Comparison of the WEKA and SVM-light based on support vector machine in classifying Alzheimer’s disease using structural features from brain MR imaging

K Tantiwetchayanon¹, Y Vichianin¹, T Ekjeen¹, K Srungboonmee², C Ngamsombat², O Chawalparit³, and Alzheimer’s Disease Neuroimaging Initiative*

¹Department of Radiological Technology, Faculty of Medical Technology, Mahidol University, Thailand.
²Center of Data Mining and Biomedical Informatics, Faculty of Medical Technology, Mahidol University, Thailand.
³Department of Radiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand.

*Data used in preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report.

E-mail: khajonsak.tantiwet@gmail.com

Abstract. The aim was to compare the WEKA and SVM-light software based on support vector machine (SVM) algorithm using features from brain T1-weighted MRI for differentiating AD patients and normal elderly subjects. The FreeSurfer software was used to extract cerebral volumes and thicknesses from T1-weighted brain MRI (100 AD patients and 100 normal elderly subjects). Seven structures were selected based on literature reviews consisting of hippocampus and amygdala volume, entorhinal cortex thickness of both hemispheres, and total gray matter volume. Relative volume of hippocampus, amygdala, and total gray matter were normalized by total intracranial volume (TIV). Fifteen combinations of seven structures were applied as input features to WEKA and SVM-light. The receiver operating characteristic (ROC) analysis and area under the curve (AUC) were used to evaluate the classification performance. The combination of hippocampus relative volume and entorhinal cortex thickness provided the highest classification performance and the AUC values were 0.913 and 0.918 for WEKA and SVM-light, respectively. There was no statistically difference of the AUC values (p-value > 0.05) between two software using the same input features. In conclusion, there was no statistically difference between the use of WEKA and SVM-light software for differentiating AD patients and normal elderly subjects.

1. Introduction

Alzheimer’s disease (AD) is the most common irreversible neurodegenerative type of dementia [1]. Currently, there is no treatment for this disease. Early detection such as morphological changes of brain structures, i.e. the gray matter atrophy of subcortical volume and thickness, plays an important role. The brain atrophy of AD begins in the medial temporal lobe, specifically in hippocampus and entorhinal cortex, and progresses to other lobes including temporal, parietal, and frontal area in the severe stage of AD [1,2].

Neuroimaging is recommended for early detection of AD especially the use of the magnetic resonance imaging (MRI). MRI is a noninvasive tool which has been used to rule out other causes of dementia. The structural MRI provides information of brain atrophy which is closely associated with the pathological features of AD [1]. It is time consuming process for specialists to manually localize the
brain features from MR images in order to evaluate the disease. To reduce this burden, an automatic brain extraction and classification algorithm has been introduced [3].

The machine learning has been widely applied in medical field such as in the computer-aided diagnosis. It is a pattern recognition technique that the classifier is trained using the labeled features data and provides the decision output based on the learning criteria [3]. The support vector machine (SVM) has been utilized as a powerful algorithm to classify normal subjects and AD patients grounded upon the different brain features using mainly SVM-light and WEKA software [4-11]. Jongkreangkrai et al. [4] had applied the SVM-light software to classify the T1-weighted brain MRI of AD patients together with normal subjects. Similar to Yamashita et al. [5], SVM-light software has been performed with the combination of functional and morphological brain MR image features for differentiating AD patients and clinically normal subjects. The WEKA software with sequential minimal optimization (SMO) technique was applied on different brain feature models [10,11], such as multiple indices from diffusion tensor imaging in order to classify healthy and mild cognitive impairment (MCI) subjects [11].

The SVM-light software [12] is an open-source C-based implementation of SVM with fast optimization algorithm. The software requires some C language programming skill to perform SVM algorithm via command line interface (CLI). On the contrary, the WEKA software [13] provides easy-to-use Weka Knowledge Explorer graphical user interface (GUI). The WEKA GUI software enables beginners to easily perform data preprocessing, clustering, classification, regression, visualization, and feature selection. Therefore, the purpose of this study was to compare the performance of the SVM algorithms from the two software, SVM-light and WEKA’s SMO software, by comparing the classification results from brain T1-weighted MRI features of the normal elderly subjects and AD patients.

2. Subjects and MRI data
The MRI data used in this study were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI is to investigate the combinations of serial MRI, positron emission tomography (PET), other biological markers, clinical data and neuropsychological assessment for seeking the progression of MCI and early AD patients.

T1-weighted brain MRI data with the scanning parameters using series of magnetization prepared rapid gradient echo (MP-RAGE) with 1.2 mm slice thickness of 1.5 Tesla (65 for AD patients and 60 for normal elderly controls) and 3 Tesla (35 for AD patients and 40 normal elderly controls) were collected based on screening period of MRI examination. 100 T1-weighted brain MRI of AD patients (48 males and 52 females, mean age 75.00 ± 7.34, age range 55-90 years) and 100 T1-weighted brain MRI of normal elderly subjects (51 males and 49 females, mean age 75.69 ± 5.68, age range 55-90 years) were used for a brain feature extraction process.

3. Method
3.1. Experimental design
All T1-weighted brain MRI series of normal elderly subjects and AD patients were extracted and processed using the FreeSurfer software. The brain features suggested by prior literatures includes hippocampus [4,7,9,14] and amygdala volume [4,7], entorhinal cortex thickness [4,14] of both brain hemispheres, and total gray matter volume [6-8]. The relative volume of hippocampus, amygdala, and total gray matter volumes were calculated in order to compensate the variation in subjects’ head size [15]. Then, the relative volume of hippocampus, amygdala and total gray matter, and entorhinal cortex thickness were utilized to develop fifteen input feature models for further used in the classification process using the SVM algorithm. Finally, the ROC analysis was performed and the AUC values were compared based on DeLong et al methods [16] with 95% confidence interval (p-value < 0.05). The process was summarized in the figure 1.
3.2. Feature extraction of brain MRI by using FreeSurfer software

The MRI raw data were transformed into NIfTI format using dcm2nii software obtained from www.nitrc.org. Then, cortical reconstruction and volumetric segmentation was performed using the FreeSurfer image analysis suite version 6.0 with "recon-all" command [17]. The software is documented and freely available for download online at http://surfer.nmr.mgh.harvard.edu/.

Since the brain volume measurements have been found to be highly correlated with head size while the thickness measurements are not [15], the relative volume of hippocampus, amygdala, and total gray matter were calculated using the total intracranial volume (TIV) as equation (1).

\[
\text{The relative volume} = \frac{\text{volume}}{\text{TIV}}
\]

The t-test with 95% confidence interval and Mann-Whitney U test were performed to investigate the difference of seven structures including hippocampus and amygdala volumes, entorhinal cortex thickness of both brain hemispheres and total gray matter volume between the two groups.

3.3. Fifteen models generation

Seven cerebral structures from feature extraction including relative volumes of hippocampus, amygdala, entorhinal cortex thickness of both brain hemispheres, and total gray matter volume were selected. Then, fifteen combinations were generated as feature models for classification step as shown in Table 1. Then, the fifteen features are normalized into the range 0-1 following equation (2).

\[
\text{Normalized feature} = \frac{x_i - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}
\]

where \(x_i\), \(x_{\text{min}}\), and \(x_{\text{max}}\) are the normalized feature \(i\), the lowest and highest value of feature, respectively.
Table 1. The feature models generated from brain structures obtained from FreeSurfer software.

| Feature models | Hippocampus relative volume | Amygdala relative volume | Entorhinal cortex thickness | Total gray matter relative volume |
|----------------|-----------------------------|--------------------------|-----------------------------|----------------------------------|
|                | Left | Right | Left | Right | Left | Right | Left | Right |
| Hip [4,7,9,14] | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     |
| Amyg [4,7]     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     |
| En [4,14]      | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     |
| GM [6-8]       | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     |
| HipAmyg [4]    | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     |
| HipEn          | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     |
| HipGM          | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     |
| AmygEn         | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     |
| AmygGM         | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     |
| EnGM           | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     |
| HipAmygEn [4]  | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     |
| HipAmygGM      | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     |
| HipEnGM        | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     |
| AmygEnGM       | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     |
| HipAmygEnGM    | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     | ✓    | ✓     |

3.4. Classification by using SVM algorithm

The SVM with radial basis function (RBF) was trained with labeled class dataset to obtain the classification pattern. Then, the model was used to predict a new dataset by putting them into a possible class based on leave-one-out cross validation. To employ SVM algorithm, SVM-light and WEKA were applied. The cost function of optimal hyperplane (C) and gamma of RBF were varied to obtain the highest classification performance using grid search technique with the values of $C = \{0.001, 0.01, 0.1, 1, 10, 100, 1000\}$ and gamma $= \{0.001, 0.01, 0.1, 1, 10, 100, 1000\}$, respectively [18].

3.4.1. SVM-light [12]. SVM algorithm with RBF in SVM-light software was applied to each feature models utilizing the leave one out cross-validation suggested by Jongkreangkrai et al. [4].

3.4.2. WEKA [13]. The minimal optimization algorithm (SMO) with RBF in WEKA software was used for training the SVM model. The cross-validation fold was set equal to the sample size (n=200) in order to perform the leave-one-out cross-validation.

3.5. ROC analysis and AUC values calculation using ROCKIT tool

The ROC analysis was used to determine the performance of SVM classification. For the SVM-light, the AUC values based on nonparametric statistics was calculated using ROCKIT, the free software package for ROC analysis developed by the University of Chicago (available online at http://metz-roc.uchicago.edu/MetzROC). For WEKA software, the ROC analysis and AUC value were included in the classifier results.

3.6. AUC values comparison using MedCalc statistical software

The AUC values were compared based on DeLong et al methods [16] with 95% confidence interval (p-value < 0.05). The calculation of the standard error of the AUC and of the difference between two AUC values was done using MedCalc for Windows, version 18.5 (MedCalc Software, Ostend, Belgium).

4. Results

There was a statistical difference in hippocampus, amygdala and total gray matter relative volumes, and entorhinal cortex thicknesses between normal elderly subjects and AD patients.
The highest AUC value was obtained from HipEn for both software (0.913 for WEKA and 0.918 for SVM-light respectively). Moreover, HipAmygEn and HipAmygEnGM also provided high AUC values (greater than 0.9) for both software. Table 2 indicates the AUC values and optimal parameters for all input features and p-value of the statistical test of AUC values from both software. There was no statistical difference of AUC values between WEKA’s SMO and SVM-light when using the same feature input (p-value > 0.05).

Table 2. AUC values from two software and the statistical different of AUC value using similar features models between two software.

| Feature models   | WEKA         | SVM-light    | p-value |
|------------------|--------------|--------------|---------|
|                  | C  | Gamma | AUC | C  | Gamma | AUC |         |
| Hip              | 100 | 0.01  | 0.886 | 1000 | 0.1  | 0.892 | 0.867 |
| Amyg            | 0.01 | 10 | 0.855 | 100 | 1 | 0.862 | 0.872 |
| En             | 1  | 1  | 0.864 | 10 | 0.1  | 0.870 | 0.882 |
| GM             | 100 | 0.1  | 0.743 | 0.1 | 10 | 0.749 | 0.842 |
| HipAmyg         | 1  | 1 | 0.891 | 1 | 1 | 0.897 | 0.862 |
| HipEn          | 100 | 0.1  | 0.913 | 100 | 0.1  | 0.918 | 0.864 |
| HipGM          | 100 | 0.01 | 0.884 | 100 | 0.01 | 0.888 | 0.902 |
| AmygEn         | 0.1 | 1 | 0.889 | 0.1 | 1 | 0.898 | 0.785 |
| AmygGM         | 10 | 0.01 | 0.855 | 0.1 | 10 | 0.863 | 0.846 |
| EnGM          | 1000 | 0.01 | 0.875 | 100 | 0.1 | 0.880 | 0.880 |
| HipAmygEn      | 0.1 | 1 | 0.909 | 1000 | 0.1 | 0.907 | 0.862 |
| HipAmygGM      | 10 | 0.1 | 0.892 | 10 | 0.1 | 0.895 | 0.918 |
| HipEnGM       | 1000 | 0.01 | 0.908 | 10 | 0.1 | 0.895 | 0.700 |
| AmygEnGM       | 0.01 | 1 | 0.893 | 0.1 | 1 | 0.902 | 0.781 |
| HipAmygEnGM    | 0.1 | 1 | 0.909 | 0.1 | 1 | 0.915 | 0.835 |

5. Discussion
This study showed that the performance of SVM classification of normal elderly subjects and AD patients, with features from T1-weighted MRI, was not statistically different between WEKA’s SMO and SVM-light (p-value > 0.05), while most of AUC values are slightly higher in SVM-light.

The combination of the features yielded better accuracy (higher AUC) compared with those using only of single features [4]. Long et al. study [7] predicted Alzheimer’s disease using quantification of MRI deformation based on SVM algorithm. Their result showed that the AUC value of classification between AD patients and normal elderly controls using total gray matter, hippocampus and amygdala are 0.995 (accuracy = 96.5%), 0.989 (accuracy = 95.5%) and 0.983 (accuracy = 93.5%), respectively. In Jongkreaengkrai et al. study [4], SVM-light with T1-weighted MRI were utilized to classify AD patients and normal subjects. The results indicated that the combination of hippocampus and amygdala volume, and entorhinal cortex thickness provided highest classification performance with AUC equal to 0.8906. However, the combination of hippocampus volume and entorhinal cortex thickness for AD classification were not investigated. In this study, the combination of hippocampus relative volume and entorhinal cortex thickness yielded highest performance in differentiating normal elderly subjects and AD patients based on SVM algorithm. More importantly, these two structures are closely related to the pathology of AD because these are the early brain atrophy areas of AD [2]. However, further study should be investigated to obtain an appropriate optimal classification model.
6. Conclusion
The SVM algorithm is a powerful classifier to differentiate AD patients and elderly subjects with the use of cerebral features from brain MRI data. The combination of hippocampus relative volume and entorhinal cortex thickness of both brain hemispheres, was found to be a powerful feature model with none statistically different between the results obtained from the WEKA and SVM-light software.

References
[1] Alzheimer’s Association 2018 *Alzheimers Dement* **14** 367-429
[2] Igor O K 2014 *MSRJ* **04** 24-33
[3] Arimura H, Magome T, Yamashita Y and Yamamoto D 2009 *Algorithms* **2** 925-52
[4] Jongkreaungkrai C, Vichianin Y, Tocharoenchai C, Arimura H Arimura H and Alzheimer’s Disease Neuroimaging Initiative 2016 *J. Phys. Conf. Ser.* **694** 012036
[5] Yasuo Y, Hidetaka A, Takashi Y, Chiaki T, Ohara T, Koji K, Yasuhiko N, Nobuyoshi O, Hiroshi H and Fukai T 2013 *J Biomed Sci Eng.* **6** 1090-8
[6] Arimura H, Yoshiura T, Kumazawa S, Tanaka K, Koga H, Mihara F, Honda H, Sakai S, Toyofuku F and Higashida Y 2008 *ACAD RADIOL* **15** 274-84
[7] Long X, Chen L, Jiang C, Zhang L and Alzheimer’s Disease Neuroimaging Initiative 2017 *PLoS ONE* **12** e0173372
[8] Klöppel S, Stonnington C M, Chu C, Draganski B, Sahill R I, Rohrer J D, Fox N C, Jack C R, Ashburner J and Frackowiak R S J 2008 *Brain* **131** 681-9
[9] Rajeev J, Moni R S, Palanikumar S and Gopalakrishnan T 2012 *Asian Biomed* **6** 87-94
[10] Orimaye S O, Wong J S M, Golden K J, Wong C P and Soyiri I N 2017 *BMC Bioinformatics* **18** 34
[11] Laurence O D et al. 2012 *PLoS ONE* **7** e32441
[12] Joachims T 2002 SVM light (Cornell University) http://svmlight.joachims.org
[13] Frank E, Hall M A and Witten I H 2002 *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"* (Morgan Kaufmann) p 7-70
[14] Rasha E, Omar H, Mohamed K, Mina R A, Sameh S, Michael B and Magdy E Z 2014 *The Egyptian Journal of Radiology and Nuclear Medicine* **45** 511-8
[15] Christopher G S et al. 2016 *NeuroImage: Clinical* **11** 802-12
[16] DeLong E R, DeLong D M and Clarke-Pearson D L 1988 *Biometrics* **44** 837-45
[17] Fischl B 2012 *NeuroImage* **62** 774-81
[18] Hsu C W, Chang C C and Lin C J 2010 A Practical Guide to Support Vector Classification http://csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf

Acknowledgement
The authors give thanks for all facilities from the Faculty of Medical Technology, Mahidol University. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) and DOD ADNI. ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions: see www.adni.loni.usc.edu/about/funding. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. The authors would like to acknowledge the Laboratory for Computational Neuroimaging at the Athinoula A. Martinos Center for Biomedical Imaging for FreeSurfer software and NITRC. NITRC-IR, and NITRC-CE have been funded in whole or in part with Federal funds from the Department of Health and Human Services, National Institute of Biomedical Imaging and Bioengineering, the National Institute of Neurological Disorders and Stroke for dcm2nii tool.