MIALab Project HS2020 - Image normalization has an important influence on the segmentation

Thomas Buchegger *University of Bern,* Carolina Duran *University of Bern,* Stefan Weber, *University of Bern*

***Abstract*—Here is the abstract.**

January 03, 2021

I. Introduction

This is a very important citation test. [1] Postprocesses of clinical diagnosis images for treatment planning often include manual segmentation of brain regions. To segment and label the brain structures in a large dataset is complicated and time-consuming by manual operator-guided segmentation. Furthermore, it is affected by user variability and prone to limiting the standardisation. This paper will depend upon automated segmentation that segments and labels five different brain regions.

To conclude we hypothesis that normalization has an important influence in the segmentation process of brain regions. We present in this paper the acquired results applying multiple normalization methods and comparing them for every segmented brain region. Furthermore, we analyse the results and conclude our findings.

Our hypothesis: "Image normalization has an important influence on the segmentation". Aims of the introduction:

- **To demonstrate importance/impact need**
- **To demonstrate novelty**
- **To justify the hypothesis / aims / investigated technology**
- **To establish expectations/scope of the report**

Some questions which could be answered: Currently, z score normalization is implemented.

- **Are there more powerful normalization methods?**
- **Is normalization really needed?**
- **Is the provided data already normalized in some way?**
- **Can you also unnormalize data to show a negative effect?**

For the research:

- **Understand the problem**
- **What have others already done / is there already a solution?**
- **Can I apply a solution to another problem to my problem?**
- **What lessons can I learn from others work?**
- **What deficiencies exist in others work?**

1) **Demonstrate importance**
   a) **Define the problem**
   b) **Explain the criticality, impact of the problem**
2) **Demonstrate Novelty**
   a) **Explain what is the state of the art / current practice**
   b) **Explain what preliminary / related work has been done towards solving the problem by you and others**
   c) **Explain what deficiencies / problems still exist (specifically the one that you will try to address)**
3) **Present and Justify the Hypothesis / aim / objective**
   a) **State Hypothesis /aim/ objective**
   b) **Describe evidence supporting hypothesis**
4) **Establish expectations of the report**
   a) **Describe scope of the presented work**
   b) **Present a summary of remainder of the report**

II. Materials and Methods

MIA pipeline in general (short overview), your experiment in detail

T. Buchegger, C. Duran and S. Weber are students from the University of Berne in Switzerland.
E-mail of T. Buchegger: thomas.buchegger@students.unibe.ch
E-mail of C. Duran: carolina.duran@students.unibe.ch
E-mail of S. Weber: stefan.weber1@students.unibe.ch

A. Medical Image Analysis pipeline

The Medical Image Analysis (MIA) pipeline taught in the MIA Lab lectures follow the sequence of firstly perform *Registration* to T1- and T2-weighted images (T1w and T2w image), then *Pre-Process*-methods. Additionally *Feature Extraction* followed by the *Classification* of the images are performed. At the end *Post-Processing*-methods are performed. Eventually, the segmentation of one patient's data set is achieved. During *Registration* a floating image is transformed with an affine transformation, such that it is similar to a given reference image.

In this project no *Registration*, *Feature Extraction*, *Classification* nor *Post-Processing* has been applied to the images. The phase of the MIA pipeline analysed and applied to the images was *Pre-Processing*. The aim of *Pre-Processing* is to improve the image quality for the subsequent classification. It includes among others bias field correction, skull stripping, intensity normalization and/or histogram matching. A skull stripping, as well as several different normalization methods were applied.

B. Medical Background

The five brain regions segmented in this paper are the white and grey matter, hippocampus, amygdala and the thalamus. All regions are visible in figure 1.

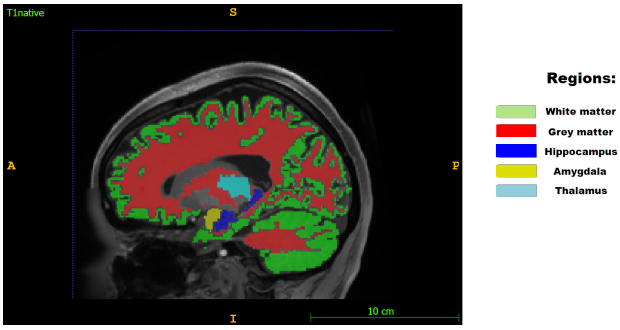


Fig. 1: Ground truth picture of all brain regions to be segmented in this project.

C. Data

The data used for this project was from the *Human Connectome Project* and has been provided by the Medical Image Analysis Lab team at the University of Bern. For anonymization the faces and ears of the skull have been blurred. Overall, the dataset consisted of 30 MRI patient images, out of which 20 were used for training and 10 for testing the model. From each patient the ground truth image, a T1w and a T2w image were available. The images were generated by a 3 Tesla MRI. The atlas used are the MNI152 standard-space T1-weighted average structural template images available from the *McConnell Brain Imaging Centre*[1] (BIC) and the *NeuroImaging & Surgical Technologies Lab*[2]. It corresponds to the MNI-ICBM atlas[2] and is derived from 152 structural images, averaged together after high-dimensional nonlineal registration into this MNI152 coordinate system. Each MRI file is of the size 118x118x217 pixels.

From MIA Lab slides: Details on data pre processing

- **T1w , T2w in native space**
- **Both bias field corrected**
- **Brain mask available**
- **Transformation to atlas space available**

D. Model

The model used was a Random Forest classifier with which different parameters with all used normalization techniques were combined. Which parameters lead to which results can be found in the results section.The following normalization techniques were used:

1) ZScore:

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

2) MinMax:

$$z = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{2}$$

3) stikN:

4) Whitestripe:

5) Fuzzy-C means: With *Fuzzy C-means* a mask for a specified tissue type given a T1w image and its brain mask is created. Then this tissue mask is used as input to the function. The tissue mask is then used to find an approximate mean of the tissue intensity in another target contrast and moves it to some standard value.

6) Histogram Matching:

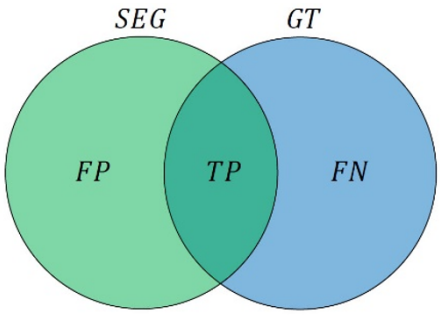
7) Gaussian Mixture Model:

E. Evaluation

For this project, the used normalization techniques were evaluated on the T1- and T2-weighted MRI data sets. For this evaluation the Dice Similarity Coefficient (DSC) as well as the Hausdorff distance was chosen. DSC returns as a quality metric how much two regions overlap and is, therefore, a useful metric for segmentation. The dice coefficient of two sets, as explained in figure 2, is a measure of their intersection scaled by their size. The result is in the range from 0 to 1, where 1 is a perfect segmentation.

[1]URL: http://www.bic.mni.mcgill.ca/ServicesAtlases/HomePage, Date: 23.12.2020.

[2]URL: http://nist.mni.mcgill.ca/?page_id=714, Date: 23.12.2020.



$$DICE(SEG, GT) = 2\frac{|SEG \cap GT|}{|SEG| + |GT|}$$

Fig. 2: Graphical representation of the Dice Similarity Coefficien. SEG stands for the achieved segmentation whereas GT means ground truth.

The Hausdorff distance measures how far two subsets of a metric space are from each other. So two sets are close if every point of either set is close to some point of the other set. Then the Hausdorff distance, explained in figure 3, is the longest distance from one of the two sets to the other set.

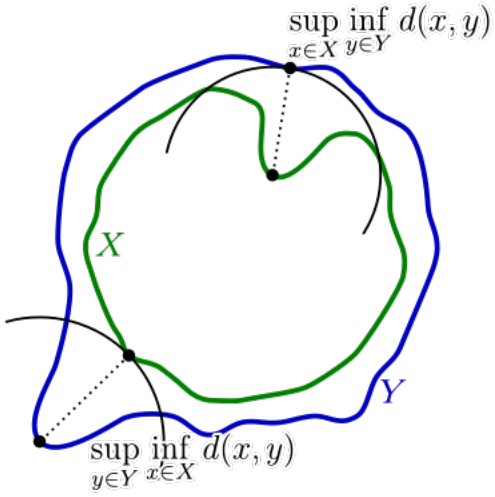


Fig. 3: Components of the calculation of the Hausdorff distance between the green line X and the blue line Y.

III. Results

In depth analysis of experiment related results

A. Data

B. Model

C. Evaluation

IV. Discussion

Aims of the Discussion part:

- **Highlight importance of your work (highlight novelty /impact etc**
- **To interpret your results in relation to your original problem**
- **To put your work into the context of existing work**
- **To present any limitations of the presented work**
- **To make future recommendations**
- **To provide a conclusion of the work**

1) **Importance of the work**

   a) **Summarise your results**
   b) **Reiterate the importance of the work (novelty , impact etc)**

2) **Interpretation of results**

   a) **Interpret your results focussing on the problem described in the introduction. What do the results mean for the described problem?**
   b) **Explain any unusual/important findings (be careful if not your original investigative subject)**

3) **Provide context**

   a) **Describe your results in relation to others and try to explain any discrepancies**
   b) **Emphasize how your results support or refute your hypotheses current thinking in the field. Were results as expected? If not why and what does this mean?**

4) **Limitations of your work**

   a) **Describe any limitations /deficiencies of your work and what impact they have on the findings**
   b) **Suggest possible future solutions**

V. Conclusion

- **Summarise your findings and relate your findings back to your hypothesis / aim / objective and to your problem.**
- **Based on your findings, suggest next steps towards solving your problem**

REFERENCES

[1] Y. H. Roohani and E. G. Kiss, "Improving Accuracy of Nuclei Segmentation by Reducing Histological Image Variability," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11039 LNCS, pp. 3–10, 2018.