The Thai Journal of Radiological Technology

Journal homepage: https://he02.tci-thaijo.org/index.php/tjrt

TSRT THE THAI SOCIETY OF RADIOLOGICAL TECHNOLOGIST

Original ***Article***

การจำแนกและทำนายโรคอัลไซเมอร์จากภาพรังสีวิทยาของสมองด้วยภาพ T1-weighted MRI โดย SVM อัลกอริทึม

Alzheimer's Disease Classification and Prediction Using T1-weighted MR Brain Imaging Based on SVM Algorithm

ชญานนท์ ภมะราภา[1] • ธวัชชัย เอกจีน[1] • วัชระ ชุ่มบัวตอง[2] • ยุทธพล วิเชียรอินทร์[1*]

[1]ภาควิชารังสีเทคนิค คณะเทคนิคการแพทย์ มหาวิทยาลัยมหิดล ประเทศไทย
[2]ศูนย์เหมืองข้อมูลและชีวการแพทย์สารสนเทศ คณะเทคนิคการแพทย์ มหาวิทยาลัยมหิดล ประเทศไทย

Chayanon Pamarapa[1] • Tawatchai Ekjeen[1] • Watshara Shoombuatong[2] • Yudthaphon Vichianin[1*]

[1]Department of Radiological Technology, Faculty of Medical Technology, Mahidol University, Thailand
[2]Center of Data mining and Biomedical Informatics, Faculty of Medical Technology, Mahidol University, Thailand

*ผู้รับผิดชอบบทความ: ดร.ยุทธพล วิเชียรอินทร์ | Corresponding author: Yudthaphon Vichianin, PhD (yudthaphon.vic@mahidol.edu)

Received: 25 October 2021 | Revised: 19 December 2021 | Accepted: 23 December 2021

Thai J Rad Tech 2021;46(1):69-79

บทคัดย่อ

บทนำ: ปัจจุบันโรคอัลไซเมอร์ถือเป็นหนึ่งในปัญหาเชิงสุขภาพของประชากรโลก โดยตัวบ่งชี้ทางชีวภาพ (biomarker) ที่สำคัญสำหรับการตรวจวินิจฉัยโรคอัลไซเมอร์โดยแพทย์คือ การใช้ภาพรังสีวิทยาของสมองด้วยภาพ T1-weighted MRI ข้อสำคัญคือการตรวจวินิจฉัยโรคอัลไซเมอร์ในระยะแรกเริ่มมีผลต่อการชะลอความรุนแรงของการดำเนินโรคให้ลดน้อยลง การศึกษานี้มีวัตถุประสงค์เพื่อสร้างปัญญาประดิษฐ์ในรูปแบบของแบบจำลองการเรียนรู้ด้วยเครื่อง (Machine learning model) สำหรับการพยากรณ์การดำเนินของโรคอัลไซเมอร์ในระยะแรกเริ่ม ประกอบไปด้วย ภาวะถดถอยทางสมองในระยะท้าย (Late mild cognitive impairment, LMCI) ภาวะถดถอยทางสมองในระยะต้น (Early mild cognitive impairment, EMCI) และ ผู้ป่วยที่ไม่มีภาวะสมองเสื่อม (Cognitive normal, CN) ซึ่งแบบจำลองถูกออกแบบและเรียนรู้จากภาพรังสีวิทยาของสมองประเภท 3DT1-weighted MRI ผ่าน Support Vector Machine (SVM) อัลกอริทึม สำหรับการพยากรณ์โรคในกลุ่มผู้ป่วยอายุ 65-75 ปี **วิธีการศึกษา:** ภาพรังสีวิทยาของสมองในงานวิจัยนี้รวบรวมข้อมูลจากสถาบัน Alzheimer's Disease Neuroimaging Initiative ประกอบไปด้วยข้อมูลภาพรังสีวิทยาสมองของผู้ป่วยภาวะถดถอยทางสมองในระยะท้าย จำนวน 61 คน ผู้ป่วยภาวะถดถอยทางสมองในระยะต้น จำนวน 95 คน และผู้ป่วยที่ไม่มีภาวะสมองเสื่อม จำนวน 92 คน โดยวิธีการสร้างแบบจำลองในการศึกษาครั้งนี้ประกอบไปด้วย 3 ขั้นตอนหลัก ได้แก่ ขั้นที่ 1 การเตรียมข้อมูล (Data preprocessing) ขั้นที่ 2 การสกัดคุณลักษณะของข้อมูล (Feature extraction) และขั้นที่ 3 การจำแนกประเภทและการทำนายข้อมูลโดยอัลกอริทึม (Algorithm classification) ในขั้นตอนที่ 1 และ 2 มีการใช้งานซอฟต์แวร์ FreeSurfer สำหรับการเตรียมข้อมูลและการสกัดข้อมูล และในขั้นตอนสุดท้ายใช้การจำแนกและทำนายข้อมูล ในลักษณะตัวจำแนกแบบไบนารี จำนวนทั้งสิ้น 3 คู่การจำแนก (คู่การจำแนก CN กับ LMCI คู่การจำแนก CN กับ EMCI และคู่การจำแนก EMCI กับ LMCI) และมีการใช้เทคนิคคัดเลือกคุณลักษณะ (Feature selection) ผ่านค่าสถิติ F1-score ร่วมกับการจำแนกและทำนายข้อมูลของแบบจำลอง **ผลการศึกษา:** คู่การจำแนก CN กับ LMCI มีค่า AUC เท่ากับ 0.79 และมีค่าความถูกต้อง (Accuracy) เท่ากับ 73.86% ในขณะที่คู่การจำแนก CN กับ EMCI มีค่า AUC เท่ากับ 0.64 และมีค่าความถูกต้อง เท่ากับ 59.89% และคู่การจำแนก EMCI กับ LMCI มีค่า AUC เท่ากับ 0.67 และมีค่าความถูกต้อง เท่ากับ 66.67% **สรุปผลการศึกษา:** ผลการศึกษาจากงานวิจัยนี้พบว่าแบบจำลอง SVM

สำหรับการจำแนกและพยากรณ์โรคอัลไซเมอร์ในระยะแรกเริ่มประสบความสำเร็จในการทำนายสำหรับคู่การจำแนก CN กับ LMCI คู่การจำแนก EMCI กับ LMCI และคู่การจำแนก CN vs. EMCI จากค่าความสามารถในการทำนายมากไปน้อย ตามลำดับ และจากผลการศึกษาครั้งนี้ทางผู้วิจัยเล็งเห็นว่าภาพรังสีวิทยาของสมองด้วยภาพ T1-weighted MRI อาจเป็นตัวบ่งชี้ทางชีวภาพที่สำคัญสำหรับการตรวจวินิจฉัยโรคอัลไซเมอร์ในระยะแรกเริ่ม

คำสำคัญ: โรคอัลไซเมอร์, ภาวะถดถอยทางสมอง, SVM อัลกอริทึม, ภาพ T1-weighted MRI

Abstract

Introduction: Nowadays, Alzheimer's disease (AD) is one of the worldwide health issues. Clinicians utilize Magnetic Resonance Brain imaging as one of the key biomarkers for AD diagnosis. Early-stage detection could prevent the high progression of the disease. The study aimed to create a machine learning model for predicting patients who are under an early stage of Alzheimer's disease (AD) which were consisting of late mild cognitive impairment (LMCI), early mild cognitive impairment (EMCI), and cognitive normal (CN) for patient aged 65-75 using 3DT1-weighted MR Brain imaging based on Support Vector Machine (SVM) classification. **Methods:** The imaging data were acquired from the Alzheimer's Disease Neuroimaging Initiative consisted of 61 LMCI patients, 95 EMCI patients, 92 cognitive normal subjects. There were three main steps of this work including 1) data preprocessing, 2) features extraction, and 3) algorithm classification. The first two steps were performed using FreeSurfer software to normalize the imaging data and extract features of interest. The final step was algorithm classification and algorithm training with three binary classification groups (CN vs. LMCI, CN vs. EMCI, and EMCI vs. LMCI) with feature selection training methodology based on F1-score, three classification models in total. **Results:** The CN vs. LMCI classifier achieved a 0.79 AUC value with 73.86% of accuracy. Meanwhile, the CN vs. EMCI classifier achieved a 0.64 AUC value with 59.89% of accuracy, and the EMCI vs. LMCI classifier achieved a 0.67 AUC value with 66.67% of accuracy. **Conclusion:** Our findings indicated that the proposed SVM classification models succeeded to classify and predict Alzheimer's disease progression for CN vs. LMCI, EMCI vs. LMCI, and CN vs. EMCI ordered by its prediction performance from high to low, respectively. Importantly, we emphasized that MR Brain imaging might be a potential biomarker for an early-stage Alzheimer's disease diagnosis.

Keywords: Alzheimer's Disease, Mild Cognitive Impairment, SVM Algorithm, T1-weighted MR image

Introduction

According to the journal of the Alzheimer's association 2021 report[1], population aging is a global phenomenon. Every country is experiencing a high growth rate of elderly in its population. There were 6.2 million persons aged 65-year-old or older in United State. The problem is an increase of prevalence in the counting and proportion of patients under Alzheimer's disease to normal cognitive condition people, especially with people aged older than 65-year-old. The percentage of people with Alzheimer's dementia is increasing when compared with the report in 2019[2] by age: 3.0% becomes 5.3% in 2021 of people aged 65-74, 17.0% reduces to 13.8 percentages of people aged 75-84, and 32.0% becomes 34.6% of people age more than 85 have Alzheimer's dementia condition which means the prevalence of Alzheimer's disease is trending to be more severe proportion in the future global population.

Alzheimer's disease is one of the various types of neurodegenerative disease. It is mainly one of the symptoms of Dementia syndrome. The signs of Alzheimer's disease are similar to overall signs of Dementia syndrome such as memory loss, uncontrol emotional expression, some physical ability, language problem, etc. Alzheimer's disease is also a degenerative disease which means that the progression of the symptoms becomes worse over time. It is thought to begin 20 years or more before symptoms arise[1,2]. The Alzheimer's disease is caused by non-effective synapses and uncommunicable synapses between neurons in the human brain[3]. The accumulation of the abnormal protein fragment called 'Beta-amyloid' and protein tau (tau tangles) happened in various areas in the brain to induce abnormal synapse[4]. The level of beta-amyloid and protein tau on some areas of the brain and neurons was significantly increased starting around 20 years before noticeable symptoms were expected to develop[5]. Glucose metabolism of those areas began to decrease up to 18 years before expected symptom onset, and brain morphological atrophy began with a period up to 13 years before expected symptom onset[6]. There are three overall stages of Alzheimer's disease consisted of preclinical Alzheimer's disease, mild cognitive impairment (MCI), and Alzheimer's dementia stage[1,2].

According to the Alzheimer's Association report, there is no single test for Alzheimer's dementia detection[1]. In current practice, obtaining a medical and family history from the individual, asking family members, conducting cognitive tests or neurologic examinations are part of the clinical examinations for diagnosis. Whereas one of the biomarkers for diagnosis and advancing research is using radiological modalities for detecting the changing of brain morphology using Magnetic Resonance Imaging (MRI)[2]. Nowadays, the term "Big data" and Artificial Intelligence (AI) is trending. The large storage of medical information can be beneficially used and analyzed with the help of AI. In the medical imaging field, there is a term of 'Computer-Aided Diagnostic' (CAD)[7] that brings the concept of Machine learning model which is a sub-type algorithm in AI to use for computer-based diagnostic on the medical imaging information[8]. There are related works using the concept of machine learning and deep learning for Alzheimer's disease staging prediction that provided a great performance on the Alzheimer's disease prediction[9-12].

This study aimed to create classification models based on a machine learning algorithm which was focusing on the early stage of Alzheimer's disease progression. Moreover, we brought the concept of feature selection in order to improve our reliability of training features, and to maximize the model prediction performance. Thus, this study aimed to create a machine learning model for supporting a prediction of patients under an early stage of Alzheimer's disease (AD) which were consisting of late mild cognitive impairment (LMCI), early mild cognitive impairment (EMCI), and cognitive normal (CN) for patient aged 65-75 yrs using T1-weighted MR Brain imaging based on Support Vector Machine (SVM) classification algorithm.

Methods

In this work, the overall study and algorithm training was divided into three main steps for machine learning model building including 1) Data acquisition, 2) Computer-Aided diagnostic for LMCI, EMCI, CN classification and 3) Output and model performance evaluation. The details were described as follows.

Data acquisition

With an inclusion criterion, this study acquired clinical imaging from the patients aged 65 to 75-year-old who had T1-weighted MR Brain Imaging. Data obtained from The Alzheimer's Disease Neuroimaging Initiative[13] – an online database of a longitudinal multicenter study including imaging which has made major contributions to AD researchers around the world. The details of data selection have details as shown in Table 1.

Moreover, with an exclusion criterion, we excluded the image with poor image quality such as the appearance of the artifact on T1-weighted MR Brain image, for the instant, wrap-around artifact, poor signal from magnetic field (B_0) susceptibility, and non-uniformity images.

In addition, the MRI protocol scanning with a repetition time (TR) was 2300 ms, echo time (TE) 3 ms, and inversion time (TI) 900 ms. Voxel dimensions were 1.200 mm ×1.015 mm × 1.015 mm. All images were acquired using 8-channel head array receiver coils. MRI feature extraction and atlas normalization were performed on all images[14]. The example of data collections based on disease progression showed in Figure 1.

Table 1. The demographic of datasets from ADNI in this work.

Demographics of ADNI Datasets						
Clinical Diagnosis	Number of Patients			Average Age (year)		
	Male	Female	Total	Male	Female	P-value
LMCI	29 (47.54%)	32 (52.46%)	61	70.75	70.85	0.82
EMCI	48 (50.53%)	47 (49.47%)	95	70.91	69.37	0.01*
CN	31 (33.70%)	61 (66.30%)	92	71.00	70.357	0.30

Note: 1) Range of all patients in this work were 65-75 years-old, 2) Independence sample T-test was used with alpha 0.05 as the statistical analysis tool

Table 2. T1-weighted MR Brain images extracted features by FreeSurfer.

T1-weighted MR Brain images: 38 extracted features by FreeSurfer				
1. Left lateral ventricle	2. Left Inferior Lateral ventricle	3. Left Cerebellum White matter	4. Left Cerebellum cortex	5. Left Thalamus proper
6. Left Caudate	7. Left Putamen	8. Left Pallidum	9. Third Ventricle	10. Fourth Ventricle
11. Brain Stem	12. Left Hippocampus	13. Left Amygdala	14. CSF	15. Left Accumbens area
16. Left Ventral diencephalon	17. Left Vessel	18. Left Choroid plexus	19. Right Lateral ventricle	20. Right Inferior Lateral ventricle
21. Right Cerebellum White matter	22. Right Cerebellum cortex	23. Right Thalamus proper	24. Right Caudate	25. Right Putamen
26. Right Pallidum	27. Right Hippocampus	28. Right Amygdala	29. Right Accumbens area	30. Right Ventral Diencephalon
31. Right vessel	32. Right choroid plexus	33. Optic Chiasm	34. Corpus Callosum posterior	35. Corpus Callosum Mid posterior
36. Corpus Callosum central	37. Corpus Callosum mid anterior	38. Corpus Callosum anterior		

Process of computer-aided diagnostic (CAD)

There are three main steps to create the machine learning models including 1) Preprocessing, 2) Feature extraction, and 3) Classification for T1-weighted MR Brain images. The steps of CAD in this study are described as follows.

The first step, preprocessing, is the most important. It started from the input for training the algorithm by the normalized data collection and provided a proper image quality, great positioning for the patient's brain (no rotation), and no artifact affected on the image. Acquired T1-weighted MR Brain images which included the scanner non-uniformity correction, at this point of preprocessing the volumetric analysis performed using freely available FreeSurfer software[15]. We may have to correct the effect from the B1 bias field from conventional MRI scanners which mostly included this effect. In the process of FreeSurfer software pipeline , the intensity normalization was performed after B1 bias field correction by the 'mri_normalize' function which converts the original brain volume into normalized white matter volume, Then, the correction of gradient inhomogeneity was performed afterward which related to the different techniques of scanning in each institute (slick thickness or length of scan dependent), non-uniformity correction for different region-pixel of signal in the image, bias field correction happened with the old MRI instrument that has large magnetic field susceptibility[15].

The second step was the feature extraction using FreeSurfer. This step of work was registering the obtained T1-weighted MR Brain images with the MNI Average Brain (305 MRI) Stereotaxic Registration Model or 'MNI305' template[15]. Then, the B1 estimating for the bias field will be processed. The output of the estimated field will show the intensity for the overall voxel on the imaging plane. After that, the intensity of each voxel will be automatically divided individually by estimating edge from the previous MNI305 template with B1 estimating calculation. The system detected the skull plane and surrounded structure to remove the skull out of the imaging plane. Afterward, the voxel of the classified 'White matter' area in the brain was defined. Finally, the intensity gradient between the area of grey matter and the cerebrovascular fluid (CSF) were defined to focus on the structure of interest to be extracted. The illustrated mentioned MR Brain feature extraction steps are shown in Figure 2.

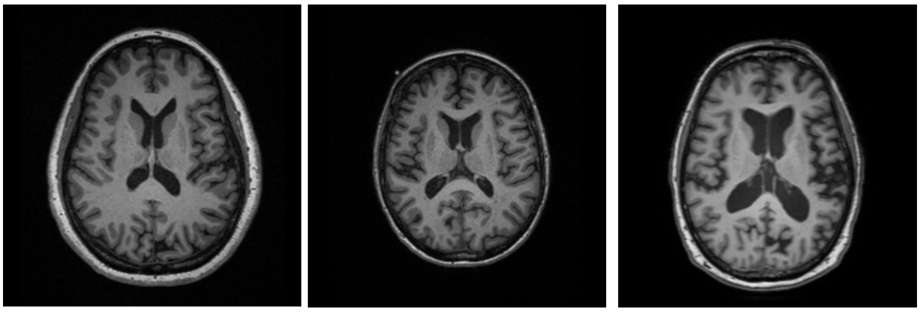

Figure 1. Example of T1-weighted MR Brain images collection in this work with CN subject, EMCI patient, LMCI patient image from left to right.

Figure 2. The result of cortical surface extraction of T1-weighted MR Brain image (left) and step of features extraction by using FreeSurfer.

The list of T1-weighted MR Brain images extracted features from FreeSurfer, 38 features in total which were considered as the relative volume of structure based on the individual intracranial volume (ICV) of a subject, are shown in Table 2.

The final step was the classification. This study used the concept of binary classification for the supervised learning technique as suggested by current literature that preferred the effectiveness of using binary classification over multiclass classification approaches. Besides, the Support vector machine (SVM) is suitable to a binary classification scenario rather than multiclass problems[16,17]. The features were selected by using F1-score for feature selection methodology[18]. The top-ten features which have the highest ranking F1-score were used for iterative training[19]. Then, the group of selected features which provided the highest prediction performance in each classification group were used to train the binary classification model individually (CN vs. LMCI, CN vs. EMCI, and EMCI vs. LMCI). By the way, each feature was calculated the relative volume of structure based on the individual intracranial volume (ICV) of a subject before passing through the classification process.

In this work, the MATLAB R2020a was used with classification learners provided various machine learning algorithms including the SVM algorithm which was the proposed method in this study. The setting parameter for all classifiers in this study was 10-fold cross-validation, automatically optimized regularization (C) parameter, automatically optimized gamma parameter, and grid-search technique applied using linear, quadratic, cubic, fine gaussian, medium gaussian, and coarse gaussian kernel.

Interpretation of model performance

All model's performance (CN vs. LMCI, CN vs. EMCI, and EMCI vs. LMCI classification model) were evaluated by interpretation tools of diagnostic ability for a binary model classification. The main tool for evaluation was the receiver operation characteristic (ROC) curve by integrating the area under the ROC curve or AUC[20]. Other tools were used for interpretation consisted of the accuracy, sensitivity, specificity, and the Matthews correlation coefficient (MCC)[21].

Results

The results of three SVM classification algorithms was evaluated for predicting the patient diagnosed with early Alzheimer's disease progression from CN, EMCI, and LMCI with mentioned interpretation tools. The results of each model and discussion were represented as following. To get the top three classification algorithms for each group in this work, we designed a process of features selection from all of 38 T1-weighted MR Brain image features to obtain the most predictive performance list of features of each classification group. The term 'ranked features' was used to define the group of high prediction performance features for each group. The prediction performance based on F1-score was listed by the first ten ranked features which are shown in Table 3.

Each classification was trained using iterative training with adding one more feature for the next iteration in order. The performance of iterative feature selection training with the first 10-rank T1-weighted MR Brain features were shown in Figure 3. The results from iterative T1-weighted MR Brain feature selection training provided the best performance of each classification with AUC values: 0.79 (73.8% of accuracy), 0.64 (59.9% of accuracy), 0.67 (66.7% of accuracy) for CN vs. LMCI, CN vs. EMCI, and EMCI vs. LMCI, respectively. The plots of the AUC of the ROC curve (false positive rate: x-axis, true positive rate: y-axis) were shown in Figure 4.

Additional performance with additional interpretation tools were shown in Table 4. According to Table 4, the data was sorted from highest to lowest classifier performance based on AUC, accuracy, and sensitivity ranging from CN vs. LMCI classifier, EMCI vs. LMCI, to CN vs. EMCI, respectively. It could be referred to the complexity of the challenge of each classification group. Meanwhile, CN vs. EMCI classifier had a higher performance than EMCI vs. LMCI classifier in specificity. The reliability of the model (MCC) went in a similar direction with AUC. Hence, we achieved the best classification model with a 0.79 AUC value for CN vs. LMCI prediction, the first runner-up with 0.67 AUC value for EMCI vs. LMCI prediction, and the second runner-up with 0.64 AUC value for CN vs. EMCI prediction. The details of each trained model were described as follows.

The model for CN vs. LMCI prediction (0.79 AUC value) was trained with 8 extracted features of T1-weighted MR Brain images which consist of Left inferior lateral ventricle, Right hippocampus, Right amygdala, Right inferior lateral ventricle, Left lateral ventricle, Left amygdala, Left hippocampus, and Right lateral ventricle. The model for EMCI vs. LMCI (0.67 AUC value) was trained with 3 extracted features from T1-weighted MR Brain images which were Right amygdala, Right hippocampus, and Left hippocampus. The model for CN vs. EMCI (0.64 AUC value) was trained with only one extracted feature from T1-weighted MR Brain images which was the Left inferior lateral ventricle.

Table 3. F1-score of features in each classification group of T1-weighted MR Brain features.

T1-weighted MR Brain image: First 10-rank features based on F_1-Score			
Feature Rank	Classification groups		
	CN vs, LMCI	CN vs. EMCI	EMCI vs. LMCI
1	Left Inferior Lateral Ventricle (0.781)	Left Inferior Lateral Ventricles (0.672)	Right Amygdala (0.759)
2	Right Hippocampus (0.780)	Right Inferior Lateral Ventricle (0.651)	Right Hippocampus (0.755)
3	Right Amygdala (0.774)	Left Hippocampus (0.609)	Left Hippocampus (0.751)
4	Right Inferior Lateral Ventricle (0.770)	3rd Ventricle (0.591)	Right Inferior Lateral Ventricle (0.741)
5	Left Lateral Ventricle (0.765)	Left Amygdala (0.564)	Left Inferior Lateral Ventricle (0.741)
6	Left Amygdala (0.761)	Right Hippocampus (0.537)	Right Accumbens area (0.732)
7	Left Hippocampus (0.757)	Left Choroid Plexus (0.516)	Left Putamen (0.718)
8	Right Lateral Ventricle (0.750)	Left Accumbens area (0.453)	CSF (0.713)
9	Left Accumbens area (0.738)	Left Thalamus Proper (0.402)	Right Cerebellum Cortex (0.713)
10	3rd Ventricle (0.734)	Corpus Callosum Central (0.229)	Right Pallidum (0.685)

Table 4. Classifier performance with interpretation tools.

Trained models (Classifiers)	AUC	Accuracy (%)	Sensitivity (%)	Specificity (%)	MCC
CN vs. LMCI	0.79	73.86	73.64	74.42	0.44
CN vs. EMCI	0.64	59.89	56.20	70.00	0.23
EMCI vs. LMCI	0.67	66.67	69.72	59.57	0.27

Discussion

According to the overall steps and results of this study, there were some limitations of work that could be improved. The details of each aspect were described as follows. In this work, the dataset had been acquired the imaging data for algorithm training from a single database, ADNI. The model was trained by the morphological or structural of the northern American patients (the USA and Canada). So, the input images for the feature extraction process may take a discrepancy of transforming the imaging data dimension into the volumetric data dimension because the original template of FreeSurfer, MNI305, was established with the western patients (European were the most proportion of dataset)[22].

The age length of designed study (65 to 75-years-old, 70.80 years of average) could be one of the performance reducing points because they were unmatched with the MNI305 atlas design (23.4±4.1 years-old)[23]. The number of training subjects in this work was limited and not large enough to achieve desired SVM prediction performance because we obtained patient data who aged 65-75 with 92 cognitive normal subjects, 95 early mild cognitive impairment patients, and 61 late mild cognitive patients, 248 subjects in total. This study recruited data with a small and unbalanced number of labeled data in each group especially the group of late mild cognitive patients. To improve the machine learning algorithm performance, labeled data should be at least 80 to 560 to achieve mean average and root mean squared error below 0.01[24].

In this study design, the SVM algorithms with grid search technique may not provide the best prediction performance for the overall three binary classification problems due to the fuzziness of data dimension across each binary classification. To handle this problem and improve the efficiency of the models, the multi-SVM fuzzy classification model[25] should be applied in future studies because it might improve the models to be more robust and generalized for overall training datasets[26].

Importantly, another point of interest was the corresponding T1-weighted MR Brain structures which were picked up for training in each classification group. Based on the F1-score for feature selection, most of the extracted features (organs) in the limbic system and nearby structures were selected for algorithm training which consisted of the Hippocampus in both hemispheres, Right amygdala for CN vs. LMCI and EMCI vs. LMCI classification groups. The left inferior lateral ventricle was the interest area for investigation because it was picked up for CN vs. LMCI and CN vs. EMCI classification groups. Based on our study, possible features for early-stage of Alzheimer's disease prediction were the Lateral ventricle of both hemispheres, Right inferior lateral ventricle, Left amygdala. The selected features for training which based on the feature selection method related to other studies[11, 12] which were picking up the features for model training by clinical relation (Hippocampus, Amygdala, Entorhinal cortex thickness). There was a study investigated that the morphological changes of presenile dementia (Alzheimer's dementia) patients notably associated with the Hippocampus and Amygdala region[27].

Conclusions

This study achieved three classifiers based on the SVM algorithm that were trained to classify the binary classification for elderly subjects aged 65 to 75-year-old with cognitive normal (CN), early mild cognitive impairment (EMCI), and late mild cognitive impairment (LMCI) condition. Our findings indicated that the proposed SVM classification models succeeded to classify and predict Alzheimer's disease progression for CN vs. LMCI, EMCI vs. LMCI, and CN vs. EMCI from high to lower prediction performance with 0.79, 0.67, and 0.64 of AUC value, respectively.

Importantly, we emphasized that the T1-weighted MR Brain image may be a potential biomarker for an early stage of Alzheimer's disease diagnosis, and the organs and nearby organs of the limbic system which are Hippocampus, Amygdala, Lateral ventricle, and Inferior lateral ventricle, were potential features to investigate the progression of early-stage of mild cognitive impairment before the late stage of Alzheimer's disease.

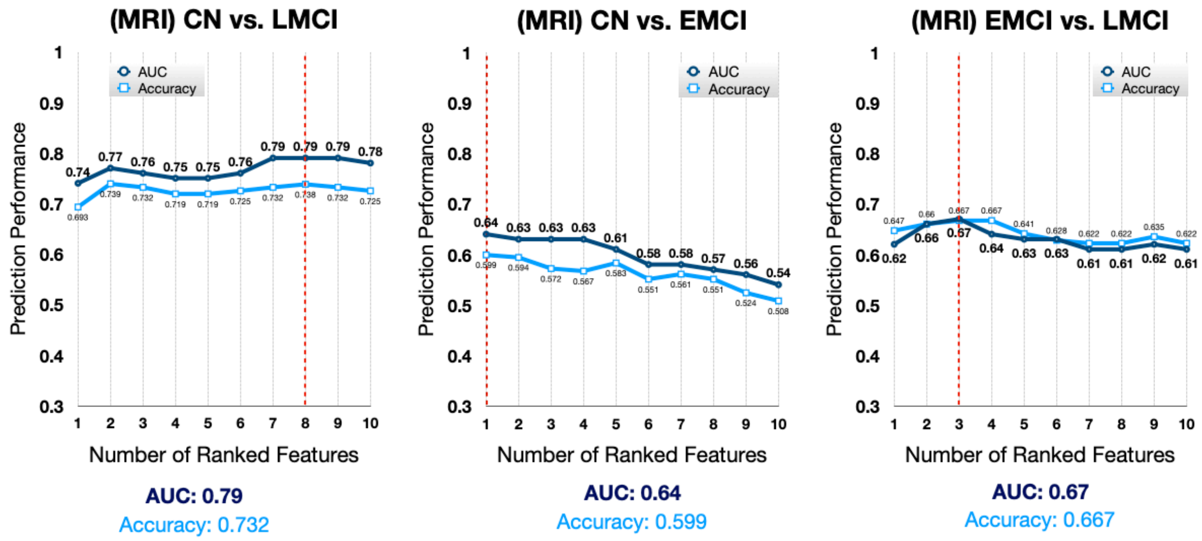

Figure 3. The performance of iterative feature selection training with first 10-rank features based on F1-score.

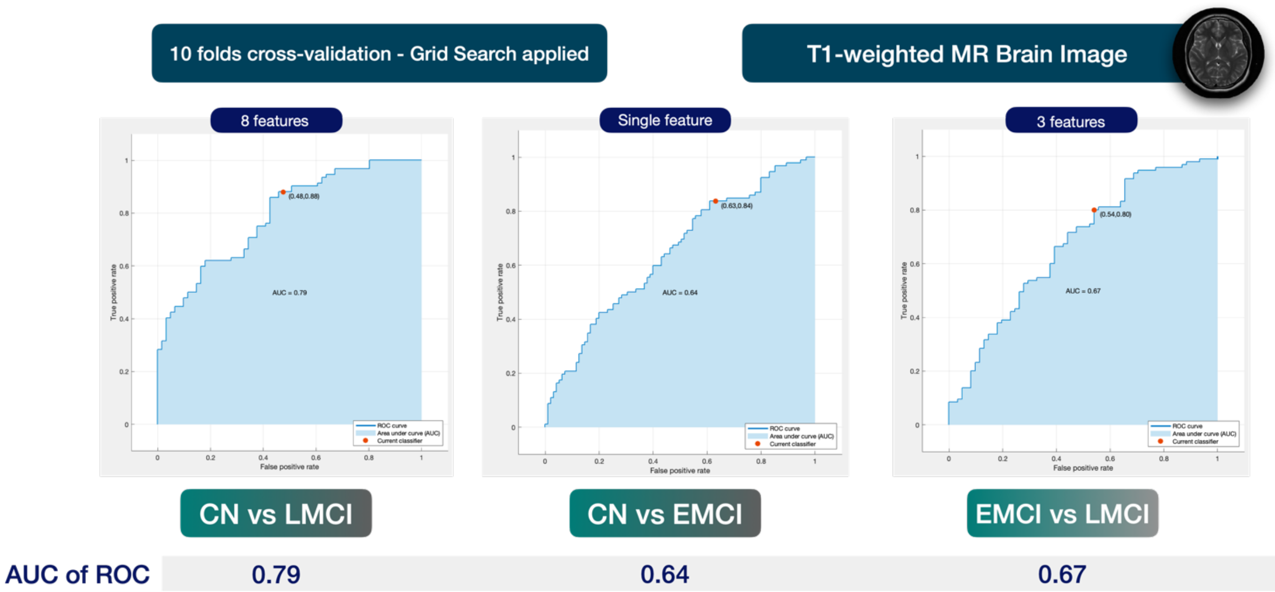

Figure 4. The result of the trained classifier with ranked T1-weighted MR Brain feature in different classification groups.

Acknowledgements

I'm so appreciated to all of my professors consisting of Assoc. Prof. Dr. Yudthaphon Vichianin, Asst. Prof. Dr. Tawatchai Ekjeen from the Department of Radiological Technology, Faculty of Medical Technology, Mahidol University, and Assoc. Prof. Dr. Watshara Shoombuatong from Center of Data mining and Biomedical Informatics, Faculty of Medical Technology, Mahidol University for their guidance in my research, the concept of machine learning in radiology especially in the MRI field of study from the started point until the day I could finish this study beneath their instructions, their willingness, and dedication. I would also to express my thanks to Mr. Khajornsak Tantiwetchayanon, who inspired me in many aspects of this work and has always been a great and kind support to me. I wish to extend my thanks and honor to Assist. Prof. Dr. Kitiwat Khamwan from the Department of Radiology, Faculty of Medicine, Chulalongkorn University who gave me an inspiration for the presentation and great suggestions for this study. Moreover, I would like to say thank you to all the lecturers in the department of Radiological technology, and all the supporting people.

Importantly, this study is partly supported by the Graduate Scholarship of the Faculty of Medical Technology, Mahidol University.

References

[1] 2021 Alzheimer's disease facts and figures. Alzheimer's & Dementia. 2021;17(3):327-406.

[2] Association As. 2019 Alzheimer's disease facts and figures. Alzheimer's & Dementia. 2019;15(3):321-87.

[3] Sheng M, Sabatini BL, Südhof TC. Synapses and Alzheimer's disease. Cold Spring Harb Perspect Biol. 2012;4(5):a005777.

[4] Bloom GS. Amyloid-β and Tau. JAMA Neurology. 2014;71(4):505.

[5] Beason-Held LL, Goh JO, An Y, Kraut MA, O'Brien RJ, Ferrucci L, et al. Changes in Brain Function Occur Years before the Onset of Cognitive Impairment. Journal of Neuroscience. 2013;33(46):18008-14.

[6] Jagust W. Imaging the evolution and pathophysiology of Alzheimer disease. Nature Reviews Neuroscience. 2018;19(11):687-700.

[7] Doi K. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. Computerized Medical Imaging and Graphics. 2007;31(4-5):198-211.

[8] Camlica Z, Tizhoosh HR, Khalvati F, editors. Medical Image Classification via SVM Using LBP Features from Saliency-Based Folded Data2015: IEEE.

[9] Basaia S, Agosta F, Wagner L, Canu E, Magnani G, Santangelo R, et al. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. NeuroImage: Clinical. 2019;21:101645.

[10] Ding Y, Sohn JH, Kawczynski MG, Trivedi H, Harnish R, Jenkins NW, et al. A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using 18F-FDG PET of the Brain. Radiology. 2019;290(2):456-64.

[11] Jongkreangkrai C, Vichianin Y, Tocharoenchai C, Arimura H. Computer-aided classification of Alzheimer's disease based on support vector machine with combination of cerebral image features in MRI. Journal of Physics: Conference Series. 2016;694:012036.

[12] Tantiwetchayanon K, Vichianin Y, Ekjeen T, Srungboonmee K, Ngamsombat C, Chawalparit O. Comparison of the WEKA and SVM-light based on support vector machine in classifying Alzheimer's disease using structural features from brain MR imaging. Journal of Physics: Conference Series. 2019;1248:012003.

[13] (ADNI) AsDNI. Study design [Available from: http://adni.loni.usc.edu/study-design/#background-container.

[14] Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): Clinical characterization. Neurology. 2010;74(3):201-9.

[15] Fischl B. FreeSurfer. Neuroimage. 2012;62(2):774-81.

[16] Rocha A, Goldenstein SK. Multiclass From Binary: Expanding One-Versus-All, One-Versus-One and ECOC-Based Approaches. IEEE Transactions on Neural Networks and Learning Systems. 2014;25(2):289-302.

[17] Rifkin R, Klautau A. In Defense of One-Vs-All Classification. J Mach Learn Res. 2004;5:101–41.

[18] Sasaki Y. The truth of the F-measure. Teach Tutor Mater. 2007.

[19] Wood T. What is the F-score? [Available from: https://deepai.org/machine-learning-glossary-and-terms/f-score.

[20] Hajian-Tilaki K. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. Caspian J Intern Med. 2013;4(2):627-35.

[21] Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics. 2020;21(1).

[22] Evans A, Collins L, Mills SR, Brown ED, Kelly RL, Peters T. 3D Statistical Neuroanatomical Models from 305 MRI Volumes1993. 1813-7 vol.3 p.

[23] Mandal PK, Mahajan R, Dinov ID. Structural Brain Atlases: Design, Rationale, and Applications in Normal and Pathological Cohorts. Journal of Alzheimer's Disease. 2012;31(s3):S169-S88.

[24] Figueroa RL, Zeng-Treitler Q, Kandula S, Ngo LH. Predicting sample size required for classification

performance. BMC Medical Informatics and Decision Making. 2012;12(1):8.

[25] Chen X, Harrison R, Zhang Y-Q. Multi-SVM Fuzzy Classification and Fusion Method and Applications in Bioinformatics. Journal of Computational and Theoretical Nanoscience. 2005;2:534-42.

[26] Guernine T, Zeroual K, editors. SVM Fuzzy Hierarchical Classification Method for Multi-class Problems. 2009 International Conference on Advanced Information Networking and Applications Workshops; 2009 26-29 May 2009.

[27] Hopper MW, Vogel FS. The limbic system in Alzheimer's disease. A neuropathologic investigation. Am J Pathol. 1976;85(1):1-20.

วารสารรังสีเทคนิค
The Thai Journal of Radiological Technology

การจำแนกและทำนายโรคอัลไซเมอร์จากภาพรังสีวิทยาของสมองด้วยภาพ T1-weighted MRI โดย SVM อัลกอริทึม

Alzheimer's Disease Classification and Prediction Using T1-weighted MR Brain Imaging Based on SVM Algorithm

ชญานนท์ ภมะราภา • ธวัชชัย เอกจีน • วัชระ ชุ่มบัวตอง • ยุทธพล วิเชียรอินทร์

Thai J Rad Tech 2021;46(1):69-79

วารสารรังสีเทคนิค
วารสารวิชาการของสมาคมรังสีเทคนิคแห่งประเทศไทย
ภาควิชารังสีเทคนิค คณะเทคนิคการแพทย์ มหาวิทยาลัยมหิดล
แขวงศิริราช เขตบางกอกน้อย กทม. 10700