Comparison of Machine Learning Techniques on MS Lesion Segmentation

Ahsen Feyza DOĞAN
Istanbul Medipol University
Department of Biomedical Engineering
Istanbul, Turkey
afdogan@st.medipol.edu.tr

Dilek GÖKSEL DURU
Istanbul Arel University
Department of Biomedical Engineering
Istanbul, Turkey
dilekgokselduru@arel.edu.tr

***Abstract*— Multiple sclerosis arises with conformational change in myelin sheath. Magnetic resonance imaging is frequently used in detection of MS. In this study, to figure out MS lesion, machine learning techniques, namely k means and support vector machine are used. K means is an unsupervised technique used to cluster data into k groups. Support vector machine is a supervised machine learning technique used as classifier. Since dataset does not contain label of images, labels are generated by pixel values adopted from original MR image. Classification results were achieved as 70.24% and 91.04% for k means and SVM respectively. According to the promising results, future research will focus on the automatization of this segmentation process via deep learning leading to medical decision support system.**

Keywords — Multiple Sclerosis; MRI; image segmentation; k means; support vector machine.

I. INTRODUCTION

Multiple sclerosis (MS) is a neurological disorder in the central nervous system occurring as a result of damages in the structure of axons and their myelin sheathes. MS lesions reflect changes in shape and location in time for a single control or between controls [1]. Therefore automatic segmentation and objective evaluation of MS lesions are of great importance. Myelin sheath have crucial importance on signal transmission between neurons. When conformational structure of myelin sheath changes, it may affect functioning of organs. For example, it effects vision system, or muscles where people face hindered capability of mobility [2].

Although there is no definite treatment of MS, early detection is important to decrease progression of disease and managing symptoms. There are several ways used to diagnose MS such as blood testing, evoked potential test, lumbar puncture etc. [3]. Manual methods lead to subjective decisions and diagnosis, which may be misleading so post processing of MS-MRI is important and preferable. In MRI, it is possible to observe how MS lesion disseminate through space and time. So MRI is the most sensitive method for determination of MS lesion [3]. Manual and subjective diagnosis of MS should be avoided in terms of misdiagnosis, unnecessary surgical interventions by different physicians with varying expertise [4]. So it is highly important to automatize this process by using image segmentation and emerging techniques [5].

978-1-7281-2420-9/19/$31.00 ©2019 IEEE

There are several types of image segmentation techniques depending on contrast, where one of the most common and popular technique is k means, which is an unsupervised technique clustering the data into k group [6]. Another promising segmentation technique in literature is support vector machine (SVM) [7,8], which is a supervised technique. However these segmentation approaches need image processing to simplify analysis of MR images, where mostly it is not easy to distinguish pathology like lesion etc. from raw MR images.

With advanced technology, diagnosis by using machine learning techniques is easier [5]. In this study, both of these supervised and unsupervised techniques, namely k means and support vector machine approaches are implemented to obtain more accurate result for MS lesion detection and to provide a comparative result in terms of supervised and unsupervised techniques. Here, in k means, clustering is used whereas thresholding and binarization are used in SVM for lesion segmentation. Aim of this study in long term is to create a decision support system for physicians by using image processing, supervised and unsupervised techniques. Future study will focus on completely automatic version of the decision support system by using deep learning techniques.

II. METHOD

A. Support Vector Machine

SVM is a supervised machine learning technique for classification. In this technique, the model tries to figure out the best hyperplane separating the data into varying classes. Wide margin has crucial importance for the accuracy. There are different application routines of SVM depending on dataset such as linear, quadratic, RBF etc. In this study, linear SVM is chosen because it was sufficient to classify the labels, where labelling the data was done for -1 and 1 in 2D space (1) (Fig. 1). The approach finds a hypersurface dividing the negative and positive samples (pixels, labels) with the largest possible margin on all sides of the hyperplane (Fig.1). Here, the classification problem results as either -1 or 1 as in equation 1. For each slice of the input dataset, each group of connected pixels are labeled and analyzed.

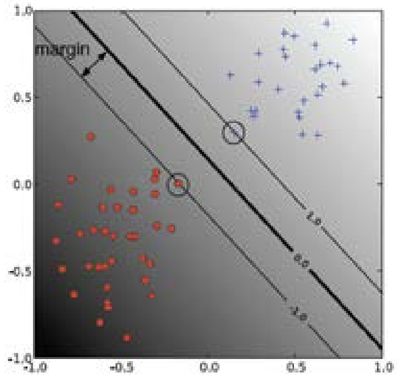

Fig.1. Linear separation of two classes in two dimension with support vector machine [1].

$$w_0{}^T x + b^0 = 1 \text{ or } w_0{}^T x + b^0 = -1 \quad (1)$$

In (1), w_0 represents weight vector. It finds the best hyperplane by minimizing cost. Location of data is represented by x and b is the bias, which is used to decrease distance between the predicted and the original data.

B. K Means

K means is an unsupervised technique. It clusters data into k groups depending on their similarities, where k cannot be equal to 1 and bigger than the number of sample in data. The main methodology and the segmentation framework of the study is represented in Fig. 2. In this study, k is adjusted to four. And their values are determined upon the histogram of the input image (Fig. 3) concerning their effect in the binary segmentation of the original MR image.

Figure 2 represents the workflow of this study generally. Firstly, data was read as an input and smoothened for further process. Then features are extracted for modified segmentation of MS lesion. After that, these features were classified. At this point pixel values have crucial importance, k means technique uses these pixel values directly whereas SVM uses label. Finally calculated segmented lesion compared to segmented lesion from original data are calculated for performance evaluation resulting a match of 70.24% for k means and 91.04% for SVM.

C. Database

The MS-MR examined in this study is open source MS dataset [9]. The two approaches, k means and SVM method, both have been applied in MATLAB environment on real brain MS-MR data (Fig. 4) to extract MS lesion depending on its pixel values. So the segmentation framework is verified and validated on original (known) input open source data.

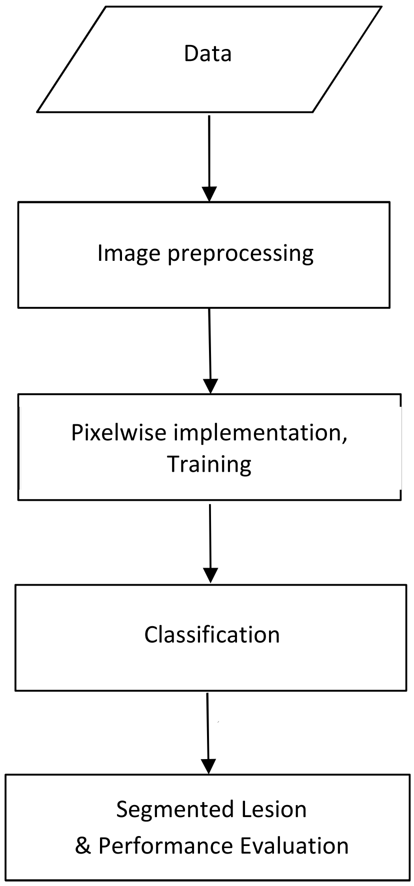

Fig. 2. MS lesions segmentation framework

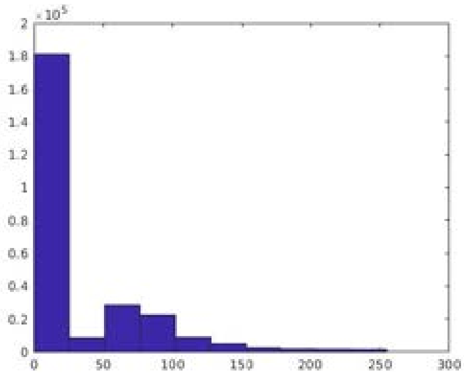

Fig. 3. Histogram of an exemplary original image

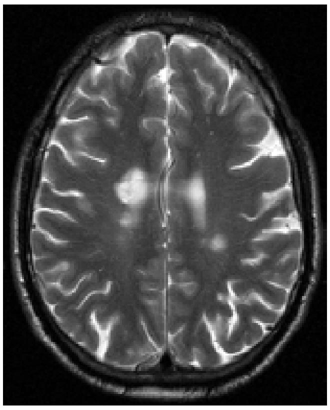

Fig. 4. Sample brain MR image with MS lesion

III. RESULTS

In this work, two machine learning techniques, namely k means and SVM are used to semi automatically detect the lesion in MR images in MS pathology. The results are promising where the findings of SVM with 91.04% accuracy is acceptable with respect to the current literature, where a detailed discussion will be given in next section.

The main steps of the methodology are interpreted in Fig. 2, where the feature extraction for k means is achieved by the values gained from the histogram (Fig. 3) of the input image (Fig. 4). And the segmented brain and the subtracted MS lesion reflects the successful k means results (Fig. 5).

Depending on the selection of the pixels upon histogram (Fig. 3), the segmentation varies. The resulting finding is a gray matter mask actually, where the lesion can be easily detected and extracted (Fig. 5). The subtracted MS lesion reflects a successful finding of k means implementation (Fig. 5).

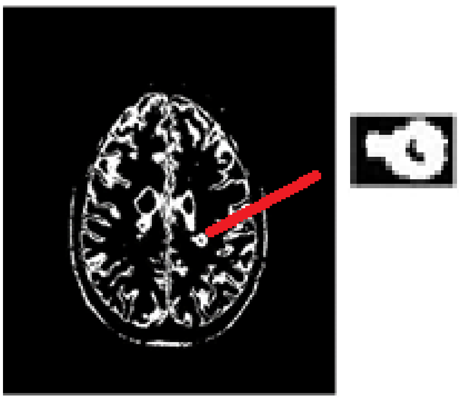

Fig. 5. Segmented brain and MS lesion (zoomed) by using k means

The k means algorithm is checked in every iteration for convergence. Figure 6 represents this behavior and successful implementation of the method.

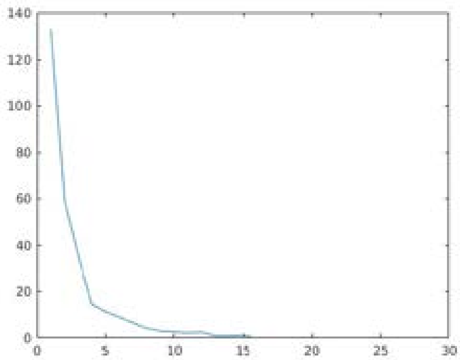

Fig.6. Error rate graph of k means across iteration

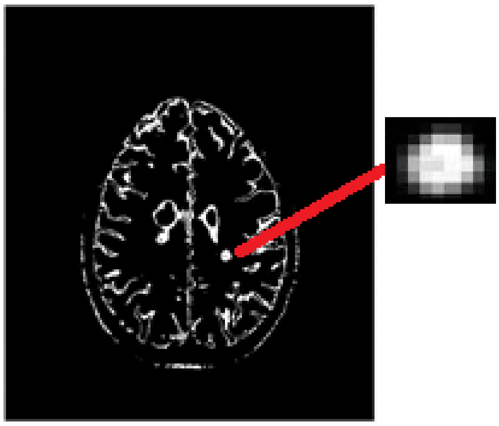

Fig. 7. Segmented MS lesion by using SVM

The SVM implementation also resulted successfully with an accuracy value of 91.04% compared to the original (known) lesion. As represented in Fig. 7, the MS lesion is clearly figured out in the pathological MR image.

IV. DISCUSSION & CONCLUSION

The aim of this study is to determine MS lesion in MRI by using image segmentation and machine learning techniques. To obtain more accurate results, and to compare supervised and unsupervised approaches in MS lesion detection, two different kinds of machine learning techniques were used. The routines are also applied to simulated data to address classification and to validate our in house delivered codes. Real data results are subject to this work, where the results are compared to the given and known input data. It is obviously seen that supervised learning gives better result for this study. Finally, 91.04% and 70.24% accuracies were obtained for support vector machine and k means implementations in MATLAB respectively.

For k means, pixel values are taken into consideration for clustering. Since there is no distinct rule about how to group pixels, it is highly important to interpret histogram graph. In

Figure 6, it is obviously seen that after 15 iterations, the error rate (loss) became zero, which addresses that the convergence is achieved successfully.

Support vector machine as interpreted in II, is a supervised technique where the dataset does not contain labels. For this reason, labels were prepared manually. There are two subcategories corresponding to pixel feature. Thereafter, depending on these subcategories binarization process was done. Finally after these processes lesion was detected easily and clearly as seen in Fig. 5 and 7. Comparison of the segmentation findings represented in Fig. 5 and 7, results that SVM is much more successful than k means for segmentation. In Fig. 5, segmented gray matter area is more than in Fig. 7. One reason for this is that k means relies on objective assignment, whereas SVM benefits from spatial coordinates of data. So SVM is more convenient for this study and accuracy values support this hypothesis. In literature, most of segmentation studies rely on k means in finding MS lesion. However, when classification accuracy obtained from SVM is being compared to the ones in literature in the field, then our results are acceptable and successful, where a study carried in Japan, researcher obtained 93.05% accuracy [10], and in another collaborative study of India and USA, researchers obtained 96.17% accuracy [11]. To conclude, the findings are acceptable and promising, and it is possible to obtain more accurate results and automatize the process with deep learning techniques [5]. Future work relies in the implementation of deep learning techniques for detection of MS lesions.

REFERENCES

[1] Abdullah, B.A. "Segmentation of Multiple Sclerosis Lesions in Brain MRI" (2012). Open Access Dissertations University of Miami, 711. https://scholarlyrepository. miami. edu/oa_dissertations/

[2] Guttmann CR, Kikinis R, Anderson MC, Jakab M, Warfield SK, Killiany RJ, et al. "Quantitative follow-up of patients with multiple sclerosis using MRI: reproducibility", *J. Magn. Reson. Imaging*, Apr;9(4):509-18, 1999.

[3] Traboulsee AL, Li DK. "The role of MRI in the diagnosis of multiple sclerosis", *Advances in Neurology*, 98, 125-46, 2006.

[4] Kazancli, E., Prchkovska, V., Rodrigues, P., Villoslada, P., Igual L, "Multiple Sclerosis Lesion Segmentation using Improved Convolutional Neural Networks," *Proc. 13th International Joint Conf. on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2018)*, vol. 4: VISAPP, pp 260-269, 2018.

[5] D. Göksel Duru and A.D. Duru, "MR-MS Image Classification based on Convolutional Neural Networks," 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), p.1-4, 2019. DOI: 10.1109/EBBT.2019.8741752

[6] E. Işıkçı and D. Göksel Duru, "Extraction of Multiple Sclerosis Lesions in MR Images Based on Active Contours," Medical Technologies National Conference (TIPTEKNO), pp. 122-125, 2015.

[7] K. Machhale, H. B. Nandpuru, V. Kapur, L. Kosta, "MRI brain cancer classification using hybrid classifier (SVM-KNN)," *International Conference on Industrial Instrumentation and Control*, p.1-4, 2015.

[8] B. Magnin, L. Mesrob, S. Kinkingnhun, M. Plgrini-Issac, O. Colliot, M. Sarazin, B. Dubois, S. Lehricy, H. Benali, "Support vector machine-based classification of Alzheimer's disease from whole-brain anatomical MRI," *Neuroradiology*, 51, 73–83, 2009.

[9] C.P. Loizou, V. Murray, M.S. Pattichis, I. Seimenis, M. Pantziaris and C.S. Pattichis, "Multi-scale amplitude modulation-frequency modulation (AM-FM) texture analysis of multiple sclerosis in brain MRI images," *IEEE Trans. Inform. Tech. Biomed.*, vol. 15, no. 1, pp. 119-129, 2011.

[10] D. Yamamoto, H. Arimura, S. Kakeda, T. Magome, Y. Yamashita, et al., "Computer-aided detection of MS lesions in brain magnetic resonance images", *Comput. Med. Imaging Graphics*,34(5), p.404-13, 2010.

[11] H. Selvaraj, S. Thamarai Selvi, D. Selvathi, L. Gewali, "Brain MRI Slices Classification Using Least Squares Support Vector Machine", International Journal of Intelligent Computing in Medical Sciences & Image Processing, volume 1, Issue 1, 2007.