



Staging of Alzheimer's Disease based on MRI using CNN

Douaa Baddour

Tishreen University
Computers and Automated Control
douaa.w.baddour@tishreen.edu.sy

Majd Ali

Tishreen University
Artificial Intelligence
majed.ali.1976@tishreen.edu.sy

Abstract— Early detection of chronic diseases and determining the stages of damage to the patient is considered one of the most important stages of treatment, as it helps doctors take important remedial measures that help the patient recover or reduce the risk of the disease to a minimum. Alzheimer's disease is one of the neurological diseases that lead to brain atrophy, which leads to the loss of its functions. MRI images of the brain are used to detect Alzheimer's disease, but it is difficult to determine both the stages of the disease and the amount of damage in a patient using this MRI technique. In this research, we aim to detect Alzheimer's disease in addition to determining its stage based on deep learning techniques by using a classifier that uses the convolutional neural network (CNN). In the research, magnetic resonance images of the brain were used, and the hippocampus region was extracted in assessing the amount of damage because it is the most important region in diagnosing damage to the disease and reducing the amount of data entered into the neural network, our results show an accuracy of 95% in estimating brain damage. The results of the classifier used were able to determine the amount of damage according to four stages of the disease.

Keywords— *deep learning - CNN - detection of Alzheimer's disease.*

I. INTRODUCTION

Early detection of diseases is one of the most important aspects of treatment that helps doctors treat diseases and save lives, especially chronic diseases. Alzheimer's disease is considered a neurological disease that affects the brain and causes nerve cells to die, which ultimately leads to a decline in the cognitive and mental functions of individuals affected by the disease [1].

The progression of Alzheimer's disease can be classified into different stages based on its severity, with each stage having distinct symptoms and characteristics [2]. Magnetic

resonance imaging (MRI) is a widely used imaging modality for the diagnosis of Alzheimer's disease as it provides detailed information about the brain's structural changes associated with the disease [2]. However, accurate staging of Alzheimer's disease based on MRI remains a challenging task due to the complexity and variability of the disease [3].

Many studies have been conducted to explore the potential of CNNs in Alzheimer's disease diagnosis and staging using MRI data. For instance, Ocasio and Duong (2021) proposed a deep learning framework based on a CNN to predict the conversion of Mild Cognitive Impairment (MCI) to AD using MRI data. The study involved 320 normal cognition patients, 554 MCI patients, and 237 AD patients. The best CNN model was able to achieve a balanced accuracy of 0.793 [4].

Another study by Arafa et al. (2023) proposed a CNN model that classified two classes and achieved a promising accuracy of 99.95% and 99.99% for the classification of the AD stage, while the fine-tuned VGG16 model achieved an accuracy of 97.44% for the same classification [5].

In addition to CNNs, other machine-learning algorithms have also been utilized in the field of Alzheimer's disease diagnosis and staging using MRI data. For example, Bucholz et al. (2023) proposed a hybrid model that combines four types of ML models based on random forest, Support Vector Machines, logistic regression, and KNN ensemble approaches. The highest recorded prediction accuracy was achieved for the ensemble and random forest models, i.e., 85.0% and 84.6% respectively, using patient data from three-time points [6].

Overall, these studies highlight the potential of machine learning algorithms, particularly CNNs, in Alzheimer's disease diagnosis, staging, and prediction using MRI data. The proposed methods could aid in the early detection and treatment of Alzheimer's disease, ultimately improving patient outcomes.

¹ Copyright © 2023 by ESS Journal

II. ALZHEIMER'S DISEASE

Alzheimer's disease is a neurological disease that causes the death of brain cells. Alzheimer's is the leading cause of dementia [7]. Researchers trying to understand the cause of Alzheimer's disease focus on the role of two types of proteins:

- **Amyloid plaques:**

When this protein aggregates, it has a toxic effect on nerve cells, causing a breakdown in communication between cells [8].

- **Tangles:**

Tau proteins play a specific role in supporting the internal structure of nerve cells and their transport system for carrying essential substances like nutrients. However, in Alzheimer's disease, the tau proteins become deformed and arranged into structures called neurofibrillary tangles, which disrupt the transport system and can be cytotoxic [9].

A. Alzheimer's disease stages

There are four stages associated with Alzheimer's disease. Below is an explanation of the stages:

1) *Very mild Alzheimer's (Mild Cognitive Impairment):* Is characterized by early injury and tissue loss in the hippocampus, an area in the temporal lobe of the brain where neurofibrillary tangles and plaques of a protein called amyloid-beta occur at this stage. These signs can be seen in magnetic resonance imaging of the brain and indicate the onset of Alzheimer's disease. This stage is undetected clinically, meaning before the patient has any symptoms [10].

2) *Mild Alzheimer's:* At this stage, the patient has a marked problem with memory and thinking, with amnesia for recent events. Patients may have difficulty remembering recent events that occurred to them [11].

3) *Medium Alzheimer's (moderate):* As Alzheimer's disease progresses to the moderate stage, symptoms of confusion and forgetfulness worsen. Individuals may lose awareness of their whereabouts, the day of the week, or the season. They may also have difficulty recognizing people's faces or identifying their names. People may forget details of their personal lives, repeat favorite stories, or make up stories to fill in gaps in memory [12].

4) *Severe Alzheimer's:* In the late stage of the disease, patients experience a loss of the ability to communicate coherently. The individual can no longer speak or speak intelligibly, although he may occasionally say some words or phrases [10]. The person may become unable to walk unaided and then unable to sit. At the end of the stage, the person loses the ability to swallow and control bladder and bowel functions [12].

B. Time rate of acceleration of Alzheimer's disease

Rates of progression to Alzheimer's disease vary greatly. On average, patients with Alzheimer's disease live between three and 11 years after diagnosis, and in some cases, patients may live 20 years or more [12].

Figure (1) shows the stages of Alzheimer's and the time rate for the development of symptoms.

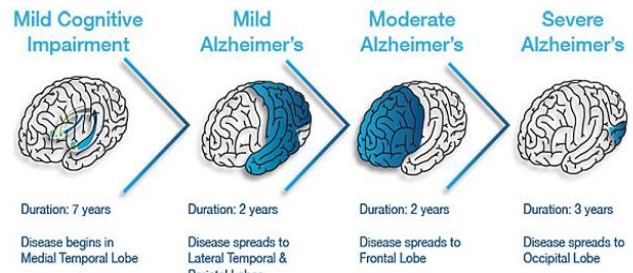


Figure 1: Alzheimer's stages and the time rate for the development of symptoms

III. CNN DEEP LEARNING NETWORKS

A Convolutional Neural Network, or CNN, is a type of deep learning that is similar to a multi-layer Perceptron. However, the difference lies in what the network learns, how it is built, and what it is. The ultimate goal of a CNN is to learn the features of an image and use them to classify it [13].

Convolutional Neural Networks (CNNs) are widely employed in computer vision applications, particularly in the field of image recognition. The initial application of CNNs dates back to 2010 when researchers employed this deep learning technique to develop a computer program capable of accurately identifying a specific object within an image [14].

CNNs comprise several important components, including the convolutional layer, activation function, pooling layer, and fully connected layer. Each of these layers performs a specific function in the network architecture [15].

The convolutional layer applies a filter, which is a matrix used to extract specific features from the input image. The design of the filter involves several conditions, the most important of which is its suitable dimensions to extract features effectively. Multiple types of filters can be used to extract features [16].

The pooling layer reduces the size of the activation maps resulting from the first stage. This layer's primary purpose is to reduce the size of the matrix resulting from the first stage by applying one of two methods: Max OR Average. Max-pooling is the most commonly used technique [13].

The fully connected layer consists of several layers of perceptron, and the neurons that form it are fully connected. This layer represents the final classification process [17].

In recent years, several approaches have been proposed for using CNNs to detect Alzheimer's disease. AbdulAzeem et al. [18] proposed a CNN model with three convolution layers and max-pooling after each layer for binary classification of AD. Their model achieved 95.6% accuracy.

Al-Khuzai et al. [19] developed a CNN model called "Alzheimer Network" (AlzNet) for binary classification of AD and CN. Their model achieved 97.99% training accuracy and 99.53% testing accuracy.

Al-Adhaileh et al. [20] proposed two pre-trained CNN models, AlexNet and ResNet50, for the diagnosis of AD. AlexNet outperformed ResNet50 with an accuracy rate of 94.53%.

Antony et al. [21] suggested VGG16 and VGG19 models for AD diagnosis, achieving accuracy rates of 81% and 84%, respectively.

Liu et al. [22] developed a CNN model from scratch and used transfer learning with AlexNet and GoogleNet models. GoogleNet achieved the highest classification accuracy rate of 93.02%.

Savaş et al. [23] used different CNN models to classify image objects from the ADNI database and found that the EfficientNetB0 model achieved the highest accuracy rate of 92.98%, while EfficientNetB2 and B3 achieved the highest precision, sensitivity, and specificity values for the AD class, respectively. These results demonstrate the effectiveness of CNNs in Alzheimer's disease detection and provide insights into the best-performing models for this task.

IV. DATASET

The Open Access Series of Imaging Studies (OASIS) is a valuable resource for researchers working on Alzheimer's disease. OASIS provides a large, publicly available database of neuroimaging data, including magnetic resonance imaging (MRI) scans and detailed clinical information about the patients [24]. OASIS V_3 is the latest version of this dataset, which includes 3D MRI images of Alzheimer's cases, along with detailed information about the patient's condition and life expectancy data for adults aged 42 to 95 years, including cognitively normal individuals and individuals with early-stage dementia.

The Clinical Dementia Rating (CDR) variable is commonly used to assess the severity of dementia in Alzheimer's disease. Table 1 shows the CDR values used in this study, ranging from 0 to 3. A CDR score of 0 indicates no Alzheimer's disease, while a score of 0.5 indicates very mild Alzheimer's, and scores of 1, 2, and 3 indicate mild, moderate, and severe Alzheimer's disease, respectively. For this study, only the first three CDR scores were used to assess the severity of Alzheimer's disease in the dataset. This allowed us to focus on early-stage Alzheimer's disease, which is crucial for developing effective diagnostic and treatment strategies.

TABLE I CLINICAL DEMENTIA RATING (CDR)

Score	Description
0	Alzheimer's Does Not Exist
0.5	Very Mild Alzheimer's
1	Mild Alzheimer's
2	Medium Alzheimer's
3	Severe Alzheimer's

To prepare the dataset for machine learning-based studies, the 3D MRI images were sliced into 2D images, and images containing the hippocampus were extracted. The hippocampus

is a critical region for determining the onset and progression of Alzheimer's disease [25], and selecting the most effective hippocampal images is crucial for developing accurate and reliable machine learning models. Figure (2) shows the stages of preparing the dataset before training, which include processing the 3D MRI images, slicing them into 2D images, identifying images containing the hippocampus, and selecting the most effective images for the final dataset. This process ensures that the dataset is optimized for developing machine learning models for the detection and diagnosis of Alzheimer's disease.

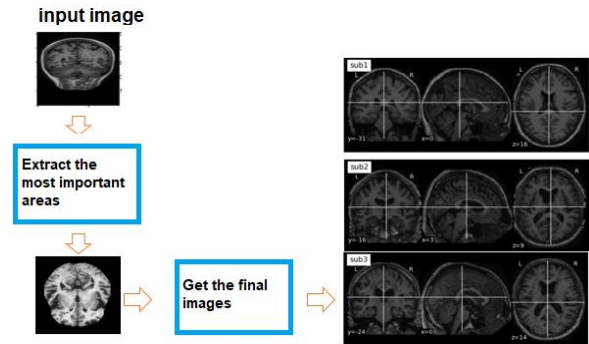


Figure 2: Preparing the dataset before training

Nibabel is a widely used Python library for working with neuroimaging data, providing tools for reading, writing, and manipulating various neuroimaging file formats, including NIfTI-1, DICOM, and MINC [26].

In this study, we utilized Nibabel to slice 3D MRI images into 2D images and selected the hippocampus images to build our dataset. The model was developed based on the following considerations and procedures:

- Applying a photo resizing process.
- Data augmentation techniques were applied to increase the number of images and improve the efficiency of the training process. These techniques included withdrawal, enlargement, and rotation of the images, which provided a large and diverse dataset for training [17, 18].

VI. OUR PROPOSED MODEL

After conducting various tests and exploring different scenarios, we developed a novel convolutional neural network (CNN) model for our study. Our proposed novel CNN model is designed to detect and diagnose Alzheimer's disease using neuroimaging data. The model is trained on preprocessed images with a size of 208x176, and the dataset is split into 80% training and 20% validation.

The model consists of several layers shown in Table 2, including three convolutional layers with a (3,3) kernel size, a ReLU activation function, and padding of type "same". After each convolution layer, we add a 2D max-pooling layer of different sizes (3,3) and (2,2) to reduce the size of the activation maps. We also use dropout layers with a ratio of 0.4 to prevent overfitting. Finally, we have a fully connected layer with a softmax activation function to classify the data into four classes.

TABLE II DESIGN OF CNN LAYERS

Layer (type)	Output Shape
Conv2d	(208,176,32)
Max pooling 2D	(104,88,32)
Conv2d	(104,88,32)
Max pooling 2D	(52,44,32)
Conv2d	(52,44,64)
Max pooling 2D	(26,22,64)
dropout	(26,22,64)
flatten	(36608)
dense	(128)
Dense_1	(4)

A. Performance evaluation parameters

In recent studies, the confusion matrix has been widely used to analyze model performance and provide insights into classification performance. The confusion matrix is robust in capturing data relationships and distributions. It offers valuable information for classification models using different metrics [27]. The primary components of the confusion matrix include True Positive (Tp), True Negative (Tn), False Positive (Fp), and False Negative (Fn) values. Based on these outcomes, various performance metrics can be calculated, including:

Accuracy (ACC): Is defined as the ratio of the number of correctly predicted samples to the total number of predicted samples as shown in eq. (1).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Sensitivity (Recall): As expressed in eq. (2), represents the proportion of samples predicted as positive out of the total number of positive samples. It is also referred to as the true positive rate.

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

Precision: As defined in equation (3), is also known as positive predictive value. It quantifies the number of samples correctly identified as positive out of the total number of samples predicted as positive.

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

F1-score: Represented by equation (4), is the harmonic mean of precision and recall. It provides a single metric that balances the trade-off between precision and recall in classification tasks.

$$F1 = \frac{2}{\frac{1}{precision} + \frac{1}{recall}} = 2 * \frac{precision * recall}{precision + recall} \tag{4}$$

ROC-AUC curve: Receiver Operating Characteristic - Area Under Curve is a widely used evaluation tool in machine learning for assessing the performance of classification models, particularly in binary classification problems. It visualizes the trade-off between the true positive rate (sensitivity) and the false positive rate (1 - specificity) at different classification thresholds.

The ROC-AUC curve is constructed by plotting the true positive rate on the y-axis against the false positive rate on the x-axis. Each point on the curve corresponds to a different threshold for classifying samples as positive or negative. By varying the threshold, we can adjust the balance between sensitivity and specificity.

The Area Under Curve (AUC) is a summary measure derived from the ROC curve. It quantifies the overall performance of the classifier in distinguishing between positive and negative samples. The AUC represents the probability that a randomly chosen positive sample will be ranked higher than a randomly chosen negative sample by the classifier. A perfect classifier has an AUC value of 1, indicating it can perfectly separate the classes, while a random classifier has an AUC of 0.5.

In the context of multi-class classification, the ROC-AUC curve is extended to evaluate the performance of models that classify samples into more than two classes. The extension involves creating a separate ROC curve for each class, treating the samples of that class as positive and the samples of other classes as negative. The performance of the model is then assessed based on the AUC values for each class.

VII. RESULTS

The training data set contains four folders for each disease category, and each category contains 1450 images, each image having dimensions 208×176×3.

A training model was reached when the model reached 95% accuracy. Given that the loss decreased to zero, the results were according to Figure (3).

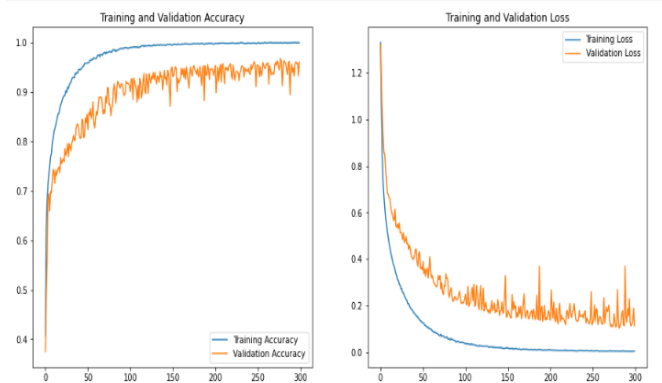


Figure 3: Accuracy and loss results

The learning results show that the success curves start to peak values, while the loss curves drop to zero values.

By relying on the ROC-AUC Curve_MULTI CLASS charts shown in Figure (4) in order to show the classifier's ability to separate and classify, it shows that maximum values

are reached for the classified cases, and thus confirm the classifier’s ability and flexibility in dealing with test images with high efficiency.

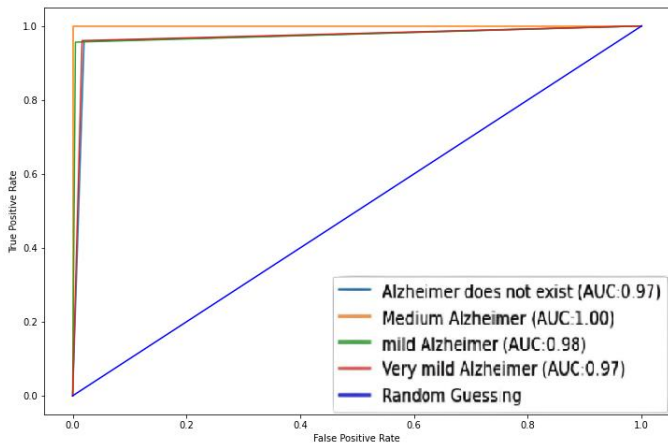


Figure 4: ROC-AUC Curve_MULTI CLASS

The confusion matrix shown in Figure (5), shows the real separation between the varieties.

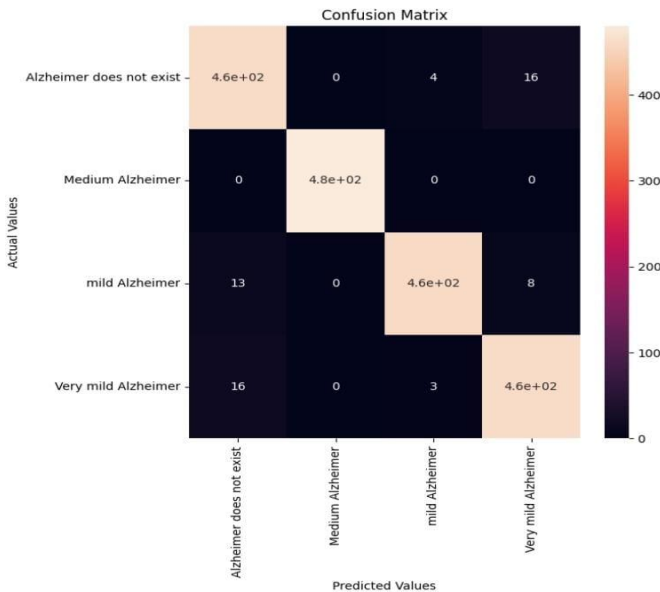


Figure 5: confusion matrix

The results from the previous confusion matrix demonstrate the model's classification ability, with each row representing a specific stage of the disease. In the Normal stage (Alzheimer Does Not Exist), the model accurately classified 460 images, while making 20 classification errors. In the Mild early stage (Very Mild Alzheimer), the model achieved perfect classification results. In the Mild stage (Mild Alzheimer), the model correctly classified 459 images but made 21 incorrect classifications. Lastly, in the Moderate stage (Medium Alzheimer), the model correctly classified 461 images and had 19 incorrect classifications. These results highlight the model's effectiveness in distinguishing between different stages of the disease.

Table 3 shows Precision, and Recall for each class.

TABLEIII F1, PRECISION, AND RECALL FOR EACH CLASS

F1	Precision	Recall	Class
0.949	0.974	0.958	Alzheimer does not exist
0.987	0.974	1.000	Medium Alzheimer
0.953	0.972	0.935	Mild Alzheimer
0.954	0.960	0.950	Very mild Alzheimer

When executing a test consisting of 15 images on the proposed model not included in the training set, the model showed an accuracy of 94.4%.

Table 4 shows a comparison with some related studies, we note the high accuracy that our model reached in classifying 4 classes.

TABLEIV COMPARE WITH RELATED STUDIES

Study	Techniques	No. Classes	Accuracy
AbdulAzeem et al. [18]	CNN	2	95.60%
Al-Khuzaiie et al. [19]	CNN (AlzNet)	2	97.88%
Al-Adhaileh et al. [20]	AlexNet, ResNet50	3	94.53%, 58.07%
Antony et al. [21]	VGG16	2	81%
Liu et al. [22]	CNN from scratch	2	78.02%
Savaş et al. [23]	EfficientNet B0	3	92.98%
Our Study	CNN from scratch	4	95%

CONCLUSION :

In conclusion, our study presents a novel convolutional neural network (CNN) model for the detection and diagnosis of Alzheimer's disease using neuroimaging data. By leveraging the power of deep learning techniques, our model demonstrates impressive performance in accurately classifying Alzheimer's disease across multiple stages. The evaluation results show high F1 scores, precision, and recall values for each class, highlighting the robustness and effectiveness of our approach. These findings suggest that our model has great potential as a reliable tool for early detection and precise classification of Alzheimer's disease, which can significantly contribute to improving patient outcomes and facilitating personalized treatment strategies. The development of accurate and efficient diagnostic models is crucial in the fight

against Alzheimer's disease, and our study contributes to the growing body of research in this field. Future work can focus on further optimizing the model and exploring the integration of multi-modal data for enhanced diagnostic accuracy.

References

- [1] Ahmad, K., Hassan Baig, M., Mushtaq, G., Amjad Kamal, M., H Greig, N., & Choi, I. (2017). Commonalities in biological pathways, genetics, and cellular mechanism between Alzheimer disease and other neurodegenerative diseases: an in silico-updated overview. *Current Alzheimer Research*, 14(11), 1190-1197.
- [2] Jack Jr, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., ... & Silverberg, N. (2018). NIA-AA research framework: toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, 14(4), 535-562.
- [3] Klöppel, S., Stonnington, C. M., Barnes, J., Chen, F., Chu, C., Good, C. D., ... & Frackowiak, R. S. (2008). Accuracy of dementia diagnosis—a direct comparison between radiologists and a computerized method. *Brain*, 131(11), 2969-2974.
- [4] Ocasio, E., & Duong, T. Q. (2021). Deep learning prediction of mild cognitive impairment conversion to Alzheimer's disease at 3 years after diagnosis using longitudinal and whole-brain 3D MRI. *PeerJ Computer Science*, 7, e560.
- [5] Arafa, D. A., Moustafa, H. E. D., Ali, H. A., Ali-Eldin, A. M., & Saraya, S. F. (2023). A deep learning framework for early diagnosis of Alzheimer's disease on MRI images. *Multimedia Tools and Applications*, 1-33.
- [6] Bucholc, M., Titarenko, S., Ding, X., Canavan, C., & Chen, T. (2023). A hybrid machine learning approach for prediction of conversion from mild cognitive impairment to dementia. *Expert Systems with Applications*, 119541.
- [7] Zvěřová, M. (2019). Clinical aspects of Alzheimer's disease. *Clinical biochemistry*, 72, 3-6.
- [8] Gouras, G. K., Olsson, T. T., & Hansson, O. (2015). β -Amyloid peptides and amyloid plaques in Alzheimer's disease. *Neurotherapeutics*, 12, 3-11.
- [9] Brion, J. P. (1998). Neurofibrillary tangles and Alzheimer's disease. *European neurology*, 40(3), 130-140.
- [10] Morris, J. C., McKeel, D. W., Storandt, M., Rubin, E. H., Price, J. L., Grant, E. A., ... & Berg, L. (1991). Very mild Alzheimer's disease: Informant-based clinical, psychometric, and pathologic distinction from normal aging. *Neurology*, 41(4), 469-469.
- [11] David S. Knopman, Helene Amieva, Alzheimer disease, *Nature Reviews Disease Primers*, 10.1038/s41572-021-00269-y, December 2021
- [12] Gomperts, S. N. (2016). Lewy body dementias: dementia with Lewy bodies and Parkinson disease dementia. *Continuum: Lifelong Learning in Neurology*, 22(2 Dementia), 435.
- [13] Wen, J., et al. (2020). Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation. *Medical image analysis*, 63, 101694.
- [14] Udofia, U. (2018). Basic overview of convolutional neural network (cnn). Retrieved May, 27, 2019.
- [15] Davis, R. M. (2021). I, Dentist: Is artificial intelligence the future of oral healthcare? (Doctoral dissertation, Massachusetts Institute of Technology).
- [16] Mohiuddin, M., Islam, M. S., & Kabir, M. H. (2021, September). Performance Analysis of Bearing fault diagnosis using Convolutional Neural Network. In 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON) (pp. 1-6). IEEE.
- [17] Saba, T., Rehman, A., Mehmood, Z., Kolivand, H., & Sharif, M. (2018). Image enhancement and segmentation techniques for detection of knee joint diseases: A survey. *Current Medical Imaging*, 14(5), 704-715.
- [18] AbdulAzeem, Y., Bahgat, W. M., & Badawy, M. (2021). A CNN based framework for classification of Alzheimer's disease. *Neural Computing and Applications*, 33, 10415-10428.
- [19] Al-Khuzai, F. E., Bayat, O., & Duru, A. D. (2021). Diagnosis of Alzheimer disease using 2D MRI slices by convolutional neural network. *Applied Bionics and Biomechanics*, 2021.
- [20] Al-Khuzai, F. E., Bayat, O., & Duru, A. D. (2021). Diagnosis of Alzheimer disease using 2D MRI slices by convolutional neural network. *Applied Bionics and Biomechanics*, 2021.
- [21] Antony, F., Anita, H. B., & George, J. A. (2022). Classification on Alzheimer's Disease MRI Images with VGG-16 and VGG-19. In *IOT with Smart Systems: Proceedings of ICTIS 2022, Volume 2* (pp. 199-207). Singapore: Springer Nature Singapore.
- [22] Liu, J., Li, M., Luo, Y., Yang, S., Li, W., & Bi, Y. (2021). Alzheimer's disease detection using depthwise separable convolutional neural networks. *Computer Methods and Programs in Biomedicine*, 203, 106032.
- [23] Savaş, S. (2022). Detecting the stages of Alzheimer's disease with pre-trained deep learning architectures. *Arabian Journal for Science and Engineering*, 47(2), 2201-2218.
- [24] Marcus, D. S., Wang, T. H., Parker, J., Csernansky, J. G., Morris, J. C., & Buckner, R. L. (2007). Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *Journal of cognitive neuroscience*, 19(9), 1498-1507.
- [25] Sabuncu, M. R., Desikan, R. S., Sepulcre, J., Yeo, B. T. T., Liu, H., Schmansky, N. J., ... & Alzheimer's Disease Neuroimaging Initiative. (2011). The dynamics of cortical and hippocampal atrophy in Alzheimer disease. *Archives of neurology*, 68(8), 1040-1048.
- [26] Nipy developers. (n.d.). Nibabel: Access a cacophony of neuroimaging file formats. Retrieved June 5, 2023, from <https://nipy.org/nibabel/>
- [27] Ruuska, S., Hämäläinen, W., Kajava, S., Mughal, M., Matilainen, P., & Mononen, J. (2018). Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle. *Behavioural processes*, 148, 56-62.
- [28] Shadi, S., Hadi, S., Nazari, M. A., & Hardt, W. (2019). Outdoor navigation for visually impaired based on deep learning. In *Proc. CEUR Workshop Proc* (Vol. 2514, pp. 97-406).