Indonesian Journal of Electrical Engineering and Computer Science
Vol. 36, No. 1, October 2024, pp. 64~73
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v36.i1.pp64-73

Alzheimer's prediction via CNN-SVM on chatbot platform with MRI

Muhammad Syaekar Kadafi[1], Ahmad Khalil Yaqubi[2], Purbandini[3], Suryani Dyah Astuti[4]
[1]Department of Engineering, Faculty of Advanced and Multidisciplinary Technology, Airlangga University, Surabaya, Indonesia
[2]Faculty of Science and Technology, Airlangga University, Surabaya, Indonesia
[3]Department of Mathematics, Faculty of Science and Technology, Airlangga University, Surabaya, Indonesia
[4]Department of Physics, Faculty of Science and Technology, Airlangga University, Surabaya, Indonesia

Article Info

Article history:

Received Feb 12, 2024
Revised May 13, 2024
Accepted Jun 5, 2024

Keywords:

Alzheimer disease
Chat robot
Convolutional neural network
Magnetic resonance imaging
Support vector machine

ABSTRACT

Artificial intelligence (AI), consisting of models and algorithms capable of concluding data to produce future predictions, has revolutionary potential in various aspects of human life. One application is an Alzheimer's disease (AD) prediction chat robot (chatbot). Only now has a method provided very accurate findings and recommendations regarding the early detection of AD using magnetic resonance imaging (MRI). Therefore, this research aims to measure AD prediction performance in four stage classes, namely very mild demented, mild demented, moderate demented, and non-demented, using brain MRI images trained in the convolutional neural network (CNN)-support vector machine (SVM) model. The research involved nine combination schemes of dataset proportions and preprocessing in the CNN-SVM model. Evaluation shows that scheme 1 produces the highest accuracy, precision, recall, and F1-score, namely 98%, 99%, 98%, and 98%. The chatbot, trained using CNN, achieved 99.34% accuracy in question responses, and was then combined with AD prediction models for improved accuracy. The test results show that the chatbot functionality runs well for each transition, with a functionality score reaching 99.64 points out of 100.00. This success shows excellent potential for early detection of AD. This research brings new hope in preventing AD through AI, with potential positive impacts on human health and quality of life.

Corresponding Author:

Suryani Dyah Astuti
Department of Physics, Faculty of Science and Technology, Airlangga University
Surabaya, Indonesia
Email: suryanidyah@fst.unair.ac.id

1. INTRODUCTION

Artificial intelligence (AI) in medicine proliferates, providing physicians with new tools to offer better patient care recommendations, diagnostic assistance, treatment suggestions, and the latest medical knowledge [1]. AI, which consists of models and algorithms capable of concluding existing data to produce future predictions [2], has the potential to revolutionize many aspects of our lives [3]. AI-driven methods can potentially enhance the prognosis and treatment of Alzheimer's disease (AD). Enhancing early identification of AD may improve patient outcomes and quality of life by using machine learning techniques and incorporating chat robot (Chatbots). To fully fulfill AI's promise in tackling this urgent healthcare issue, however, future research must concentrate on validation, longitudinal studies, ethical issues, user experience, and multidisciplinary cooperation. Chatbot, one of the many applications of AI, is a virtual conversational agent that allows users to engage with AI-based computer programs through voice or writing [4], [5].

Recently, chatbots have been launched in various fields, including business, retail, and services such as healthcare [6]–[9]. Existing research has demonstrated the effectiveness of medical chatbots in activities such as ensuring proper medication adherence. Cancer patient satisfaction has been shown to increase after conversations with chatbots. Interacting with an empathetic chatbot has a mitigating impact on patients' mental health [9]. Interactions with medical chatbots reduced symptoms of hopelessness and anxiety [10], while medical chatbots correctly identified more than 96% of infected patients during the COVID-19 pandemic [11]. The burden on healthcare practitioners can be reduced and simplified with the help of medical chatbots [12], such as AD prediction chatbots.

AD is a degenerative brain disease that causes memory loss and cognitive impairment, including difficulty speaking, thinking, and completing tasks. AD was named after Alois Alzheimer, who discovered it first in 1906 [13]. AD causes 60–80% of all dementia cases [14]. In 2020, approximately 57.4 million people were diagnosed with dementia [15]. In other words, 2 out of 3 people with dementia have AD. The increase in the number of AD sufferers is estimated to reach 152 million by 2050 [16]. Individuals affected by AD experience a severe decline in cognitive function, and this has a significant impact on their quality of life and general health [14]. In addition, the average lifespan of Alzheimer's sufferers after diagnosis is 7.6 years and 5.8 years [17]. Mild cognitive impairment (MCI) is a promising stage because it is still in the preclinical stages of AD, serving as a niche target for early treatment with the potential to stop or slow the progression of AD [18]. This suggests that MCI is an effective early-stage intervention to reverse or halt the pathological progression of AD. Magnetic resonance imaging (MRI) can provide comprehensive 3D images of internal body components such as the brain [19]. MRI has been widely used to understand morphological and functional brain changes in vivo, including AD [20], schizophrenia [21], and others. Therefore, structured MRI can provide information about the brain's anatomical structure, aiding in detecting and measuring AD brain shrinkage patterns [22].

This work expands on earlier studies that used brain MRI imaging and machine learning to forecast the onset of AD. For improved prediction accuracy, it combines convolutional neural network (CNN) and support vector machine (SVM) algorithms in a novel way inside a chatbot platform. It also looks at how preprocessing methods affect the performance of the model. In general, it advances the area of AI in healthcare by investigating cutting-edge methods for improving patient care and early AD diagnosis. The study recommends investigating other data modalities besides MRI images to improve prediction accuracy and comprehend the course of AD. Predictive models may be improved, and early biomarkers can be found through longitudinal research. It is essential to validate the model on various datasets and enhance its interpretability. Integrating patient viewpoints and addressing ethical issues are essential for translation into therapeutic practice. To advance research on AD prediction and enhance patient outcomes, interdisciplinary cooperation, and financing are crucial. The possibility of forecasting the course of AD by combining brain MRI image processing with a CNN-SVM model is shown in this study. The model attains good accuracy, precision, recall, and F1-score metrics, especially when resizing and rescaling MRI images, by assessing different preprocessing procedures and dataset proportions. For people looking for information about AD, integrating this paradigm into a chatbot platform improves accessibility and usability. Nonetheless, more verification and moral deliberations are required to guarantee the model's dependability and moral use in the medical field. This study emphasizes how crucial AI-driven methods are for early AD diagnosis and patient treatment. This study aimed to assess the utility of machine learning (ML)-based classification algorithms in overcoming limiting factors associated with the pathological differentiation of the various stages involved in the AD developmental process. ML and multivariate pattern analysis are powerful conventional tools for building image-based predictive models in computer-assisted diagnostics [23]. The research uses a Kaggle dataset and a collection of chatbot queries for AD prediction. The assessment procedure includes preprocessing techniques, several data divisions, and black box testing. The implications suggest further research into optimization, ethical concerns, and possible therapeutic applications. This research design begins with implementing an ML algorithm that receives brain MRI data from patients diagnosed with AD and brain MRI data from healthy patients. After that, ML preprocesses the data to identify the most significant feature differences between the two patient groups. The next step is to integrate the chatbot with ML algorithms, allowing the chatbot to predict a person's likelihood of developing AD based on brain MRI data provided by the user.

2. MATERIALS AND METHODS

This research utilizes the Alzheimer's dataset (4 classes of Images) from Kaggle, consisting of 6,400 MRI images as X1. The dataset includes the four initial stages of AD in the form of 1-slice coronal or axial images, comprising non-demented, very mild demented, mild demented, and moderate demented stages. Detailed information regarding this research material can be found in Table 1.

The Kaggle Alzheimer's dataset, which consists of 6,400 MRI pictures divided into four stages of AD, is used in the experimental setup. There is a 40-item chatbot question list divided into six categories. Three approaches are used to evaluate the AD prediction algorithm, with data division and preparation differences. There are nine different assessment schemes used. Black box testing evaluates the chatbot's functionality using the state transition approach. This thorough technique aims to comprehend chatbot reliability and algorithmic model performance in AD prediction. A curated question list is generated for the chatbot, including many AD-related subjects. The algorithm's performance is measured using accuracy, precision, recall, and F1-score metrics, with assessment done via multiple preprocessing techniques and dataset proportions. The dataset's distribution and the questions' classification are described in Tables 1 and 2, respectively.

Data for a question list is also needed to develop a chatbot text-processing model. The chatbot question list consists of six sections, involving dialogue, essential, risk, diet, keeping active, and other medical, with 40 questions. The last five sections of this questionnaire are sourced from [24], [25]. Evaluation of the four main metrics in classification, namely accuracy, precision, recall, and F1-score, will be carried out see in Table 2.

Table 1. Research materials

Stages of AD	Number of images
Non-demented	3,200
Very mild demented	2,240
Mild demented	896
Moderate demented	64
Total	6,400

Table 2. Number of questions

Dialog	Basic	Risk	Diet	Keeping active	Other medical
12	8	2	10	2	6

The AD prediction algorithm system will be evaluated through three different methods. First, the algorithm model will be evaluated without preprocessing. Second, evaluation will be carried out on the algorithm model with first-order preprocessing, namely resizing and rescaling. Finally, algorithm models with second-order preprocessing, namely rescaling and resizing, will also be evaluated. Variations in the proportion of training: testing data division, namely 60%:20%, 45%:40%, and 37.5%:50%, will be explored in the evaluation.

It is hoped that the results of this system evaluation can provide legitimacy to the reliability of the algorithm model in making predictions related to Alzheimer's. This comprehensive evaluation aims to provide essential insights into the performance and effectiveness of algorithmic models in the context of Alzheimer's prediction. In addition, a table that includes the model evaluation scheme is presented to provide a more detailed understanding.

Table 3 was designed by considering three dataset proportions and preprocesses used in the data processing. The use of three proportions of datasets provides the diversity necessary to train, test, and validate the model well. Meanwhile, the three preprocessing stages mentioned include critical steps in preparing the data before input into the model. By detailing the model evaluation scheme, readers can understand that the evaluation results are based on a thorough and appropriate framework.

Table 3. List of system evaluation schemes

Name	Dataset proportion	Preprocess
Scheme 1	Train 60%: test 20%: valid 20%	No preprocessing
Scheme 2	Train 45%: test 40%: valid 15%	
Scheme 3	Train 37.5%: test 50%: valid 12.5%	
Scheme 4	Train 60%: test 20%: valid 20%	Preprocess 1: rescale (0-1) and resize (150×150)
Scheme 5	Train 45%: test 40%: valid 15%	
Scheme 6	Train 37.5%: test 50%: valid 12.5%	
Scheme 7	Train 60%: test 20%: valid 20%	Preprocess 2: resize (150×150) and rescale (0-1)
Scheme 8	Train 45%: test 40%: valid 15%	
Scheme 9	Train 37.5%: test 50%: valid 12.5%	

In addition, tests commonly used in Alzheimer's prediction chatbot systems, especially black box testing, will be carried out. This testing phase explores the chatbot's benefits, drawbacks, and significance to ensure its quality and reliability. Testing will adopt the state transition technique, which involves providing input in images and text to observe the chatbot's behaviour. Further explanation is shown in Table 4, a test table based on the state transition diagram. This transition (T) table will outline the results of observations and analysis regarding the chatbot's response to each transition, allowing for a detailed evaluation of the chatbot's performance in various contexts and interactions.

Table 4. Testing based on state transition diagram

T	Domain	Action	Objective	Mark
T1	Enter	Click login	Enter home page	100 points if action and goal match. 0
T2	Search	Click search, type "ALCHA", Click ALCHA	Enter Chatbot	points if action and goal do not match.
T3	Home page	Click start	Enter home	
T4	Identity	Type "Identity"	Display identity	
T5	Order	Type "Command"	The command list appears	
T6	Service	Click logout	Quit Chatbots and Telegram	
T7	Service	Send text	Answer or text	Chatbot model accuracy
T8		Send picture	Diagnose or respond to image	Accuracy of AD prediction models

3. RESULTS AND DISCUSSION

The study combined the feature extraction capabilities of CNN with the discriminative power of SVM to create a hybrid CNN-SVM model for AD prediction. In order to prepare the data, the initial dataset was then randomly divided into three main parts: training data, testing data, and validating data. The process aims to ensure that the developed model can learn from large amounts of data. The model is tested on data unseen during training and verified on a small portion of the data used for validation. In this step, by default, 25% of the training data is taken to form validating data. The entire process of dividing this dataset was carried out carefully to ensure sufficient representation of each stage of AD in each data group.

Three preprocessing strategies were investigated, consisting of no preprocessing, preprocessing 1, and preprocessing 2. The resizing and rescaling of brain images significantly improved the model's performance. In preparing to train a model for an Alzheimer's prediction task using a hybrid CNN-SVM, it is crucial to involve every step of data processing. The brain image of AD patients' needs to be resized using linear interpolation, which connects two data points with a straight line. By selecting an appropriate X3 preprocessor, the model is hoped to obtain better representation, improve performance, and make learning easier.

The hybrid model exhibited adaptability and robust performance across different data proportions and preprocessing methods. Hybrid CNN-SVM with TensorFlow has proven to be a powerful combination in data analysis for image classification. The hybrid technique intends to combine the feature extraction capabilities of CNN and the discriminative power of SVM to produce a robust model. The model consists of two convolution layers with max-pooling, a flattened layer, and a dense layer designed to extract features from images hierarchically. Classification uses random fourier features to replace the SVM output layer with a radial basis function (RBF) kernel. The following is a diagram of the AD prediction model shown in Figure 1.

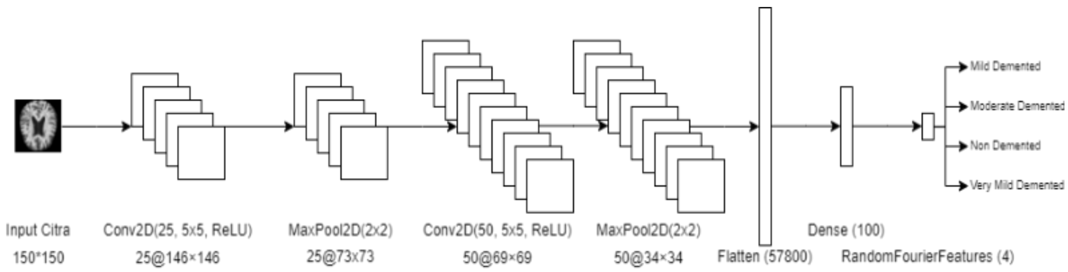

Figure 1. AD prediction model diagram

By integrating the AD prediction model, text processing model, and questionnaire dataset through a Telegram bot, smooth communication was enabled, improving the sharing of information. The use of random fourier features in the model architecture shows the integration of SVM into the CNN model. This feature is used as an SVM representation with a RBF kernel, allowing the model to capture non-linear relationships between features in the data. Traditionally, SVM is often used for classification because of its ability to handle non-linear problems [26]. With the random fourier features layer, the model attempts to approximate complex non-linear mapping while retaining the advantages of CNN in extracting features from image data. The CNN-SVM hybrid innovation reflects a holistic approach by combining the advantages of two ML paradigms. This allows the model to handle complex challenges with non-linear relationships [27].

Leveraging the advantages of SVM while still utilizing the power of hierarchical learning features from CNN, implementing the early stopping callback shows an understanding of the risk of overfitting during model training. Callbacks reflect an effort to ensure the model fits the training data to generalize unseen data well. By providing a patience limit of 5, the model will automatically stop training if there is no improvement in the accuracy of the validation data after reaching the patience limit. They were followed by model training for 100 epochs, with each epoch completing the training data processing.

Preprocessing played a critical role in boosting model accuracy, precision, recall, and F1-score, particularly in handling non-linear relationships in AD prediction. After the training process is complete, the next step is to evaluate the performance of the AD prediction system model through data analysis of images that were not seen during the training and validation stages. The following is the model evaluation results for three data proportion schemes and three preprocessing schemes. Table 5 presents the results of the evaluation without preprocessing, while Table 6 displays the outcomes of preprocess evaluation 1, and Table 7 showcases the findings of preprocess evaluation 2.

Table 5. Evaluation without preprocessing

Name	Accuracy	Precision	Recall	F1-score
Scheme 1	26%	26%	30%	22%
Scheme 2	24%	24%	22%	20%
Scheme 3	77%	66%	71%	62%

Table 6. Preprocess evaluation 1

Name	Accuracy	Precision	Recall	F1-score
Scheme 4	98%	99%	98%	98%
Scheme 5	95%	95%	95%	95%
Scheme 6	89%	92%	87%	89%

Table 7. Preprocess evaluation 2

Name	Accuracy	Precision	Recall	F1-score
Scheme 7	97%	98%	95%	97%
Scheme 8	93%	84%	91%	86%
Scheme 9	90%	92%	87%	90%

Based on Table 5, the performance of the AD stage prediction model shows a stagnant tendency in various data proportion schemes of more than 60%. This phenomenon occurs due to preprocessing needing to be applied. However, increasing model complexity impacts accuracy by 3% to 80% for a dataset proportion of 37.5%:50% without preprocessing. In addition, the model produces an accuracy of 89% and 88% for a dataset proportion of 37.5%:50% with preprocess one and a dataset proportion of 37.5%:50% with preprocess 2. Therefore, a more complex model is likely suitable for preprocessing 1.

On the other hand, when the image is not resized, the model has difficulty handling the spatial structure of the 208×176-pixel image. Resizing the image to smaller dimensions and scaling pixel values to a smaller range are preprocessing steps that help train and better converge more efficiently. After carrying out the rescale and resizing process on the image, the model evaluation results show a consistent increase in accuracy, precision, recall, and F1-score for each data proportion scheme, as depicted in Tables 6 and 7. Specifically, the absence of scaling and resizing in the image negatively impacts accuracy, precision, recall, and F1-score, which is only around 20% to 30% on training data of more than 37.5%.

Large pixel values trigger significant weight differences during training if the image is not scaled. This results in unstable convergence and overfitting. Based on Tables 6 to 7, the difference in preprocessing order does not have a significant effect, namely 1%. No preprocessing significantly impacts model performance. This is because both sequences essentially do the same thing, namely, changing the size and scale of the data. Both operations have a similar goal: adjusting the data to suit the model’s needs. Since the results of the two preprocesses are very similar, there is no significant difference in the final results. The hybrid CNN-SVM model is less sensitive to small resize and resale preprocessing changes.

The model can adapt to the numerical representations generated from both preprocesses, and its impact on model performance is insignificant. It is important to note that no essential changes occur to the data after applying these two preprocesses. This means that no transformations affect feature extraction or the model's deep understanding of the data. Therefore, differences in preprocessing order have a minimal impact on the final results of model accuracy.

Based on the data proportion scheme and preprocessing, the best model is a model with a data proportion scheme of 60%:20% and preprocessing 1 with 98% accuracy, 99% precision, 98% recall, and 98% F1-score. On the other hand, the accuracy, precision, recall, and F1-score results for each data proportion scheme in preprocessing scheme 1 show slightly better results than in preprocessing scheme 2. The model results increase with each increase in the amount of training data. The analysis results in Table 8 show that preprocessing plays an essential role in improving the performance of the AD model. Model performance can be stagnant and not optimal without optimal preprocessing, especially rescaling and resizing, mistreated by the evaluation results in Table 6. The following is a detailed table evaluating the best model.

Table 8. Details of best scheme test data evaluation

Stages	Accuracy	Precision	Recall	F1-score	Amount of data
Mild demented	97.74%	98.87%	97.22%	98.04%	180
Moderate demented		100%	100%	100%	13
Non-demented		97.08%	98.59%	97.83%	640
Very mild demented		98.19%	96.65%	97.41%	448
Macro average		98.53%	98.12%	98.32%	1,281
Weighted average		97.75%	97.74%	97.74%	1,281

The model evaluation performs excellently for various classes in the best model scheme. The "moderate demented" class stands out with accuracy, precision, recall, and an F1-score that reaches 100%. Meanwhile, the "mild demented," "non-demented," and "very mild demented" classes also showed excellent performance, with accuracy, precision, recall, and F1-score metrics in the range of 97.85%-99.36%.

A text processing model was developed to enhance user interaction, effectively processing queries and providing relevant responses. These schemes provide a solid basis for evaluating model performance in predicting and classifying AD stages in brain MRI images. Apart from that, question list data is also needed in developing a chatbot text processing model. Trusted information sources from Alzheimer's research UK's information services add value to the chatbot service, ensuring users receive accurate and up-to-date information. After creating a list of questions, the next step is modelling the text processing algorithm. Algorithm modelling aims to enable the chatbot to provide the correct answer based on the level of similarity of the user's question to the list of previously compiled questions. In this context, if the probability of similarity between the user's question and the list of questions is less than 25%, the chatbot will take the answer from the ChatGPT model. This modelling process involves processing the intents in the dataset, where each intent represents a purpose or topic of conversation. Figure 2 shows a diagram of the text processing model.

The model measures the similarity between texts. This process allows the chatbot to provide more relevant answers to the user's questions. Using the intent information in the dataset, the system can match the user's question intent with the list of questions and provide an appropriate response. This step involves extracting the word patterns of each intent, and the results are organized during model training and, based on Figure 2, forming and training a neural network (NN) model using Keras. First, the sequential NN model is built sequentially with appropriate layers, starting with an input layer with 128 neurons.

Then, a rectified linear unit (ReLU) activation function is applied, followed by a dropout layer to reduce overfitting by randomly setting several neurons to 0 during training. Next, a hidden layer with 64 neurons and a ReLU activation function is added, followed by another dropout layer. The process of model formation and training continues. The output layer is added with the number of neurons according to the number of classes and the Softmax activation function. After that, the model is configured with the previously created stochastic gradient descent (SGD) optimizer, using categorical cross-entropy as the loss function and accuracy as the monitoring metric. Finally, the model is trained using training data for 200 epochs with a batch size of 5. These steps form a neural network model that can be used for classification with an accuracy of up to 99%. The following is a diagram of model accuracy and loss against epoch. Figure 3 displays a graphic representation of the text processing model.

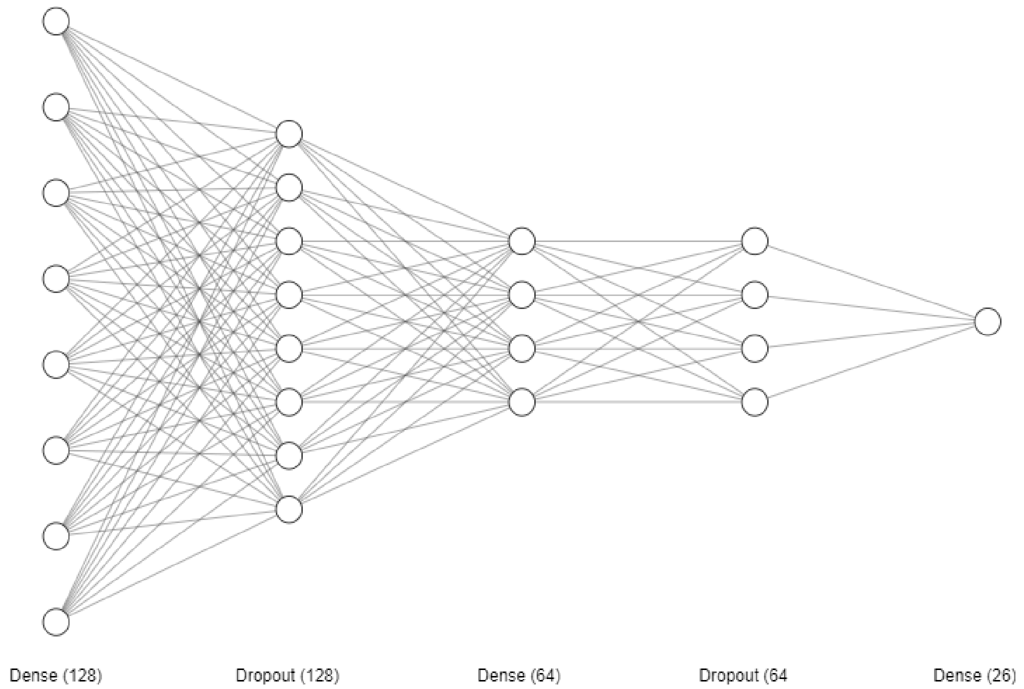

Figure 2. Text processing model diagram

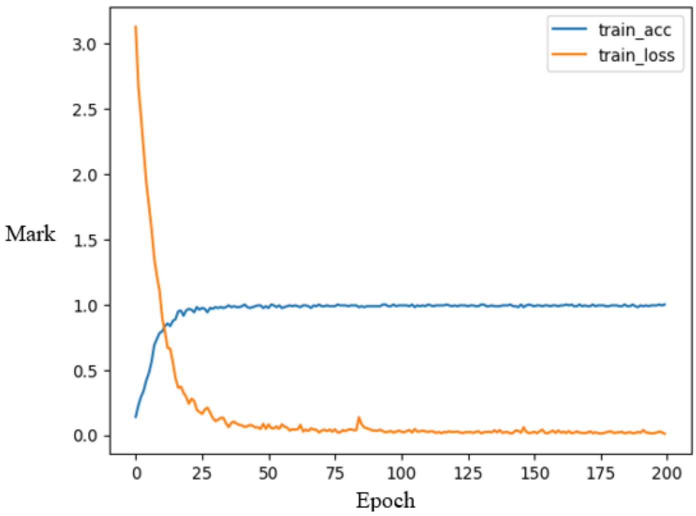

Figure 3. Graphic of text processing model

After completing the development of the AD prediction model for data analysis, the text processing model for chatbot design, and the questionnaire dataset for data collection, the next step is to integrate all these components via a Telegram bot. The integration enables a seamless flow of information between the AD prediction models, text processing models, and questionnaire datasets, creating a reliable and responsive chatbot. To achieve this, testing needs to be carried out to provide an in-depth explanation of the results obtained regarding chatbot functionality by referring to the test table in Table 4. This table provides an understanding of the situation or conditions being tested, making it easier to interpret the test results. Column T (transition) in Table 6 describes the transition number corresponding to the transition column in Table 4, clearly identifying the transition step being tested. In the context of results interpretation, the output column provides a detailed description of the results of the scenario transition test. Next, the scenario column describes the scenarios designed for testing. Table 9 presents the results of the chatbot functionality testing.

Table 9. Chatbot functionality testing

T	Scenario	Output	Mark
T1	User logs in to their Telegram account	User successfully logs in to Telegram account	100.0
T2	User searches for ALCHA	Users discovered ALCHA	100.0
T3	Users start using ALCHA	Users get basic information	100.0
T4	User types "/Identity"	User gets ALCHA identity	100.0
T5	User types "/Command"	Users get service information	100.0
T6	User send text	Users get Information	99.34
T7	Users send photos	Users can estimate AD	97.74
T8	User logs out of Telegram account	User successfully logs out of Telegram account	100.0
	Total		797.08

Chatbot functionality can be measured mathematically based on the technical transition values. The following is the calculation.

$$\text{Fungsionalitas} = \frac{\sum_{i=1}^{7} Mark\ Ti}{Amount\ of\ data} = \frac{797{,}08}{8} = 99.64\ points$$

These results indicate that the chatbot can respond and handle various scenarios effectively, meeting expected performance expectations with a functionality score reaching 99.64 points out of 100. Success in each scenario provides a positive picture of the chatbot's ability to provide appropriate and accurate responses to questions or situations submitted by the user.

Large-scale validation studies may be carried out in the future to validate the efficacy of AI models in actual clinical settings, building on these results. Clinical trials assessing the effect of AI-driven therapies on patient outcomes and longitudinal studies following people over time are crucial. Furthermore, to advance AI-driven healthcare solutions for AD prediction, multidisciplinary cooperation, user experience research, comparisons versus conventional diagnostic techniques, and ethical implications must be explored.

4. CONCLUSION

This research explores nine combination schemes between 3 preprocessing methods and 3 dataset proportions in the CNN-SVM model. The three preprocessing methods are without preprocessing one and preprocessing 2. The evaluation results show that the best performance is obtained in the model with scheme four based on accuracy, precision, recall, and F1-score, reaching values of 98%, 99%, 98%, and 98%, respectively. In other words, the best preprocessing mechanism among the three schemes studied involves resizing the image to a size of 150 pixels×150 pixels, followed by rescaling the image to a range of 0-1. The choice of this preprocessing scheme indicates that the specific steps implemented in preprocessing one positively contributes to the model's performance. Resizing images to smaller dimensions and rescaling to change pixel values to a more concentrated range have improved the model's ability to extract features and understand brain MRI image patterns. On the other hand, the chatbot has demonstrated excellent functionality in providing AD-related responses after transition testing, with a functionality score reaching 99.64 points out of 100.

REFERENCES

[1] P. Lee, S. Bubeck, and J. Petro, "Benefits, limits, and risks of GPT-4 as an AI Chatbot for medicine," *New England Journal of Medicine*, vol. 388, no. 13, pp. 1233–1239, Mar. 2023, doi: 10.1056/nejmsr2214184.

[2] Ş. Yaşar, C. Çolak, and S. Yoloğlu, "Artificial intelligence-based prediction of COVID-19 severity on the results of protein profiling," *Computer Methods and Programs in Biomedicine*, vol. 202, p. 105996, Apr. 2021, doi: 10.1016/j.cmpb.2021.105996.

[3] C. Wang, T. S. H. Teo, and M. Janssen, "Public and private value creation using artificial intelligence: an empirical study of AI voice robot users in Chinese public sector," *International Journal of Information Management*, vol. 61, p. 102401, Dec. 2021, doi: 10.1016/j.ijinfomgt.2021.102401.

[4] M. Almalki and F. Azeez, "Health chatbots for fighting COVID-19: a scoping review," *Acta Informatica Medica*, vol. 28, no. 4, pp. 241–247, 2020, doi: 10.5455/AIM.2020.28.241-247.

[5] J. Zhang, Y. J. Oh, P. Lange, Z. Yu, and Y. Fukuoka, "Artificial intelligence chatbot behavior change model for designing artificial intelligence chatbots to promote physical activity and a healthy diet: viewpoint," *Journal of Medical Internet Research*, vol. 22, no. 9, p. e22845, Sep. 2020, doi: 10.2196/22845.

[6] A. R. Dennis, A. Kim, M. Rahimi, and S. Ayabakan, "User reactions to COVID-19 screening chatbots from reputable providers," *Journal of the American Medical Informatics Association*, vol. 27, no. 11, pp. 1727–1731, Nov. 2020, doi: 10.1093/jamia/ocaa167.

[7] X. Luo, S. Tong, Z. Fang, and Z. Qu, "Frontiers: machines vs. humans: the impact of artificial intelligence chatbot disclosure on customer purchases," *Marketing Science*, vol. 38, no. 6, pp. 937–947, Sep. 2019, doi: 10.1287/mksc.2019.1192.

[8] A. Rese, L. Ganster, and D. Baier, "Chatbots in retailers' customer communication: how to measure their acceptance?," *Journal of Retailing and Consumer Services*, vol. 56, p. 102176, Sep. 2020, doi: 10.1016/j.jretconser.2020.102176.

[9] M. de Gennaro, E. G. Krumhuber, and G. Lucas, "Effectiveness of an empathic chatbot in combating adverse effects of social exclusion on mood," *Frontiers in Psychology*, vol. 10, Jan. 2020, doi: 10.3389/fpsyg.2019.03061.

[10] R. Fulmer, A. Joerin, B. Gentile, L. Lakerink, and M. Rauws, "Using psychological artificial intelligence (tess) to relieve symptoms of depression and anxiety: Randomized controlled trial," *JMIR Mental Health*, vol. 5, no. 4, p. e64, Dec. 2018, doi: 10.2196/mental.9782.

[11] A. Martin *et al.*, "An artificial intelligence-based first-line defence against COVID-19: digitally screening citizens for risks via a chatbot," *Scientific Reports*, vol. 10, no. 1, p. 19012, Nov. 2020, doi: 10.1038/s41598-020-75912-x.

[12] L. Xu, L. Sanders, K. Li, and J. C. L. Chow, "Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review," *JMIR Cancer*, vol. 7, no. 4, p. e27850, Nov. 2021, doi: 10.2196/27850.

[13] C. Birkenbihl *et al.*, "ANMerge: a comprehensive and accessible Alzheimer's disease patient-level dataset," *Journal of Alzheimer's Disease*, vol. 79, no. 1, pp. 423–431, Jan. 2021, doi: 10.3233/JAD-200948.

[14] "2022 Alzheimer's disease facts and figures," *Alzheimer's and Dementia*, vol. 18, no. 4, pp. 700–789, Apr. 2022, doi: 10.1002/alz.12638.

[15] E. Nichols *et al.*, "Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019," *The Lancet Public Health*, vol. 7, no. 2, pp. e105–e125, Feb. 2022, doi: 10.1016/S2468-2667(21)00249-8.

[16] L. Liu, S. Zhao, H. Chen, and A. Wang, "A new machine learning method for identifying Alzheimer's disease," *Simulation Modelling Practice and Theory*, vol. 99, p. 102023, Feb. 2020, doi: 10.1016/j.simpat.2019.102023.

[17] C. S. Liang *et al.*, "Mortality rates in Alzheimer's disease and non-Alzheimer's dementias: a systematic review and meta-analysis," *The Lancet Healthy Longevity*, vol. 2, no. 8, pp. e479–e488, Aug. 2021, doi: 10.1016/S2666-7568(21)00140-9.

[18] L. Parnetti, E. Chipi, N. Salvadori, K. D'Andrea, and P. Eusebi, "Prevalence and risk of progression of preclinical Alzheimer's disease stages: a systematic review and meta-analysis," *Alzheimer's Research and Therapy*, vol. 11, no. 1, 2019, doi: 10.1186/S13195-018-0459-7/FIGURES/4.

[19] A. Muzamil, S. D. Astuti, Kamelia, and Suhariningsih, "Fat suppression spectral adiabatic inversion recovery (SPAIR) to optimize the quality of MRI pelvis image," *Malaysian Journal of Medicine and Health Sciences*, vol. 17, pp. 74–77, 2021.

[20] A. Chandra, G. Dervenoulas, and M. Politis, "Magnetic resonance imaging in Alzheimer's disease and mild cognitive impairment," *Journal of Neurology*, vol. 266, no. 6, pp. 1293–1302, Jun. 2019, doi: 10.1007/s00415-018-9016-3.

[21] Y. Xiang, J. Wang, G. Tan, F. X. Wu, and J. Liu, "Schizophrenia identification using multi-view graph measures of functional brain networks," *Frontiers in Bioengineering and Biotechnology*, vol. 7, Jan. 2020, doi: 10.3389/fbioe.2019.00479.

[22] A. Zuhriyah, A. Muzzamil, S. D. Astuti, and Suhariningsih, "Determination the ischemic stroke of brain MRI based on apparent diffusion coefficient (ADC) with b value variation," *Journal of Physics: Conference Series*, vol. 1505, no. 1, p. 012041, Mar. 2020, doi: 10.1088/1742-6596/1505/1/012041.

[23] S. Nam, D. Kim, W. Jung, and Y. Zhu, "Understanding the research landscape of deep learning in biomedical science: scientometric analysis," *Journal of Medical Internet Research*, vol. 24, no. 4, p. e28114, Apr. 2022, doi: 10.2196/28114.

[24] "Alzheimer's research UK's information services," *Dementia: Your Questions Answered*, 2020.

[25] "Alzheimer's research UK's information services," *Dementia: Your Questions Answered*, 2022. https://www.alzheimersresearchuk.org/wpcontent/plugins/mof_bl_0.2.9/downloads/FAQ_1220_1222_WEB .pdf.

[26] S. D. Astuti *et al.*, "Gas sensor array to classify the chicken meat with E. coli contaminant by using random forest and support vector machine," *Biosensors and Bioelectronics: X*, vol. 9, p. 100083, Dec. 2021, doi: 10.1016/j.biosx.2021.100083.

[27] S. D. Astuti, Y. Mukhammad, S. A. J. Duli, A. P. Putra, E. M. Setiawatie, and K. Triyana, "Gas sensor array system properties for detecting bacterial biofilms," *Journal of Medical Signals and Sensors*, vol. 9, no. 3, pp. 158–164, 2019, doi: 10.4103/jmss.JMSS_60_18.

BIOGRAPHIES OF AUTHORS

Muhammad Syaekar Kadafi born in Makassar, Indonesia, on December 23, 2002, he embarked on a research journey marked by unwavering dedication to academic achievement. The highlight of his journey was achieving a Bachelor's degree in Robotics and Artificial Intelligence Engineering from Airlangga University located in Surabaya, East Java, Indonesia, in 2024. He has demonstrated a firm commitment to advancing knowledge in computer science. Previous publications include significant contributions to Alzheimer's Disease Prediction using Machine Learning in Chatbots. Current and past research interests include Alzheimer's, CNN, SVM, MRI, and Chatbots, demonstrating continued dedication to scientific endeavours. He can be contacted at email: syekhadafi@gmail.com.

Ahmad Khalil Yaqubi Ph.D. Medical Physics, M.Sc. Experimental Condensed Matter Physics B.Sc. in Physics and Mathematics, he has 11 publication in scopus, as well as certificates in leadership, project management, research, monitoring and data analyzing. 5 years of diverse work experience in human resources, project management, supervision, data analyzing and field visits with various international/national organizations. He can be contacted at email: ahmad.khalil.yaqubi.359720-2021@fst.unair.ac.id.

Purbandini is a senior lecturer of Information Systems program study and Robotics and Artificial Intelligence Engineering program study at Airlangga University, with a Master of Computer (M. Kom) in Informatics Engineering from Institut Teknologi Sepuluh Nopember Surabaya, Indonesia, in 2006. She obtained his bachelor of Computer Science in Institut Teknologi Sepuluh Nopember Surabaya, in 1997. She has 15 years of experience in business intelligence and computer vision. Her scientific production counts over 14 articles indexed by Scopus, 8 articles indexed by GShcolar, and 8 articles indexed by WoS. She produced inventors of 1 patent (both national), and he gave around ten invited talks at national and international conferences. In professional organizations, she has recently been a member of the Institute of Electrical and Electronics Engineers (IEEE) and Indonesian Informatics and Computer Higher Education Association (APTIKOM). Her current research includes computer vision, data mininig, artificial intelligence, analyze sentiment, and decision support system. He can be contacted at email: purbandini@fst.unair.ac.id.

Suryani Dyah Astuti is a full professor of Biophysics at Airlangga University, with a Ph.D. in Biophysics from Universitas Airlangga, Indonesia, in 2011. She obtained his B.Sc. in Gadjah Mada University and M.Sc. in Bandung Institute of Technology Indonesia in 1994 and 1999, respectively. She has 15 years of experience in Photodynamic Therapy (PDT), medical instrumentation and computation analysis. Her scientific production counts over 277 papers, 84 of which have been published in peer-reviewed, high-rank international journals. She produced inventors of 12 patents (both national), and he gave around ten invited talks at national and international conferences. In professional organizations, she has recently been a member of the Phisycal of Society Indonesia (PSI) and the Alliance of Indonesian Medical Physicists (AFISMI). Her current research includes photobiomodulation, photosensitizer for PDT, ozone technology for decontamination of food and herbal medicines, diagnostic and therapy (MRI, CT scan, Linac), electronic nose technology for food quality detection, and biomedical applications. He can be contacted at email: suryanidyah@fst.unair.ac.id.