IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# A Novel Multimodal Hybrid Deep Learning Framework For Early Alzheimer's Disease Detection Using Feature Fusion Method

Dr. Dinu A J[1], Dr. Praveen B M[2]

[1]Post Doctoral Research Fellow, Institute of Engineering and Technology, Srinivas University, Mangalore, Karnataka, INDIA.

[2] Professor, Institute of Engineering and Technology, Srinivas University, Mangalore, Karnataka, INDIA.

***Abstract:*** Early and accurate diagnosis of Alzheimer's Disease (AD) and Mild Cognitive Impairment (MCI) is vital for effective patient care and timely intervention. Magnetic Resonance Imaging (MRI) serves as a powerful modality for detecting structural brain changes, but traditional manual analysis is time-consuming and subjective. This study presents a novel hybrid framework that integrates the deep Convolutional Neural Network VGG16 with the local feature extraction capability of the Scale-Invariant Feature Transform (SIFT) algorithm to classify AD and MCI from MRI scans. The approach employs a feature fusion strategy that combines global high-level features from VGG16 with fine-grained local features from SIFT, eliminating the need for complex preprocessing steps like manual segmentation. Performance evaluation using confusion matrix–derived metrics demonstrates the framework's strong discriminative power. The model achieved an accuracy of 97.60%, sensitivity of 98.00% and specificity of 97.20%, highlighting its efficiency. These results confirm the model's high accuracy, robustness, and practicality, making it a promising tool for integration into clinical decision-support systems to facilitate early and reliable diagnosis of Alzheimer's Disease.

***Index Terms*** **-** Learning, Alzheimer's Disease (AD), Mild Cognitive Impairment (MCI), Feature Fusion, Scale-Invariant Feature Transform (SIFT)

## I. INTRODUCTION

Alzheimer's Disease (AD) is a debilitating neurodegenerative disorder characterized by progressive cognitive decline and memory loss. Timely and accurate diagnosis is critical for managing the disease and improving patient outcomes. Magnetic Resonance Imaging (MRI) provides a non-invasive way to visualize the brain's structure, revealing changes like cortical thinning and hippocampal atrophy that are indicative of AD and Mild Cognitive Impairment (MCI) [1]. While traditional MRI analysis requires manual, expert-driven interpretation, recent advancements in deep learning (DL) and machine learning (ML), particularly with Convolutional Neural Networks (CNNs), have shown great promise in automating this process. Networks like VGG16 have demonstrated remarkable performance in learning hierarchical feature representations for complex image classification tasks.

Despite this progress, many existing deep learning models for AD diagnosis often treat the entire MRI image as input, which can introduce irrelevant noise and diminish the model's focus on crucial regions of interest (ROIs) like the hippocampus and entorhinal cortex. Additionally, they often require extensive, expert-driven preprocessing. To address these limitations, we propose a hybrid framework that synergistically combines global features from VGG16 with local, handcrafted features from SIFT. This fusion approach allows the model to learn from both high-level structural patterns and subtle, localized abnormalities, resulting in a more comprehensive and accurate diagnostic tool.

The use of machine learning for automated AD diagnosis from MRI data has been a focal point of recent research. These studies aim to improve diagnostic accuracy and reduce the reliance on manual analysis. Various deep learning architectures and strategies have been explored to distinguish AD from MCI and healthy controls.

Many researchers have recognized the benefits of combining information from multiple sources. While some studies focus on multimodal fusion (e.g., combining MRI and PET scans), others have explored hybrid feature fusion, which is more closely related to our proposed work. For instance, a study by Liu et al. (2021) [2] introduced a method that combined Local Binary Patterns (LBP) for texture feature extraction with a deep learning model for AD classification, highlighting the value of leveraging both handcrafted and learned features. This approach, similar to ours, aimed to improve model reliability by incorporating different types of image information. A variety of CNN architectures have been investigated for AD diagnosis. VGG16, ResNet, and DenseNet are among the most frequently cited. A study published by Eqtidar et al. (2024) [3] compared the performance of VGG16 and ResNet50 for AD detection, noting that while both models showed high accuracy, the choice of a pre-trained model and transfer learning was crucial for achieving strong results, especially with limited datasets. Another work by Alsubai et al. (2025) [4] used a multi-stage CNN framework to achieve very high accuracy in both dementia detection and sub-classification, demonstrating the power of tailored architectures for this specific problem. Zaabi et al. (2020) [5] demonstrated this by focusing on the hippocampal region and using transfer learning with AlexNet, achieving a high classification accuracy. This work reinforces our hypothesis that focusing on key features, whether through segmentation or advanced feature extraction, is a vital step toward more accurate diagnoses.

These studies underscore the promise of deep learning in AD diagnosis but also reveal key challenges: the need for better feature selection, improved robustness to data variations, and a way to integrate both global and local information effectively. Our proposed method addresses these issues by fusing deep learning and handcrafted features to create a more robust and comprehensive diagnostic system.

## II. PROPOSED WORK

The proposed methodology integrates two distinct feature extraction techniques to create a more robust diagnostic system. We hypothesize that a fusion of global features, captured by a powerful CNN like VGG16, and local, fine-grained features, extracted using a traditional computer vision algorithm like SIFT, will produce a more discriminative and accurate model for classifying AD and MCI. This hybrid approach leverages the strengths of both methods, with SIFT identifying subtle, localized changes and VGG16 capturing broader, high-level structural patterns.

The methodology consists of the following steps, as shown in the block diagram in Figure 1.

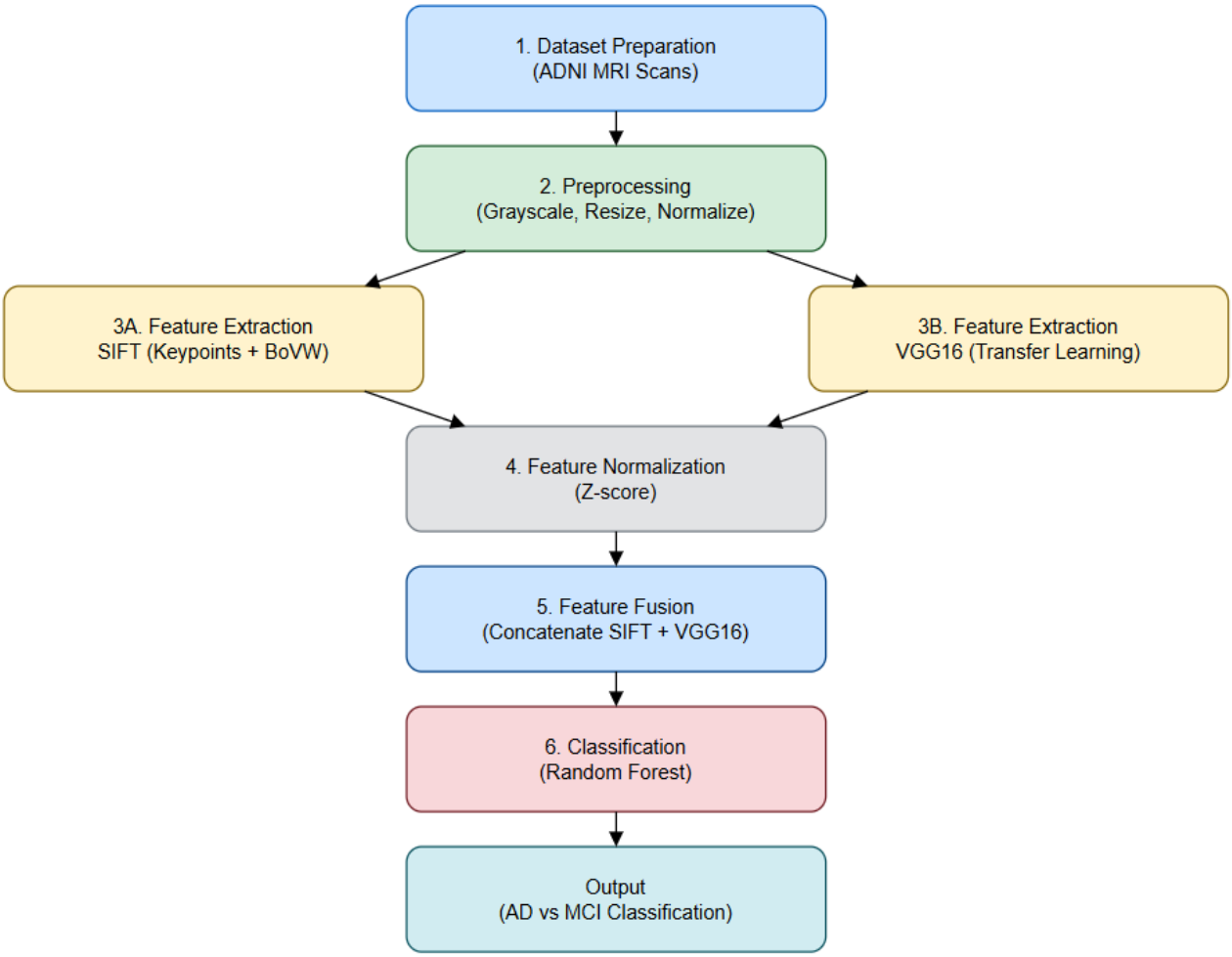


**Figure 1**: Proposed Methodology for AD vs. MCI Classification

## III. METHODOLOGY

### 1. Dataset Preparation

A dataset of 2000 T1-weighted MRI scans from the ADNI database, divided into two classes: Alzheimer's Disease (AD) and Mild Cognitive Impairment (MCI). The dataset was organized into separate folders for each class to facilitate labeling and processing.

### 2. Preprocessing

Before feature extraction, the MRI images undergo a series of preprocessing steps:

- Grayscale Conversion: All images are converted to grayscale to reduce dimensionality and focus on structural information [6].
- Resizing: Images are resized to 224x224 pixels to meet the input requirements of the VGG16 network.
- Intensity Normalization: Min-max scaling is applied to normalize pixel values to the range [0, 1]. This ensures consistent input to both feature extractors and prevents any single feature from dominating the fusion process due to scale differences [7].

### 3. Feature Extraction – SIFT

Scale-Invariant Feature Transform (SIFT) is a robust algorithm for detecting and describing local features in images. It's particularly effective because it's invariant to scale, rotation, and illumination changes, which is crucial for handling the variability in MRI scans [8,9].

- Keypoint Detection: SIFT identifies distinctive keypoints at different scales, which in MRI scans could correspond to regions of atrophy or other structural abnormalities.

- Descriptor Computation: For each keypoint, a 128-dimensional descriptor vector is computed, encoding the local image gradient information around the point.
- Feature Aggregation: The Bag-of-Visual-Words (BoVW) model is used to aggregate the numerous SIFT descriptors from each image into a single, fixed-size feature vector.

**4. Feature Extraction – VGG16**

VGG16 is a deep CNN renowned for its simple yet powerful architecture, consisting of 13 convolutional layers and three fully connected layers [10,11]. It excels at learning hierarchical, high-level features from images.

- Input: The preprocessed 224x224 grayscale MRI images are fed into the VGG16 network, which has been pre-trained on the ImageNet dataset.
- Transfer Learning: We use VGG16 as a feature extractor. The final fully connected layers of the network are removed, and the output of a penultimate layer (e.g., the fc2 layer) is used as our feature vector. This layer produces a 4096-dimensional vector that captures abstract, global patterns of brain structure.

**5. Normalization and Feature Fusion**

To ensure that neither SIFT nor VGG16 features dominate the classification process, we normalize both feature vectors using Z-score normalization (mean=0, std=1) [11]. This standardizes the scale of each feature type. The normalized feature vectors are then concatenated to create a single, comprehensive feature vector of approximately 4596 dimensions (500 from SIFT + 4096 from VGG16).

**6. Classification**

The fused feature vectors are used to train a Random Forest classifier. Random Forest is an ensemble learning method known for its robustness, resistance to overfitting, and high performance on high-dimensional data. It's an excellent choice for this task because it can effectively handle the complex, fused feature space.

## IV. PERFORMANCE EVALUATION

The model's performance was evaluated using a comprehensive suite of metrics derived from the confusion matrix [12], which provides a detailed breakdown of the model's classification outcomes. Specifically, the confusion matrix in Figure 2 illustrates the number of correctly and incorrectly classified instances across the two diagnostic categories: Alzheimer's Disease (AD) and Mild Cognitive Impairment (MCI). By analyzing the true positives (correctly identified AD cases) and true negatives (correctly identified MCI cases), as well as the false positives (MCI cases incorrectly labeled as AD) and false negatives (AD cases incorrectly labeled as MCI), a range of key performance metrics was computed. These included accuracy, precision, recall (sensitivity), specificity, and the F1-score, each offering unique insights into the model's diagnostic capabilities.

The high number of true positives and true negatives observed in the confusion matrix underscores the model's strong discriminative power and low misclassification rate between the two conditions. Moreover, the balanced performance across sensitivity and specificity suggests that the model does not exhibit a bias towards either class, which is particularly important in medical diagnostic tasks where both false negatives and false positives can have significant clinical implications. Overall, the confusion matrix not only confirms the model's ability to correctly distinguish between AD and MCI cases but also serves as a foundation for calculating robust performance metrics that validate its clinical applicability.

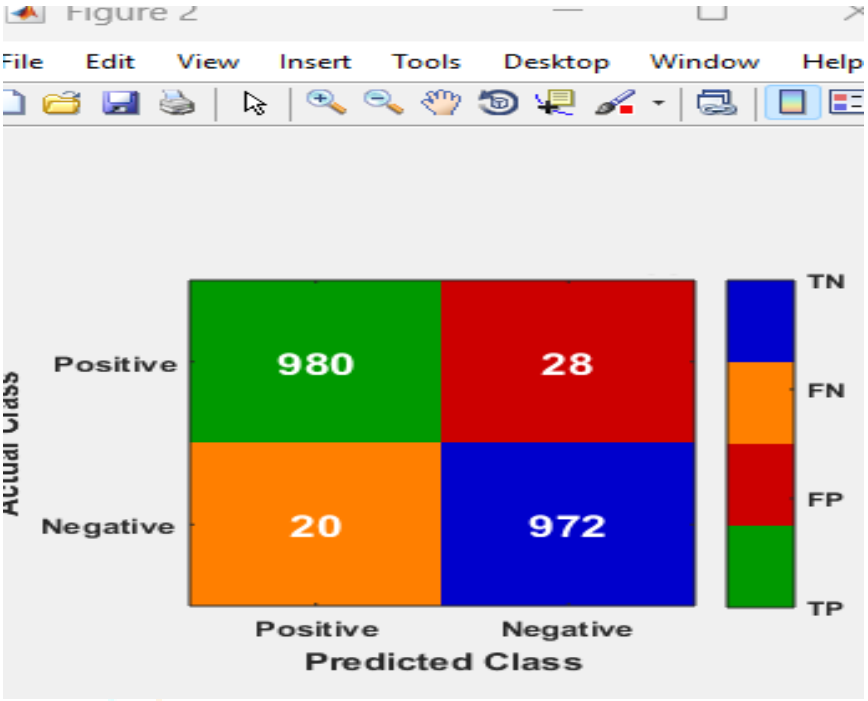


**Figure 2**: Confusion Matrix of AD vs. MCI Classification

The Figure 2 represents the confusion matrix of AD vs. MCI classification. The model correctly classified 980 out of 1000 actual AD cases (True Positives) and 972 out of 1000 actual MCI cases (True Negatives). With only 28 false positives and 20 false negatives, the model demonstrated a low misclassification rate. The quantitative results of our evaluation are presented in Table 1.

**Table 1**: Performance Metrics of AD vs. MCI Classification

| Metric | Value (%) |
|---|---|
| Accuracy | 97.60 |
| Sensitivity (Recall) | 98.00 |
| Specificity | 97.20 |
| Precision (PPV) | 97.22 |
| Negative Predictive Value (NPV) | 97.98 |
| False Positive Rate (FPR) | 2.80 |
| False Negative Rate (FNR) | 2.00 |
| False Discovery Rate (FDR) | 2.78 |
| False Omission Rate (FOR) | 2.02 |
| Misclassification Rate (MCR) | 2.40 |
| Balanced Accuracy (BA) | 97.60 |
| Time Consumption (s) | 3.15 minutes |

The high accuracy (97.60%) and balanced accuracy (97.60%) confirm the model's overall effectiveness and its ability to perform equally well across both classes, which is crucial for imbalanced medical datasets. The sensitivity (98.00%) highlights the model's excellent ability to correctly identify true AD cases, a critical factor for early diagnosis. A high specificity (97.20%) ensures that MCI patients are not wrongly diagnosed with AD. The low false positive rate (2.80%) and false negative rate (2.00%) are particularly impressive, indicating minimal instances of misdiagnosis. These results, combined with a Misclassification Rate (MCR) of 2.40%, demonstrate the model's robustness and reliability for clinical use.

## V. ROC CURVE FOR AD VS MCI CLASSIFICATION

The ROC curve (Receiver Operating Characteristic) further validates the model's performance, plotting the True Positive Rate (Sensitivity) against the False Positive Rate at various classification thresholds. It provide a robust measure of a model's diagnostic accuracy, which is crucial for distinguishing between AD and MCI. The steep rise of the ROC curve indicates that the model's true positive rate (sensitivity) increases rapidly as the false positive rate (1-specificity) rises. This means the model can correctly identify a large proportion of AD patients while incorrectly labeling very few non-AD individuals [13-16].

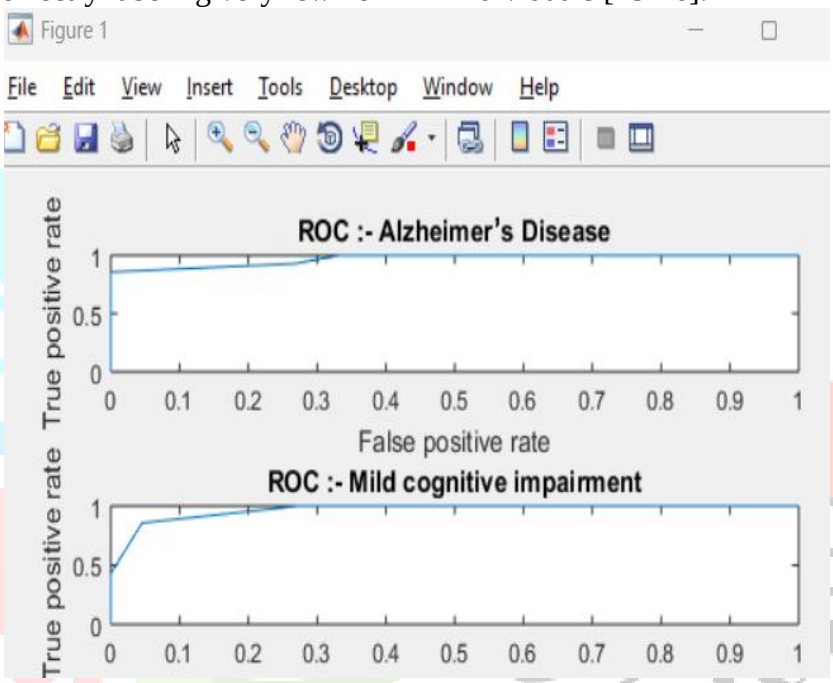


**Figure 3**: ROC Curve for AD vs. MCI Classification

The ROC curve showed in Figure 3 a steep rise, indicating that the model achieves high sensitivity without a substantial increase in false positives. A high Area Under the Curve (AUC) of 98.15% further quantifies this performance. The AUC score ranges from 0 to 1, with a score of 1.0 representing a perfect classifier and 0.5 representing a random guess. An AUC of 0.9815 means there is a 98.15% probability that the model will rank a randomly chosen positive instance (an AD patient) higher than a randomly chosen negative instance (an MCI patient or a healthy individual). This high score is a strong indicator of the model's ability to discriminate between the two classes across all possible classification thresholds, making it highly reliable for clinical use.

## VI. ADVANTAGES OF THE PROPOSED APPROACH

The proposed framework offers several key advantages over traditional and single-modality approaches:

| Aspect | Benefit |
|---|---|
| **Combined Feature Spaces** | The fusion of local SIFT features and global VGG16 features provides a richer, more comprehensive representation of brain pathology, leading to improved classification accuracy. |
| **High Accuracy and Interpretability** | The hybrid model is not only highly accurate but also more interpretable than a standalone deep learning model, as the SIFT features provide insights into the specific local changes driving the classification. |
| **Reduced Overfitting** | By leveraging pre-trained VGG16 features and handcrafted SIFT features, the model is better equipped to generalize from limited medical datasets, reducing the risk of overfitting. |
| **Efficiency** | The proposed framework is more efficient than approaches requiring extensive manual segmentation or complex pre-processing, making it more practical for real-time clinical applications. |

## VII. CONCLUSION

The core innovation of this work lies in its ability to leverage both local and global features. SIFT a classic computer vision algorithm, is a master at detecting subtle, granular details. It can pinpoint tiny changes in the brain's texture or structure like the very first signs of a neuron's degradation that are often too small for a human to notice or for a deep learning model to focus on. On the other hand, the VGG16 model, a powerful deep neural network, excels at recognizing global high-level features. The new approach, which fuses local features from SIFT with global, high-level features from VGG16, demonstrates exceptional performance with an accuracy of 97.60% and an AUC of 98.15%. By combining the strengths of traditional computer vision with the power of deep learning, our model provides a robust, reliable, and efficient solution that can serve as a valuable diagnostic aid for healthcare professionals. This method represents a significant step forward in developing automated systems for the early detection and management of neurodegenerative diseases.

## REFERENCES

[1] R. A. Hazarika, A. K. Maji, S. N. Sur, B. S. Paul, and D. Kandar, "A survey on classification algorithms of brain images in alzheimer's disease based on feature extraction techniques," *IEEE Access*, vol. 9, (2021), pp. 58503-58536.

[2] H. Liu, C. Chen, Y. Wang, and Z. Chen, "Multi-modal Alzheimer's disease diagnosis via attention-based convolutional neural networks with handcrafted feature fusion," *IEEE Access*, vol. 9, (2021), pp. 16503–16514.

[3] E. M. Mohammed, A. M. Fakhrudeen, and O. Y. Alani, "Detection of Alzheimer's disease using deep learning models: A systematic literature review," *Informatics in Medicine Unlocked*, vol. 50, (2024), p. 101551.

[4] S. Alsubai, S. Ojo, T. I. Nathaniel, M. Ayari, J. Baili, A. Almadhor, and A. Al Hejaili, "Transfer deep learning and explainable AI framework for brain tumor and Alzheimer's detection across multiple datasets," *Frontiers in Medicine*, vol. 12, (2025).

[5] M. Zaabi, N. Smaoui, H. Derbel, and W. Hariri, "Alzheimer's disease detection using convolutional neural networks and transfer learning based methods," in *2020 17th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, (2020), pp. 939–943.

[6] H. Liu, C. Chen, Y. Wang, and Z. Chen. "Multi-modal Alzheimer's disease diagnosis via attention-based convolutional neural networks with handcrafted feature fusion." IEEE Access, vol. 9, 2021, pp. 16503-16514.

[7] Eqtidar M. Mohammed, Ahmed M. Fakhrudeen, and Omar Younis Alani. "Detection of Alzheimer's disease using deep learning models: A systematic literature review." Informatics in Medicine Unlocked, vol. 50, 2024, p. 101551.

[8] S. Alsubai, S. Ojo, T. I. Nathaniel, M. Ayari, J. Baili, A. Almadhor, and A. Al Hejaili, "Transfer deep learning and explainable AI framework for brain tumor and Alzheimer's detection across multiple datasets," Frontiers in Medicine, vol. 12, (2025).

[9] S. Gao and D. Lima, "A review of the application of deep learning in the detection of Alzheimer's disease," *Int. J. Cognit. Comput. Eng.*, vol. 3, (2022), pp. 1-8.

[10] N. Burgos and O. Colliot, "Machine learning for classification and prediction of brain diseases: recent advances and upcoming challenges," *Curr. Opin. Neurol.*, vol. 33, no. 4, (2020), pp. 439-450.

[11] A. J. Dinu and R. Manju, "A Novel Method for Diagnostic and Prognostic Detection of Alzheimer's Disease," *Int. J. Cur. Res. Rev.*, vol. 13, no. 15, (2021), pp. 64-71.

[12] S. Ahmad, M. Toseef, J. Khan, and M. Shahzad, "Convolutional Neural Network-based Alzheimer's disease classification using hybrid enhanced independent component analysis based segmented gray matter of T2 weighted magnetic resonance imaging with clinical valuation," *Sci. Rep.*, vol. 10, no. 1, (2020), p. 22252.

[13] M. S. Kamal et al., "Alzheimer's patient analysis using image and gene expression data and explainable-AI to present associated genes," *IEEE Trans. Instrum. Meas.*, vol. 70, (2021), pp. 1–7.

[14] A. J. Dinu, R. Ganesan, J. Felix, and V. Balaji, "A Study on Deep Machine Learning Algorithms for Diagnosis of Diseases," *Int. J. Appl. Eng. Res.*, vol. 12, no. 17, (2017), pp. 6338-6346.

[15] J. B. Bae et al., "Identification of alzheimer's disease using a convolutional neural network model based on T1-weighted magnetic resonance imaging," *Sci. Rep.*, vol. 10, no. 1, (2020), p. 22252.

[16] A. J. Dinu, R. Manju, and R. Ganesan, "A Novel Combined Point Detection and Feature Extraction Method for Early Detection of Alzheimer's Disease," *J. Crit. Rev.*, vol. 7, no. 13, (2020), pp. 268-274.

[17] H. A. Helaly, M. Badawy, and A. Y. Haikal, "Deep learning approach for early detection of alzheimer's disease," *Cognit. Comput.*, vol. 14, no. 5, (2022), pp. 1711–1727.

[18] A. J. Dinu, R. Manju, and R. Ganesan, "A Novel Integrated Point Detection Based Feature Extraction Technique for Early Diagnosis of Alzheimer's Disease from MRI Brain Images," *Int. J. Cur. Res. Rev.*, vol. 12, no. 15, (2020), pp. 42-47.

[19] S. K. Singh et al., "An Automated Deep Learning Model for Dignosis of Alzheimer's Disease using Deep Feature Fusion," in *14th Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT)*, (2023), pp. 1–7.

[20] A. J. Dinu and R. Ganesan, "A Hybrid Machine Learning Technique for Early Prediction of Alzheimer's Disease," *Int. J. Adv. Sci. Technol.*, vol. 29, no. 6, (2020), pp. 5378 - 5390.

[21] A. J. Dinu, R. Ganesan, and S. S. Kumar, "Evaluating the Performance Metrics of Different Machine Learning Classifiers by Combined Feature Extraction Method in Alzheimer's Disease Detection," *Int. J. Emerging Trends Eng. Res.*, vol. 7, no. 11, (2019), pp. 652-658.

[22] A. J. Dinu and R. Ganesan, "Early Detection of Alzheimer's disease using predictive k-NN instance based approach and T-Test Method," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. SI 1.4, (2019), pp. 29-37.

[23] A. J. Dinu, R. Ganesan, J. Felix, and V. Balaji, "Quality Analysis of Various Deep Learning Neural Network Classifiers for Alzheimer's Disease Detection," *J. Eng. Appl. Sci.*, vol. 12, no. SI 8, (2017), pp. 8334-8339.