhomogeneity by information minimization. *Med. Imag., IEEE Trans.* 20(12), 1398–1410 (2001)
6. Maier, O.: MedPy. <https://pypi.python.org/pypi/MedPy>, accessed: 2015-03-29
7. Maier, O., Wilms, M., et al.: Extra tree forests for sub-acute ischemic stroke lesion segmentation in MR sequences. *Journal of Neuroscience Methods* 240(0), 89–100 (2015)
8. Mitra, J., Bourgeat, P., Fripp, J., Ghose, S., Rose, S., Salvado, O., Connelly, A., Campbell, B., Palmer, S., Sharma, G., et al.: Lesion segmentation from multimodal MRI using random forest following ischemic stroke. *NeuroImage* (2014)

9. Nyul, L., Udupa, J., Zhang, X.: New variants of a method of mri scale standardization. *Medical Imaging, IEEE Transactions on* 19(2), 143–150 (Feb 2000)
10. Pedregosa, F., Varoquaux, G., et al.: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12, 2825–2830 (2011)
11. Rekik, I., Allassonniere, S., Carpenter, T.K., Wardlaw, J.M.: Medical image analysis methods in MR/CT-imaged acute-subacute ischemic stroke lesion: Segmentation, prediction and insights into dynamic evolution simulation models. A critical appraisal. *NeuroImage: Clinical* 1(1), 164–178 (2012)
12. Seghier, M.L., Ramlackhansingh, A., Crinion, J., Leff, A.P., Price, C.J.: Lesion identification using unified segmentation-normalisation models and fuzzy clustering. *NeuroImage* 41(4–3), 1253–1266 (Jul 2008)
13. Wilke, M., de Haan, B., Juenger, H., Karnath, H.O.: Manual, semi-automated, and automated delineation of chronic brain lesions: a comparison of methods. *NeuroImage* 56(4), 2038–2046 (Jun 2011)

Ischemic Stroke Lesion Segmentation Using Local Gradient and Texture Features

Syed M. S. Reza, Linmin Pei, and K. M. Iftekharuddin

Vision Lab, Department of Electrical and Computer Engineering
Old Dominion University, Norfolk, VA 23529, USA.
 {sreza002, lxei001, kiftekha}@odu.edu

Abstract. This work proposes fully automatic ischemic stroke lesion segmentation in multimodality brain MRI by extending our prior brain tumor segmentation (BTS) work [1]. The extensions of the BTS method include development of relevant MR image intensity inhomogeneity correction, several new features and feature ranking methods. We characterized brain lesions with multiple features such as piece-wise triangular prism surface area (PTPSA), multi-fractal Brownian motion (mBm), structure tensor based local gradient, regular intensities and intensity differences of MRI modalities. As in BTS, we used classical Random Forest (RF) [2] to classify the brain tissues as lesion or background. The method is evaluated on 28 patients' images having sub-acute ischemic stroke lesions from ISLES2015 SISS challenge dataset [3].

1 Methods

The success of our prior texture based BTS works[1] [4] [5] had driven the motivation of this work. Although our prior BTS work showed excellent performance in tumor segmentation, the methods needed brain lesion specific adaptation such as a new intensity inhomogeneity correction [6] technique for MR images, and introduction of a new local gradient feature information [7] and as well as a minimum redundancy maximum relevance (mRMR) feature selection [8] method. The overall flow diagram of our proposed lesion segmentation method is shown in Figure 1.

The detail description of most of the above methods can be found in our prior BTS publications[1] [4]. We briefly discuss the additional steps as follows:

Local Texture Feature Extraction: In our prior works we extract the local texture (PTPSA, mBm) features after the preprocessing. However, the preprocessing (skull-stripping and slice co-registration) step is skipped in this work since the images are already preprocessed. Note the local texture features are extracted before intensity inhomogeneity correction. As described in [9], the multi-scale wavelets do not require these corrections.

Intensity Inhomogeneity correction: It comprises of two steps. In the first step we perform 10 point histogram matching, where the reference images of four modalities are arbitrarily set from a single patient (in this case, the first patient pat001) data. Next step is normalizing all the intensity values around the

mean intensity value of cerebrospinal fluid (CSF). The method is described in [6]. However, instead of performing a two-class classification (e.g., CSF vs. Rest), we simply threshold the intensity differences among the modalities and obtain the CSF mask. We consider the histogram matched images, the normalized image differences among the modalities and the CSF mask as features for subsequent processing.

Structure Tensor based Local Gradient feature: Eigen value decomposition of the 2D structure tensor matrix [7] is performed to capture the local gradient information. From all four modalities Eigen values (λ_1, λ_2) are used as additional new features which allow a more precise description of the local gradient characteristics.

Feature ranking and selection: Since all the features are not equally important and redundancy among the features degrades the classifier's performance, a mutual information based mRMR[8] feature ranking technique is implemented. From the feature list, we choose 19 top ranked features out of 35 total features. The feature selection method show that intensity features in Flair and DWI MRI modalities as well as mBm, intensity difference, and local gradients features extracted from Flair and DWI modalities show the most discriminative properties respectively. This observation is also confirmed in other relevant works [10].

2 Results and Discussions

We obtain 2D segmented tissues using the predicted pixel labels from RF. These 2D segments are then stacked to generate volume image. Example lesion segments using two slices are shown in Figure 2.

Quantitative evaluation: We evaluate our preliminary lesion segmentation results using 28 training patients' obtained from the ISLES-2015 SISS challenge dataset. The performance efficiency is evaluated by across patient cross-validation, where odd numbered patients are used for training while the even numbered patients' are used for testing, and vice versa. On an average 59% Dice score overlap with 23% standard deviation is obtained from across the patient cross-validation. We notice comparatively better performance for lesions with larger size. Patientwise quantitative Dice overlap are presented in Table 1.

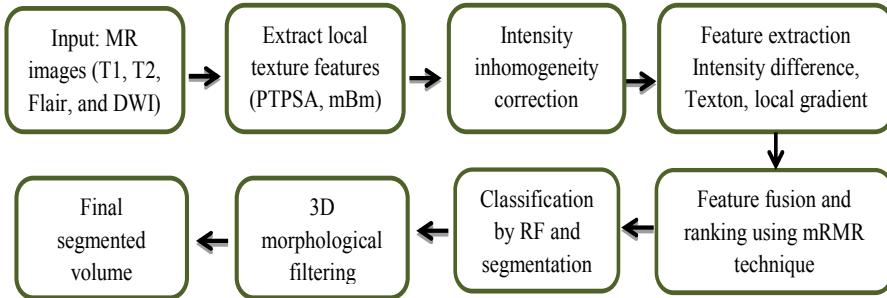


Fig. 1. Generic flow diagram of the proposed method

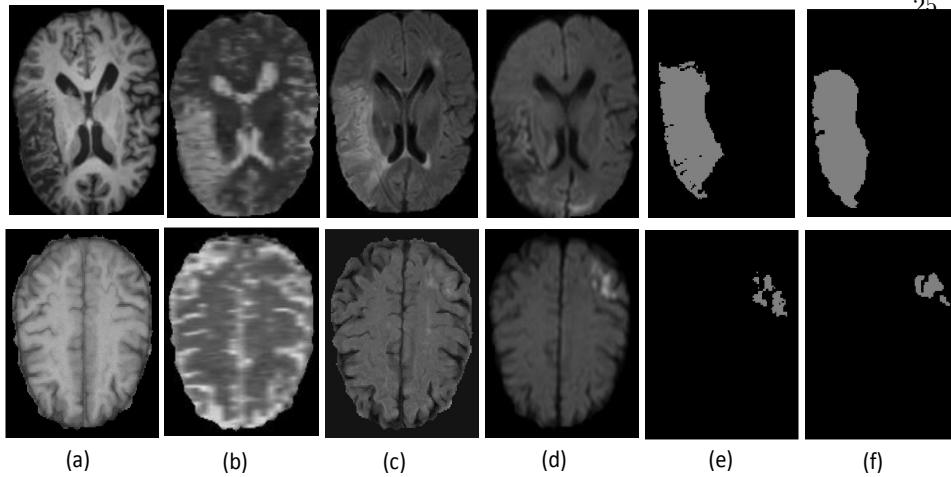


Fig. 2. Segmented lesions with corresponding input and ground-truth images. Each row represents an example set of multimodality MRI slices; Input: (a) T1, (b) T2, (c) Flair (d) DWI (e) Segmented lesion (f) ground-truth.

Pat	P01	P02	P03	P04	P05	P06	P07	P08	P09	P10	P11	P12	P13	P14
DC	0.89	0.75	0.35	0.75	0.81	0.89	0.81	0.77	0.83	0.53	0.60	0.65	0.13	0.80
Pat	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25	P26	P27	P28
DC	0.74	0.47	0.51	0.55	0.26	0.73	0.08	0.55	0.56	0.56	0.62	0.27	0.26	0.79

Table 1. Summary of quantitative Dice(DC) score for 28 patients: Avg. 59%, std. 23%

3 Conclusion

This work proposes an automatic lesion segmentation method and cross validation using ISLES-2015 SISS challenge dataset. Experimental results with 28 clinical patient data confirm the efficacy of our method for sub-acute ischemic stroke lesion segmentation. The training results show comparable performance when compared to other state-of-the art works posted on the VSD website[3]. Note the evaluation results reported here are obtained using the ground-truth provided on the VSD web on our local machines. We notice a considerable number of false positives in our detections that compromise the overall results. Our future works include study of more effective features and sophisticated feature selection techniques. We also plan to study deep neural network based segmentation technique to develop a generalize method for both sub-acute and acute ischemic lesion segmentation. At the time of writing this report, the evaluation process had been in-process in the online evaluation tool posted on the VSD website [3]. The cross-validation results from VSD web will be reported in the next iteration of this submission.

Acknowledgments

This work is partially supported through a grant from NCI/NIH (R15CA115464) and uses brain MR images from the ISLES-2015 SISS challenge dataset.

References

1. S. Reza and K. M. Iftekharuddin, "Multi-class abnormal brain tissue segmentation using texture features," *in Proceedings MICCAI-BRATS*, pp. 38-42, 2013.
2. A. Cutler and L. Breiman, "Random forests-classification description," Technical report, University of California, Berkeley, 2004.
3. <https://www.virtualskeleton.ch/ISLES/Start2015>
4. S. Reza and K. M. Iftekharuddin, "Multi-fractal texture features for brain tumor and edema segmentation," *Proc. SPIE Med. Imag. Conf.*, vol. 9035, 2014.
5. A. Islam, S. Reza, and K. M. Iftekharuddin, "Multi-fractal texture estimation for detection and segmentation of brain tumors," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 11, pp. 3204-15, 2013.
6. J. Kleesiek, A. Biller, G. Urban, U. Kothe, M. Bendszuz, and F. Hamprecht, "ilastik for multi-modal brain tumor segmentation," *in Proceedings MICCAI-BRATS*, pp. 12-17, 2014.
7. <http://www.cs.cmu.edu/~sarsen/structureTensorTutorial/>
8. H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *Pattern Analysis and Machine Intelligence, IEEE Transactions*, vol. 27, no. 8, pp. 1226-1238, 2005.
9. A. Quddus and O. Basir, "Semantic image retrieval in magnetic resonance brain volumes," *IEEE Trans Inf. Technol. Biomed.*, vol. 16, no. 3, pp. 348-355, May 2012.
10. O. Maier, M. Wilms, J. von der Gabeltz, U. M. Krmer, T. F. Mnte, and H. Handels, , "Extra Tree forests for sub-acute ischemic stroke lesion segmentation in MR sequences," *Journal of neuroscience methods*, vol. 240, pp. 89-100, 2015.

ISLES Challenge 2015: A voxel-wise, cascaded classification approach to stroke lesion segmentation

David Robben^{1,3}, Daan Christiaens^{1,3}, Janaki Raman Rangarajan^{1,3},
Jaap Gelderblom^{1,3}, Philip Joris^{1,3}, Frederik Maes^{1,3}, and Paul Suetens^{1,2,3}

¹ KU Leuven, Dept. of Electrical Engineering, ESAT/PSI, Medical Image Computing
² iMinds, Medical IT Department
³ UZ Leuven, Medical Imaging Research Center
 Herestraat 49 – 7000 Leuven, Belgium

Abstract. We propose a supervised method based on cascaded extremely randomised forests for lesion segmentation, and evaluate the pipeline in the MICCAI Ischemic Stroke Lesion Segmentation (ISLES) challenge.

1 Introduction

In ischemic stroke, reduced blood flow to part of the brain results in localised tissue damage and eventual necrosis. Automated localisation and segmentation of the stroke lesion in patients is of great interest to clinicians and researchers alike, enabling them to differentiate potentially salvageable and permanently damaged tissue, identify effective treatments, and follow progression of the ischemic lesion [5]. The MICCAI Ischemic Stroke Lesion Segmentation (ISLES) challenge aims to evaluate and compare state-of-the-art methods, by providing two multi-modal MRI datasets for sub-acute ischemic stroke lesion segmentation (SISS) and for acute stroke outcome and penumbra estimation (SPES).

In this paper, we propose a supervised method based on cascaded extremely randomised forest classifiers for stroke lesion segmentation, and describe a single pipeline to be used for both datasets. After nested cross-validation on the training data, we obtained an average Dice score of 57 % for the SISS data and 82 % for the SPES dataset, which is on par with other contestants.

2 Method

2.1 Preprocessing

At first, the non-parametric images in both datasets were corrected for RF inhomogeneity. We estimated the bias field on the T1w-images using FSL FAST [2], using a 3-tissue model and a bias field smoothing filter of 40 mm full-width half maximum. The elevated smoothing parameter (default is 20 mm) was chosen to improve robustness to the pathology. The estimated bias field was subsequently

applied to correct all T1w- and T2w-images, as well as the Flair and DWI images in the SISS dataset. The ADC images and the perfusion measures in the SPES dataset were not corrected, as these images are already normalised or assumed to be in physical units.

Secondly, cross-subject histogram normalisation was done for each dataset and each modality. To this end, we used a linear intensity rescaling based on two percentile intensities of the histogram. These were heuristically determined based on the histogram profile of a given modality across all subjects of each dataset. For SISS we used 20 % and 99 % for T1, T2, and DWI, and 30 % and 90 % for Flair. For SPES we used 30 % and 90 % for T1, 20 % and 99 % for T2, 20 % and 90 % for DWI (ADC), and 20 % and 50 % for TTP. No intensity normalisation was applied to Tmax, CBF, and CBV.

Additionally, we wish to include spatial features in the classifier as well. Therefore, we registered all subjects T1w-images to the MNI152 template using a 12 degrees of freedom affine transformation and normalised mutual information, as implemented in FSL FLIRT [2]. The resulting transformation matrices are converted to (affine) deformation fields which provide, for each voxel in native space, the corresponding coordinate in MNI space. As such, no image interpolation is needed and the subsequent classifier training can be done in native space.

2.2 Classifier

We decided to use a voxel-wise classification approach for both segmentation tasks. That is, we build a classifier that, given a set of features of a voxel, estimates the probability that this voxel is part of a lesion. To increase computational efficiency and spatial consistency, we use a cascaded approach. First, the to-be-classified voxel is given to a classifier that uses a limited set of features. If this classifier decides with very high probability that the voxel is non-lesion, then this probability is the final answer. Else, the voxel is given to the second classifier that uses a large set of features. Then, the voxels which were not classified as non-lesion with very high probability, are given to a third classifier. This last classifier uses the same features as the second classifier and additionally the earlier computed probabilities of that voxel and its neighbouring voxels.

We use extremely randomised trees [1] as a base classifier. This classifier builds an ensemble of decision trees, but by randomising the selection of cut-point in the decision tree nodes, its training is significantly faster than the training of random forests while achieving comparable accuracy. We use the implementation provided by scikit-learn [4].

2.3 Features

Since the classifier is the same for both challenges, the features are constructed in a similar fashion.

For the SPES sub-challenge, the first cascade uses the T1c intensity. The second cascade uses the intensity of the T1c, T2, TTP, Tmax and DWI images smoothed with a sigma of 0 – 6 mm. It also has for TTP and Tmax the 0.5, 0.75

and 0.9 percentiles and for DWI the 0.1 and 0.25 percentiles of its neighbourhood for varying radii (4 – 12 mm). Finally, it has the MNI-coordinates. The third cascade uses the same features as the latter and additionally it has the earlier estimated probabilities smoothed with a sigma of 0 – 8 mm and the 0.5, 0.75 and 0.9 percentiles of its neighbourhood for varying radii (4 – 8 mm).

For the SISS sub-challenge, the first cascade uses the T1 intensity. The second cascade uses the intensity of the T1, T2, Flair and DWI images smoothed with a sigma of 0 – 8 mm. It also has for Flair and DWI the 0.5, 0.75 and 0.9 percentiles of its neighbourhood for varying radii (4 – 8 mm). Finally, it uses the MNI-coordinates. The third cascade uses the same features as the latter and additionally it has the earlier estimated probabilities smoothed with a sigma of 0 – 8 mm and the 0.5, 0.75 and 0.9 percentiles of its neighbourhood for varying radii (4 – 8 mm).

2.4 Probability threshold

After the voxel-wise classification, we have for every voxel a probability of belonging to a lesion. However, the challenge requires a binary segmentation and hence we need to threshold the resulting probabilities. Instead of using a fixed threshold for all images, we use a novel technique to find the optimal threshold.

A voxel \mathbf{x} is part of the lesion with probability $P(\mathbf{x})$, as estimated by the classifier. Given that the probability estimates are correct, the Dice score obtained with threshold P_t will be:

$$Dice(P_t) = \frac{2 |GT \cap \text{segmentation}|}{|GT| \cup |\text{segmentation}|} = \frac{2 \sum_{\mathbf{x}} I[P(\mathbf{x}) > P_t]P(\mathbf{x})}{\sum_{\mathbf{x}} P(\mathbf{x}) + \sum_{\mathbf{x}} I[P(\mathbf{x}) > P_t]}, \quad (1)$$

with I the indicator function. We exhaustively search for the optimal threshold.

3 Results

The performance of the proposed segmentation method is evaluated in the online submission system of the challenge, and relies on the average symmetric surface distance (ASSD), the Dice overlap coefficient, and the Hausdorff distance. Additionally, precision and recall are reported to discriminate between over- and under-segmentation respectively. The results of cross-validation on the training data are reported in Table 1. Example segmentations of median and maximum overlap are shown in Fig. 1.

4 Discussion and Conclusion

We presented a supervised method for stroke lesion segmentation, based on cascaded extremely randomised forests. The cascaded approach showed strong improvement over a single voxel-wise classifier, and allows to take neighbourhood information into account while still limiting the number of features and the

Table 1. Segmentation results on the training data, reported as average symmetric surface distance (ASSD), Dice coefficient, Hausdorff distance, precision, and recall.

	ASSD (mm)		Dice		Hausdorff (mm)		Precision		Recall	
	avg	std	avg	std	avg	std	avg	std	avg	std
SISS	9.36	13.85	0.57	0.28	53.88	34.58	0.58	0.33	0.68	0.21
SPES	2.03	1.35	0.82	0.07	44.29	27.59	0.81	0.14	0.85	0.07



Fig. 1. Comparison between the ground-truth labels (*green*) and the predicted segmentation (*red*), shown for selected examples with median and maximum Dice coefficient. SISS Flair dataset on the left; SPES DWI (ADC) dataset on the right.

required computation time. The method works well on both datasets, although the inter-subject variability is rather large in the SISS data. Given that this is the case for other contestants as well, it would be interesting to have access to the inter-observer variability of the ground-truth segmentations.

Future work may improve upon this method by revising the histogram normalisation. A threshold-based classifier such as ours is sensitive to the intensity scaling, and the current linear approach is sub-optimal. More advanced, non-linear approaches such as Meier et al. [3] could help in this regard.

References

1. Geurts, P., Ernst, D., Wehenkel, L.: Extremely randomized trees. *Machine Learning* 63(1), 3–42 (Mar 2006)
2. Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., Smith, S.M.: FSL. *NeuroImage* 62(2), 782–790 (aug 2012)
3. Meier, D.S., Guttmann, C.R.: Time-series analysis of MRI intensity patterns in multiple sclerosis. *NeuroImage* 20(2), 1193–1209 (oct 2003)
4. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12, 2825–2830 (2011)
5. Rekik, I., Allouanière, S., Carpenter, T.K., Wardlaw, J.M.: Medical image analysis methods in MR/CT-imaged acute-subacute ischemic stroke lesion: Segmentation, prediction and insights into dynamic evolution simulation models. a critical appraisal. *NeuroImage: Clinical* 1(1), 164–178 (2012)

ISLES (SISS) challenge 2015: segmentation of stroke lesions using spatial normalization, Random Forest classification and contextual clustering

Hanna-Leena Halme^{1,2}, Antti Korvenoja¹, and Eero Salli¹

¹ HUS Medical Imaging Center, Radiology, University of Helsinki and Helsinki University Hospital, P.O. Box 340, FI-00029 HUS, Finland

² Department of Neuroscience and Biomedical Engineering NBE, Aalto University School of Science, P.O. Box 12200, FI-00076 Aalto, Finland

1 Statistical template

In order to enable voxel-wise comparisons across subjects, all T1, T2, Flair and DWI volumes, as well as the volumes containing manual segmentations, were warped to a common template. The template was generated by the Advanced Normalization Tools (ANTs) software version 2.1.0rc3 [1] using *buildtemplateparallel* script, greedy SyN transformation model, cross-correlation similarity metric, $30 \times 90 \times 20$ interations and the T1 images of the training dataset. All T1 images were deformed to this template with ANTs software using Affine transformation model for rigid registration and SyN transformation model for warping. The detailed parameters are given in Appendix. This defined a transformation which was applied to T2, Flair, DWI and manual segmentation volumes. All MR images were masked with a volume containing only in-brain voxels to minimize the effect of background to the automatic segmentation of ischemic stroke lesion volumes.

Images in the common template space representing mean and standard deviation (std) of voxel intensities over subjects were calculated voxel-by-voxel, separately for T1, T2, Flair and DWI images. We call these images *statistical* templates from now on. Note that the lesion voxels were not included in the calculation of average and std. Because most subjects in the training dataset had lesions on the left hemisphere, there were fewer voxels contributing to the mean and std on the left; as a result, the left hemisphere appeared slightly distorted on the template images. In order to compensate for this left-right bias, the mean and std images were additionally averaged over left and right hemispheres. Furthermore, the images were smoothed with a 3D Gaussian kernel (FWHM 3 mm) to decrease the effect of registration inaccuracies.

2 Random Forest classification algorithm

The initial segmentation was predicted with an ensemble learning method. A set of features was derived from the training data and fed to a Random Forest

[2] classification algorithm implemented in Scikit-learn version 0.16.dev [3]. The Random Forest algorithm combines classification results from a number of decision trees. Several trees are constructed and fitted to the data during training phase, using a random subset of features to train each tree. The final classification is the mode of the classes obtained from all individual trees. Random Forest classification greatly reduces overfitting, which is a common problem for simple decision tree classifiers [2]. In this study, the performance of the classifier was tested with leave-one-out cross-validation, in which one subject from the training dataset was used for testing and the rest for training the classifier, and the procedure was repeated for all subjects. The classifier returned both the binary classifications and probabilities (in range 0-1) that a voxel belonged to the lesion area.

3 Classifier training

16 features were extracted from the MR images for subsequent classification. Z-score normalized voxel intensities of T1, T2, Flair and DWI images constituted features 1-4. Features 5-8 represented the Z-score deviations from global average images, calculated separately for each sequence by subtracting the global mean and dividing with the global std. The purpose of these features was to find regions showing large deviations from the normal brain, which likely indicates presence of a lesion. Features 9-12 were obtained by smoothing the original images with a 3D Gaussian kernel (FWHM 3 mm), thus including information from the local neighborhood of each voxel. Smoothing was expected to improve classification since it may reduce the effect of registration inaccuracies. Features 13-16 represented local asymmetry, obtained by comparing voxel intensities on one hemisphere to the corresponding voxel intensities on the other hemisphere. The motivation for calculating local asymmetry was the fact that lesions rarely occur symmetrically on both hemispheres. The asymmetry measure was calculated simply by subtracting the original smoothed image from the left-right-mirrored smoothed image.

In order to decrease computational time and avoid classifier overfitting, we only collected the aforementioned features from a randomly selected subset of voxels. The maximum number of lesion voxels sampled from each subject was set to 300, and in cases where the lesion was smaller than that, all lesion voxels were sampled. The ratio of lesion and non-lesion voxels per subject was kept constant, such that twice as many voxels were sampled from non-lesion area as from lesion area; thus, the maximum number of non-lesion voxels per subject was 600.

For the Random Forest classifier, the training set was resampled to train a total of 300 decision trees. 4 features were used to obtain the best split at each individual tree. The quality of each split was described by Gini impurity [2]. The trees were grown unlimitedly, i.e. until each leaf contained only samples of a single class. All parameters and default values used by Scikit-learn's RandomForest classifier are listed in Appendix.

4 Contextual clustering

The segmentation results obtained with the Random Forest method were further improved with contextual clustering (CC). The clustering method was based on a Markov random fields (MRF) prior and iterated conditional modes (ICM, [4]) algorithm, which were previously used for analysis of functional magnetic resonance imaging (fMRI) data [5]. The basic assumption in contextual clustering is that neighboring voxels tend to belong to the same class. Furthermore, it is assumed that the intensity distribution of background voxels (in present case, non-lesion voxels) is standard normal, but the distribution of lesion voxels is unknown.

The CC algorithm modified for this study consisted of the following steps:

1. Fit a gamma distribution to all nonzero voxels of the probability map given by the random forest classifier. Fitting is done using MATLABs (R2015a) function *fitdist* with default parameters.
2. Transform the probability map values to standard normal distribution by calculating the inverse normal distribution function (MATLAB function *norminv*) from the cumulative distribution function (MATLAB function *cdf*) of gamma distribution. This gives image N .
3. Define the parameter T for contextual clustering [5] using the fitted gamma distribution:

$$T = -\text{norminv}(\text{cdf}(\text{gamma}, D))$$
, where *gamma* is the fitted gamma distribution, and D some threshold. In this study we used empirically chosen $D = 0.6$, which gave reasonable results with training data.
4. Run the CC algorithm [5] using the image N , neighborhood weight coefficient $\beta = T^2/6$ and a threshold $D = 0.6$. The voxels will be reclassified to 0 or 1, corresponding to non-lesion and lesion, respectively.
5. Repeat steps 1-4 with only the voxels classified as non-lesion in the first run of CC.

Finally, all automatically segmented images were transformed back to each subject's native space using inverse transformation and nearest neighbor interpolation. After transformation it was possible to compare the automatic segmentations with the manual lesion segmentations. The classification accuracy was evaluated with the script provided at ISLES web page (<http://www.isles-challenge.org>), including measurements for Dice coefficient, average symmetric surface distance (ASSD), Hausdorff distance, precision and recall.

5 Testing phase

The test data will be spatially normalized to the common template using the parameters listed in Appendix, but without the lesion images (-x option). After that, the trained Random Forest classifier and contextual clustering will be applied to the data.

References

1. Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage*, 54(3), 2033-2044.
2. Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
3. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, 12, 2825-2830.
4. Besag, J. (1986). On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 259-302.
5. Salli, E., Aronen, H. J., Savolainen, S., Korvenoja, A., & Visa, A. (2001). Contextual clustering for analysis of functional MRI data. *Medical Imaging, IEEE Transactions on*, 20(5), 403-414.

Appendix

Generation of templates

The common template was done in two phases. First, the initial template was formed:

```
buildtemplateparallel.sh -d 3 -m 1x0x0 -n 0 -r 1 -t GR -s CC -o [initial template image] -c 0 -j 1 [T1 images]
```

After that, the final template was built using the initial template:

```
buildtemplateparallel.sh -d 3 -m 30x90x20 -n 0 -r 0 -t GR -s CC -o [template image] -z [initial template image] -c 0 [T1 images]
```

Warping of T1 images to common template was done using *antsRegistration* tool and the following parameters:

```
-metric MI[template image, T1 image,1,32] -transform affine[0.25] -convergence 10000x10000x10000x10000 -shrink factors 5x4x3x2x1 -smoothing-sigmas 4x3x2x1x0 -metric CC[template image, T1 image,1,5] -transform SyN[0.25,3.0,0.0] -convergence 50x35x15 -shrink factors 3x2x1 -smoothing-sigmas 2x1x0 -use-histogram-matching 1 -x [lesion image]
```

Parameters for Random Forest classifier

Scikit-learn's function `sklearn.ensemble.RandomForestClassifier` was used with the following parameters:

```
n_estimators=300, criterion='gini', max_depth=None, min_samples_split=2, min_samples_leaf=1, min_weight_fraction_leaf=0.0, max_features=4, max_leaf_nodes=None, bootstrap=True, oob_score=False, n_jobs=1, random_state=None, verbose=0, warm_start=False, class_weight=None
```

Stroke lesion segmentation of 3D brain MRI using multiple random forests and 3D registration

Ching-Wei Wang*, Jia-Hong Lee

Graduate Institute of Biomedical Engineering, National Taiwan University of Science and Technology, Taipei City, Taiwan
Email: cweiwang@mail.ntust.edu.tw

Abstract. Stroke is a common cause of sudden death and disability worldwide. In clinical practice, brain magnetic resonance (MR) scans are used to assess the stroke lesion presence. In this work, we have built a fully automatic stroke lesion segmentation system using 3D brain magnetic resonance (MR) data. The system contains a 3D registration framework and a 3D multi-random forest model trained from the data provided by the Ischemic Stroke Lesion Segmentation (ISLES) challenge of the 18th International Conference on Medical Image Computing and Computer Assisted Intervention. The preliminary test results show that the presented system is capable to detect stroke lesion from 3D brain MRI data.

1 Introduction

Stroke is a common cause of sudden death and disability worldwide. In clinical practice, brain magnetic resonance (MR) scans are used to assess the stroke lesion presence. A fully automatic random forest based stroke lesion 3D segmentation approach is built. A 3D segmentation framework with backward registration and forward registration is developed for processing the 3D brain data. A machine learning model is trained using the training data provided by the Ischemic Stroke Lesion Segmentation (ISLES) challenge of the 18th International Conference on Medical Image Computing and Computer Assisted Intervention. The preliminary test results show that the presented system is capable to detect stroke lesion from 3D brain MRI data. The outline of this paper is as follows. Section 2 presents the proposed method with the preliminary test results, and section 3 concludes the paper.

2 Method with Preliminary Test Results

A fully automatic machine learning based stroke lesion 3D segmentation approach is built. The data is trained using random forest [2] with the parameters as presented in the Table 1. A large multiple random forest model is developed

to generate potential candidates, and for every five stacks in Z direction, a random forest is trained. In data preparation, the regions of interest (ROI) are extracted to produce training data with two classes using the contour tracing algorithm [1], and 275 image features categorized in 24 types are extracted for training, as illustrated in figure 1.

Obtaining the potential candidates from the machine learning model mentioned above, a 3D registration framework with backward and forward searching is applied to produce optimal 3D predictions, and in the experiments, the 80th stack is selected as the referenced frame for the 3D registration framework. The system flowchart is presented in the figure 2. Figure 3 shows the inputs and system outputs of a 3D brain MRI sample.

Table 1. The parameters used for the multiple random forest model

for each random forest

The maximum depth of the trees	50
The number of trees to be generated	50
The number of features to used	275
The random number seed to be used	1

for the entire 3D segmentation system

The number of stacks to be used to train one random forest	5
The total number of random forests to be built	$N/5$

N: The maximum number of stacks in Z direction.

No.	Features	Attributes	13	Maximum	5
1	Original	1	14	Median	5
2	Hue	1	15	Anisotropic diffusion	10
3	Saturation	1	16	Bilateral	4
4	Brightness	1	17	Lipschitz	5
5	Gaussian blur	5	18	Kuwahara	3
6	Sobel filter	6	19	Gabor	44
7	Hessian	48	20	Derivatives	20
8	Difference of gaussians	10	21	Laplacian	5
9	Membrane projections	6	22	Structure	20
10	Variance	5	23	Entropy	20
11	Mean	5	24	Neighbors	40
12	Minimum	5	Total Feature		

Fig. 1. Features used for training

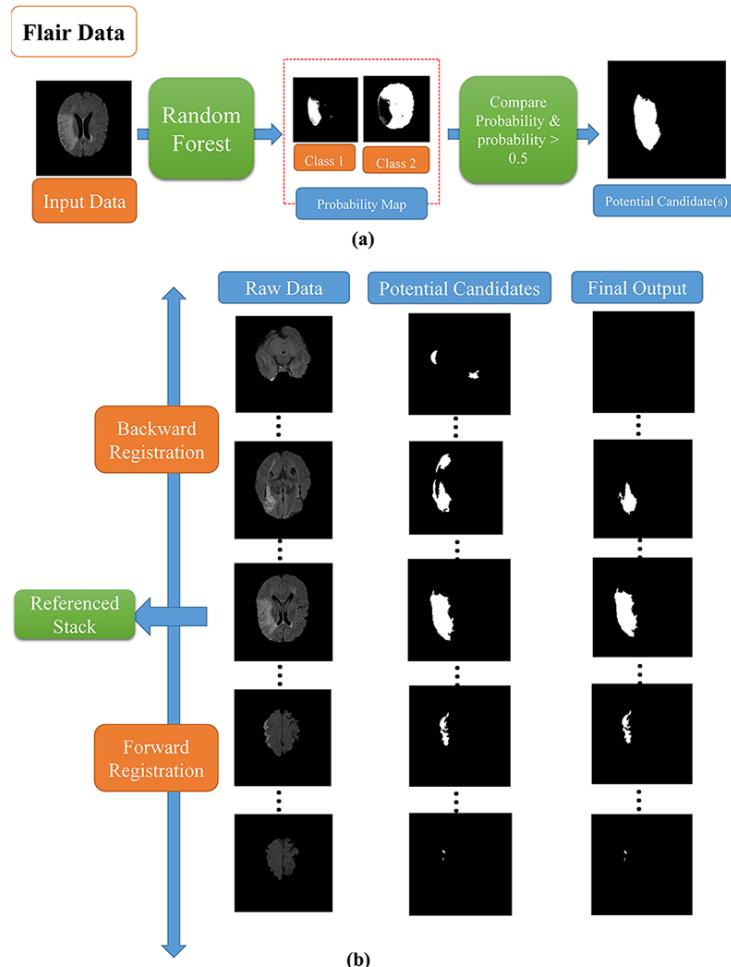


Fig. 2. System flowchart

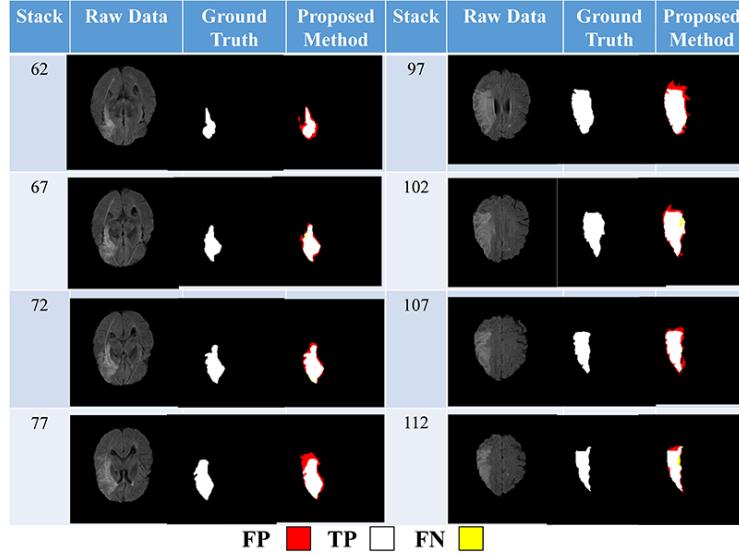


Fig. 3. Inputs and System outputs of a 3D brain MRI sample

3 Conclusion

We have presented a fully automatic stroke lesion segmentation system using 3D brain magnetic resonance (MR) data. The preliminary test results show that the presented system is capable to detect stroke lesion from 3D brain MRI data.

Acknowledgments

Authors would like to thank the Ministry of Science and Technology of Taiwan under Grant No. MOST104-2221-E-011-085 for the financial support.

References

1. Burger, W., Burge, Mark J. Digital Image Processing - An Algorithmic Introduction Using Java. Springer (2008).
2. Breiman, L. Random Forests. *Machine Learning* **45**(1), 5-32 (2001).

Input Data Adaptive Learning (IDAL) for sub-acute Ischemic Stroke Lesion Segmentation

Michael Goetz¹, Christian Weber¹, and Klaus Maier-Hein¹

Juniour Group Medical Image Computing, German Cancer Reserach Center (DKFZ)
Heidelberg, Germany

Abstract. In machine learning larger databases are usually associated with higher classification accuracy due to better generalization. In some medical applications with highly variable expressions of pathologies exactly this generalization may lead to non-optimal classifiers. Here approaches incorporating these varying manifestations of a disease in a flexible way to improve the classification are needed. This paper therefore presents a method to learn from a large training base consisting only of sparsely annotated data, adaptively selecting optimal training samples for given input data. This way heterogeneous databases are supported two-fold. First, by being able to deal with sparsely annotated data allows a quick inclusion of new data sets. Secondly, by adapting the classifiers according to the input data the heterogeneity of the database is further exploited.

Keywords: Adaptive Learning, Lesson Segmentation, Machine Learning, Random Forest

1 Motivation

Learning from large datasets becomes more and more important in computer vision. This is also true for machine learning in the context of medical image computing, but poses special challenges. The data, that are used here are usually highly individual – not only because of the variety of imaging modalities and imaging configurations but also because pathological changes have a great variety of appearances. This leads to new challenges for machine learning in medical imaging: a) how do we create large training sets covering the diversity of a pathology and b) how to incorporate such heterogeneity in a beneficial way. We propose a new algorithm that faces these challenges. Building on prior work we use an algorithm that allows learning from sparse and unambiguous regions (SURs) and enhance it with a new method were classifiers are only trained on similar data. We call this approach ‘Input Data Adapted Learning’ (IDAL).

2 Overview

Instead of training a single classifier that is used to predict all unseen images we propose to adaptively train a new classifier for every new image. This allows to

use only few, but similar images during training. While such an approach makes a classifier less general, we expect that the so-trained classifier is better suited to deal with the afore mentioned heterogeneity.

We realized this approach with a three-staged algorithm (Fig. 1). During the first stage, that is performed offline like traditional classifier training, we train a similarity classifier (SC) which can group similar images based on a similarity measure.

The offline trained SC is used in the second stage – the online training – to find images that are similar to the new, unlabeled images. Based on this individual, input-dependent subset of training images, a new voxel-based classifier (VC) is trained. For this, we used the approach presented in [5] which allows to train a voxel-based classifier (VC) from sparsely and unambiguously labeled regions (SURs). This VC is then used in the last stage to label each voxel of the new image, leading to the prediction mask.

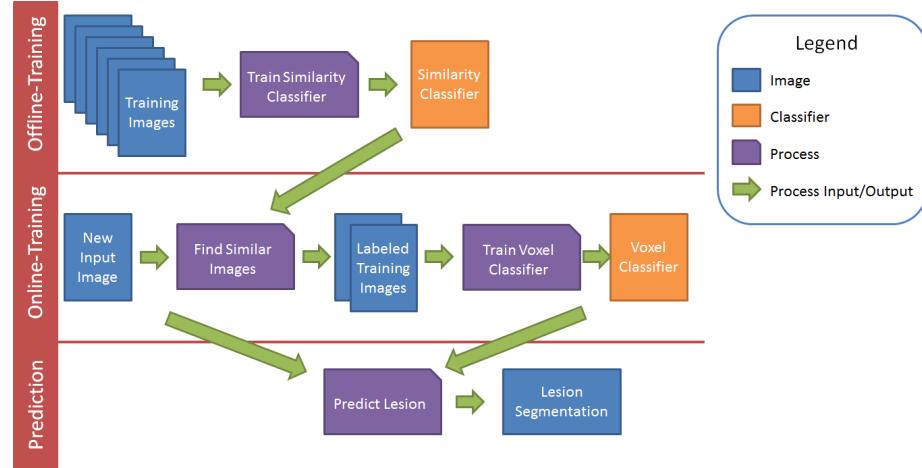


Fig. 1. Overview of the three stages used for our algorithm.

3 Preprocessing

A simple preprocessing is applied before the images where used for training or prediction. Our brainmask includes all voxels for which neither T1 nor T2 are zero.

We normalize all MR-images to a CSF-mode of 0 and an overall brain matter mode of 1. We found that using the mode instead of mean provides a more robust normalization since the mode is less affected by the size of the lesions. We obtained the CSF-area by training a simple classifier using only the pure voxel intensities.

4 Similarity Classifier (SC)

The main idea of our work is to find similar images which are then used to train a voxel classifier. Therefore, the similarity between two images is defined by the ability to successfully use them to train a classifier. Accordingly, we define the similarity $\rho(I_0, I_1)$ of two image I_0 and I_1 as the Dice score that a voxel classifier trained with I_0 scores if the mask for I_1 is predicted.

While it is possible to directly calculate the similarity of two images with known voxel labels, it needs to be estimated for new images with unknown voxel labels. We chose Neighbourhood approximating Forests (NAF) for this task [1]. NAF are trained to find the most similar images based on a high-dimensional representation by training trees that group the training data in a way that maximizes the similarity within each leaf node. For the prediction, the new images are then passed below each tree and for every sample the number of leaf nodes that had been reached and contains this sample type is returned. A high number of samples therefore indicates a high similarity between a training sample and the sample that is used to predict.

To train the NAF and use it as SC we first calculated the similarity of all training images according to the previously given definition of $\rho(\cdot, \cdot)$. We then built a feature vector for every patient based on the normalized T₁, T₂, DWI, and FLAIR image by calculating the first order statistics for the whole brain (Intensity minimum, maximum, range, mean, variance, sum, median, std. deviation, mean absolute deviation, root means square, uniformity, entropy, energy, kurtosis, skewness and the number of voxels). Although these are all image-derived values, the proposed approach also allows the use of additional information like patient age, diagnosis, etc., which are not included in the challenge data.

We trained the NAF with 100 trees, a minimum of two samples at each leaf, 30 random tests for best split at each node during the training and a maximum tree depth of 12. After predicting a new patient (Online Training stage, see Fig. 1) we chose the three highest ranked training images to train the new VC.

5 Voxel Classifier (VC)

The estimation of voxel labels is done by a voxel-wise classification. For this task extremely randomized trees (ExtraTrees)-based classifiers are used [2]. Previous work showed that ExtraTrees usually perform slightly better than canonical Random Forests [4] and were already successfully applied in lesion segmentation [6]. Voxel features were derived from the normalized MR-images. We used the intensity and the differences between each of the modalities. Additionally, the Gaussian, Difference of Gaussian, Laplacian of Gaussian (3 directions), and Hessian of Gaussian were calculated with Gaussian sigma of 1 mm, 3 mm, and 5 mm, leading to 82 features per voxel.

Instead of training on the given full labels we created sparsely and unambiguous annotated regions (SURs) which allows the fast labeling of new training data [3, 5]. The necessary labeling for the SISS 2015 challenge dataset was done

in less than $2^{1/2}$ hours. The sampling error introduced by SUR-based labeling was corrected by using DALSA-learning [5].

For this, every training sample x is weighted with an correction weight w which is selected to ensure that the probability for this sample in the training data equals the probability P for this sample in the complete image, i.e.

$$w(x) = \frac{P_{\text{Complete Image}}(x)}{P_{\text{SURs}}(x)} \quad (1)$$

We estimate the unknown $w(x)$ by training a parameter-less logistic regression that differentiates between voxels that are labeled by SURs and voxels that are within the brain mask. By using the probabilistic output of this method, w can be estimated [5].

Each ExtraTrees classifier was trained with 50 trees and the Gini purity as optimization measurement. The maximum tree depth was not limited. During each training (during similarity calculation and final VC training) the best class weights, and minimum samples at leaf nodes were independently estimated by using cross validation.

Acknowledgment: This work was carried out with the support of the German Research Foundation (DFG) within project I04, SFB/TRR 125 Cognition-Guided Surgery

References

- [1] E. Konukoglu, B. Glocker, D. Zikic, A. Criminisi: 'Neighbourhood approximation using randomized forests' Medical Image Analysis, Vol. 17, Issue 7, Pages 790-804, 2013
- [2] P. Geurts, D. Ernst, L. Wehenkel: 'Extremely randomized trees', Machine Learning, Vol. 63, Pages 3-42, 2006
- [3] M. Goetz, C. Weber, B. Stieltjes, K.H. Maier-Hein: 'Learning from Small Amounts of Labeled Data in a Brain Tumor Classification Task' NIPS 2014 Workshop on Transfer and Multi-task learning: Theory Meets Practice
- [4] M. Goetz, C. Weber, J. Bloecher, B. Stieltjes, H.P. Meinzer, K.H. Maier-Hein: 'Extremely randomized trees based brain tumor segmentation' Proceedings of BRATS Challenge-MICCAI, 2014
- [5] M. Goetz, C. Weber, F. Binczyk, J. Polanska, R. Tarnawski, B. Bobek-Billewicz, U. Koethe, J. Klesiek, B. Stieltjes, K.H. Maier-Hein: 'DALSA: Domain Adaptation for Supervised Learning from Sparsely Annotated MR Images', IEEE Transactions on Medical Imaging, vol.PP, 2015 doi: 10.1109/TMI.2015.2463078
- [6] O. Maier, M. Wilms, J. Gablentz, U.M. Krämer, T.F. Münte, H. Handels, 'Extra Tree forests for sub-acute ischemic stroke lesion segmentation in MR sequences' Journal of Neuroscience Methods, Volume 240, Pages 89-100, 2015

Distributed Deep Learning Framework for Large-Scale 3D Medical Image Segmentation

T. Klein^{1,2}, N.K. Batmanghelich², and William M. Wells III^{1,2}

¹ Brigham and Women’s Hospital, Harvard Medical School, Boston, USA

² Massachusetts Institute of Technology, Cambridge, USA

Abstract. In this paper, we propose a novel distributed framework that makes the excellent performance of the Deep Learning (DL) available for large-scale medical applications. More specifically, we propose a distributed optimization approach based on Alternating Direction Method of Multipliers (ADMM) enabling us to optimize the cost function of the Convolutional Neural Network (CNN) across multiple machines. This framework makes the computational power and memory of multiple machines available to our algorithm, which facilitates computationally and memory intensive tasks such as medical image segmentation. We apply our method on the Ischemic Stroke Lesion Segmentation (SPES) task.

Keywords: Distributed Computing, Deep Learning, Convolutional Neural Network

1 Method

The general idea of the paper is as follows, we divide the image into set of volumetric patches denoted by x_i . For each patch, we train a DL discriminative model to predict the labels of central voxels, denoted by y_i . We use Convolutional Neural Networks (CNN) [3] as the discriminative model. CNN consists of stacked up layers, combining a convolution followed by a downsampling step referred to as pooling. As the convolution is applied by shifting a filter element over the input domain, translational invariance is achieved. This also restricts the complexity of the model by reducing the number of free parameters. Finally, we collect the class labels of all central patches and use a voting scheme to get the final segmentation results.

1.1 Softmax Regression

In the following, we first revisit the preliminaries of multi-class classification, *i.e.* softmax regression, in the context of the deep learning. The input to the prediction model is the patch data of the D image channels. The output is the class labels (segmentation labels) of L voxels in the center of the patch. Therefore, it is multi-task classification problem. We denote the input patch data with $x \in \mathbb{R}^{W \times W \times W \times D}$, where W is the width of the patch and D is number

of image channels and the corresponding labels with $y \in [1,..,K]^L$ where $[1..K]$ is set of all integers between 1 to K , each denoting one tissue type. We use y_i^ℓ to denote element ℓ of the patch i . For notational brevity and future reference, we use $\mathcal{X} = \{x_i\}_{i=1}^N$, $\mathcal{Y} = \{y_i\}_{i=1}^N$. We use $(\mathcal{X}_m, \mathcal{Y}_m)$ to represent a batch of training data.

Given input patch image x , the proposed method estimates the probability of the label map of the central voxels, $P_\theta(y_i^\ell|x_i)$ where $\theta \in \mathbb{R}^n$ are the parameters of our model. We use a CNN to model $P_\theta(y_i^\ell|x_i)$. Given y_i^ℓ takes discrete values, with abuse of notation, we can represent $P_\theta(y_i^\ell|x_i)$ as a vector:

$$P_\theta(y_i^\ell|x_i) = \begin{bmatrix} P_\theta(y_i^\ell = 1|x_i) \\ P_\theta(y_i^\ell = 2|x_i) \\ \vdots \\ P_\theta(y_i^\ell = K|x_i) \end{bmatrix} = \frac{1}{\sum_{j=1}^K \exp[f_j(x_i; \theta)]} \begin{bmatrix} \exp[f_1(x_i; \theta)] \\ \exp[f_2(x_i; \theta)] \\ \vdots \\ \exp[f_K(x_i; \theta)] \end{bmatrix} \quad (1)$$

For $P_\theta(y_i^\ell|x_i)$, we propose the use of a 3D deep convolutional neural network [3]. Given the CNN, the parameters θ consist of filter maps as well as weights (and biases) for fully connected layers, representing the CNN. Finally, we learn the optimal θ by solving an optimization problem, where the objective function is the cross entropy between $P_\theta(y_i|x_i)$ and given class label y_i , defined as

$$J(\theta; \mathcal{X}, \mathcal{Y}) = - \left[\sum_{i=1}^N \sum_{k=1}^K \mathbf{1}\{y_i = e_k\} \log \frac{\exp[f_k(\theta, x_i)]}{\sum_{j=1}^K \exp[f_j(\theta, x_i)]} \right]. \quad (2)$$

Here $\mathbf{1}\{\cdot\}$ denotes the indicator function and e_k is the base vector where the k 'th element is 1.

For solving θ , one has to resort to an iterative optimization algorithm such as gradient descent. Specifically, one has to use back-propagation of the neural network. We opted for a modified version of gradient descent, AdaGrad [2], which allows to automatically determine optimal learning rates for the gradient descent scheme.

1.2 Distributed Optimization

Finding optimal θ in Eq. 2 entails solving very large-scale optimization problems that are computational and memory intensive. Due to memory and computational resources limitations, this amounts to a problem that is not suitable for sequential processing rather requiring a high degree of parallelism. One option is to use stochastic gradient descent (SGD). However, it is slow and cannot be operated efficiently in a distributed fashion, e.g. in a cluster environment. In order to scale the optimization for the large-scale problem presented in this paper, we adopted the alternating direction method of multipliers (ADMM) [1]. This framework has the advantageous properties of readily providing decomposability of the minimization, while providing comparably favorable convergence properties. We assume to have a cluster of $m \in [1..M]$ computational nodes. Every

node (worker) processes a subset of data, \mathcal{X}_m , and updates its local parameters, θ_m . Workers communicate their local parameters with the master node, which holds its own parameter θ . At convergence, the workers and the master should reach *consensus* in term of parameters, which can be written as the following optimization problem:

$$\theta = \arg \max_{\theta, \theta_1, \dots, \theta_M} \sum_{m=1}^M \left(\underbrace{J(\theta_m; \mathcal{X}_m, \mathcal{Y}_m) + (\theta_i - \theta)^T u_m + \frac{\rho}{2} \|\theta_i - \theta\|_2^2}_{L(\theta_m; \theta, u_m, \mathcal{X}_m, \mathcal{Y}_m)} \right),$$

subject to: $\theta_1 = \dots = \theta_M = \theta$

where we refer to θ and u_m as master and dual variables, respectively. $L(\theta_m; \theta, u_m, \mathcal{X}_m, \mathcal{Y}_m)$ is the objective function processed by node m that optimizes local parameters θ_m given the batch data, $(\mathcal{X}_m, \mathcal{Y}_m)$, and the parameters of the master node θ , and the dual parameter u_m .

At the iteration t , the worker m updates its parameters as,

$$\theta_m^{t+1} = \arg \max_{\theta_m} L(\theta_m; \theta, u_m, \mathcal{X}_m, \mathcal{Y}_m) \quad (3)$$

and the master and the dual variables are updated according to

$$\theta^{t+1} = \frac{1}{M} \sum_{m=1}^M \left(\theta_m^{t+1} + \frac{1}{\rho} u_m^t \right) \quad u_m^{t+1} = u_m^t + \rho (\theta_m^{t+1} - \theta^{t+1}). \quad (4)$$

The variable $\rho \in \mathbb{R}$ is referred to the ADMM penalty parameter and determines the speed of convergence. Following [1], a dynamic scheme was chosen to adjust ρ .

2 Parameters

For the 3D patch of the following dimensions are set: $41 \times 41 \times 41$ with the central voxels number $L = 3$. In terms of CNN, a 9 layer network was chosen - see Tab. 1 for parameters. In terms of ADMM parameters, the following are set: $\mu = 10$, $\tau^{inc} = \tau^{dec} = 1.01055$. For learning a batch size of 50 was chosen. The networks was trained for 5000 iterations.

CNN	
1. Convolution, Max-Pooling	Filter: $12 \times 12 \times 12$, Pooling Size: 2, Stride: 2
2. Convolution, Max-Pooling	Filter: $8 \times 8 \times 8$, Pooling Size: 2, Stride: 1
3. Convolution, Max-Pooling	Filter: $5 \times 5 \times 5$, Pooling Size: 2, Stride: 2
4. Convolution, Max-Pooling	Filter: $3 \times 3 \times 3$, Pooling Size: 2, Stride: 1
5.	Rectified Linear Unit (ReLU)
6. Inner Product	Neurons: 500
7.	Rectified Linear Unit (ReLU)
8. Inner Product	Neurons: $L \times K$
9.	Softmax Regression Layer

Table 1. Parameters of the convolutional deep neural network.

References

1. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* 3(1), 1–122 (Jan 2011), <http://dx.doi.org/10.1561/2200000016>
2. Duchi, J., Hazan, E., Singer, Y.: Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.* 12, 2121–2159 (Jul 2011), <http://dl.acm.org/citation.cfm?id=1953048.2021068>
3. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. In: *Proceedings of the IEEE*. vol. 86, pp. 2278–2324 (1998)

Automatic Ischemic Stroke Lesion Segmentation in Multi-Spectral MRI images using Random Forests Classifier

Qaiser Mahmood^{1,2} and A. Basit²

¹ Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden

² Pakistan Institute of Nuclear Science and Technology, Islamabad, Pakistan

qaiserm@chalmers.se, abdulbasit1975@gmail.com

Abstract. This paper presents an automated segmentation framework for ischemic stroke lesion segmentation in multi-spectral MRI images. The framework is based on a random forests (RF), which is an ensemble learning technique that generates several classifiers and combines their results in order to make decisions. In RF, we employ several meaningful features such as intensities, entropy, gradient etc. to classify the voxels in multi-spectral MRI images. The segmentation framework is validated on MICCAI 2015 ISLES challenge training data sets. The performance of the framework is evaluated relative to the manual segmentation (ground truth). The experimental results demonstrate the robustness of the segmentation framework, and that it achieves reasonable segmentation accuracy for segmenting the sub-acute ischemic stroke lesion in multi-spectral MRI images.

Keywords: Segmentation, automatic, MRI, ischemic stroke lesion, random forests

1 Introduction

Multi-spectral magnetic resonance imaging (MRI) [1] can be used for detecting the ischemic stroke lesion and can provide quantitative assessment of lesion area. It can be established as an essential paraclinical tool for diagnosing stroke as well as for monitoring the efficacy of experimental treatments.

For a quantitative analysis of stroke lesion in MRI images, expert manual segmentation is still a common approach and has been employed to compute the size, shape and volume of the stroke lesions. However, it is time-consuming, tedious, and labor-intensive task. Moreover, manual segmentation is prone to intra-and inter-observer variabilities [2].

Therefore, the development of fully automated and accurate stroke lesion segmentation method has become an active research field. In literature [2–4], several automated segmentation methods have been proposed for stroke lesion segmentation over the years. However, the automated stroke lesion segmentation is still a challenging task because of the gradual changes of stroke lesion appearance in multi-spectral MRI images.

Herein, we present a fully automated framework for sub-acute ischemic stroke lesion segmentation in multi-spectral MRI images. The framework is based on a supervised classification method called random forests. The main contribution in the framework is employing a set of meaningful features and the choice of steps for pre-processing the MRI images and post-processing of segmented data.

2 Method

The schematic procedure of the segmentation framework is shown in Fig.1. The framework takes the multi-spectral MRI brain images as input and it includes two-step pre-processing: (1) Correction of bias field using the N3 bias field correction algorithm [5] and (2) normalization of intensity values of each MRI modality to the interval [0 1], done by applying the linear histogram stretching. For each voxel of multi-spectral MRI images, the following set of meaningful features is extracted.

1. MRI scans intensities: These features comprise the intensity in the 4 MRI scans (T1, T2, DWI, and FLAIR) provided by the data sets and the difference between each two scans. The total number of these features was 16.
2. MRI scans smooth intensities: A Gaussian filter with size $7 \times 7 \times 7$ was employed to each MRI scan in order to extract the smooth intensities. The total number of these features was 4.
3. MRI scans median intensities: A median filter with size $5 \times 5 \times 5$ was applied to each MRI scan to obtain the median intensities. The total number of these features was 4.
4. The gradient and magnitude of the gradient: A gradient in the x, y and z direction and their magnitude was computed in order to get the information about the lines and edges in each MRI scan. The total number of these features was 16.
5. Local entropy: The entropy for each voxel in the MRI scans was computed using the neighborhoods size $9 \times 9 \times 9$. The total number of these features was 4.

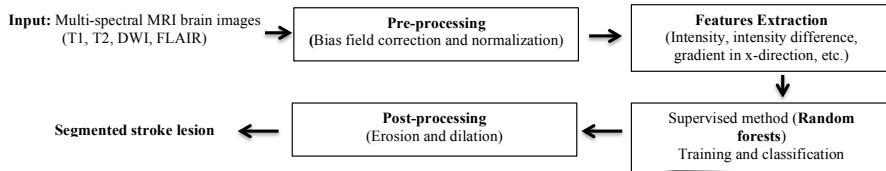


Fig. 1. Schematic procedure of the segmentation framework.

All features, mentioned above, were normalized to zero mean and unit deviation. These features are then employed to train the RF [6, 7] classifier and

classifying the sub-acute ischemic stroke lesion. In RF the training is performed using labeled data sets provided from the ground truths by building multiple decision trees, wherein every node except the leaves is a decision node that contains a feature and its corresponding threshold. Every leaf node contains a probabilistic class distribution (histogram of class labels for the voxels that have reached that node). Moreover, in RF the building of multiple decision trees is based on a random selection of a subset of features (called bootstrap aggregating), which makes RF robust to overfitting. The classification is done by traversing voxels over the trees starting from the root of each tree to a leaf node. The voxels are split at a given node based on the classification of the feature/threshold at that node. The average probabilistic decision of the class distribution from all trees is considered the final probabilistic class distribution (voxel label in this scenario). The two important parameters that affect the efficiency of RF are a number of trees and depth of each tree. In our work, we set the RF parameters: number of trees =150 and depth of each tree=50. For training, a total of 999,000 data samples (37,000 samples per training data) were used to train the RF classifier. These samples were obtained by down sampling the majority class (non ischemic stroke) data in each training data set in order to make their frequencies closer to the minority class (ischemic stroke) data. The sampling was done randomly. Finally, the post-processing is performed using the dilation followed by an erosion operation by employing the 2D 5×5 square structuring elements in order to remove the small objects classified as stroke lesion.

3 Results

The evaluation is performed on MICCAI 2015 ISLES challenge training data sets using leave-one-out cross validation. The data sets comprise 28 sub-acute ischemic stroke lesion cases. The evaluation is done using the online evaluation system provided by the challenge organizers. Table 1 presents the average quantitative results of our segmentation framework in terms of average symmetric surface distance (ASSD), Dice, Hausdorff distance, precision and recall respectively. An example of the segmentation result for sub-acute ischemic stroke lesion for the training data set “01” is shown in Fig.2.

4 Conclusions

In this paper, we present an automated framework based on the RF classifier for segmenting the sub-acute ischemic stroke lesion using multi-spectral MRI images. We employ a set of meaningful features to train the RF and classify the ischemic stroke lesion. The experimental results show the efficacy of the segmentation framework and that it can segment the sub-acute ischemic stroke lesion with reasonable accuracy. For future work, we will explore more set of features in order to improve the accuracy of our segmentation framework. The total execution time of our segmentation framework is about 25 to 30 minutes

for segmenting the stroke lesion for each training data set using the MATLAB on a MacBook Pro with an Intel processor (i5, 2.5 GHz) and 4 GB RAM.

Table 1. Average quantitative results of the segmentation framework in terms of ASSD, Dice, Hausdorff distance, precision and recall.

ASSD (mm)	Dice	Hausdorff Distance (mm)	Precision	Recall
10.30 ± 11.11	0.54 ± 0.26	82.78 ± 23.95	0.67 ± 0.33	0.50 ± 0.25

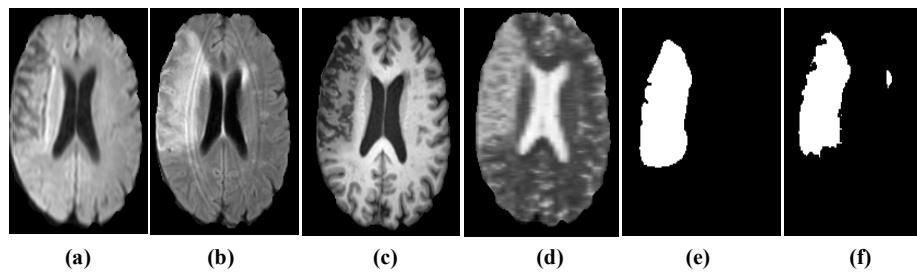


Fig. 2. Example of segmentation result for axial slice number 88 for the training data “01”: (a) DWI (b) Flair (c) T1 (d) T2 (e) ground truth and (f) automatic segmentation.

References

1. Ball, J. B. Jr, Pensak, M. L.: Fundamentals of Magnetic Resonance Imaging. *Am J Otol.*, 8, 81–85 (1987)
2. Oskar, M., Matthias, W., Janina von der, G., Ulrike, M. K., Thomas F. M., Heinz H.: Extra Tree forests for sub-acute ischemic stroke lesion segmentation in MR sequences, *Journal of Neuroscience Methods*, 89–100 (2014)
3. Rekik, I., Allassonniere, S., Carpenter, T.K., Wardlaw, J.M.: Medical image analysis methods in MR/CT-imaged acute-subacute ischemic stroke lesion: Segmentation, prediction and insights into dynamic evolution simulation models. A critical appraisal. *NeuroImage: Clinical*, 1, 164–78 (2012)
4. Mitra, J., Bourgeat, P., Fripp, J., Ghose, S., et al.: Lesion segmentation from multi-modal MRI using random forests following ischemic stroke. *NeuroImage*, 98, 324–335 (2014)
5. Sled, J.G., Zijdenbos, A.P., Evans, A.C.: A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging*, 17, 87-97 (1998)
6. Breiman, L.: Random forests. *Machine Learning*, 45, 5-32 (2001)
7. Criminisi, A., Shotton, J: Decision forests for Computer Vision and Medical Image Analysis, *Advances in Computer Vision and Pattern Recognition*, Springer-Verlag (2013)

ISLES Challenge 2015: Automated Model-Based Segmentation of Ischemic Stroke in MR Images

Tom Haeck^{1,2}, Frederik Maes^{1,2}, and Paul Suetens^{1,2}

¹ KU Leuven, Dept. of Electrical Engineering, ESAT/PSI
² UZ Leuven, Medical Imaging Research Center

Abstract. We present a novel fully-automated generative ischemic stroke lesion segmentation method that can be applied to individual patient images without need for a training data set. An Expectation Maximization-approach is used for estimating intensity models for both normal and pathological tissue. The segmentation is represented by a level-set that is iteratively updated to label voxels as either normal or pathological, based on which intensity model explains the voxels' intensity the best. A convex level-set formulation is adopted, that eliminates the need for manual initialization of the the level-set. The performance of the method for segmenting the ischemic stroke is summarized by an average Dice score of 0.78 and 0.51 for the SPES and SISS 2015 training set respectively.

1 Introduction

The MICCAI Ischemic Stroke Lesion Segmentation (ISLES) challenge comprises the automatic segmentation of ischemic stroke lesions acquired in the sub-acute stroke development stage (SISS) and automatic segmentation of acute ischemic stroke lesions for stroke outcome prediction (SPES).

Discriminative segmentation methods require a set of manually annotated training images from which the appearance of the brain structures of interest is implicitly learned by the algorithm. Generative models on the other hand do not require a set of annotated training images. Explicit prior knowledge of anatomy or intensity appearance is directly incorporated into the algorithm [1]. In clinical practice the availability of annotated training data may be limited or non-existent, such that a generative method that does not rely on training data may be preferred. We present a novel fully-automated generative ischemic stroke segmentation method that only makes use of a probabilistic atlas of white matter (WM), grey matter (GM) and cerebrospinal fluid (CSF) and for which no manual initialization is needed. The probabilistic prior guides the global search for voxel outliers that cannot be explained by the normal tissue model. The lesion boundary is represented as a level-set that spatially regularizes the segmentation.

2 Method

Classification is based on an Expectation Maximization (EM)-estimation of normal and pathological intensity models. An evolving level-set determines which of both intensity models applies to what regions in the image (Fig. 1).

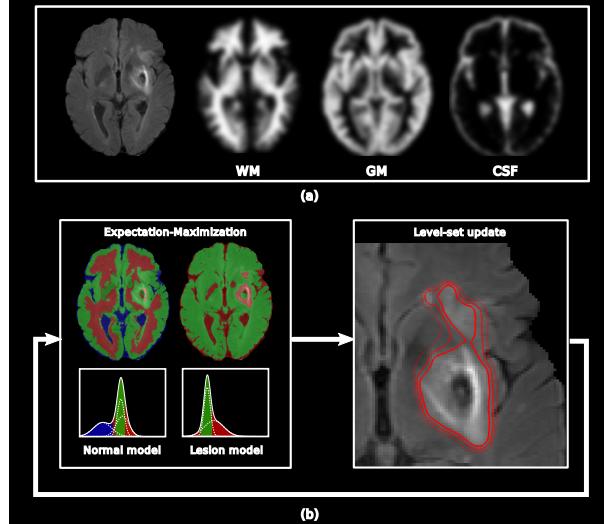


Fig. 1. (a) Spatial priors are non-rigidly registered to the patient image. (b) A full EM-estimation of the normal and pathological intensity models is done, after which a level-set is updated. This process is repeated until convergence.

Prior Registration Spatial priors of WM, GM and CSF are non-rigidly registered to the patient image. Although registration of a healthy atlas to a patient image is still an active field of research, this problem is ignored for now and standard non-rigid registration methods are used. The prior information is relaxed by smoothing the spatial priors with a Gaussian kernel.

Intensity models and the Expectation-Maximization algorithm Normal and pathological tissue intensities are modeled separately. Let G_{Σ_j} be a zero-mean multivariate Gaussian with covariance matrix Σ_j , then normal and pathological tissue are both modeled by a Gaussian mixture model

$$p(\mathbf{y}_i|\theta) = \sum_j^K G_{\Sigma_j}(\mathbf{y}_i - \mu_j)p(\Gamma_i = j), \quad (1)$$

with $\mathbf{y}_i = (y_{i1}, \dots, y_{iN})$ the intensity of voxel i and $\Gamma_i = \{j|j = 1 \dots K\}$ the tissue class. The intensity model parameters $\theta = \{(\mu_j, \Sigma_j)|j \in 1 \dots K\}$ are iteratively updated using an EM-approach [1]. For normal tissue, $K = 3$ and $p(\Gamma = j) = \pi_j$ are the spatial priors for WM, GM and CSF.

Convex level-set formulation The image I is subdivided into regions labeled Ω_{in} (pathological tissue) and Ω_{out} (normal tissue) for which the intensities are modeled by the probability distributions described in the previous paragraph

[2]. The regions are separated by a boundary $\partial\Omega$ that is implicitly represented by a level-set function. The boundary and intensity model parameters are found by minimizing the energy functional

$$\operatorname{argmin}_{\theta_{in}, \theta_{out}, \partial\Omega} \lambda_1 \int_{\Omega_{in}} -\log p_{in}(I|\Omega_{in}, \theta_{in}) d\mathbf{x} + \lambda_2 \int_{\Omega_{out}} -\log p_{out}(I|\Omega_{out}, \theta_{out}) d\mathbf{x} + \kappa L(\partial\Omega), \quad (2)$$

where $L(\cdot)$ is the length. The first two terms penalize the negative loglikelihood of the image I evaluated in respectively the pathological and normal region. The third term penalizes the length of the boundary. Parameters λ_1 , λ_2 and κ determine the relative importance of the energy terms. For each iteration to update the level-set, a full EM-estimation of the parameters θ_{in} and θ_{out} is done.

The energy functional is non-convex and the gradient flow finds a solution that depends on a manual initialization of the level-set. This initialization typically has significant impact on the segmentation result. In this work, this problem is overcome by using a convex level-set formulation that performs a global search over the image and makes a manual initialization superfluous. A global minimum is guaranteed by replacing the gradient flow by another gradient flow with the same steady-state solution and by restricting the level-set to lie in a finite interval [3]. The problem is thus reformulated as an L_1 -minimization problem that is solved by the Split Bregman-numerical scheme [3]. It is important to note that, by using spatial priors of WM, GM and CSF, the global optimum coincides with the clinically meaningful notion of normal and pathological regions.

3 Experiments and Results

The SPES and SISS training data are already skull-stripped and registered intra-patient. No further pre-processing is done. Prior registration is based on the T1-weighted MNI-Colin27 atlas (2008) that is registered to the patient volume with a cross-correlation similarity measure (radius 4 voxels) by the Advanced Normalization Tools (ANTs) toolbox [4]. The spatial priors are relaxed by a Gaussian kernel with $\sigma = 3$ voxels. For segmentation of the SPES data, we use the T2-weighted and TTP-weighted MR images and for SISS the diffusion weighted and FLAIR-weighted MR images. For SPES, the modalities are used in a completely multivariate way, i.e. with bivariate Gaussian models. For SISS, the modalities are segmented separately and a voxel is only labeled as lesion if it is a lesion in both modalities. The number of Gaussians for modeling the lesion intensities is set to 1. The energy functional hyperparameters are $\lambda_1 = \lambda_2 = 1e1$ and $\kappa = 1e1$. Performance of the algorithm for both SPES and SISS is evaluated by means of the ASSD, Dice overlap coefficient, Hausdorff distance and precision and recall (Table 1). The median Dice scores for the SPES and SISS training sets are 0.79 and 0.60 respectively (Fig. 2).

Table 1. Performance of the presented method on the SPES and SISS training set

	ASSD		Dice		Hausdorff		Precision		Recall	
	avg	std	avg	std	avg	std	avg	std	avg	std
SPES	3.51	2.13	0.78	0.08	46.31	25.17	0.78	0.11	0.80	0.12
SISS	14.43	25.88	0.53	0.26	69.67	30.77	0.62	0.31	0.56	0.29

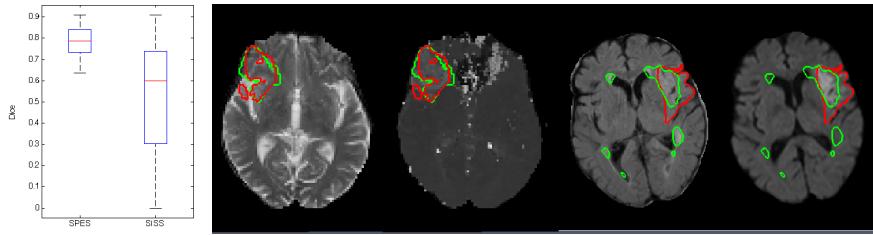


Fig. 2. **Left:** Boxplots for the SPES and SISS Dice scores. **Right:** T2- and TTP-weighted MR example image from SPES and FLAIR- and diffusion weighted MR example image from SISS with ground truth segmentations (red) and the resulting segmentations (green) for a typical segmentation (Dice score 0.79 and 0.50 for SPES and SISS).

4 Discussion and Conclusion

In plenty of clinical settings only a handful of patient images needs to be processed without the availability of an annotated training set. Generative methods have therefore an enormous practical value. We have presented a generative method for segmenting the ischemic stroke lesion in the SPES and SISS training set. The method is abundantly flexible to detect any intensity abnormality, and therefore also suitable to detect other lesions like tumor or MS.

References

1. K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated Model-Based Tissue Classification of MR Images of the Brain. *IEEE Transactions on Medical Imaging*, 18:897–908, 1999.
2. M. Rousson and R. Deriche. A variational framework for active and adaptative segmentation of vector valued images. In *Proceedings of the Workshop on Motion and Video Computing*, MOTION ’02. IEEE Computer Society, 2002.
3. T. Goldstein, X. Bresson, and S. Osher. Geometric applications of the split bregman method: Segmentation and surface reconstruction. *Journal of Scientific Computing*, 45(1-3):272–293, 2010.
4. B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, and J. C. Gee. A reproducible evaluation of ants similarity metric performance in brain image registration. *Neuroimage*, 54(3):2033–2044, Feb 2011.

A Convolutional Neural Network Approach to Brain Lesion Segmentation

Francis Dutil¹, Mohammad Havaei¹, Chris Pal², Hugo Larochelle¹, and Pierre-Marc Jodoin¹

¹ Université de Sherbrooke, Sherbrooke, QC, Canada

² École Polytechnique de Montréal, Canada

Abstract. Deep Neural Networks (DNNs) are often successful at solving problems for which useful high-level features are not obvious to design. This document presents how DNNs can be used for automatically segment brain lesions for the MICCAI Ischemic Stroke Lesion Segmentation (ISLES) challenge. We experimented several DNN architectures leveraging the recent advances in the field such as convolutional layers, max pooling, maxout units, dropout regularization, and various training strategies.

We present the results of our best performing network on the SISS and SPES training datasets. The results are obtained from the evaluation tool available on the Virtual Skeleton database. As of today, empirical results show that our approach is the most accurate one.

1 Introduction

Brain lesions are abnormalities in the tissue of an organism, usually caused by disease or trauma. The delineation and quantification of brain lesions is critical to establishing patient prognosis, and for understanding the development of pathology over time. Typically, this is performed manually by a medical expert through investigation of several Magnetic Resonance Imaging (MRI) modalities. To alleviate the tedious, time consuming manual delineation, computerised methods can be very useful.

Recently, Convolutional Neural Networks (CNNs) have proven particularly successful in many computer vision applications. For instance, the so-called AlexNet architecture [7] was the first to establish CNNs as the *de facto* state-of-the-art methodology for object recognition in natural images. The main appeal of convolutional networks comes with their end-to-end training nature [6]. That is, their ability of learning low, medium, and high-level features (which involve linear and non-linear operators) as well as the classification function. The potential of CNNs for segmentation in medical imaging however is not well understood, and has only been the subject of preliminary investigations (see workshop publications [3, 10, 9]). In other work [5], alternative to the standard CNN framework have also been explored for more general image segmentation tasks, with the argument that CNN training is overly computationally intensive.

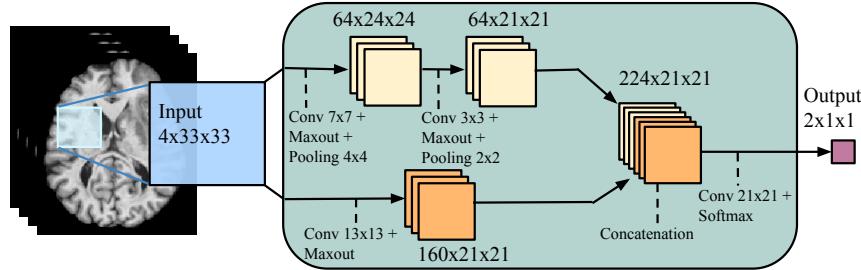


Fig. 1: Our CNN model. The input patch goes through two convolutional networks each comprising of a local and a global path. The feature maps in the local and global paths are shown in yellow and orange respectively.

In this document, we propose a successful, efficient, and automatic CNN architecture for brain lesion segmentation. Note that the proposed CNN is similar to the one used by our team on the 2015 MICCAI brain tumor segmentation (BRATS) challenge. We report results on SPES and SISS training datasets and confirm that our method is leading on both datasets.

2 Convolutional Neural Network Architecture

We approach the problem of brain lesion segmentation by solving it slice by slice, from the axial view. Let \mathbf{X} be one such 2D image (slice), where each pixel is associated with multiple channels, one for each image modality. We treat the problem of segmentation as one of taking any patch it contains and predicting the label of the pixel at its center. The problem is thus converted into an image classification problem.

In the context of this work, we tested a large number of CNN architectures and the most effective one is shown in Figure 1. As can be seen, our method uses a two-pathway architecture in which each pathway is responsible for learning about either the local details or the larger context of tissue appearances (e.g. whether or not it is close to salient regions of the brain like the skull or the CSF). The pathways are joined by concatenating their feature maps immediately before the output layer.

Finally, a prediction of the class label is made by stacking a final output layer, which is fully convolutional to the last convolutional hidden layer. The number of feature maps in the last layer matches the number of class labels and prediction is made with the *softmax* non-linearity.

2.1 Efficient Two-Phase, Patch-Wise Training

By interpreting the output of our CNN as a model for the distribution over segmentation labels, a natural training criteria is to maximize the probability

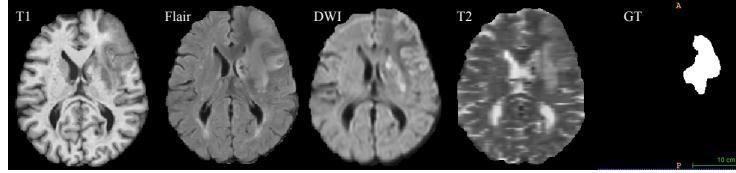


Fig. 2: SISS MRI modalities. The images show the MRI modalities used as input channels to the CNN model for SISS dataset.

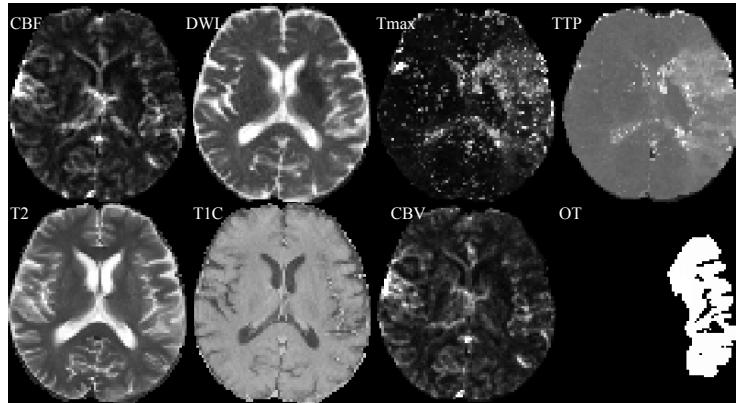


Fig. 3: SPES MRI modalities. The images show the MRI modalities used as input channels to the CNN model for SPES dataset.

of all labels in our training set or, equivalently, to minimize the negative log-probability $-\log p(\mathbf{Y}|\mathbf{X}) = \sum_{ij} -\log p(Y_{ij}|\mathbf{X})$ for each labeled brain. To do this, we follow a stochastic gradient descent approach by repeatedly selecting labels Y_{ij} at a random subset of positions (i.e. patches) within each brain, computing the average negative log-probabilities for this mini-batch of positions and performing a gradient descent step on the CNNs parameters.

Care must be taken however to ensure efficient training. Indeed, since the distribution of labels is very imbalanced (e.g. more than 98% of the brain is healthy), selecting patches from the true distribution would cause the model to be overwhelmed by healthy patches. It is well known that neural network training algorithms such as stochastic gradient descent perform poorly in cases of strong class imbalances. To avoid these issues, we initially construct our patches dataset such that all labels are equiprobable. This is what we call the *first* training phase. Then, in a *second* phase, we account for the unbalanced nature of the data and re-train only the output layer (i.e. keeping the kernels of all other layers fixed) with a more representative distribution over the labels. Using this approach, we were able to fully train CNNs in less than 6 hours.

3 Implementation details

Our implementation is based on the Pylearn2 which supports GPU's and can greatly accelerate the execution of deep learning algorithms [4].

To test the ability of CNNs to learn useful features from scratch, we employed only minimal preprocessing. We applied N4ITK bias correction [2] and clamp the 1% highest intensities to the maximum grayscale value of the 99% remaining pixels as done in [8]. These choices were found to work best in our experiments. The data was normalized within each input channel by subtracting the channel mean and dividing by its standard deviation.

The hyper-parameters of the model (kernel and pooling size for each layer) are illustrated in Figure 1. The learning rate α is decreased by a factor $\gamma = 10^{-1}$ at every epoch. The initial learning rate was set to $\alpha = 0.1$. A post processing method based on connected components was also implemented to remove flat blobs which might appear in the predictions due to bright corners of the brains close to the skull.

4 Experiments and Results

We conducted our experiments on the SISS and SPES datasets. The SISS dataset contains 28 brains with four modalities namely: T1, Flair, Diffusion Weighted Image (DWI) and T1. SPES dataset contains 30 brains with 7 modalities namely: CBF, CBV, DWI, T1c, T2, Tmax and TTP. Both datasets provide pixel-accurate level ground truth of the abnormal areas. Although the ground truth for SPES dataset contains three classes (healthy, stroke, and edema), according to the challenge website the evaluation is done by merging the two unhealthy classes. Figure 3 and Figure 2 show examples from the SPES and SISS datasets.

The [virtualskeleton](#) webpage provides a quantitative evaluation of the model [1]. It reports the dice, precision and recall coefficient, as well as the average symmetric surface distance (ASSD) and the Hausdorff distance (HD).

Table 1 and Table 2 show the results obtained from the [virtualskeleton](#) webpage on both SISS and SPES datasets and how we compare with other methods applied on these datasets. As one can see, our method (dutifl) is well in front the other methods. Our approach provides the best score on 4 of the 5 metrics on the SISS dataset, and on 3 of the 5 metrics on the SPES dataset. Also, each time our method is not rank first, it is ranked second. Let us underline the fact that since the Hausdorff distances of our method (31.75 and 23.28) is significantly lower than the ones obtained by the other methods, we may conclude that our approach is less prone at detection outliers in the brain.

Figure 4 shows visual segmentation maps produced by our model on both datasets. The first two rows show segmentation results on SPES dataset and the two bottom rows show segmentation results on SISS dataset. It takes on average 25 seconds to produce a segmentation result. The larger receptive field in the two-pathway method allows the model to have more contextual information of the lesion. At the same time, the smaller receptive field make model flexible enough to recognize the fine details of the lesion as opposed to making very smooth segmentation as in the one path method. By allowing for a second phase training and learning from the true class distribution, the model corrects most of the misclassifications produced in the first phase.

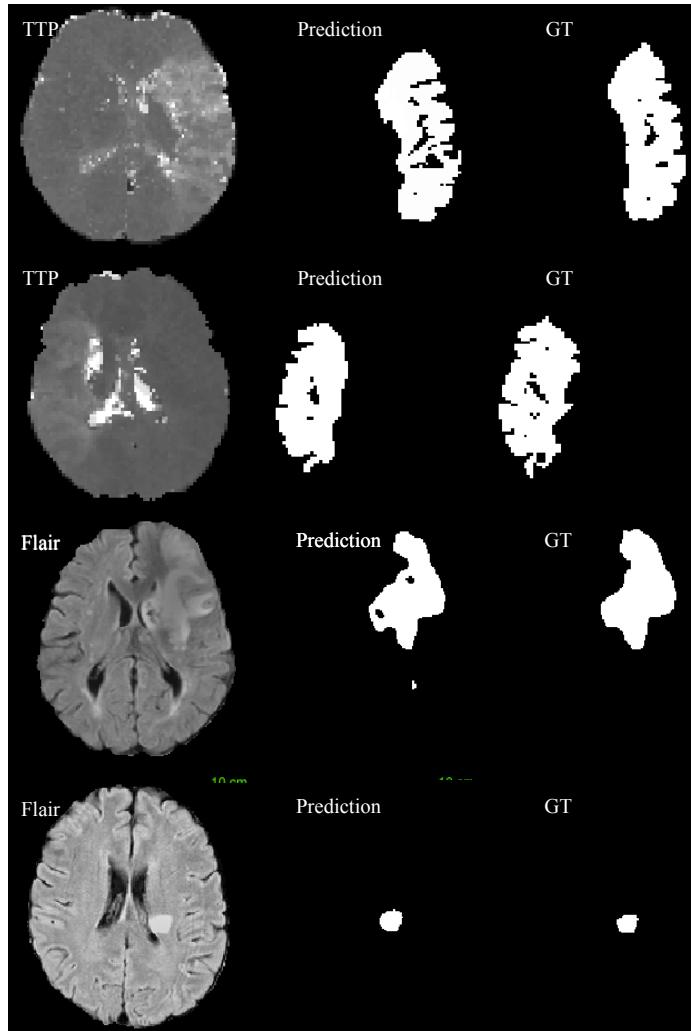


Fig. 4: Results obtained by our model on SPES (top row) and SISS (bottom row) datasets.

Method	ASSD		Dice		Hausdorff Distance		Precision		Recall	
	average	std	average	std	average	std	average	std	average	std
dutif1	8.92	19.23	0.69	0.30	31.75	28.52	0.72	0.31	0.67	0.31
halmh1	6.77	13.17	0.63	0.23	36.16	36.46	0.68	0.24	0.64	0.26
jessa1	11.59	18.34	0.45	0.24	39.23	30.70	0.52	0.26	0.51	0.31
mahmq2	10.30	11.11	0.54	0.26	82.78	23.95	0.67	0.33	0.50	0.25
maieo1	12.36	12.30	0.36	0.25	56.94	40.98	0.65	0.41	0.35	0.21
muscj1	56.77	79.90	0.48	0.38	76.88	81.77	0.57	0.43	0.44	0.37
pintal	12.18	22.59	0.50	0.31	43.21	30.50	0.61	0.34	0.55	0.33
robdd1	9.36	13.85	0.57	0.28	53.88	34.58	0.58	0.33	0.68	0.21

Table 1: Results on the SISS training dataset showing how our method compares with other methods.

Method	ASSD		Dice		Hausdorff Distance		Precision		Recall	
	average	std	average	std	average	std	average	std	average	std
dutif1	1.76	0.94	0.85	0.08	23.28	14.13	0.83	0.11	0.88	0.08
haect1	3.51	2.13	0.78	0.08	46.31	25.17	0.78	0.11	0.80	0.12
mckir1	1.42	1.01	0.85	0.06	30.71	18.91	0.84	0.10	0.87	0.07
robbd1	2.03	1.35	0.82	0.07	44.29	27.59	0.81	0.14	0.85	0.07

Table 2: Results on the SPES training dataset showing how our method compares with other methods.

5 Conclusion

In this document, we proposed a brain lesion segmentation method based on deep convolutional neural networks. Results on the SISS and SPES datasets reveal that our method is clearly the most accurate one. The high performance is achieved with the help of a novel two-pathway architecture which can model both the local details and global context. Note that the proposed CNN is close to the one used by our team on the 2015 MICCAI brain tumor segmentation (BRATS) challenge. Since there also our approach produced the most accurate results, we are inclined to believe that CNN is a promising technology for brain segmentation applications.

References

1. Virtual skeleton database. <http://www.virtualskeleton.ch/>
2. Avants, B.B., et al.: Advanced normalization tools (ants). *Insight J* (2009)
3. Davy, A., et al.: Brain tumor segmentation with deep neural networks. *proc of BRATS-MICCAI* (2014)
4. Goodfellow, I., et al.: Pylearn2: a machine learning research library. *arXiv preprint arXiv:1308.4214* (2013)
5. Huang, G.B., Jain, V.: Deep and wide multiscale recursive networks for robust image labeling. *ICLR*, *arXiv:1310.0354* (2014)
6. Jaderberg, M., Vedaldi, A., Zisserman, A.: Speeding up convolutional neural networks with low rank expansions. In: *in proc of BMVC* (2014)
7. Krizhevsky, A., et al.: ImageNet classification with deep convolutional neural networks. In: *NIPS* (2012)
8. Menze, B., et al: The multimodal brain tumor image segmentation benchmark (brats). *Medical Imaging* (2014)
9. Urban, G., et al.: Multi-modal brain tumor segmentation using deep convolutional neural networks. *proc of BRATS-MICCAI* (2014)
10. Zikic, D., et al.: Segmentation of brain tumor tissues with convolutional neural networks. *proc of BRATS-MICCAI* (2014)

Hierarchical Segmentation of Normal and Lesional Structures Combining an Ensemble of Probabilistic Local Classifiers and Regional Random Forest Classification

Andrew Jesson and Tal Arbel

Centre for Intelligent Machines, McGill University

1 Overview

We present a hierarchical framework for the simultaneous segmentation of normal and pathological structures in brain MRI. The framework starts with the ensemble decision made by a number of probabilistic local classifiers distributed throughout a fixed reference space (EPLC). The EPLC provides consistent smooth segmentations for both normal and pathological structures which are then grouped into regions and passed on to a high-level regional random forest classifier (RRF).

2 Implementation

The class of a given voxel in medical image segmentation problems depends on its location within the image reference space, its intensity value, neighborhood context in both the intensity and class label domain, and regional context between continuously labeled structures. Here we model location by an ensemble of local classifiers populated throughout a common reference space. Each local classifier models the class posterior probability of the class given voxel and neighbourhood context in the intensity domain by taking a filtered patch of the image as input. Neighbourhood context in the label domain is modeled by a global Markov Random Field (MRF). Finally, regional context is captured by a random forest classifier.

2.1 Ensemble of Probabilistic Local Classifiers

Reference space. The EPLC requires a common reference space and so we use the reference space defined by the MNI Linear ICBM Average Brain Stereotaxic Registration Model [2,3]. Each subject is normalized into this space using the rigid transform determined by the antsRegistration tool [4] with subject/template T1 image pair as input. The spatial centres of each local classifier are distributed throughout this space in a hexagonal close-packed lattice and are given by

$$\mathbf{x}_c = \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} 2i + ((j+k) \bmod 2) \\ \sqrt{3}(j + \frac{1}{3}(k \bmod 2)) \\ \frac{2\sqrt{6}}{3}k \end{bmatrix} r, \quad (1)$$

where i, j, k are the voxel indices and r is chosen to be 8.5mm.

Local classifiers. In total there are $N = 554$ local classifiers as we retain only those with spatial centres that overlap with the reference space model's brain mask. Each local classifier has a radially decaying spatial responsibility in the reference space given by

$$w_n(\mathbf{x}) = \frac{1}{Z} \exp\left(-c \frac{(\mathbf{x} - \mathbf{x}_c)^2}{r^2}\right), \quad (2)$$

where c is a constant controlling the decay of the spatial responsibility chosen here to be 0.693 and Z is a normalization constant defined by the sum of all spatial responsibilities at \mathbf{x} .

Each classifier builds distributions for $K = 11$ classes in total, namely background, cerebrospinal fluid, lateral ventricles, other ventricles, deep gray matter, cortical gray matter, cerebellar gray matter, cerebral white matter, cerebellar white matter, brain stem, and lesion. Given that ground truth is only available for lesion, we generate a set of atlases from the training data using the Multi-Atlas Label Fusion approach that was developed in previous work in the context of MS lesion segmentation to get training samples for normal structures [1].

To model neighbourhood context in the intensity domain each classifier takes a 15x7x3x3 patch centred around a given voxel as input. The first three dimensions of the patch are spatial and the fourth dimension is comprised of the T1, FLAIR, and DWI contrasts. To spatially decorrelate the input patch and reduce dimensionality we filter the patch with principal component analysis (PCA) determined kernels. Mixtures of gaussians (GMM) then model the distribution of each class given the filtered input.

For each test subject, 5000 samples are drawn randomly, with replacement, and weighted by $w_n(\mathbf{x})$. The patches are reshaped to 1x945 observation vectors $\mathbf{d}(\mathbf{x})$ to form a 5000x945 observation matrix D and each column of D is standardized. GMM inputs are given by $\hat{\mathbf{d}}_n(\mathbf{x}) = \mathbf{d}(\mathbf{x})U_n$, where U_n is a 945xp matrix of principal components determined during training. U_n is unique for each local classifier and p is determined by taking the principal components with largest explained variance such that the total retained variance is $\sim 90\%$ of the total variance from the training set.

The probability density functions $p_n(\hat{\mathbf{d}}_n(\mathbf{x}) | C_k)$ are estimated during training using a GMM for each class C_k and the number of components for each model are determined iteratively using Bayesian information criterion (BIC). The class prior probabilities $p_n(C_k)$ are estimated as the relative frequency of class C_k sampled by the local classifier during training. The posterior probability of observing class C_k given $\hat{\mathbf{d}}_n(\mathbf{x})$ by the n th classifier is given by

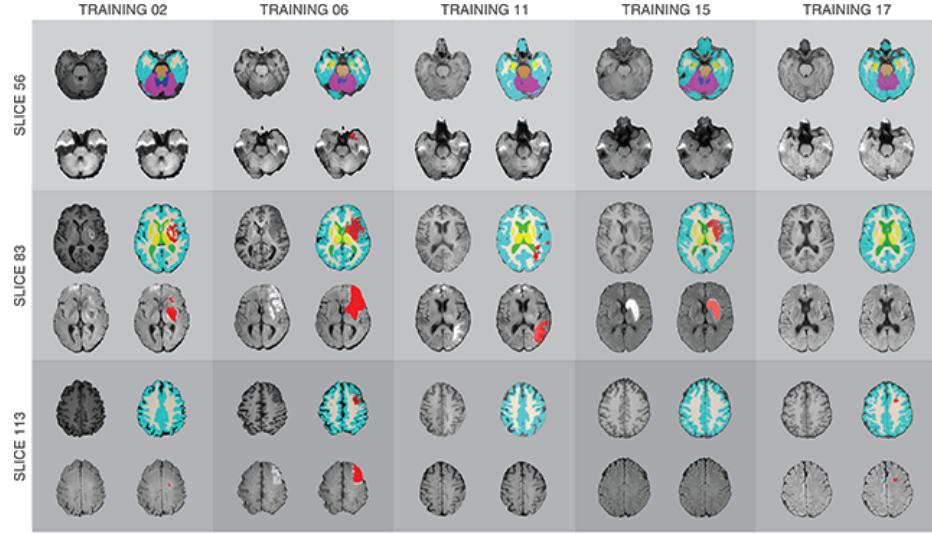


Fig. 1. Example segmentations after ensemble decision. Each square contains an axial slice from a given subject: T1, top left; Segmentation, top right; DWI, bottom left; Ground truth lesion, bottom right. Lesion, red; CSF, green; deep gray matter, yellow; cortical gray matter, cyan; cerebellar gray matter, purple; cerebral white matter, white; cerebellar white matter, blue; brain stem, beige. Results are based on 7-fold cross validation on the training data.

$$p_n \left(C_k \mid \hat{\mathbf{d}}(\mathbf{x}) \right) \propto p_n \left(\hat{\mathbf{d}}_n(\mathbf{x}) \mid C_k \right) p_n (C_k). \quad (3)$$

The ensemble confidence for a given class at voxel \mathbf{x} is then given by

$$f(w_n(\mathbf{x}), \hat{\mathbf{d}}_n(\mathbf{x}), C_k) = \sum_{n=1 \dots N} w_n(\mathbf{x}) p_n \left(C_k \mid \hat{\mathbf{d}}_n(\mathbf{x}) \right). \quad (4)$$

Global MAP-MRF. Given that samples for each local classifier are drawn randomly with replacement there is no guarantee that all voxels will be visited by the EPLC. We use a MRF solution to yield a smooth labelling. The prior energy is given by a Potts model with $\beta = 0.1$ and the observation energy is given by $-\log(f(w_n(\mathbf{x}), \hat{\mathbf{d}}_n(\mathbf{x}), C_k))$. The optimal labelling solution is found using ICM. Figure 1 shows example segmentations from several subjects.

Utilizing probabilistic outliers. As can be seen in figure 1 the healthy tissue segmentations are qualitatively consistent and acceptable; however, lesions appear consistently under segmented. Fortunately, by using a probabilistic model for each local classifier we can obtain a quantitative outlier measure using the Mahalanobis distance (computing CDF for each high dimension GMM is prohibitively time consuming). An outlier mask is generated for each subject by

thresholding the Mahalanobis distance corresponding to the maximum posterior class value by 2 standard deviations above the global mean measure. This mask is then added to the proposed lesion segmentation.

2.2 Regional Random Forest Classifier

Lesion segmentations are finally refined using a random forest classifier. Candidate lesions are defined by morphological 18-connected regions. Features for the random forest are the distance minimum, maximum, mean and variance from each normal tissue label excluding background; the volume, and solidity of the candidate lesion; the convex hull inertial tensor and principal moments of the candidate lesion; and 32 bin histograms from normalized T1, FLAIR, and DWI contrasts over the candidate lesion. In total there are 146 features and each candidate split in the forest randomly chooses 12 of these features. We use MATLAB's 'treebagger' class to implement the random forest. Candidate lesions for which the confidence of the random forest is greater than around 40% are retained for the final classification. Figure 2 shows the effect of the high level refinement on the lesion segmentations provided by the ensemble of local classifiers.

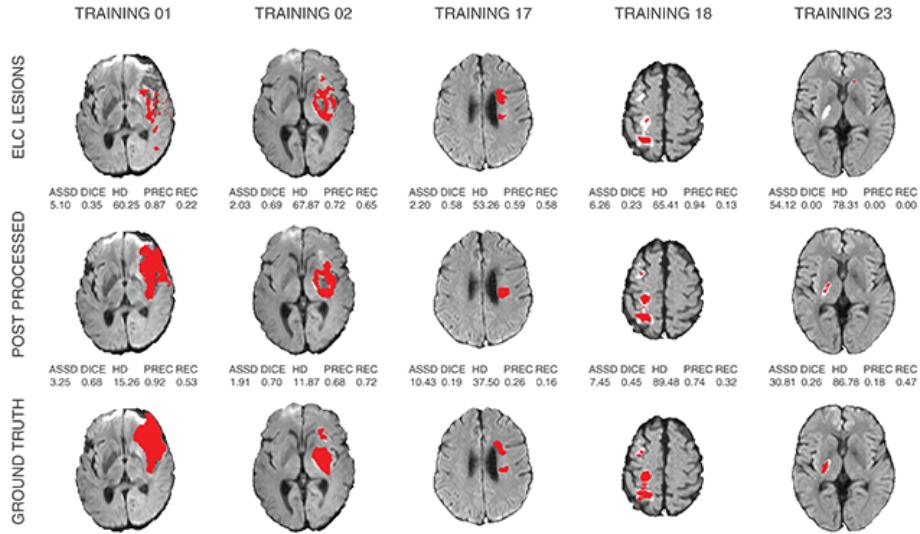


Fig. 2. Example lesions segmentations. From top to bottom: EPLC output, context based classifier output, ground truth segmentation. From left to right: different subjects in training dataset. Below rows 1 and 2 are the computed challenge metrics for the subject. Results are based on 7-fold cross validation on the training data.

References

1. Jesson, A., Arbel, T.: Hierachical MRF and Random Forest Segmentation of MS Lesions and Healthy Tissues in Brain MRI, presented at the Longitudinal Multiple Sclerosis Lesion Segmentation Challenge held in conjunction with ISBI 2015, Brooklyn, New York, April, 16, 2015. (http://www.iac1.ece.jhu.edu/w/images/7/72/Andrew_Jesson.pdf)
2. Foynov, V., Evans, A., et al: Unbiased Average Age-appropriate Atlases for Pediatric Studies. *NeuroImage*, vol. 54, no. 1. (2011)
3. Collins, L., Zijdenbos, A., et al: ANIMAL+INSECT: Improved Cortical Structure Segmentation. *IPMI LNCS*, vol. 1613/1999, 210223 (1999)
4. Avants, B., Epstein, C., et al: Symmetric Diffeomorphic Image Registration with Cross-Correlation: Evaluating Automated Labeling of Elderly and Neurodegenerative Brain. *Medical Image Analysis*, vol. 12, no. 1, 2641 (2008)

A Vascular Territory Prior for Bayesian Sub-Acute Ischemic Stroke Lesion Segmentation

Senan Doyle¹, Florence Forbes², and Michel Dojat³

¹ Pixyl Medical, Grenoble, France, senan.doyle@pixyl.io,

² INRIA Rhône-Alpes, Grenoble, France,

³ INSERM, U836, Grenoble, France

Abstract. The complex task of quantifying sub-acute stroke lesion volume is addressed by introducing prior information on the vascular territory in a Bayesian framework. Image data is described using a hidden Markov random field model, which accounts for spatial dependencies and allows for the inclusion of a priori spatial information. The algorithm is designed to perform without the need for a preliminary training phase by availing of this prior knowledge, and by imposing constraints on the estimation of MRF interaction parameters.

1 Introduction

Stroke is a global epidemic with levels of mortality and disability that entail a high societal cost. Approximately 70% of stroke survivors have significant sensorimotor, language and cognitive dysfunction requiring long-term special care and rehabilitation [3]. Ischemic stroke lesion quantification facilitates the development of new biomarkers for improved diagnosis, prognosis, and rehabilitation. Automatic methods are motivated by the demand for large-scale multi-center clinical research studies that require precise, consistent and cost-efficient analysis.

Existing methods frequently avail of complementary information from multiple sequences. For example, lesion voxels may appear atypical in one modality and normal in another. This is implicitly used by neurologists when examining data. In an automatic segmentation framework, multiple sequences increase the discriminative capacity of the model to identify different tissues and structures. The proposed statistical model is an implementation of a hidden Markov random field (HMRF) with a number of innovations to address the challenges posed by sub-acute ischemic stroke MR scans.

A general HMRF formulation is employed that encodes complex interactions between neighboring voxels. In particular, we encode the fact that certain tissue combinations in the neighborhood are penalized more than others, whereas the standard Potts model penalizes dissimilar neighboring classes equally, regardless of the tissues they represent.

Prior probabilistic spatial information is incorporated via the Markov prior to express our *a priori* knowledge on the location of tissues or lesion. The accurate segmentation of sub-acute stage ischemic stroke volume demands a more

comprehensive vascular understanding of the human brain. The statistical model includes *a priori* vascular information from a probabilistic vascular territory atlas derived from [2]. This atlas improves the classification by modeling the potential progression and delimitation of vascular accidents.

A solution to the model is found using the Expectation-Maximization (EM) framework [4] combined with a variational approximation for tractability in the presence of Markov dependencies. In particular, the mean field principle provides a deterministic way to deal with intractable MRF models[5].

2 Sub-Acute Stroke Lesion Model

The classification task is performed using a multi-sequence hierarchical Bayesian model that allows us to identify, not only the total lesion volume, but also sub-regions such as necrotic regions, that may be of clinical relevance.

We consider a finite set V of N voxels on a regular 3D grid. The intensity values observed at each voxel are denoted by $\mathbf{y} = \{\mathbf{y}_1, \dots, \mathbf{y}_N\}$. Each $\mathbf{y}_i = \{y_{i1}, \dots, y_{iM}\}$ is itself a vector of M intensity values corresponding to M different MR sequences. The classification task is to assign each voxel i to one of K classes. The latent data, denoted by $\mathbf{z} = \{\mathbf{z}_1, \dots, \mathbf{z}_N\}$, better describes the observed data in the context of the statistical model. The \mathbf{z}_i 's correspond to class memberships, and assume values in $\{e_1, \dots, e_K\}$, where e_k is a K -dimensional binary vector whose k^{th} component is 1, all other components being 0. We will denote by $\mathcal{Z} = \{e_1, \dots, e_K\}^N$ the set in which \mathbf{z} takes its values. The set of voxels V has an associated neighborhood system. Spatial dependencies between voxels are modeled by assuming a Markov random field (MRF) prior. Denoting $\psi = \{\beta, \phi\}$ additional parameters, we assume that the joint distribution $p(\mathbf{y}, \mathbf{z}; \psi)$ is a MRF with the following energy function:

$$H(\mathbf{y}, \mathbf{z}; \psi) = H_{\mathbf{Z}}(\mathbf{z}; \beta) + \sum_{i \in V} \log g(\mathbf{y}_i | \mathbf{z}_i; \phi), \quad (1)$$

where the $g(\mathbf{y}_i | \mathbf{z}_i; \phi)$'s are probability density functions of \mathbf{y}_i . For brain data, the data term $\sum_{i \in V} \log g(\mathbf{y}_i | \mathbf{z}_i; \phi)$ in (1) corresponds to the modeling of tissue dependent intensity distributions. For our multi-dimensional observations, we consider M -dimensional Gaussian distributions with diagonal covariance matrices. For each class k , $(\mu_{k1}, \dots, \mu_{kM})$ is the mean vector and $\{s_{k1}, \dots, s_{kM}\}$ the covariance matrix components. We will use the notation $\mu_m = {}^t(\mu_{km}, k = 1 \dots K)$ and $s_m = {}^t(s_{km}, k = 1 \dots K)$. When $\mathbf{z}_i = e_k$ then $\mathcal{G}(y_{im}; \langle \mathbf{z}_i, \phi_m \rangle)$ and $\mathcal{G}(y_{im}; \langle \mathbf{z}_i, \mu_m \rangle, \langle \mathbf{z}_i, s_m \rangle)$ both represent the Gaussian distribution with mean μ_{km} and variance s_{km} . The whole set of Gaussian parameters is denoted by $\phi = \{\phi_{km}, k = 1, \dots, K, m = 1, \dots, M\}$. Our data term is then defined by setting $g(\mathbf{y}_i | \mathbf{z}_i; \phi) \propto \prod_{m=1}^M \mathcal{G}(y_{im}; \langle \mathbf{z}_i, \phi_m \rangle)$. The missing data term $H_{\mathbf{Z}}(\mathbf{z}; \beta)$ in (1) describes the dependencies between neighboring Z_i 's, and is specified by further assuming that the joint distribution on $\{Z_1, \dots, Z_N\}$ is a discrete MRF:

$$P(\mathbf{z}; \beta) = W(\beta)^{-1} \exp(-H_{\mathbf{Z}}(\mathbf{z}; \beta)) \quad (2)$$

where β is a set of parameters, $W(\beta)$ is a normalizing constant and $H_{\mathbf{Z}}$ is a function restricted to pair-wise interactions,

$$H_{\mathbf{Z}}(\mathbf{z}; \beta) = - \sum_{i \in S} z_i^t \alpha - \sum_{\substack{i,j \\ i \sim j}} z_i^t \mathbb{B} z_j,$$

where we write z_i^t for the transpose of vector z_i and $i \sim j$ when i and j are neighbors. The set of parameters β is decomposed into $\beta = (\alpha, \mathbb{B})$. Parameter α is a K -dimensional vector which acts as weights for the different values of z_i . \mathbb{B} is a $K \times K$ matrix that encodes interactions between the different classes. If in addition to a null α , $\mathbb{B} = b \times I_K$ where b is a real scalar and I_K is the $K \times K$ identity matrix, parameters β reduce to a single scalar interaction parameter b and we get the Potts model traditionally used for image segmentation.

Note that the standard Potts model is often appropriate for classification since it tends to favor neighbors that are in the same class. However, this model penalizes pairs that have different classes with the same penalty, regardless of the tissues they represent. In practice, it may be more appropriate, to encode higher penalties when the tissues are known to be unlikely neighbors. For example, the penalty for a white matter and extraventricular CSF pair is expected to be greater than that of a grey matter and extraventricular CSF pair, as these two classes are more likely to form neighborhoods.

We adopt a data model comprising of five normal tissue classes; *white matter*, *grey matter*, *ventricular CSF*, *extraventricular CSF*, and *other*. The lesion is modeled by a further two classes representing abnormal tissue state; sub-necrotic and necrotic. In the absence of sufficient data to robustly and accurately estimate a full free \mathbb{B} with $K = 7$, further constraints are imposed on the MRF interaction matrix. The two lesion classes are considered sub-classes of a single *structure*, whose interaction with the normal tissue classes is not dependent on the specific lesion sub-class. Letting τ be the set of classes comprising the lesion structure, \mathbb{B} is a matrix defined by:

$$\begin{aligned} \mathbb{B}(k, k') &= b_t \quad \forall k, k' \in \tau \\ \mathbb{B}(k, k') &= b_{\{k, k'\}} \quad \text{otherwise.} \end{aligned} \quad (3)$$

Prior knowledge on the expected neighborhoods can be encoded in \mathbb{B} . For example, given two classes that are likely to be adjacent, the matrix entries for this pair can be initialized at, or even fixed to, a higher value. Conversely, when there is enough information in the data, a full free \mathbb{B} matrix can be estimated and will reflect the class structure (*i.e.* which class is next to which as indicated by the data) and will then mainly serve as a regularizing term to encode additional spatial information.

For the distribution of the observed variables \mathbf{y} given the classification \mathbf{z} , the usual conditional independence assumption is made. It follows that the conditional probability of the hidden field \mathbf{z} given the observed field \mathbf{y} is

$$P(\mathbf{z}|\mathbf{y}; \psi, \beta) = W(\beta)^{-1} \exp \left(-H_{\mathbf{Z}}(\mathbf{z}; \beta) + \sum_{i \in S} \log g(y_i | z_i, \phi) \right).$$

Parameters are estimated using the variational EM algorithm, which provides a tractable solution for non trivial Markov models.

3 Vascular Territory Atlas

We introduce a novel probabilistic vascular territory atlas to model the potential progression and delimitation of vascular accidents, and therefore overcome misclassification due to artefacts. The vascular territory atlas was derived from [2] and refined by a neurologist. The atlas is discrete, and so Gaussian blur is applied to express uncertainty inherent to patient-specific analysis. In effect, the territory prior does not forbid the realization of lesion labels at any location in the image, but expresses the lower probability of solutions that contain lesions in multiple vascular territories. The vascular territory structure is subject to an affine transformation as the iterative segmentation-registration framework executes. This also helps to overcome patient-specific bias. The discrete vascular territory atlas is shown in Figure 1a alongside an example of the Gaussian blurred territory for use in the joint segmentation & registration step (Fig.1b). The final territory (Fig.1c) is produced by the transformation of the initial territory in the joint model, and shows how the territory is adapted to the individual patient data.

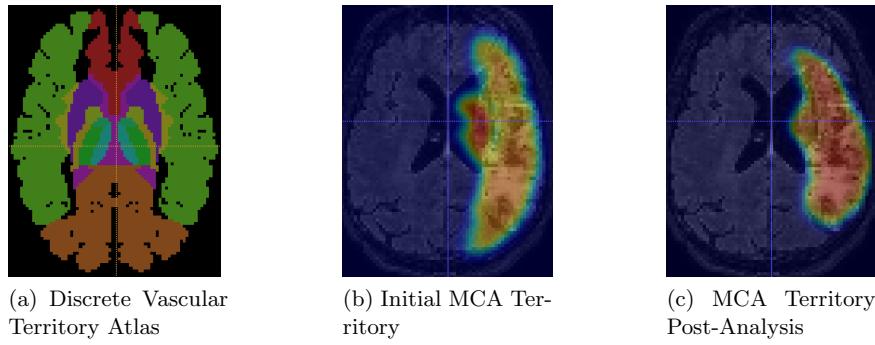


Fig. 1: Prior Vascular Territory Construction and Transformation

4 Method

The MR images are first preprocessed. The images are corrected for intensity inhomogeneities using the N4 algorithm [1]. The deformable transform that describes the mapping between the International Consortium for Brain Mapping (ICBM) template and the data space is found using tools provided by the Insight Segmentation and Registration Toolkit (ITK). The transform is used to register both the probabilistic tissue atlases, and the vascular territory atlas, to the MR sequences.

An initial preprocessing step identifies the vascular territory to be used in the subsequent segmentation step. Outlier intensities are quantified in each vascular territory grouping based on the DWI and Flair images, and the maximum is chosen.

Adopting Bayesian principles, the algorithm performs joint segmentation and registration of the a priori probabilistic vascular territory structure, as described in [6]. Bias caused by commitment to the initial registration is thus alleviated by refining the registration as the algorithm executes.

5 Discussion & Conclusion

We proposed an adaptive scheme of multiple MR sequences for sub-acute stroke lesion segmentation. Our approach is fully automatic and requires no training. The model parameters are instead estimated using a variational EM algorithm with MRF constraints and the inclusion of a priori probabilistic maps to provide a stable parameter trajectory during optimization. Vascular information is included by way of a prior vascular territory atlas that is modified in a joint segmentation & registration framework to adapt to patient-specific data.

6 The References Section

References

1. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: Improved N3 Bias Correction. *IEEE transactions on medical imaging*. 2010;29(6):1310-1320. doi:10.1109/TMI.2010.2046908.
2. Tatu, L., Moulin, T., Bogousslavsky, J., Duvernoy, H. "Arterial territories of human brain Brainstem and cerebellum." *Neurology* 47, no. 5 (1996): 1125-1135.
3. Hommel, M., Miguel, S.T., Naegele,B., Gonnet,N., Jaillard, A.: Cognitive determinants of social functioning after a first ever mild to moderate stroke at vocational age.. *Journal of Neurology, Neurosurgery and Psychiatry*, 2009, 80 (8), pp.876-80.
4. Dempster, A., Laird, N., Rubin, D.: Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B* 39 (1977) 1-38
5. Celeux, G., Forbes, F., Peyrard, N.: EM procedures using mean field-like approximations for Markov model-based image segmentation. *Pat. Rec.* 36(1) (2003) 131-144
6. Scherrer, B., Forbes, F., Garbay, C., Dojat, M. Fully Bayesian joint model for MR brain scan tissue and structure segmentation. *Medical image computing and computer assisted intervention : MICCAI .. International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2008, 11 (Pt 2), pp.1066-74.

Segmentation of Ischemic Stroke Lesions Using Dictionary Learning

Jean Huang, Lawrence Staib, and James Duncan

Department of Biomedical Engineering, Yale University, USA,
 {jean.huang, lawrence.staib, james.duncan}@yale.edu

Abstract. The segmentation of sub-acute ischemic stroke lesions in the brain is a challenging task, typically with laborious manual or semi-automatic methods performed by practiced clinicians. Thus, the search for an automatic segmentation process versatile enough to be widely used clinically continues. Furthermore, automatic multimodal magnetic resonance imaging (MRI) segmentation would improve results by allowing us to use information provided by several different imaging modalities. Here, we propose a supervised learning model using dictionary learning, sparse coding, and support vector machines on several MR sequences: DWI, FLAIR, T1 and T2. We

1 Introduction

A stroke is characterized by restricted blood flow to the brain, depriving brain tissue of vital oxygen and nutrients and resulting in cell death. In developed countries, stroke is the second most frequent cause of death and a major cause of disability. There are two main types of stroke: ischemic, due to lack of blood flow, and hemorrhagic, due to bleeding. Ischemic stroke, caused by interruption of the blood supply to the brain, is much more common than hemorrhagic stroke. Magnetic resonance imaging (MRI), due to its sensitivity and specificity in depicting alterations in brain water content, is often used for diagnosis to detect and assess the stroke lesion presence [7][9][10]. An automatic method to segment the stroke lesions from multimodal MRI would have great clinical use.

2 Materials and Methods

Our approach can be divided into several different sections. The first section consists of preprocessing, which includes image patch extraction and patch normalization. The second section consists of dicionary learning and sparse coding using these image patches [1]. The third section focuses on the relative reconstruction error map based on our learned dictionary. The last section shows how we use support vector machines on our dictionary coefficients to segment the lesions.

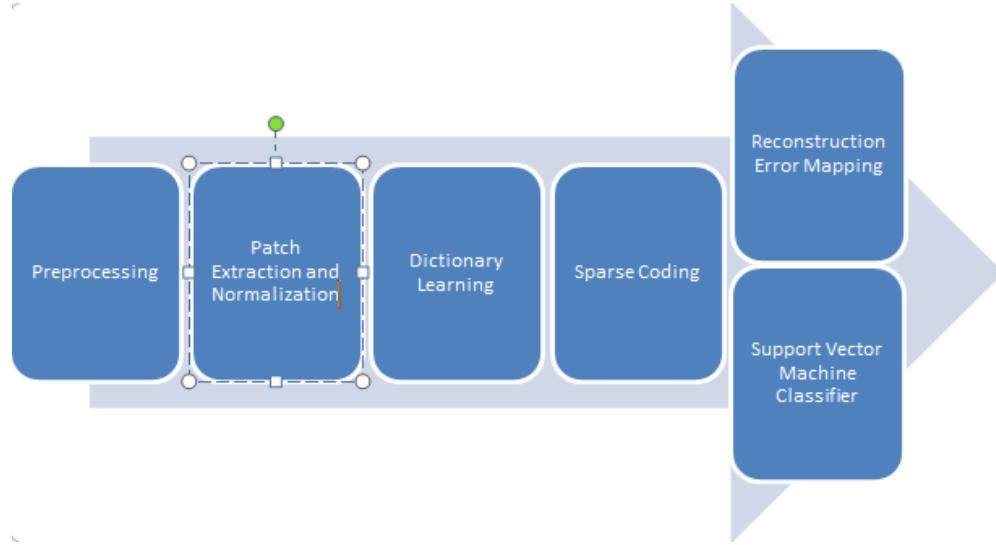


Fig. 1. Flowchart of our proposed ischemic stroke lesion segmentation method. The final step can be either through thresholding of the reconstruction error map or through support vector machine classification on the dictionary coefficients for the image patches.

2.1 Preprocessing

The training data, comprised of T1, T2, FLAIR, and DWI sequences, has already been skull-stripped and co-registered to the FLAIR images. For each MR modality, we extract m image patches from inside the brain and realign them as one-dimensional vectors $x_1, \dots, x_m \in R^k$, where we choose $k = 27$ for a three-dimensional image patch of size $3 \times 3 \times 3$. For patch normalization, we divide each patch by the value of the highest L_1 norm of all the patches. We then concatenate the corresponding vectors from the different image sequences (T1, T2, FLAIR, and DWI) and re-normalize so that $\|x_i\|_2^2 \leq 1$ and $x_i \in R^{zk}$, where z is the number of image sequences. For experiments here, image patches are divided into training and test sets, and experiments are performed using leave-one-out cross-validation.

2.2 Dictionary Learning and Sparse Coding

Using the image patches as the training set, we search for a dictionary basis $D \in C = \{D \in R^{k \times l} s.t. \forall j : \|x_i\|_2^2 \leq 1\}$, with l atoms d_j , so that we solve the optimization problem

$$\min_D \sum_{i=1}^m \frac{1}{2} \|x_i - D\alpha_i\|_2^2 \text{ s.t. } \|\alpha_i\|_1 \leq \lambda_1. \quad (1)$$

We can use one, two, or several dictionaries, based on the number of classes we want (for example, healthy versus lesion patches). The simplest case is with a single dictionary, learned from both the healthy and lesion class patches. This is the method we will describe here; in the future, we can adapt to multi-class dictionaries. For a single dictionary, the number of lesion patches is very small compared to the number of healthy patches, so this dictionary is mainly representative of the healthy brain image patches. The basis D for this dictionary satisfies $x_i \approx D\alpha_i$ for most image patches.

The equation in (1) is a sparse coding problem, where the L_1 constraint and the λ_1 sparsity induced regularizer ensures that only a few atoms of dictionary D will be used to represent an image patch. Studies have shown that this balances the trade-off between sparsity of the coefficients α_i and the reconstruction error. We reconstruct the image patches in the following step

$$\min_{\alpha_1} \|x_i - D\alpha_i\|_2^2 \text{ s.t. } \|\alpha_i\|_1 \leq \lambda_2. \quad (2)$$

and obtain a reconstruction error for each image patch based on the λ_2 constraint. Compared to the healthy tissue intensities, the lesions are outliers, so the reconstruction error for those patches would be higher than those for the healthy tissue patches [15].

2.3 Reconstruction Error

For each test patch, we calculate the sparse coefficients α_i given the sparsity constraint λ_2 . The relative reconstruction error of each patch is given as

$$err(x_i, \alpha_i) = \frac{\|x_i - D\alpha_i\|_2}{\|x_i\|_2}. \quad (3)$$

We can map the reconstruction error at the position of the centered voxel within each patch, and the result is an error map throughout the whole brain. We then chose a threshold, for which the test patches with reconstruction error larger than the threshold are classified as lesion patches [6][13].

2.4 Support Vector Machine

Beyond thresholding the reconstruction error, we can take our method a step further. Given our trained overcomplete dictionary, we can further use supervised learning classification techniques on dictionary coefficients to obtain a better segmentation. Here, we take our coefficients α_i and input them as feature vectors into the support vector machine (SVM) framework. From our training set, we have our input feature vector α_i and corresponding target $y_i \in \{-1, +1\}$, where $y_i = +1$ is for lesion tissue and $y_i = -1$ is for healthy tissue. In the feature space, the SVM models are of the form

$$\mathbf{y}(\alpha) = \omega^T \phi(\mathbf{x}) + \mathbf{b} \quad (4)$$

where the $\phi()$ maps the input vector into higher dimensional feature space, \mathbf{b} is the bias, and ω is the weight vector. A separating hyperplane, generalized to the nonlinear case, that separates the data into healthy and lesion patches can be described as

$$\begin{aligned}\omega^T \phi(\mathbf{x}_i) + \mathbf{b} &\geq +1 & \text{if } y_i = +1 \\ \omega^T \phi(\mathbf{x}_i) + \mathbf{b} &\leq -1 & \text{if } y_i = -1\end{aligned}$$

We want to find, out of all the possible separating hyperplanes, the one with maximum margin $\frac{2}{\|\mathbf{w}\|}$ between the two classes. This hyperplane will separate our healthy patches from our lesions patches. After we have trained our SVM classifier on the coefficients α_i from our dictionary \mathbf{D} , we can use it on test data. From our reconstructed image patches, instead of thresholding based on reconstruction error, we plug the coefficients into the SVM classifier[5].

2.5 Implementation

Our method is implemented in MATLAB. We also use the SParse Modeling Software (SPAMS) [3] from Inria and OMP-Box [2] toolbox from Technion. Many values have been tested using parameter sweeps, and we have found that a good image patch size is $3 \times 3 \times 3$. After testing dictionary sizes ranging from 100 to 5000, we found that a dictionary size $l = 1000$ atoms produced the best results. Empirical results for parameter values showed that sparsity constraints of $\lambda_1 = 0.7$ and $\lambda_2 = 0.85$ tended to work best.

3 Discussion and Conclusion

We proposed an automatic segmentation method for ischemic stroke lesions using dictionary learning on multimodal MRI. Our method uses a single dictionary on all the modalities treated equally, but in the future, we can extend to two dictionaries, one for the lesion class and one for the healthy class of image patches, or even more dictionaries, one each for the white matter, gray matter, and cerebrospinal fluid. We can adapt the size of each dictionary based on the complexity of each class [6].

We can also better utilize the DWI sequence, with its wealth of information about water diffusion directionality and magnitude. We can calculate tensors from the DW images and use invariants such as fractional anisotropy and mean diffusivity to include in our image patches. We can include diffusion directionality information such as principal eigenvectors or eigenvalues. These scalar and vector values we can either concatenate in our image patch vectors, or directly concatenate them in our feature vector that we plug into the support vector machines. Furthermore, instead of SVM, we could use other types of classifiers on the dictionary coefficients, such as neural networks or random forests. With better computational power, we could also try to work with larger image patches, which could capture more textural information, and potentially further improve results.

References

- [1] ICCV 2009 Tutorial.
- [2] Ron Rubinstein - Software.
- [3] SParse Modeling Software.
- [4] M. Aharon, M. Elad, and A. Bruckstein. -SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation. *IEEE Transactions on Signal Processing*, 54(11):4311–4322, November 2006.
- [5] Christopher J. C. Burges. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2:121–167, 1998.
- [6] Daniel Garca-Lorenzo, Simon Francis, Sridar Narayanan, Douglas L. Arnold, and D. Louis Collins. Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. *Medical Image Analysis*, 17(1):1–18, January 2013.
- [7] Yacine Kabir, Michel Dojat, Benot Scherrer, Florence Forbes, and Catherine Garbay. Multimodal MRI segmentation of ischemic stroke lesions. *Conference Proceedings*, 2007:1595–1598, 2007.
- [8] Julien Mairal, Francis Bach, Jean Ponce, and Guillermo Sapiro. Online Dictionary Learning for Sparse Coding. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML ’09, pages 689–696, New York, NY, USA, 2009. ACM.
- [9] D. M. Mezzapesa, M. Petruzzellis, V. Lucivero, M. Prontera, A. Tinelli, M. Sancilio, A. Carella, and F. Federico. Multimodal MR examination in acute ischemic stroke. *Neuroradiology*, 48(4):238–246, April 2006.
- [10] Islem Rekik, Stphanie Allassonnire, Trevor K. Carpenter, and Joanna M. Wardlaw. Medical image analysis methods in MR/CT-imaged acute-subacute ischemic stroke lesion: Segmentation, prediction and insights into dynamic evolution simulation models. A critical appraisal. *NeuroImage: Clinical*, 1(1):164–178, 2012.
- [11] Ron Rubinstein, Tomer Peleg, and Michael Elad. *Analysis K-SVD: A Dictionary-Learning Algorithm for the Analysis Sparse Model*. 2012.
- [12] Tong Tong, Robin Wolz, Pierrick Coup, Joseph V. Hajnal, and Daniel Rueckert. Segmentation of MR images via discriminative dictionary learning and sparse coding: Application to hippocampus labeling. *NeuroImage*, 76:11–23, August 2013.
- [13] Nick Weiss, Daniel Rueckert, and Anil Rao. Multiple sclerosis lesion segmentation using dictionary learning and sparse coding. *Medical image computing and computer-assisted intervention: MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 16(Pt 1):735–742, 2013.
- [14] Shaoting Zhang, Yiqiang Zhan, and Dimitris N. Metaxas. Deformable segmentation via sparse representation and dictionary learning. *Medical Image Analysis*, 16(7):1385–1396, October 2012.

Prediction of Ischemic Lesions using Local Image Properties and Random Forests

John Muschelli

Johns Hopkins Bloomberg School of Public Health

1 Methods

1.1 Registration with Symmetric Normalization

Although the images were registered to the FLAIR sequence, some local differences between spatial locations were observed. We therefore re-registered each sequence to the FLAIR image using Symmetric Normalization (SyN) included in the ANTs software.

1.2 Inhomogeneity Correction

Although, there may be non-uniform image intensities spatially in the brain, we did not perform inhomogeneity correction, for example, the N4 bias correction available within ANTs. We believe that as some lesions may be very large, using inhomogeneity corrections may induce lesion areas to be similar to non-lesion areas. We use estimated smoothed images in the prediction, which should help correct for some inhomogeneity.

1.3 Intensity Normalization

As MRI are acquired in arbitrary units, we performed an intensity-based normalization. Let $x_{s,v}$ represent the intensity of image sequence s for voxel v . Let $Z_{s,v}$ represent the intensity-normalized data. We subtract an estimated mean (μ_s) and standard deviation (σ_s) for each sequence and normalize the data as follows:

$$Z_{s,v} = \frac{x_{s,v} - \mu_s}{\sigma_s}$$

Estimation of the mean and standard deviation The mean and standard deviation are actually estimated from the trimmed distribution of the data of all voxels within the brain mask. The trimming procedure takes all voxels within the brain mask, removes the upper and lower 20% of the data, and estimates the mean and standard deviation from these intensities. This standardizes voxels to the number of (trimmed) standard deviations above the trimmed mean. The goal of the trimming is to delete high-intensity voxels from the lesion or low-intensity voxels from edema.

1.4 Imaging Predictors

We derived a set of imaging predictors from each scan. We will describe each here with their rationale for use. These features make up the potential set of predictors for image segmentation. Each operation uses the normalized images.

Normalized intensity information The normalized voxel intensity value in z-units was included, as it is the main predictor used in visual inspection; high values are indicative of lesion.

1.5 Flipped Difference Image

As most lesions are only on one side of the brain, we calculated the difference in intensity between a voxel and the voxel on its contralateral side. We obtained this image by first rigidly registering the image to the MNI template to account for any head tilt, then flip the image over the left-right axis, and then take the difference of the flipped image and the original image.

Local Moment Information For each voxel, we extracted a neighborhood, denoted N_v , of all adjacent neighboring voxels in 3 dimensions and the voxel itself. Let $I_{s,v}(k)$ denote the normalized voxel intensity in for voxel neighbor k , where $k = 1, \dots, 27$. We created the voxel neighborhood mean intensity ($\bar{x}_{s,v}$):

$$\bar{x}_{s,v} = \frac{1}{N_v} \sum_{k \in N_v} x_{s,v} \quad (1)$$

We calculated the voxel neighborhood standard deviation (SD), skew, and kurtosis using the following method of moments estimators:

$$\begin{aligned} \text{SD}_{s,v} &= \sqrt{\frac{1}{N_v} \sum_{k \in N_v} (x_k(v) - \bar{x}(v))^2} \\ \text{Skew}(v) &= \frac{\frac{1}{N_v} \sum_{k \in N_v} (x_k(v) - \bar{x}(v))^3}{\left[\frac{1}{N_v} \sum_{k \in N_v} (x_k(v) - \bar{x}(v))^2 \right]^{3/2}} \\ \text{Kurtosis}(v) &= \frac{\frac{1}{N_v} \sum_{k \in N_v} (x_k(v) - \bar{x}(v))^4}{\left(\frac{1}{N_v} \sum_{k \in N_v} (x_k(v) - \bar{x}(v))^2 \right)^2} \end{aligned}$$

We acknowledge that we did not divide by $N_v - 1$ for standard deviation and skewness, nor did we subtract by 3 for kurtosis. As N_v should be the same per

voxel, this should not affect the estimates for prediction and will be accounted for in any generalized linear model in the estimated coefficient. We also estimated the local gradient of the normalized intensity for each voxel neighborhood $\nabla_{v,s}$:

$$\nabla_{v,s} = \sqrt{\nabla_{v,s,x}^2 + \nabla_{v,s,y}^2 + \nabla_{v,s,z}^2}$$

Voxels higher in their local mean correspond to voxels adjacent to higher HU voxels on average, which are more likely to be lesion. The higher order moments can provide information about how homogeneous the intensities in the neighborhood are and where edges occur.

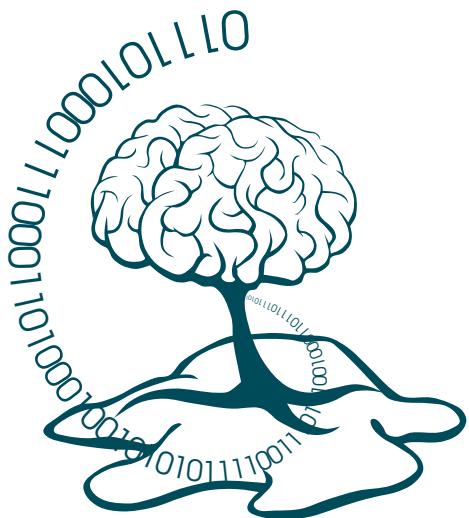
Global Head Information We created 3 images which were obtained by smoothing the original image using large Gaussian kernels ($\sigma = 5mm^3, 10mm^3, 20mm^3$), which can capture any potential homogeneity throughout the scan.

1.6 Model

To train an algorithm, we used 9 images and downsampled 300,000 voxels. We then used a random forest on these to predict lesion, using 500 trees. From the random forest, we obtained the probability of lesion and determined the threshold for these probabilities using the out-of-sample voxels from the training images, optimizing for the Dice Similarity Index (DSI). We then predicted lesions on the test dataset of 19 scans.

2 Discussion

We believe that our method allows for a robust procedure for segmentation of large ischemic lesions. This is due to intensity normalization and the set of features that can differentiate lesion areas from healthy tissue. Using local properties, we can leverage spatial information of the image. As such, we can use methods that can use the voxel information without having a more complicated multi-variate framework. The random forest allows for a flexible framework for prediction, especially as some features are highly correlated.



Acute Stroke Penumbra Estimation

ISLES 2015: SPES

Proceedings
5th October 2015
Munich, Germany



UNIVERSITÄT ZU LÜBECK



Technische Universität München

u^b

UNIVERSITÄT
BERN

MICCAI15
MUNICH

Segmenting the ischemic penumbra: a spatial Random Forest approach with automatic threshold finding

Richard McKinley¹, Levin Häni¹, Roland Wiest¹, Mauricio Reyes²

¹ Department of Diagnostic and Interventional Neuroradiology, Inselspital, Bern University Hospital, Switzerland

² Institute for Surgical Technology and Biomechanics, University of Bern, Switzerland

Abstract. We propose a fully automatic method for segmenting the ischemic penumbra, using image texture and spatial features and a modified Random Forest algorithm, which we call Segmentation Forests, which has been designed to adapt the original Random Forests algorithm of Breiman to the segmentation of medical images. The method is fast, taking approximately six minutes to segment a new case, and has yields convincing results (An out-of-sample average Dice coefficient of 0.85, with a standard deviation of 0.06).

1 Introduction

In patients presenting with acute stroke, it is important to be able to quickly identify hypoperfused tissue-at-risk, in order to assess the suitability of intra-arterial therapy. Thresholding maps derived from perfusion-weighted imaging provides a usable but crude assessment of this volume of tissue: the technique is prone to artifacts in the processing of the perfusion maps, leading to, for example, identification of tissue as at-risk on the contralateral side of the brain. Fast, automatic methods for identifying the tissue at risk which improve on thresholding are therefore needed.

In this paper we introduce such a method, based on a modification of the standard Random Forest approach [1].

Random Forests are a popular machine learning algorithm in medical imaging applications, but standard implementations of the algorithm are not optimal for medical imaging data, which is a) highly correlated at the patient level, and b) unbalanced, with the target class often having a prevalence of 1% or less. Segmentation forests avoid these problems by bootstrapping training data first at the patient level, and the second by using out-of-sample patients to empirically discover a threshold at which the Dice coefficient of the segmentation is maximized, avoiding the need for holding out training data to tune the classifier. Preliminary results of applying this technique to the ISLES acute stroke dataset are reported.

2 Method

Our segmentation algorithm uses pre-processing and texture and spatial features inspired by the features used in the BraTumIA brain tumour segmentation tool [2] and a previous pilot study on stroke segmentation [3]

2.1 Standardization and feature processing

Prior to model construction, features were extracted from the multimodal imaging volumes using the Insight Segmentation and Registration Toolkit, available from itk.org.

Before feature extraction, the structural image modalities (T2 and T1 contrast) are smoothed (using the GradientAnisotropicDiffusionImageFilter from ITK version 7.4.2) and a window filter is applied to the TMax map (with minimum value 0 and maximum value 100) to suppress abnormally high values. All feature maps were then rescaled with ITK version 7.4.2 to lie within the range [0,256]. The T1c image is coregistered to an atlas to allow extraction of atlas coordinates and the location of the mid-sagittal plane.

We then extract, for each voxel of the volume, and each image modality, a feature vector, consisting of the following features: - local texture features, extracted over both 3-by-3-by-3 and 5-by-5-by-5 voxel volumes - mean intensity - intensity variance, skewness and kurtosis - signal to noise ratio, entropy and energy - local intensity percentiles - local image gradient features (gradient magnitude computed using GradientMagnitudeRecursiveGaussianImageFilter from ITK version 7.4.2, with a sigma of 1.0) - point intensity of the gradient magnitude - mean of the gradient magnitude over 3-by-3-by-3 and 5-by-5-by-5 volumes - variance of the gradient magnitude over 3-by-3-by-3 and 5-by-5-by-5 volumes - a symmetry feature computed using a corresponding point on the contralateral side of the brain (found using the previously computed atlas coordinates): the difference between the voxel intensity and a smoothed intensity (computed using a SmoothGaussFilterType from ITK version 7.4.2) from the contralateral side.

Additional features were the unscaled image modalities, atlas coordinates, and an indication of whether the voxel is on the ispi- or contralesional (inferred by comparing the means of the scaled TMax on each side of the brain.)

2.2 The Segmentation Forest classifier

Random forests have been successfully used in numerous medical imaging applications, either alone or together with a conditional random field regularization. Random Forests is an example of bagging, in which a number of weak classification algorithms are trained, each on a random sample of the training data: their outputs are then averaged, yielding a better classification than each individual

classifier. In the case of Random Forests, the weak classifiers are decision trees. Each tree is built on a different bootstrap sample of the training data, and at each stage of building the tree only a subset of the features are available for classification: this prevents the trees from being too closely correlated, which would spoil the benefits of ensembling the weak learners.

Bagging is less helpful when, as is the case in medical imaging, training data is stratified into correlated clusters: the training examples extracted from a single patient's imaging data are closely correlated to one another, meaning that bagging fails to decorrelate the weak learners. In addition, the validity of out-of-bag measures of performance is diminished (since the out-of-bag samples are closely correlated to the training data, the performance on those samples will in general be much higher than on new data). To mitigate this, we introduce a classifier we call Segmentation Forests. The algorithm works by training a number of small Random forest classifiers (for example, with 50 trees), each one trained on the data given by a bootstrap sample of the training cases. The final classifier is then formed by averaging the output of each individual random forest, meaning that the final classifier is, in the end, still an ensemble of trees, as with an ordinary random forest.

Given a new case, the output of this classifier is then a score, between 0 and 1, for each voxel. To generate a segmentation from this map, we set a threshold on this map. Rather than choosing a threshold of 0.5, or calculating a theoretically correct threshold based on the relative incidence of the background and lesion classes, we instead derive an empirical threshold during training. We apply each small forest built on a bootstrap sample of the training cases to the patient cases not selected by the bootstrap sample. We calculate an estimate of the threshold which optimizes the Dice coefficient over these out-of-bag cases, and the average of these estimates is then used as the threshold for the final classifier.

3 Preliminary Results and Discussion

The segmentation forest classifier was implemented using the SpeedRF Random Forest implementation of the H2O machine learning package (Version 2.8.4.4), and the accompanying R package, both acquired from CRAN or via h2o.ai. This implementation of Random Forests is faster and more memory-efficient than the standard R implementation, allowing for the use of all data in the training sets, without downsampling of the background class.

The initial results were generated using a segmentation forest setup, with ten bootstrap samples of the training data, each of which consisted of 50 trees, with an `mtries` parameter of 30 and a maximum depth of 40.

We used the out-of-bag results to generate segmentations of each test case, with an average Dice score of 0.85 (sd 0.06). The mean average symmetric surface distance (ASSD) of our method was 1.42 (sd 1.01), and the mean Hausdorff distance was

30.71 (sd 18.91). Currently our method does not involve any post-processing to remove isolated outlier voxels: we could expect the Hausdorff distance in particular to benefit from such post-processing. Time to run the algorithm on a single case, including feature extraction, was six minutes.

Further work to improve the classifier will consist in optimizing, using the out-of-bag error, the parameters of the individual random forest models, and incorporating post-processing to improve the Hausdorff distance.

References

- [1] L. Breiman, “Random forests,” *Mach. Learn.*, pp. 5–32, 2001.
- [2] N. Porz, S. Bauer, A. Pica, P. Schucht, J. Beck, R. K. Verma, J. Slotboom, M. Reyes, and R. Wiest, “Multi-Modal Glioblastoma Segmentation: Man versus Machine,” *PLoS One*, vol. 9, no. 5, p. e96873, May 2014.
- [3] S. Bauer, P. P. Gratz, J. Gralla, M. Reyes, and R. Wiest, “Towards automatic MRI volumetry for treatment selection in acute ischemic stroke patients.” *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2014, pp. 1521–4, Aug. 2014.

Lesion Segmentation of the Penumbra in Acute Stroke in the MICCAI 2015 ISLES Challenge

Elias Kellner¹, Karl Egger², Maddalena Strumia^{1,3}, Valerij G Kiselev¹, Horst Urbach²
and Marco Reisert¹

¹ Department of Radiology, Medical Physics, University Medical Center Freiburg, Germany

² Department of Neuroradiology, University Medical Center Freiburg, Germany

³ German Cancer Research Center (DKFZ) and the German Cancer Consortium (DKTK),
Heidelberg, Germany

Abstract. The proposed method is based on a left-to-right comparison of the brain hemispheres to (i) restrict the segmentation only to the affected one and (ii) to pre-select the potential lesion by a threshold-free comparison with the contralateral, healthy side as a reference.

1 Method

In almost all cases of acute embolic anterior circulation stroke only one hemisphere is affected (corresponding to the ground truth segmentations of the training dataset). We exploit this feature to (i) restrict the segmentation to only the affected hemisphere and (ii) to pre-select the potential lesion by comparing local histograms of the affected side with the contralateral, healthy counterpart used as reference values. The algorithm is illustrated in Fig. 1 below.

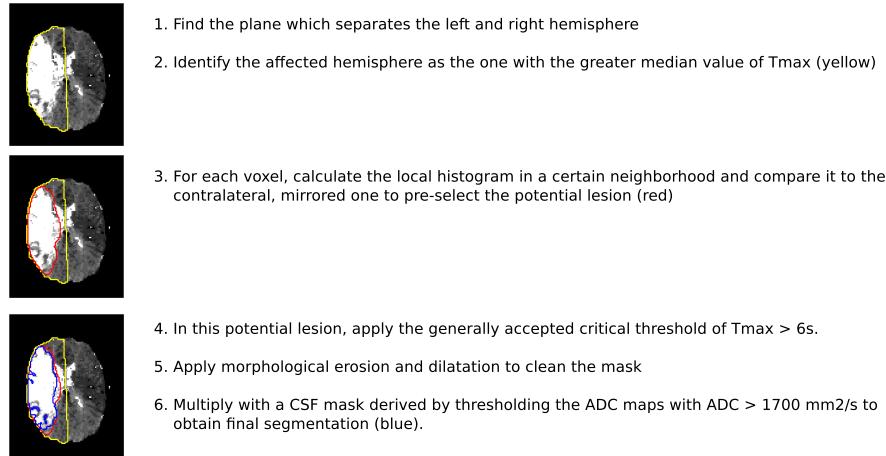


Fig. 1. Illustration and description of the proposed algorithm.

Due to the perfusion-diffusion-mismatch conceptm, our approach is based on the evaluation of just the Tmax and ADC-values. The important steps of the proposed method are described below in more detail.

1.1 Separation of the Hemispheres

For the separation of the hemispheres, we use a simple, but fast and robust algorithm. Since the brain extraction was already performed in the training datasets, the brain mask can easily be generated by thresholding the Tmax maps with $\text{Tmax} > 6\text{s}$. A additional, mirrored mask is generated as by flipping the original mask on the L-R direction. The separation plane of the brain hemispheres can be found by the optimal overlap of the mirrored and the original masks. To achieve this, a global search on the grid of all possible rotations and translations between the to masks is performed, and the correlation value is computed. The highest correlation indicates the best transformation, from which the separation plane can be derived.

1.2 Potential Lesion Segmentation

The affected hemisphere is identified as the one with the highest median value of Tmax. For each voxel at position \mathbf{x} , a normalized, regional histogram $H(\mathbf{x}, t_i)$ is calculated in a $20 \times 20 \times 12 \text{ mm}^3$ neighborhood with a bin-width of $t_{i+1} - t_i = 1.5 \text{ s}$. The difference to the corresponding contralateral histogram $\tilde{H}(\mathbf{x}, t_i)$ taken from the mirrored part of the brain is calculated via

$$D(\mathbf{x}) = \frac{1}{2} \sum_i |H(\mathbf{x}, t_i) - \tilde{H}(\mathbf{x}, t_i)| \quad (1)$$

This resulting map of histogram differences is thresholded by 0.5 to find the regions with unusual Tmax values. The distance map does not show whether the unusual value is higher or lower than the one of the corresponding healthy side. Therefore, only voxels with a Tmax median value greater than the one of the healthy side are selected. Further, to remove potentially noise voxels, only those with a sufficient number of non-zero values in their neighborhood – corresponding to 1/2 of the neighborhood volume – are considered. To clean the mask, we apply morphological erosion with a kernel of $6 \times 6 \times 4 \text{ mm}^3$, then we omit connected components with a volume smaller than 20 mL, and finally apply morphological dilation with the same kernel used for erosion.

1.3 Thresholded Lesion Segmentation

The previous step already provides us with a good segmentation of the potentially abnormal region. To select only the penumbra region with critical perfusion deficit, an additional threshold has to be applied. We applied a thresholding with the generally accepted value of $\text{Tmax} > 6\text{s}$, as performed for the ground truth segmentation of the MICCAI ISLES challenge. To clean the mask, we again apply the same morphological operation as described above.

1.4 CSF removal

The ADC maps are thresholded with $\text{ADC} > 1700 \text{ mm}^2/\text{s}$ to define a simple, but effective CSF mask. This mask is multiplied with the Tmax segmentation.

2 Results and Discussion

The following discussion refers to the 30 examples of the training dataset of the MICCAI ISLES challenge.

In all examples, the plane separating the left and right hemisphere of the brain could always be determined. With this, the affected hemisphere was always identified correctly.

The generally applied hard threshold of $\text{Tmax} > 6\text{s}$, but not free from limitations, as the critical perfusion delay might depend on individual factors such as patient age and previous diseases, and measurement noise.

The proposed method can help overcome this problems, as it includes a pre-selection of the potential lesion based on a comparison with the contralateral, healthy regions of the brain. This can be observed in e.g. example 10, Fig. 2, where the brain extraction did not remove some regions outside of the brain where the Tmax map show an enhancement. This is falsely included in the ground truth, but excluded in our approach, as it appears equally on both sides of the brain.

With this pre-selection, the final segmentation could be restricted to a reasonable region: on average, the ground truth was included by a rate of 85% with a false positive rate of 67%. With a final thresholding by the standard value of $\text{Tmax} \geq 6\text{s}$, followed by morphological operations, the false positives could be reduced to 21% by a false negative rate on an equal order of 20%. This is satisfiable, as for the diffusion-perfusion mismatch ratio, not the precise shape of the segmentation is crucial, but the volume.

We limited the evaluation to the Tmax and ADC maps only in order to keep the method as practical as possible. Further, in many cases stroke patients are unsettled and move heavily, such that the image quality does not allow for a proper coregistration of different modalities on the level of the voxel-size. Finally, the computational costs of the proposed method are very low, which is crucial in acute stroke. The evaluation is performed in a couple of seconds on a standard desktop PC.

3 Conclusions

We presented a simple and fast algorithm for penumbra segmentation in acute stroke. The key features of the method are the low computational costs, the restriction to the affected hemisphere only, and the threshold-free pre-selection of abnormal Tmax values.

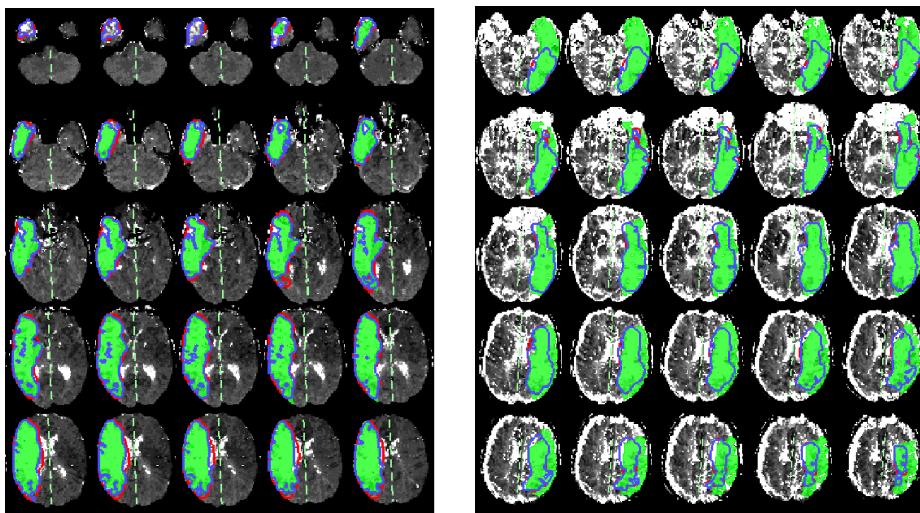


Fig. 2. Left: Training set 16. This example shows a good agreement with ground truth (green). The pre-selection of the potential lesion (red) defines an accurate restriction to the final segmentation (blue). **Right: Training set 10.** The significant discrepancy to the ground truth can partly be explained by the apparently insufficient brain extraction, which retains vessels outside of the skull with a strong Tmax enhancement. This might be falsely included in the ground truth, but excluded in the proposed method, as the enhancement is also present in the contralateral, healthy reference regions. The frontal part in the lower slices was likewise excluded in our segmentation, since the contralateral region also shows a significant enhancement.

Random forests for acute stroke penumbra estimation

Oskar Maier^{1,2}, Matthias Wilms¹, and Heinz Handels¹

¹ Institute of Medical Informatics, Universität zu Lübeck

² Graduate School for Computing in Medicine and Life Sciences, Universität zu Lübeck

maier@imi.uni-luebeck.de

Abstract. Ischemic stroke treatment decisions are time-critical and depend largely on the potentially salvageable tissue. This rises the need for accurate, reproducible and reliable segmentation of acute ischemic stroke lesions from brain MR scans. This article details a contribution to the Acute Stroke Penumbra EStimation (SPES) sub-task of the Ischemic Stroke Lesion Segmentations Challenge (ISLES), organized in conjunction with the MICCAI 2015. The proposed method bases on previous works, which showed the approach to handle the tasks well and to be applicable to potentially flawed data acquired in clinical routine. The method is described in detail and all chosen parameter values are disclosed. Preliminary results on the training data places the approach among the highest ranking contributions.

Keywords: acute ischemic stroke, lesion segmentation, penumbra estimation, magnetic resonance imaging, brain MRI, random forest, RDF

1 Introduction

Ischemic stroke is caused by an obstruction of the blood supply to the brain and the subsequent death of brain tissue. Its diagnosis often involves the acquisition of brain magnetic resonance (MR) scans to assess the strokes presence, location, extent, evolution and other factors. If diagnosed early, part of the under-perfused tissue could still be salvaged by re-establishing the blood flow. Since the available treatment options are not risk-free and can e.g. lead to inter-cranial bleeding, the decision has to be made individually, depending on the potential gain and under great time restriction. An automated method to distinguish the already necrotic from the potentially salvageable tissue (furthermore termed penumbra, although this term is disputed) would be highly beneficial for the clinical routine and reduce incorrect decisions. The ISLES 2015 challenge offers the first platform for researchers to compare their methods directly and fair. Our contribution draws its base from a previously published method targeted towards sub-acute ischemic stroke lesions [3], which showed good results. It is based on carefully selected features extracted from the MR sequences and used to train a random forest (RF).

2 Method

The challenge's training data consists of multi-spectral (T1c, T2, DWI, CBF, CBV, TTP, Tmax) scans of 30 patients displaying acute ischemic stroke. For training the manual segmentations of a single expert rater has been provided. More details on the data can be found on www.isles-challenge.org.

2.1 Pre-processing

The image data is provided with a 1 mm isotropic resolution, already co-registered and skull-stripped. Nevertheless, the training cases of the challenge display high intensity differences, a normal occurrence for MRI, where intensity ranges are not standardized. With a learning based intensity standardization method implemented in MedPy [2] and based on [4] we harmonize each T1c, T2 and DWI sequences intensity profile without a prior bias-correction step. The Tmax sequence, which is considered the most discriminative for penumbra estimation, is cut-off at an upper value of 100, which corresponds to 10 s.

2.2 Forest classifier

We employ the RF classifier implemented in [5], which is similar to the propositions made by [1]. The classification of brain lesions in MRI is a complex task with high levels of noise [3], hence a sufficiently large number of trees must be trained.

2.3 Features

The primary distinction criteria for identifying pathological tissue of stroke lesions is the MR intensity in the different sequences. The bulk of our voxel-wise features therefore bases on the intensity values.

intensity First feature is the voxel's intensity value.

gaussian Due to the often low signal-to-noise ratio in MR scans and intensity inhomogeneities of the tissue types, we furthermore regard each voxel's value after a smoothing of the volume with a 3D Gaussian kernel at three sizes: $\sigma = 3, 5, 7\text{ mm}$.

hemispheric difference Gliomas mostly affect a single hemisphere, therefore we extract the hemispheric difference (in intensities) after a Gaussian smoothing of $\sigma = 1, 3, 5\text{ mm}$ to account for noise. For simplicity, the central line of the sagittal view is taken as sufficiently close approximation of the sagittal midline.

center distance Finally, we extract the distance to the image center (assumed here to coincide roughly with the brain’s center of mass) in *mm* as final feature. Note that this is not intensity based, but rather discloses each voxel’s rough location inside the brain.

All features are extracted from each of the MR sequence, hence in total we obtain 52 values per multi-spectral voxel. Note that all of these features are implemented in MedPy [2].

2.4 Post-processing

After thresholding the a-posteriori class probability maps for a crisp segmentation, all but the largest connected component are removed. No other post-processing steps are employed.

3 Experiments

3.1 Training choices and parameter values

For training our RF, we sample 1,000,000 voxels randomly from all training cases. The ratios between classes in each case are kept intact (i.e. stroke class samples will be highly under-represented). A total of 100 trees are trained for the forest. As split criteria the Gini impurity is employed, a maximum of $\sqrt{52}$ features is considered at each node. No growth restrictions are imposed. The a-posteriori class probabilities produced by the forest are thresholded at a value of 0.35 to counter under-segmentation.

3.2 Preliminary results

Online evaluation is provided with the Dice’s coefficient (DC), the average symmetric surface distance (ASSD) and the Hausdorff distance (HD) as quality metrics. Using a leave-one-out evaluation scheme, we have obtained the scores presented in Tab. 1.

Table 1. Mean evaluation results and standard deviation on 28 training cases. See the text for details on the abbreviations employed.

	DC	ASSD	HD
30 cases	0.83 ± 0.06	1.38 ± 0.66	23 ± 13

4 Discussion and conclusion

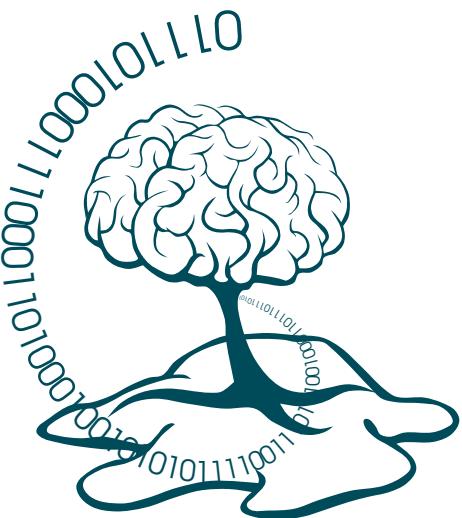
After the challenge date, the final results can be found on the challenge web-page www.isles-challenge.org. Preliminary results rank our method among the best performing. An advantage of our approach is its flexible design, that allows an application for a large number of brain lesion segmentation tasks (see e.g. [3]). Slightly adapted versions have been handed in to the sibbling challenges ISLES 2015: SISS and BRATS 2015.

By employing RFs, we have a powerful classifier at our hand that is robust against uninformative features, generalizes well and produces good results for a wide range of parameters. Mixing widely used with specially designed features, we can successfully learn to discriminate between the acute stroke area and healthy brain tissue.

On the downside, they suffer from the same drawbacks as all other machine learning based methods: The training set must be carefully chosen and types of cases not present in the training data can not be processed.

References

1. Criminisi, A., Shotton, J., Konukoglu, E.: Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning. *Foundations and Trends® in Computer Graphics and Vision* 7(2–3), 81–227 (2012)
2. Maier, O.: MedPy. <https://pypi.python.org/pypi/MedPy>, accessed: 2015-03-29
3. Maier, O., Wilms, M., et al.: Extra tree forests for sub-acute ischemic stroke lesion segmentation in MR sequences. *Journal of Neuroscience Methods* 240(0), 89–100 (2015)
4. Nyul, L., Udupa, J., Zhang, X.: New variants of a method of mri scale standardization. *Medical Imaging, IEEE Transactions on* 19(2), 143–150 (Feb 2000)
5. Pedregosa, F., Varoquaux, G., et al.: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12, 2825–2830 (2011)



Oskar Maier

Institute of Medical Informatics
Universität zu Lübeck
Germany

Mauricio Reyes

Institute for Surgical Technology & Biomechanics
University of Bern
Switzerland

Björn Menze

Computer Science
TU Munich
Germany

www.isles-challenge.org
Ischemic Stroke Lesion Segmentation



UNIVERSITÄT ZU LÜBECK



Technische Universität München



UNIVERSITÄT
BERN

