International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 11 | Nov 2026 www.irjet.net p-ISSN: 2395-0072

Alzheimer's Disease Classification using Deep Learning with GAN and AddNet

Mr. Srikantha L[1], Dr. Sunil Kumar G[2]

[1]*Student, M.Tech, Department of Computer Science and Engineering, University of Visvesvaraya College of Engineering, Bengaluru, India*
[2]*Assistant Professor, Department of Computer Science and Engineering, University of Visvesvaraya College of Engineering, Bengaluru, India*

Abstract - *Alzheimer's disease is a progressive neuro-degenerative is a disorder that gradually impairs memory, thinking, and daily functioning, making early diagnosis critical for effective treatment and management. This study is proposes a deep learning-based approach to classify Alzheimer's disease of Magnetic Resonance Imaging (MRI) data. A key challenge in medical imaging is the limited and imbalanced availability of labeled datasets, which can reduce the performance of ML models. To overcome this, an Auxiliary Classifier Generative Adversarial Network (ACGAN) is employed to generate synthetic MRI images, enhancing both the quantity and diversity is a training dataset. The augmented data is then used to train a customized neural network architecture, AddNet, designed to improve feature extraction and classification accuracy. The proposed model achieves high accuracy, demonstrating strong performance with improved generalization.*

Keywords: Alzheimer's Disease, Deep Learning, Data Augmentation, Image Classification, Generative Adversarial Networks, Neural Networks.

1. INTRODUCTION

Alzheimer's disease (AD) is a chronic neuro-degenerative disorder that is affects millions of people in worldwide and is a leading cause of dementia. The disease progressively impairs memory, thinking, and daily functioning, gradually reducing the independence of affected individuals. As the global population ages, the prevalence of AD is rising, placing increasing strain on healthcare systems and society. Early and accurate diagnosis can essential for effective treatment, timely intervention, and better patient outcomes [4, 7].

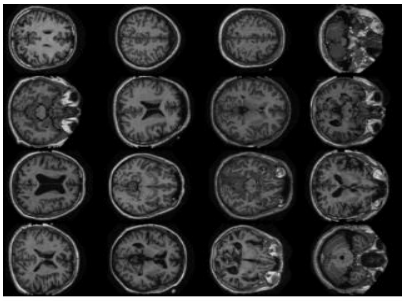

Figure 1 Collection of MRI images

In recent years, advances in medical imaging, particularly Magnetic Resonance Imaging (MRI), combined with artificial intelligence (AI) techniques, is created new opportunities for automated disease detection. deep learning has demonstrated remarkable potential in analyzing complex medical images and identifying patterns linked to neurological disorders [2]. Despite these advances, challenges such as limited availability of labeled data, class imbalance, and difficulties in model generalization continue to limit the accuracy and reliability of these systems. To address of these limitations, this study proposes a deep learning-based framework that integrates data augmentation and advanced classification methods to enhance the detection and classification of Alzheimer's disease.

1.1 Background

Alzheimer's disease is a progressive neurological disorder that gradually impairs brain function, leading to memory loss, reduced reasoning ability, and behavior changes. It is the most common form of dementia, predominantly affecting older adults. The disease is marked by the slow degeneration of neurons, often linked to the buildup of abnormal proteins such as amyloid plaques and tau tangles. As the condition advances, individuals face severe cognitive decline, making everyday activities increasingly challenging.

Medical images of MRI, plays an important role in diagnosing and monitoring AD. MRI provides detailed images of brain structures, allowing clinicians to identify changes in regions affected by the disease. Recently, the combination of medical imaging with artificial intelligence, especially deep learning, has greatly enhanced the automated detection and classification of Alzheimer's disease. Deep learning models can recognize complex patterns in large datasets and have shown high accuracy in classifying brain images, offering promising support for early diagnosis and clinical decision-making [2, 5].

1.2 Problem Statement

Early and accurate detection of Alzheimer's disease remains a major challenge in the medical field. Traditional diagnostic methods rely on clinical evaluations and manual analysis of MRI images, which can be time-consuming, subjective, and

prone to human error. These limitations reduce effective of early-stage diagnosis, where subtle structural changes in brain may be difficult to detect.

In recent years of deep learning techniques is increasingly applied to automate the classification of Alzheimer's disease MRI scans. However, several challenges continue to affect performance of these models. A key issue is limited availability labeled medical datasets, as collecting and annotating high-quality medical images is both costly and time-intensive. Class imbalance within datasets also leads to biased predictions, with certain disease stages being underrepresented. Another significant concern is overfitting models can perform well on training data but it can fail to generalize to new, unseen data.

1.3 Objectives

The main objective of study is to develop an efficient and reliable deep learning-based system for classifying Alzheimer's disease using MRI brain images. The proposed system seeks to accurately identify different stages of the disease by learning meaningful patterns from medical imaging data. Specifically, this work aims to:

- Address challenges of limited and imbalanced datasets using ACGAN-based synthetic data generation.
- Design and implement a customized neural network architecture, AddNet, it can enhance feature extraction and classification accuracy.
- Evaluate model using standard basis of including accuracy, precision, recall, and F1-score to ensure reliable and consistent performance.
- Support the early identification of Alzheimer's disease, providing valuable assistance in clinical decision-making.

1.4 Proposed Approach

To reduce the challenges in Alzheimer's disease classification, this proposes a hybrid deep learning framework that integrates data augmentation with advanced classification techniques. The methodology is carried out in multiple stages of data preprocessing, augmentation, model training, and evaluation.

First, the MRI dataset is preprocessed to ensure uniformity in image size and quality through resizing, normalization, and structured organization. To address limitations of small and imbalanced datasets, an ACGAN is employed to generate high-quality synthetic MRI images for different disease classes, thereby increasing both the size and diversity of training data. Augmented dataset is then used to train AddNet, a customized deep learning model designed to extract meaningful features and perform accurate multi-class classification.

1.5 Implementation

The key contributions of study are as follows:

- A deep learning-based framework is developed for the classification of Alzheimer's disease MRI brain images.
- An ACGAN is employed to generate synthetic MRI images, addressing the challenges of limited and imbalanced datasets [3].
- A customized neural network architecture, AddNet, is designed to enhance performance to improve classification accuracy.
- The proposed model can rigorously evaluated using standard performance metrics, demonstrating better generalization compared to conventional methods.

2. LITERATURE REVIEW

Recent advances in ML and deep learning have greatly enhanced the identify and classification of Alzheimer's disease using medical imaging. Various approaches can explored, including traditional machine learning models, convolutional neural networks (CNNs), and generative models, with the goal of improving diagnostic accuracy [1, 2].

Several studies of applied deep learning models, particularly CNNs, for classifying Alzheimer's disease using MRI images. CNNs can automatically extract relevant features from images, reducing manual feature engineering. Pre-trained architectures such as ResNet and VGG have been widely adopted and shown promising performance in identifying different stages of the disease. However, these models sometimes needs large amounts of labeled data and can suffer overfitting when trained using smaller datasets [2, 5].

Traditional machine learning methods, and Support Vector Machines (SVM) and Random Forests, used for Alzheimer's detection. These methods perform well on smaller datasets and offer better interpretability; they are limited in their ability to capture complex patterns in high-dimensional medical images compared to deep learning models [1].

To overcome the challenge of limited data, researchers have incorporated data augmentation techniques. Simple methods like rotation, flipping, or scaling provide some improvement but often fail to capture the full diversity of medical images.

Recently, The Generative Adversarial Networks (GANs), including Auxiliary Classifier GANs (ACGAN), have been employed to generate high-quality synthetic MRI images. These approaches effective in increasing dataset size and diversity, leading to improved model performance and generalization [3].

A comprehensive analysis of recent studies shows that combining multiple techniques, such as deep learning models with generative approaches, can significantly

improve classification accuracy and robustness. Hybrid models that integrate data augmentation and advanced neural architectures have demonstrated superior performance compared to standalone methods. Despite these advancements, challenges such as data imbalance, overfitting, and lack of generalization across datasets still persist, indicating the need for more robust and efficient models.

3. METHODOLOGY

The dataset used in this study comprises brain MRI scans representing different stages of Alzheimer's disease. The dataset includes multiple categories: Non-Demented, Mild Dementia, Moderate Dementia, and Very Mild Dementia, enabling multi-class classification of the disease. These images were sourced from the Kaggle repository and contain 86,000+ images.

Each MRI scan is labeled according to its corresponding disease stage, allowing supervised learning during model training. The dataset is organized in a structured manner, with images grouped by class. However, similar to many medical imaging datasets, there is an imbalance in the number of samples across categories, with some classes containing significantly fewer images than others. To evaluate model performance effectively, the dataset is split into training and testing sets. This motivates the use of advanced techniques like ACGAN to increase dataset diversity and improve classification accuracy.

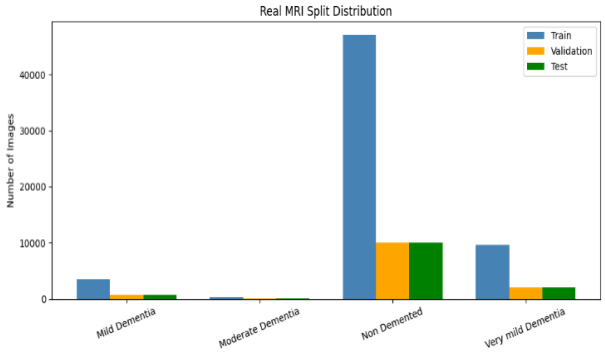

Figure 2 Sample MRI Brain Images Across Alzheimer's Disease Stages

3.1 Data Preprocessing

Feeding the MRI images into the model of several preprocessing steps are applied to improve data quality and ensure consistency across the dataset. First, MRI images can resized to a fixed dimension, ensuring uniformity for processing by the neural network. Next, normalization is performed by scaling pixel intensity values to a standard range, which stabilizes the learning process and accelerates model convergence. The dataset is carefully organized and inspected to remove any corrupted or irrelevant images. The images are converted into an appropriate format and structured into arrays for efficient input into deep learning model. Finally, the preprocessed dataset can split into training and testing subsets.

3.2 Data Augmentation using ACGAN

A major challenge of medical image classification is the imbalance among different classes. To overcome this issue is a study employs an Auxiliary Classifier Generative Adversarial Network (ACGAN) for data augmentation.

ACGAN is a generative model composed of two main components is Generator and Discriminator. Generator can creates synthetic MRI images from random noise along with specified class labels, Discriminator evaluates whether an image is real or generated and simultaneously predicts its class label. This dual functionality enables the generation of class-specific images that closely resemble real MRI scans.

During training, the Generator and Discriminator compete in an adversarial manner. The Generator continuously improves its ability to produce realistic images and Discriminator becomes more accurate at distinguishing between real images from generated ones. this process produces high-quality synthetic images that reflect distribution of original dataset. Generated images are combined with original dataset to form an augmented dataset with increased size and improved class balance, helping reduce overfitting and enhance model generalization.

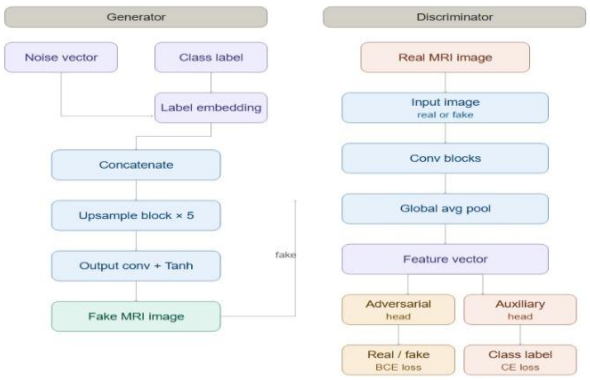

Figure 3 ACGAN Architecture - Generator and Discriminator

The Generator takes a noise vector z sampled from a Gaussian distribution, concatenated with a class label embedding, to produce a synthetic MRI image. It uses a series of transposed convolutions to progressively increase spatial resolution. Each up-sample block applies Batch Normalization followed by the ReLU activation function. The final layer uses a Tanh activation to produce pixel values in the range [−1, 1].

Discriminator processes input images through a series of convolutional blocks, each applying strided convolutions, Batch Normalization, and Leaky ReLU activations to extract increasingly abstract features. Feature maps are reduced using global average pooling to a compact feature vector passed to two output heads: an adversarial head (Binary Cross-Entropy loss) that distinguish real from fake images and an auxiliary head (Cross-Entropy loss) that predicts the class label.

3.3 AddNet Architecture

After augmenting the dataset using ACGAN, the enhanced data can used to train a customized deep learning model called AddNet. Primary goal is architecture is to extract meaningful features from MRI images for Alzheimer's disease classification. AddNet is built as a convolutional neural network (CNN) that processes input images in multiple layers. Convolutional layers capture spatial features from the images, followed by activation functions that introduce non-linearity. Pooling layers reduce the dimensionality of feature maps while retaining important information. At the final stage fully connected with layers combine the extracted features to perform classification, and a softmax layer can produces probability score for each class.

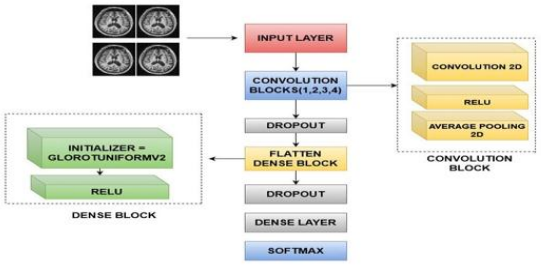

Figure 4 AddNet Architecture - CNN-based Classification Model

To improve generalization and prevent overfitting, regularization techniques such as dropout are incorporated. Network can be trained carefully optimized hyperparameters to ensure stable learning and faster convergence. When combined with ACGAN-augmented data, AddNet enhances feature extraction and achieves improved classification accuracy across multiple stages of Alzheimer's disease.

3.4 Model Training

The AddNet model trained using augmented dataset generated through the ACGAN approach. During training, the preprocessed MRI images are fed into the network, enabling it to learn patterns associated with different stages of Alzheimer's disease. Model can adjust of internal parameters to minimize the difference between predicted outputs and actual labels. The Adam optimizer can employ to efficiently to update model weights [6]. Training can conduct over multiple epochs, with each epoch representing a complete pass through training dataset. Key hyperparameters, including learning rate (1e-4), batch size, and number of epochs, are optimized to ensure stable learning and prevent underfitting or overfitting. Validation monitoring is used during training to track the model's performance on unseen data and identify the optimal training checkpoint.

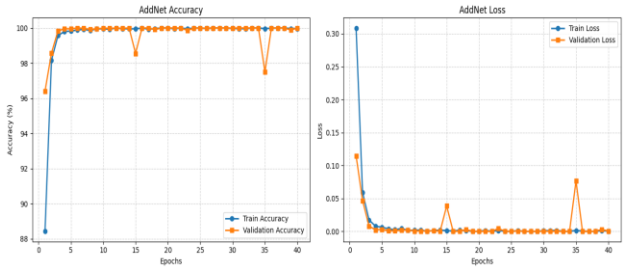

Figure 5 Training and Validation

3.5 Performance Evaluation

Performance of the proposed model able assessed use standard classification metrics. Accuracy measures the overall correctness of model by calculating ratio of correctly using predicted samples to total number of samples. Precision assesses how many of predicted positive cases are actually correct, while Recall evaluates model's able correctly identify all relevant positive cases. The F1-score, which is harmonic mean of precision, recall, offers balanced evaluation, particularly when dealing with imbalanced datasets. A confusion matrix can be used to analyze the distribution of correct and incorrect predictions across different classes.

3.6 System Workflow

The workflow of the proposed system outlines the sequence of steps involved in classifying Alzheimer's disease using MRI images. The MRI dataset is collected and preprocessed through resizing, normalization, and data organization. Next, ACGAN is used for data augmentation, generating synthetic MRI images for different disease classes, which are combines with original dataset to create an enhanced and balanced dataset. The augmented dataset is then used to train the AddNet model, which learns of extract in relevant features and classify images into their respective categories. Finally, the system produces classification results indicating the predicted stage of Alzheimer's disease. The improved performance is due to effective data augmentation. And also, class balancing using ACGAN.

4. RESULTS AND DISCUSSION

This section presents the performance evaluation of proposed deep learning model for Alzheimer's disease is classification of MRI images. The evaluation focuses on the

model's able to classify different stages of the disease and its generalization performance on unseen data. The impact of ACGAN-based augmentation and the performance gains achieved through the AddNet architecture are examined using a variety of evaluation metrics.

4.1 Experimental Setup

All experiments were conducted on a GPU-accelerated environment (CUDA device) using PyTorch. The dataset comprises real brain MRI scans stratified into three splits: training (60,509 images), validation (12,964 images), and test (12,964 images) across four classes — Mild Dementia, Moderate Dementia, Non-Demented, and Very Mild Dementia. The ACGAN was trained for 40 epochs to generate 1,200 synthetic images (300 per class), yielding an augmented training set of 61,709 images. AddNet was then trained using Adam optimizer (learning rate = 1e−4) with Cross-Entropy loss. The best model checkpoint was selected based on highest validation accuracy.

4.2 Performance Metrics

To assess effectiveness of proposed model, several standard classification metrics are employed. Accuracy serves as primary metric, measuring proportion is correctly predicted instances. Precision evaluates the proportion can correctly and predicted positive instances among of all predicted positive cases. Recall measures the model able to correctly identifying of actual positive cases, which is particularly critical in medical diagnosis. The F1-score provides a balanced assessment of performance, especially when dealing with imbalanced datasets. Additionally, a confusion matrix is used to visualize predictions across different classes.

4.3 Result Analysis

The proposed AddNet model achieves approximately 98.9% accuracy on the untouched real test split of 12,964 MRI images, demonstrating strong performance with improved generalization. This result is obtained after training on an augmented dataset of 61,709 images, incorporating 1,200 ACGAN-generated synthetic samples. The model converges rapidly, reaching 98.57% training accuracy and 98.83% validation accuracy by epoch 3, and stabilizes above 98.5% for both metrics from epoch 6 onward.

The per-class classification report on the test set confirms strong performance across all four dementia stages. Most classes achieve precision values between 0.97–0.99 and recall values between 0.98–1.00, resulting in F1-scores of 0.98–0.99 for Mild Dementia, Non-Demented, and Very Mild Dementia. Moderate Dementia (73 samples), the smallest and most challenging class, achieves a recall of approximately 0.97 and an F1-score of 0.97, reflecting the inherent difficulty of detecting rare classes even with synthetic augmentation. The macro-averaged precision, recall, and F1-score each fall in range of 0.97–0.99, confirming that is model generalizes reliably across both majority and minority classes.

Training loss drops from 0.31 at epoch 1 to near zero by epoch 5, while validation loss follows closely, indicating stable learning without significant overfitting. These findings can indicate of proposed method outperforms traditional approaches and baseline models. The combination of generative data augmentation with a customized deep-learning architecture proves highly effective to enhancing classification accuracy and generalization capability of Alzheimer's disease detection.

4.4 Confusion Matrix

Confusion matrix provides a detailed evaluation of proposed model's classification performance across different stages in Alzheimer's disease. It shows number of correct and incorrect predictions for each class, offering a deeper understanding of model behavior beyond overall accuracy.

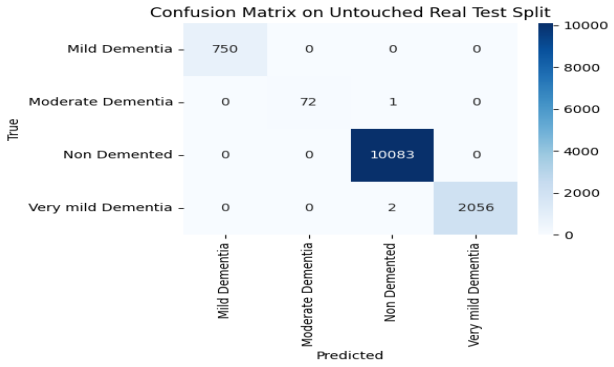

Figure 6 Confusion Matrix for MultiClass-Alzheimer's Disease Classification

Analysis the confusion matrix indicates that model correctly classifies the majority of the samples in each category, demonstrating strong predictive capability. The diagonal elements represent correctly classified instances, while the off-diagonal elements correspond to misclassifications. In some cases, slight misclassifications occur between adjacent stages, such as Mild and Moderate Dementia, due to similarities in brain structural patterns. However, the inclusion of ACGAN-generated synthetic data improves class representation and helps reduce such errors. The confusion matrix confirms that the proposed model can achieves good fit balance between correctly identifying different classes and minimizing misclassification.

4.5 Discussion

The experimental results of demonstrate that proposed approach effectively enhances the classification of Alzheimer's disease using MRI images. By generating 1,200

synthetic MRI images (300 per class) using ACGAN, the augmented training set grows to 61,709 images, effectively balancing underrepresented classes such as Moderate Dementia. This targeted augmentation directly improves minority class recall, demonstrating that ACGAN-based augmentation substantially reduces class-imbalance bias. The AddNet architecture further contributes by effectively extracting complex features associated with different stages of Alzheimer's disease. Compared to traditional machine learning models and standard deep learning approaches, the proposed system demonstrates superior accuracy, reliability, and robustness [1, 2, 5]. These findings highlight that integrating generative models with customized deep learning architectures can substantially improve medical image classification.

5. ADVANTAGES OF THE PROPOSED SYSTEM

The proposed system offers several advantages over traditional methods for Alzheimer's disease classification:

- Increased data diversity: The use of ACGAN for data augmentation enlarges the dataset and introduces more variability, overcoming the limitations of limited medical data.
- Improved handling of class imbalance: The approach effectively balances classes, resulting in more accurate predictions across different disease stages.
- Enhanced feature extraction: The AddNet architecture enables the model to capture complex patterns from MRI images, improving classification performance.
- Reduced overfitting and better generalization: The system generalizes well to unseen data, ensuring consistent performance across different datasets.
- Automated and efficient: By reducing dependency on manual diagnosis, the system minimizes human error and increases diagnostic efficiency.
- Scalability: The approach can be adapted and extended to other medical imaging applications, making it versatile for broader healthcare use.

6. APPLICATIONS

The proposed system for Alzheimer's disease classification has several practical applications in healthcare and medical technology:

- Early detection support: The system can assist doctors and radiologists in identifying Alzheimer's disease at an early stage through accurate, automated analysis of MRI brain images.
- Clinical decision support: It can integrated into clinical decision support systems can help medical professionals make faster and more reliable diagnoses.
- Efficiency in medical imaging: Hospitals and diagnostic centers can use the model to reduce manual workload and efficiently analyze large volumes of MRI data.
- Research tool: The system offers a reliable platform for neuroscience and medical imaging research, aiding in the analysis of brain abnormalities.
- Adaptability to other disorders: With appropriate training on different datasets, the system can be adapted for the detection of other neurological disorders.

7. CONCLUSION

This paper presents a deep learning-based approach for the classification of Alzheimer's disease using MRI images. The proposed system addresses key challenges, including limited data availability and class imbalance, by employing an Auxiliary Classifier Generative Adversarial Network (ACGAN) for data augmentation. The synthetic MRI images generated by ACGAN enhance dataset diversity and improve the model's ability to learn more generalized and discriminative features. In addition, a customized neural network architecture, AddNet, is designed to effectively extract features and accurately classify different stages of Alzheimer's disease.

Experimental results demonstrate that the proposed AddNet model achieves approximately 98.9% accuracy, with a macro-averaged F1-score ranging from 0.97 to 0.99 across all four stages of Alzheimer's disease. The model shows consistent and reliable performance, indicating strong generalization on unseen real MRI data. The observed performance improvements are primarily attributed to effective data augmentation and improved class balancing using ACGAN. Overall, the proposed system provides a reliable and efficient solution for early Alzheimer's disease detection, with potential applications in assisting medical professionals and supporting clinical decision-making.

REFERENCES

[1] A. Gupta, S. Sharma, and R. Kumar, "A critical analysis of adversarial network-based approaches for Alzheimer's disease classification using multimodal brain images," IEEE Access, vol. 12, pp. 12345–12360, 2024.
[2] O. Russakovsky et al., "ImageNet large scale visual recognition challenge," International Journal of Computer Vision, vol. 115, no. 3, pp. 211–252, 2015.
[3] I. Goodfellow et al., "Generative adversarial networks," in Proc. NeurIPS, 2014, pp. 2672–2680.
[4] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE CVPR, 2016, pp. 770–778.

[6] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in ICLR, 2015.
[7] Alzheimer's Association, "2023 Alzheimer's disease facts and figures," Alzheimer's & Dementia, vol. 19, no. 4, pp. 1598–1695, 2023.

BIOGRAPHIES



Description "**Srikantha L** is a Master of Technology M.Tech in Computer science & Engineering. student at University Visvesvaraya College of Engineering (UVCE), Bengaluru, India. "

2nd Author Photo

Description "**Dr. Sunil Kumar G** is an Assistant Professor in the Department of Computer Science and Engineering at University Visvesvaraya College of Engineering (UVCE), Bengaluru, India. "