Improved Classification of Alzheimer's Disease With Convolutional Neural Networks

Conference Paper · December 2023

DOI: 10.1109/SPMB59478.2023.10372725

8 authors, including:

Jiblal Upadhya
Middle Tennessee State University
10 PUBLICATIONS **21** CITATIONS
SEE PROFILE

Samir Poudel
Middle Tennessee State University
16 PUBLICATIONS **46** CITATIONS
SEE PROFILE

Mitu Hasan
University of Dhaka
13 PUBLICATIONS **34** CITATIONS
SEE PROFILE

kiran kumar Donthula
Middle Tennessee State University
9 PUBLICATIONS **39** CITATIONS
SEE PROFILE

Classification of Alzheimer's Disease With Convolutional Neural Networks

S. Muhammed[1], J. Upadhya[1], S. Poudel[3], M. Hasan[1],
K. Donthula[4], J. Vargas[1], J. Ranganathan[2] and K. Poudel[3]

[1]Computational and Data Science, Middle Tennessee State University, Murfreesboro, TN, USA
[2]Engineering Technology, Middle Tennessee State University, Murfreesboro, TN, USA
[3]Computer Science, Middle Tennessee State University, Murfreesboro, TN, USA
[4]Senior Software Engineer, Optum Inc, Murfreesboro, TN, USA
{sam2et, ju2i, sp2ai, mh2ay, jorge.vargas}@mtmail.mtsu.edu,
jaishree.ranganathan@mtsu.edu,
kiran.donthula@optum.com,
khem.poudel@mtsu.edu

***Abstract*— This work focuses on classifying MRI images using machine learning models to identify Alzheimer's disease (AD), the most common form of dementia, at an early stage [1]. It is now possible to identify and forecast the onset of AD by analyzing brain scans collected through Magnetic Resonance Imaging (MRI) and using artificial intelligence (AI) technologies, classifying patients as either at risk or not. The main goal is to make precise predictions.**

Improved prediction and detection tools for radiologists, physicians, and caregivers will be made available as a result of the study's determination of the likelihood that individuals would develop AD and proper categorization of them. Using machine learning methods such as Convolutional Neural Network (CNN), Support Vector Machine (SVM), and fastai—a user-friendly deep learning library and framework—we developed models using a dataset of 6400 MRI images from the Alzheimer's Disease Neuroimaging Initiative (ADNI) 4 class for the early detection and classification of AD [2]. The models achieve an impressive accuracy rate of 84.4%, with the addition of a pre-trained Visual Geometry Group (VGG) layer to the sequential model, outperforming other algorithms. This study demonstrates a viable method for classifying and diagnosing Alzheimer's disease early on, utilizing MRI images and machine learning models. The acquired accuracy rates show the models' potential to assist doctors and other caregivers in AD diagnosis and management, thereby maximizing overall effectiveness.

Keywords*— *MRI Images, SVM classifier, Convolutional Neural Network, Deep Learning, fastai, Alzheimer's Disease.

I. Introduction

Alzheimer's disease (AD) is a degenerative condition that progressively impairs cognitive functions, particularly memory, and is the most common cause of dementia in older individuals [3, 4]. It ranks as the seventh leading cause of death worldwide.

Symptoms include memory loss, confusion, speech, and reasoning impairment, which worsen over time. In advanced stages, individuals become unable to communicate or perform basic tasks, requiring constant care. Importantly, Alzheimer's can affect the brain for a decade or more before symptoms manifest. Aberrant protein accumulation in the brain leads to the loss of neuronal functionality and viability. While there is currently no cure, therapies can slow the disease's progression. Diagnosis relies on magnetic resonance imaging (MRI) to determine disease phases and rule out other potential causes of symptoms. Early detection is crucial for timely intervention to maintain daily functioning [5–7].

Consequently, research focuses on leveraging MRI images to identify Alzheimer's early, using machine learning methods including Convolutional Neural Networks, Support Vector Machines, and fastai. The goal is to develop reliable prediction and detection tools to assist healthcare professionals and caregivers, ultimately improving the lives of those affected by this devastating disease.

II. Literature Review

Recent advancements in the field of medical imaging, particularly concerning the identification of Alzheimer's disease through MRI images, have been predominantly centered around leveraging Convolutional Neural Network (CNN) models. These CNN architectures, notably ResNet, have gained considerable attention for their efficacy in analyzing such complex data. However, it's important to note that alongside CNNs, researchers have explored a spectrum of alternative models to address this task. Among these models are Deep Boltzmann Machines (DBM) introduced by Salakhutdinov and Hinton in 2009 [8], which have shown promise in handling intricate patterns within MRI data. Additionally, Support Vector Machines (SVM) proposed by Cortes and Vapnik in 1995 [9] have also been employed in several studies focused on Alzheimer's identification using MRI images. The evolving landscape of research in this domain has culminated in a diverse range of publications, each employing various models

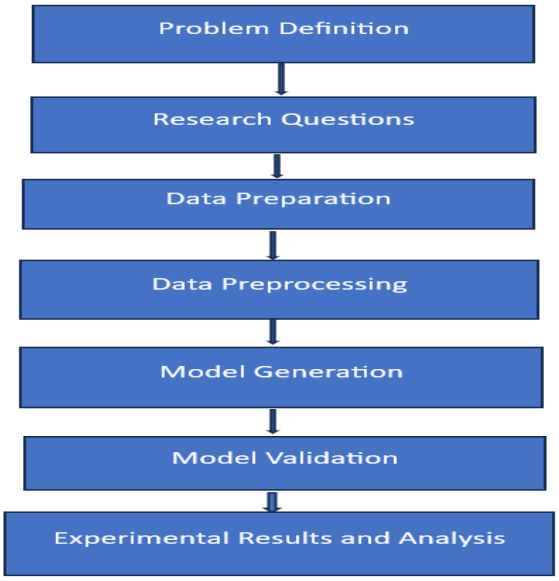

Figure 1. Method Chart

and methodologies. Table 1 provides a comprehensive summary of these publications, detailing the specific models utilized and the reported accuracies achieved in the identification of Alzheimer's disease from MRI images. The amalgamation of these findings presents a comprehensive overview of the efficacy of different models in addressing this critical healthcare challenge.

III. Materials and Methods

The method proposed in this work represents a systematic approach, guiding the progression from problem definition to experimental results and data analysis. This process is described in the following sections and is presented as a flowchart in Figure 1.

III-A. Data Acquisition

We obtained MRI images from the ADNI Dataset [10]. The dataset comprises a total of 6,400 images, categorized as follows: 3,200 are labeled with mild cognitive impairment, 2,240 indicate very mild cognitive impairment, 896 suggesting mild cognitive impairment, and 64 showing moderate cognitive impairment. These images are organized into two directories: one for training images and the other for testing images. Within these directories, images of individuals with Alzheimer's and their respective levels of cognitive impairment are evenly distributed [11].

III-B. Image Pre-Processing

Preprocessing Alzheimer's image data using the Image Data Generator involves a systematic series of steps essential for optimizing data for training or evaluating a convolutional neural network (CNN) model. These steps are chosen with specific objectives in mind and are grounded in best practices:

1) **Import Libraries:** Import essential libraries, such as TensorFlow or PyTorch, for image data manipulation.
2) **Create Image Data Generator:** Instantiate an Image Data Generator to load and augment data, enhancing model generalization.
3) **Rescale Data:** Rescale pixel values to a typical range of 0 to 1 to aid model convergence.
4) **Load and Preprocess Images:** Load and preprocess images from a specified directory, following best practices for feature extraction.
5) **Generate Batches:** Create batches of preprocessed images, customizing parameters like image size and batch size based on task requirements.
6) **Data Splitting:** If separate datasets for training, validation, and testing exist, use the Image Data Generator to split data into training and validation subsets, ensuring independent evaluation.
7) **Convert Images to Suitable Input Data:** Translate pixel data from image files into numerical values (ranging from 0 to 255 for grayscale images), which can be stored in data structures for model input.

These preprocessing steps are chosen based on established principles and can be adapted to specific research goals and dataset characteristics. The systematic approach ensures that data is optimized for effective CNN model training and evaluation.

III-C. Transfer Learning in Convolutional Neural Network

Transfer learning assists in extracting generic features from sizable datasets, enabling the model to generalize effectively to new, smaller datasets [12]. Pre-trained models are regularized on a large dataset to reduce the risk of overfitting new tasks [13].

Differentiated Transfer Learning: In this method, the pre-trained CNN functions as a fixed feature extractor. The features of a new dataset are extracted using the remaining layers after removing the final few layers. These features are then used to feed a newly trained classifier or fully connected layers for the current task.

Fine-tuning: Along with feature extraction, fine-tuning entails unfreezing and retraining a few of the pre-trained model's top layers at a slow learning rate. This allows the model to modify the learned weights to better suit the new assignment.

The choice of a pre-trained model is influenced by several variables, including the size and similarity of the

Table 1. Papers on Alzheimer's disease using machine learning

Paper Title	Authors	Year	Datasets	Algorithms	Results
Diagnosis of Alzheimer's disease using PET images with deep learning	Li, et al.	2021	ADNI	CNN	AUC: 0.973
Machine Learning Techniques for Diagnosis of Alzheimer's Disease: A Review	Balachandar et al.	2020	ADNI, OASIS	SVM, Random Forest, KNN	Acc: 91%, AUC: 0.98
Early Detection of Alzheimer's Disease: A Machine Learning Approach Using Combination of EEG Features and Neurocognitive Tests	Khan et al.	2020	ADNI	SVM	Acc: 96%
A Deep Learning Model for Alzheimer's Disease Diagnosis from Brain MRI Images using Transfer Learning	Rajan et al.	2020	ADNI	VGG16, ResNet50	Acc: 98%, AUC: 0.99
A Machine Learning-Based Method for the Diagnosis of Alzheimer's Disease from Structural MRI Images	Cheng et al.	2019	ADNI	SVM, RF, MLP	Acc: 91%, AUC: 0.95
Automated Diagnosis of Alzheimer's Disease via an Ensemble of Convolutional Neural Networks using Structural and Functional MRI	Talo et al.	2019	ADNI	CNN	AUC: 0.94
Automatic diagnosis of Alzheimer's disease via unsupervised feature learning-based on a deep belief network with compressed sensing MRI	Zhang et al.	2018	ADNI	DBN	Acc: 90%, AUC: 0.92
Predictive analytics with gradient boosting in clinical medicine and healthcare	Al-Jumaili et al.	2018	ADNI, OASIS, AIBL	GBM	Acc: 80%, AUC: 0.86
Diagnosis of Alzheimer's disease using a combination of deep learning and feature extraction from structural MRI and PET images	Li et al.	2018	ADNI	CNN, SVM	Acc: 92%, AUC: 0.96
Deep learning based imaging data completion for improved brain disease diagnosis	Wang et al.	2016	ADNI	DBN	Acc: 93%

new dataset to the pre-training dataset, the complexity of the new task [14, 15], and the available computational resources. Well-known pre-trained models include VGGNet, Inception, and MobileNet.

III-D. The Softmax activation function

The softmax activation function, as shown in equation (1), is employed in the output layer of neural networks for multiclass classification tasks. In such tasks, the objective is to assign an input to one of several possible classes. The softmax function interprets the network's output as class probabilities, facilitating decision-making based on the class with the highest probability.

$$\text{softmax}(\mathbf{z})_i = \frac{e^{z_i}}{\sum_j e^{z_j}} \tag{1}$$

III-E. VGG16 in Convolutional Neural Network

The VGG16 architecture consists of 16 layers, including 13 convolutional layers and 3 fully connected layers [16]. The convolutional layers primarily utilize 3x3 filters with a stride of 1, followed by max-pooling layers with a 2x2 pool size and a stride of 2. This use of smaller filters and pooling layers helps capture detailed features at different spatial scales.

The VGG16 network has a fixed input size of 224x224 pixels, meaning input images must be preprocessed to this size before feeding them into the network. For convolution operations, strides 1, 2, and 3, and a first filter with 32 3x3 kernels are employed on an image consisting of 8 blocks with a kernel size of 45x45x45. These kernels act as feature finders, convolving with the image to produce a set of features that have been convolved. The size of the kernel signifies a neuron's receptive field, enforcing regional connections between neurons and the previous volume in the neural network.

The pooling layer: The pooling layer utilizes the max-pooling aggregation function. It selects the maximum value within an area defined by the kernel size (k) in an input of size h x w with a stride (s). This layer effectively summarizes the outputs of adjacent sets of

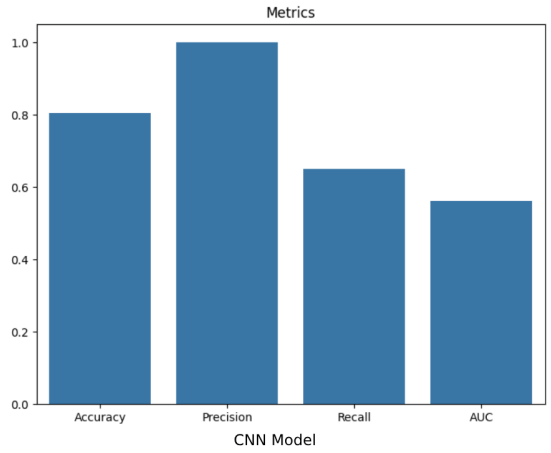

Figure 2. Evaluation graph

inputs while reducing the dimensionality of the inputs. It performs downsampling when dealing with high-resolution images, reducing spatial dimensions while preserving important information and minimizing the number of parameters. In this work, only one pooling layer is employed.

Dropout: Dropout serves as a regularization method in neural networks, randomly deactivating neurons in hidden layers during training using a specific probability 'r', effectively zeroing their output. Implemented post the pooling layers, this approach involves two dropout layers with respective dropout rates of 0.25 and 0.5 within the architecture. Notably, dropout has no bearing on the forward or backward steps of backpropagation. It aids in curbing overfitting by introducing redundancy and enhancing resilience, all while minimally impacting the training procedure.

IV. Results and Discussion: CNN Model

In this study, we utilized MRI scan images categorized into three groups: mild dementia, non-dementia, and very mild dementia. For training and validation, 80

The CNN architecture utilized in our study comprises several layers, including a convolutional layer, an activation layer, a pooling layer, and a fully connected layer, similar to the architecture described in [17, 18]. The initial layer, known as the convolutional layer, employs a kernel or filter to analyze the input image and extract features that aid in determining whether the image corresponds to an Alzheimer's patient. Subsequently, the activation layer is the second component of the architecture, utilizing a Rectified Linear Unit (ReLU) to enhance the model's capacity to capture nonlinear relationships in the data, thereby improving training speed. Finally, flattening the feature matrix into a vector format and passing it through the fully connected layer further enhances the model's training performance.

To assess the performance of our models, we examined confusion matrices to compare their performance. Evaluation metrics such as sensitivity, specificity, and accuracy were used. The formulas for these parameters are as follows:

$$Sensitivity(Recall) = \frac{TP}{TP+FN} \quad (2)$$

$$Specificity(Selectivity) = \frac{TN}{TN+FP} \quad (3)$$

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP} \quad (4)$$

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

The model's accuracy in predicting Alzheimer's disease using VGG16 transfer learning on CNN is 80.4%. Precision, which measures the proportion of correctly predicted positive cases out of all predicted positive cases, is not specified here. Recall, also known as sensitivity or the true positive rate, is 62%. In the context of Alzheimer's disease prediction, this means that the model successfully identified 62% of the patients who have the disease. The AUC (Area Under the Curve) is 0.56, which is a measure of the model's ability to distinguish between positive and negative cases. In this case, the AUC of 0.56 indicates that the model has a moderate level of discriminative power. Overall, these metrics suggest that the VGG16 transfer learning model on a CNN has achieved a reasonable level of accuracy in predicting Alzheimer's disease. However, there is still room for improvement, particularly in terms of recall and AUC.

V. Results and Discussion: Support Vector Machine

Support Vector Machines (SVMs) in image classification involve the extraction of features from images, training an SVM on these features, and then using the trained model to classify new images. The primary objective of SVM is to discover the optimal hyperplane that maximizes the separation between classes in the feature space. It accomplishes this by establishing a decision boundary that maximizes the margin (distance) between data points from different classes. SVMs can incorporate different kernel functions to capture various patterns and adapt to diverse data distributions.

The accuracy score obtained represents the proportion of correctly classified instances out of the total number of instances. In this case, the SVM model achieved an accuracy of 70%, indicating that 70% of the instances were correctly classified. Precision, with a score of 0.49, suggests that out of all the instances predicted as positive, only approximately 49% were true positive cases. A recall score of 0.7 indicates that the model identified 70% of the actual positive cases correctly. The F1 score, which is the harmonic mean of precision and recall,

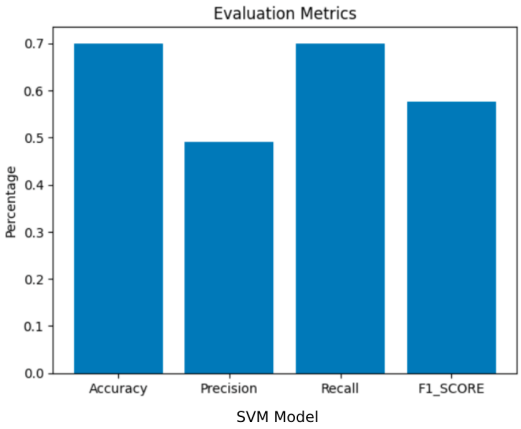

Figure 3. SVM Evaluation Graph

provides a single metric that balances both precision and recall for the classification task. The F1 score of 0.576 indicates a reasonably balanced performance between precision and recall for the classification task.

In predicting Alzheimer's disease, achieving high values for accuracy, precision, recall, and F1 score is desirable, as it indicates a good ability to correctly classify both positive and negative instances. Since we have experimented with limited data resources, the obtained metrics are significant enough for interpreting the efficacy of the models.

VI. Adding a pre-trained VGG Model as a Layer to the Sequential Model

The model architecture consists of several dense layers with batch normalization, dropout layers for regularization, and activation functions to introduce non-linearity. The final output layer uses the softmax activation function for multi-class classification. A sequential model was utilized and initialized, enabling us to stack layers sequentially.

Accuracy measures the overall correctness of the model's predictions. An accuracy of 0.844 indicates that the model's predictions match the true labels for approximately 84.4% of the samples. However, it is important to consider precision, recall, and AUC as well, as they provide more insights into the model's performance, especially in cases of imbalanced datasets or when specific class predictions are of greater importance.

Precision is the measure of how well the model predicts positive samples correctly. A precision of 0.903 indicates that out of all the samples, the model predicted as positive, approximately 90.3% of them are truly positive.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 vgg16 (Functional)          (None, 7, 7, 512)         14714688

 dropout (Dropout)           (None, 7, 7, 512)         0

 flatten_1 (Flatten)         (None, 25088)             0

 batch_normalization (BatchN  (None, 25088)            100352
 ormalization)

 dense (Dense)               (None, 2048)              51382272

 batch_normalization_1 (Batc  (None, 2048)             8192
 hNormalization)

 activation (Activation)     (None, 2048)              0

 dropout_1 (Dropout)         (None, 2048)              0

 dense_1 (Dense)             (None, 1024)              2098176

 batch_normalization_2 (Batc  (None, 1024)             4096
 hNormalization)

 activation_1 (Activation)   (None, 1024)              0

 dropout_2 (Dropout)         (None, 1024)              0

 dense_2 (Dense)             (None, 4)                 4100

=================================================================
Total params: 68,311,876
Trainable params: 53,540,868
```

Figure 4. Model Summary

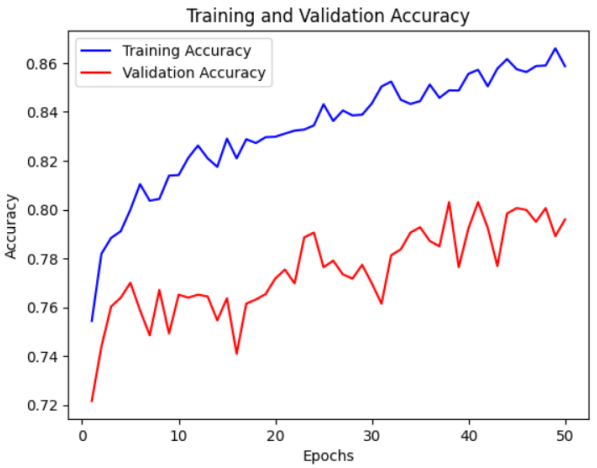

Figure 5. Training and Validation Accuracy Model

Recall, also known as sensitivity or true positive rate, measures the model's ability to identify positive samples correctly. A recall of 0.709 indicates that the model correctly identified approximately 70.9% of the positive samples.

AUC (Area Under the Curve) represents the performance of the model across all possible classification thresholds. An AUC of 0.638 suggests that the model has a moderate level of discriminatory power.

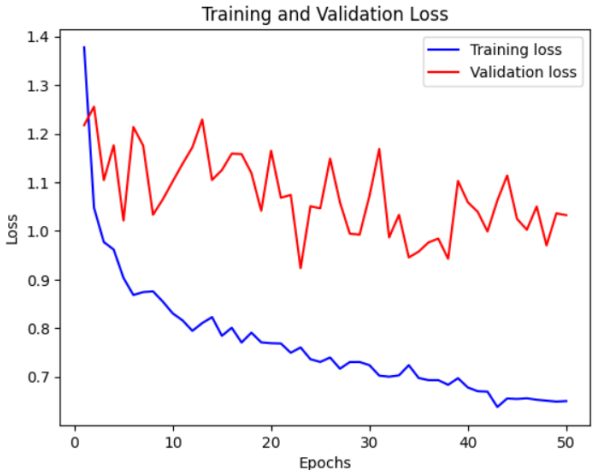

Figure 6. Training and Validation Loss Model

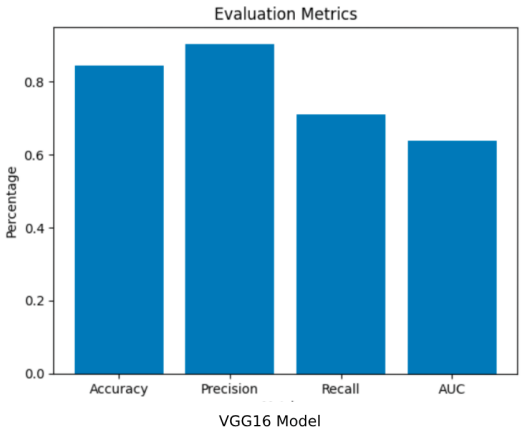

Figure 7. VGG16 Evaluation Graph

In our study, we also ventured into the realm of deep learning, utilizing the fastai framework with the overarching goal of democratizing the field of artificial intelligence. fastai, as introduced by Howard et al. in their seminal work [19], is a powerful tool for simplifying and accelerating the process of training deep learning models. While we employed fastai for experimentation in our research, it's important to note that we did not present specific results stemming from this endeavor in this paper. Our exploration with fastai served as a valuable component of our methodology, laying the foundation for potential future investigations and shedding light on the immense potential of democratizing deep learning for broader applications in the field of medical imaging and beyond.

VII. Conclusions

The primary objective of this research was to develop accurate algorithms for the early detection of Alzheimer's disease. We are pleased to report that our models achieved an accuracy of 80.4%, with the convolutional neural network (CNN) and the incorporation of a pre-trained VGG layer into the sequential model emerging as the top-performing algorithms with 84.4% accuracy. Our approach involved a comprehensive exploration of various machine learning algorithms to identify those capable of timely and precise disease detection.

Model	Accuracy	Precision	Recall	AUC
CNN Model	80.4%	-	62%	0.56
SVM	70%	0.49	70%	-
VGG Model	84.4%	0.903	70.9%	0.638

Table 2. Model Performance for Alzheimer's Disease Prediction

One key factor contributing to the success of our models was their ability to effectively address the overfitting challenge commonly associated with machine learning techniques. Overcoming overfitting is crucial in maintaining both the accuracy of disease classification and computational efficiency. The incorporation of a pre-trained VGG layer into the sequential model, in particular, demonstrated promising results for early disease detection. This success can be attributed to the transfer learning capabilities of pre-trained models, allowing us to harness knowledge gained from extensive training on large datasets.

As we look ahead, there remains substantial potential for further improvement in our classification accuracy. This can be achieved by delving deeper into advanced machine learning models, perhaps in isolation, to explore their capabilities for enhancing Alzheimer's disease detection. The outcome of our research underscores the promise of machine learning in this vital area of healthcare and paves the way for continued advancements. By continually refining and innovating these algorithms, we can make significant strides in early Alzheimer's disease detection, ultimately benefiting patients and healthcare providers alike.

Acknowledgements

We would like to thank the Computer Science, Computational and Data Science, and Engineering Technology departments for providing the necessary resources, facilities, and research environment that facilitated the smooth progress of this study.

References

[1] A. W. Salehi, P. Baglat, B. B. Sharma, G. Gupta, and A. Upadhya, "A cnn model: earlier diagnosis and classification of alzheimer disease using mri," *2020 International Conference on Smart Electronics and Communication (ICOSEC)*. IEEE, 2020, pp. 156–161.

[2] M. Tanveer, B. Richhariya, R. U. Khan, A. H. Rashid, P. Khanna, M. Prasad, and C. Lin, "Machine learning techniques for the

diagnosis of alzheimer's disease: A review," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 16, no. 1s, pp. 1–35, 2020.

[3] M. Prince, E. Albanese, M. Guerchet, and M. Prina, "World alzheimer report 2014. dementia and risk reduction: An analysis of protective and modifiable risk factors." Ph.D. dissertation, Alzheimer's Disease International, 2014.

[4] A. Association *et al.*, "2016 alzheimer's disease facts and figures," *Alzheimer's & Dementia*, vol. 12, no. 4, pp. 459–509, 2016.

[5] C. L. Masters, R. Bateman, K. Blennow, C. C. Rowe, R. A. Sperling, and J. L. Cummings, "Alzheimer's disease," *Nature reviews disease primers*, vol. 1, no. 1, pp. 1–18, 2015.

[6] C. A. Lane, J. Hardy, and J. M. Schott, "Alzheimer's disease," *European journal of neurology*, vol. 25, no. 1, pp. 59–70, 2018.

[7] C. Kavitha, V. Mani, S. Srividhya, O. I. Khalaf, and C. A. Tavera Romero, "Early-stage alzheimer's disease prediction using machine learning models," *Frontiers in public health*, vol. 10, p. 853294, 2022.

[8] R. Salakhutdinov and G. Hinton, "Deep boltzmann machines," *Artificial intelligence and statistics*. PMLR, 2009, pp. 448–455.

[9] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, pp. 273–297, 1995.

[10] C. R. Jack Jr, M. A. Bernstein, N. C. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, P. J. Britson, J. L. Whitwell, C. Ward *et al.*, "The alzheimer's disease neuroimaging initiative (adni): Mri methods," *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 27, no. 4, pp. 685–691, 2008.

[11] M. Odusami, R. Maskeliūnas, R. Damaševičius, and S. Misra, "Explainable deep-learning-based diagnosis of alzheimer's disease using multimodal input fusion of pet and mri images," *Journal of Medical and Biological Engineering*, pp. 1–12, 2023.

[12] S. Poudel, R. Paudyal, B. Cankaya, N. Sterlingsdottir, M. Murphy, S. Pandey, J. Vargas, and K. Poudel, "Cryptocurrency price and volatility predictions with machine learning," *Journal of Marketing Analytics*, 08 2023.

[13] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[14] M. Hasan, S. Hamdan, S. Poudel, J. Vargas, and K. Poudel, "Prediction of length-of-stay at intensive care unit (icu) using machine learning based on mimic-iii database," 06 2023, pp. 321–323.

[15] N. J. Dhinagar, S. I. Thomopoulos, P. Rajagopalan, D. Stripelis, J. L. Ambite, G. Ver Steeg, and P. M. Thompson, "Evaluation of transfer learning methods for detecting alzheimer's disease with brain mri," *18th International Symposium on Medical Information Processing and Analysis*, vol. 12567. SPIE, 2023, pp. 504–513.

[16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[17] M. Chowdhury, C. Li, and K. Poudel, "Combining deep learning with traditional machine learning to improve phonocardiography classification accuracy," *2021 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*. IEEE, 2021, pp. 1–5.

[18] O. Papadimitriou, A. Kanavos, P. Mylonas, and M. Maragoudakis, "Classification of alzheimer's disease subjects from mri using deep convolutional neural networks," *Novel & Intelligent Digital Systems Conferences*. Springer, 2023, pp. 277–286.

[19] J. Howard and S. Gugger, "Fastai: A layered api for deep learning," *Information*, vol. 11, no. 2, p. 108, 2020.