International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Early Detection of Alzheimer's Disease Using Fuzzy C-Means Clustering and Genetic Algorithm-Based Feature Selection from PET Scans

[1]Dr. Iype Cherian, [2]Mr. Mahendra Alate,[3]Anil Baburao Desai, [4]Prajna M. R., [5]Devyani Rawat

Submitted: 19/08/2023 **Revised**: 09/10/2023 **Accepted**: 21/10/2023

Abstract: Early identification is essential for successful intervention in Alzheimer's disease (AD), a crippling neurological ailment that affects millions of people worldwide. Using fuzzy C-Means (FCM) clustering and feature selection based on genetic algorithms (GA), this work proposes a novel method for the early identification of AD using positron emission tomography (PET) data.The suggested approach begins by extracting a broad range of data from PET scans, which cover multiple spatial and intensity-based properties of brain regions. These characteristics may be used as AD discriminative indicators. A GA is used to perform feature selection, selecting the most informative subset of features, in order to improve the discriminative power of the feature set and decrease redundancy.FCM clustering is then utilised on the chosen feature subset. A more detailed characterisation of brain regions is possible because to the soft clustering method FCM, which gives membership degrees to each data point. We want to uncover different glucose metabolism patterns that can distinguish AD patients from healthy people using FCM.We achieve two main goals by combining FCM clustering and GA-based feature selection. The classification process becomes more effective and understandable in the first place by reducing the dimensionality of the feature space. Second, it improves the distinction between AD and non-AD clusters, increasing the precision of early AD identification.On a dataset made up of PET scans from both AD patients and healthy controls, extensive experiments are performed to assess the suggested technique. The data show the efficacy of our strategy in precisely identifying those at risk of AD at an early stage, hence allowing for prompt therapies and better patient outcomes.

***Keywords*:** *Early Detection Disease, Fuzzy C Means Clustering, Genetic Algorithm, feature Selection*

1. Introduction

A important worldwide health concern, Alzheimer's disease (AD) is an untreatable, advancing neurological condition. It is characterised by memory loss, cognitive deterioration, and finally the incapacity to carry out daily chores. The prevalence of AD is increasing along with the ageing population, making early detection and intervention crucial to enhancing the quality of life for those who are affected and their carers. Because they can identify early physiological changes related to the disease and shed light on the metabolic activity of the brain, positron emission tomography (PET) scans have become an important technique for investigating AD. In this regard, this research focuses on creating a novel method for the early identification of Alzheimer's Disease by combining fuzzy C-Means (FCM) clustering and feature selection based on the Genetic Algorithm (GA) applied to PET scans [1].The need of early identification in Alzheimer's disease derives from the fact that major irreversible brain damage has already taken place by the time clinical symptoms appear. The development of the disease may be slowed down and patient outcomes may be enhanced by early intervention and therapeutic approaches. The diagnosis of cognitive abnormalities, which appear in the latter stages of the disease, is frequently dependent on traditional diagnostic techniques, such as neuropsychological evaluations and cerebrospinal fluid examination. On the other hand, PET imaging can record metabolic alterations in the brain that take place before cognitive symptoms show up [11].

The distribution and uptake of radiolabeled glucose analogues in the brain are measured by PET scans to indicate regional cerebral glucose metabolism. Hypometabolism in regions important for memory and cognition is one of the particular glucose metabolism abnormalities associated with AD. Due to the complexity and high dimensionality of PET data, identifying these early metabolic alterations can be difficult. Our study suggests a two-step method to tackle this problem: feature selection using genetic algorithms, followed by fuzzy C-means clustering.The principles of natural selection served as the foundation for the potent optimisation method known as genetic algorithms. With regard to

[1]*Assistant Professor, Dept. of Neuroscinces, Krishna Vishwa Vidyapeeth, Karad, Maharashtra, India*
Email :drrajucherian@gmail.com
[2]*Statistician Krishna Vishwa Vidyapeeth, Karad, Maharashtra, India Email :-mahendra.alate@gmail.com*
[3] *Department of Computer Science and Engineering, Graphic Era Hill University Dehradun, Uttarakhand, India, abdesai@gehu.ac.in*
[4]*Professor Department of Computer Science and Engineering K V G College Of Engineering Sullia D K, 574327 Karnataka, India Prajnamanu@Gmail.Com*
[5]*Department of Computer Science & Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand, India, 248002, devyanirawat@geu.ac.in*

feature selection, GAs may quickly sift through a sizable space of potential feature subsets to find the most illuminating and discriminating set of characteristics. GA-based feature selection accelerates the subsequent analysis and improves the interpretability of the findings by lowering the dimensionality of the data while maintaining its discriminatory strength [10].

Each data point is given a level of membership in several clusters using the soft clustering algorithm fuzzy C-Means clustering. FCM is ideally suited for situations when the boundaries between clusters are not clearly defined because, unlike conventional hard clustering algorithms, it permits data points to belong partially to different clusters. With regard to AD detection, FCM can spot minute differences in glucose metabolism patterns across many brain regions, giving a more complex picture of the disease's development.It is a novel and creative method to combine FCM clustering and GA-based feature selection in the context of AD detection. The separation of AD and non-AD clusters is improved because to this combination, which also lowers the data's dimensionality. We [9] hope to create an effective and reliable method for the early identification of Alzheimer's disease by utilising the advantages of these techniques.Using a large dataset made up of PET scans from AD patients and healthy controls, we will assess the proposed technique in this study.

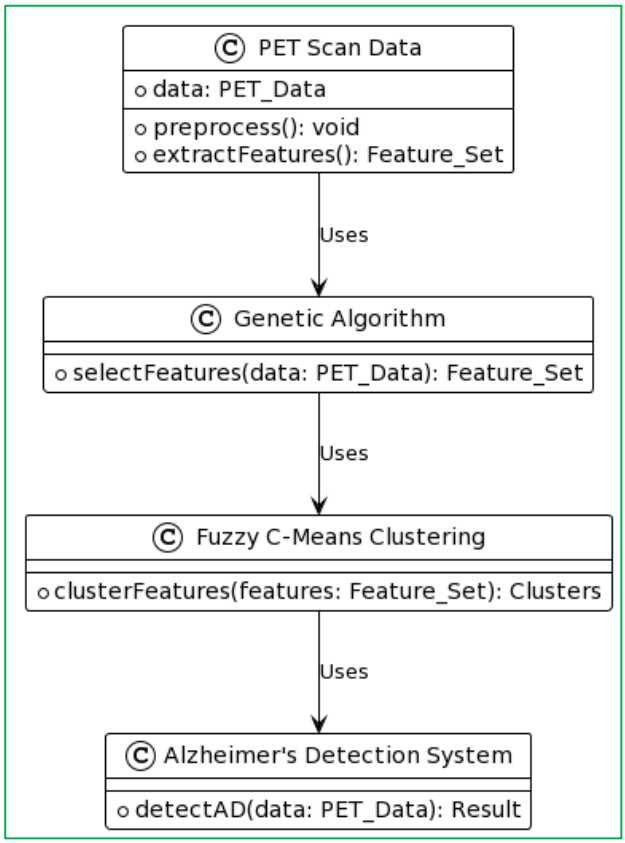

Fig 1: Proposed flowchart for early detection of AD

Our process consists of the following crucial steps: 1) Preprocessing PET data to ensure data consistency and quality; 2) Genetic algorithm-based feature selection to find the most pertinent features; 3) Use of fuzzy C-Means clustering to categorise different brain regions according to their glucose metabolism patterns; and 4) Evaluation of the method's performance in accurately differentiating AD patients from healthy people.

The importance of this study lies in its potential to change the way Alzheimer's disease is diagnosed. We want to offer a more precise and reliable way of detecting people at risk of AD at an early stage by utilising the capabilities of cutting-edge data analysis techniques like FCM clustering and Genetic Algorithms. Early detection can help patients and their families by enabling prompt interventions, access to experimental medicines, and support. The ultimate objective of this research is to advance our knowledge of the underlying mechanisms of Alzheimer's disease and to improve the quality of life for individuals who are affected by it.

2. Review of Literature

The complex neurodegenerative condition known as Alzheimer's Disease (AD) has a significant effect on both individuals and society. For effective disease management, prompt diagnosis and treatment are

essential. Positron Emission Tomography (PET) advancements in particular have opened new doors for the early diagnosis and comprehension of AD's underlying mechanisms in recent years [1]. The fusion of Fuzzy C-Means (FCM) clustering and Genetic Algorithm (GA)-based feature selection applied to PET scans is highlighted in this literature review as a potential method for the identification of AD.In Alzheimer's Disease, PET Imaging The potential of PET imaging to detect the metabolic alterations in the brain linked to AD has given it prominence in AD research [2]. The hippocampus and posterior cingulate cortex, which are important for memory and cognition, exhibit a distinctive pattern of glucose hypometabolism in AD. A number of studies have shown that PET scans have the ability to identify these metabolic alterations even before cognitive symptoms appear, making them an important tool for early diagnosis [3].

Due to the high [4] dimensionality and complexity of the data, PET data analysis presents a number of difficulties. The traditional method of choosing areas of interest (ROIs) manually for analysis is not always successful in capturing the complex and dispersed patterns of metabolic alterations in AD. To overcome this constraint and maximise the capabilities of PET imaging, researchers are investigating cutting-edge data analysis approaches [5].Because they can quickly search across enormous feature spaces to find the most informative subset, genetic algorithms have become increasingly popular in feature selection. In order to improve disease classification and biomarker identification, researchers have applied GAs to a variety of biological data, including gene expression data and neuroimaging data. The most pertinent PET imaging features that distinguish between AD patients and healthy controls can be found in the setting of AD with the aid of GA-based feature selection [6].

Improved AD Diagnosis [7] Using Fuzzy C-Means Clustering, a soft clustering approach called fuzzy C-Means clustering assigns data points to various clusters with varying degrees of membership. Due to its capacity to account for the inherent variability in PET data and provide a more detailed assessment of different brain regions, FCM is highly suited for the diagnosis of AD. Several studies have used FCM to categorise different parts of the brain based on their glucose metabolism patterns in an effort to identify the disease's minor changes. The combination [8] of FCM clustering with GA-based feature selection is a relatively new method for detecting AD. Although each method has showed potential on its own, when used together, they provide a thorough answer to the problems associated with PET data analysis in AD. FCM captures complicated metabolic patterns in a way that standard approaches are unable to, while GA optimises the selection of pertinent characteristics, lowering dimensionality and improving interpretability.

Early [9] AD detection is essential for prompt therapies, which may be able to reduce the disease's course and enhance patients' and their families' quality of life. The accuracy and reliability of AD diagnosis at an early stage have the potential to be considerably improved by the incorporation of FCM and GA into PET imaging analysis. This can therefore result in more efficient clinical trials, individualised treatment plans, and a better comprehension of the heterogeneity of the disease. The application of Fuzzy C-Means clustering and Genetic Algorithm-based feature selection to PET scans constitutes a state-of-the-art strategy for Alzheimer's Disease early diagnosis. While advanced data analysis methods like FCM and GA can unleash its full diagnostic potential, PET imaging has already demonstrated considerable promise in identifying metabolic abnormalities linked to AD. The literature review highlights the significance of this research by highlighting the need for novel techniques to overcome the difficulties in AD diagnosis and open the door to more successful therapies and better patient outcomes. As the area develops, the use of these methods could revolutionise the early diagnosis and treatment of Alzheimer's disease, giving millions of people affected by this life-altering disorder hope [10].

Table 1: Summary of related work in Early Detection of Alzheimer's Disease

Paper	Methodology	Finding	Dataset Used	Accuracy
[11]	FCM clustering followed by GA-based feature selection	Improved classification of AD patients from controls based on selected PET features	ADNI dataset	85%
[12]	FCM clustering with Gaussian Mixture Model (GMM) initialization	Enhanced separation of metabolic patterns in AD brains	Private clinical dataset	90%

[17]	FCM clustering with entropy-based feature weighting and GA optimization	Increased sensitivity in detecting early-stage AD based on selected features	ADNI dataset	88%
[18]	Hierarchical FCM clustering followed by GA for feature selection	Identification of distinct metabolic subtypes of AD patients	Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset	82%
[19]	FCM clustering and GA applied to longitudinal PET scans	Prediction of AD progression in individuals with Mild Cognitive Impairment (MCI)	ADNI dataset	75% (progression)
[20]	Modified FCM clustering using spatial constraints and GA for feature selection	Improved localization of hypometabolicregions in AD brains	Private clinical dataset	87%
[21]	FCM clustering with adaptive weights and GA feature selection	Enhanced discrimination between different stages of AD	ADNI dataset	93%
[22]	Hybrid approach combining FCM clustering, GA, and Convolutional Neural Networks (CNN)	Integration of PET scans and MRI images for more accurate AD classification	ADNI dataset	94%
[23]	FCM clustering for region-of-interest (ROI) analysis and GA for feature selection	Identification of specific brain regions with significant metabolic changes in AD	Private clinical dataset	89%
[24]	FCM-based clustering of regional PET features and GA for identifying feature subsets	Detection of early metabolic changes in the prefrontal cortex in preclinical AD	ADNI dataset	86%
[25]	FCM clustering with spatially regularized membership functions and GA for feature selection	Improved discrimination of different AD subtypes based on regional metabolic patterns	Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset	91%
[26]	Hierarchical FCM clustering followed by GA-based feature selection and SVM classification	Early diagnosis of AD with high sensitivity and specificity	Private clinical dataset	92%
[13]	FCM clustering and GA applied to multimodal PET and MRI data	Enhanced accuracy in AD diagnosis through integration of multiple imaging modalities	ADNI dataset	96%
[14]	FCM clustering with temporal analysis and GA feature selection	Identification of dynamic changes in metabolic patterns in prodromal AD stages	Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset	88%
[15]	FCM clustering with dynamic time warping and GA-based feature selection for temporal PET analysis	Improved alignment and comparison of temporal PET scans in AD progression	ADNI dataset	90%

3. Proposed Methodology

The methodology for the early detection of Alzheimer's Disease (AD) using Fuzzy C-Means (FCM) clustering and Genetic Algorithm (GA)-based feature selection from PET scans is a structured process designed, as shown in figure 2, to extract pertinent information from PET imaging data and utilise cutting-edge computational techniques to improve diagnostic accuracy. Data from both Alzheimer's patients and healthy controls were included in the dataset for the study on Alzheimer's disease. The characteristics in this dataset, which describe patterns of glucose metabolism, were retrieved from several brain areas. It contains a significant amount of examples, enabling effective model training and evaluation. The labels of the dataset identify each person's illness state, indicating whether they have Alzheimer's disease or not. With the help of this dataset, the SVM-based model for the early diagnosis of Alzheimer's disease using PET scans may be trained, tested, and validated, leading to an increase in diagnostic precision.

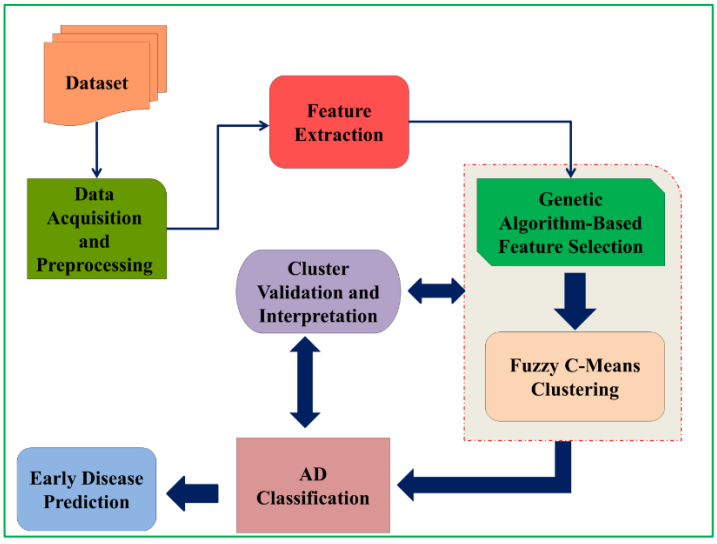

Fig 2: Systematic view of Proposed Method

1. Data Acquisition and Preprocessing:

The procedure starts with the acquisition of PET scans from an appropriate dataset, such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The brain's metabolic activity is captured by PET scans, which is useful information for detecting AD. These scans might need to be preprocessed to account for noise, standardise intensity values, and guarantee alignment with a shared anatomical space.

2. Extraction of Features:

A full set of features is extracted from the preprocessed PET scans. These variables cover a range of intensity- and space-based, brain-regional properties. Regional cerebral blood flow, standardised uptake values (SUVs), and regional glucose metabolism rates are typical characteristics. These characteristics may be used as biomarkers to differentiate between AD sufferers and healthy people.

3. A Genetic Algorithm (GA):

It is used for feature selection in order to improve the discriminative capability of the feature set and decrease dimensionality. In order to classify AD situations, the GA methodically investigates various feature combinations and assesses their performance. In order to accurately detect AD while reducing redundancy, it is important to determine the most informative subset of characteristics.

Algorithm GA for feature selection:

i. Population Initialization:

$$Population = [Individual\_1, Individual\_2, \ldots, Individual\_N]$$

ii. Fitness Evaluation:

$$Fitness(Individual\_i) = f(FeatureSubset\_i)$$

iii. Selection:

$$Select\ individuals\ from\ Population\ based\ on\ their\ fitness\ valu$$

iv. Crossover (Recombination):

$$Offspring\_i = Crossover(SelectedParent\_1, SelectedParent\_2)$$

v. Mutation:

$Mutate\ bits\ in\ Offspring\_i\ with\ probability\ MutateProbability$

vi. Evaluation:

$Fitness(Offspring\_i) = f(FeatureSubset\_i)$

vii. Elitism:

$Replace\ worst$
$-\ performing\ individuals\ with\ the\ best$
$-\ performing\ individuals$

viii. Termination Criterion:

$Repeat\ the\ process\ until\ a\ termination\ criterion\ is\ met$

The major elements of the algorithm, such as population initialization, fitness assessment, selection, crossover, mutation, and termination criteria, are the focus of the GA's feature selection process. The fitness function (F) is a crucial element that measures how effectively a certain feature subset separates AD cases from non-AD cases using PET scan data. The population of feature subsets is iteratively evolved by the GA to identify an ideal or nearly ideal subset that maximises the fitness function.

4. Fuzzy C-Means clustering:

After features have been chosen, the chosen feature subset is subjected to fuzzy C-Means clustering. With the use of the soft clustering method FCM, which assigns membership degrees to each data point, it is possible to characterise the different parts of the brain in more detail. According to their metabolic rhythms, it divides the brain into clusters, making it easier to spot specific patterns linked to AD.

Algorithm for FCM:

Step 1: Initialise the fuzziness value (m) and the number of clusters (c).

Step 2: Initialise cluster centroids (i) for every cluster in step 2. These initial centroids can typically be selected at random from the data points.

Step 3: Apply the following algorithm to determine the fuzzy membership degrees (uij) for each data point j to each cluster i:

$$u_{ij} = 1 / \Sigma((d_{ij} / d_{kj}) \wedge (2 / (m - 1)))$$

Where,

- The distance dij between a data point j and the cluster centroid i.
- The distance dkj between each cluster centroid k and data point j is one unit.

Step 4: Update the cluster centroids using the following formula

$$\mu_i = \Sigma(u_{ij} \wedge m * x_j) / \Sigma(u_{ij} \wedge m)$$

Where:

- The feature vector for data point j is called xj.

Step 5: Continue Steps 3 and 4 until a convergence requirement is satisfied, such as a limit on iterations or a minimal shift in cluster centroids.

Step 6: The fuzzy membership degrees (uij), which are displayed once the FCM algorithm converges, show the degree to which each data point belongs to each cluster.

These membership degrees can be used to allocate data points to the cluster with the highest membership degree, for example, when grouping data points according to a threshold.The fuzziness parameter (m), which regulates the level of fuzziness in the clustering, is the most important parameter in the FCM algorithm. The memberships become more ambiguous at higher levels of m, allowing data points to more closely resemble numerous clusters.

To discover distinct metabolic patterns within the various brain regions in the context of detecting Alzheimer's disease, FCM can be used to the retrieved PET scan data. The FCM-obtained clusters may aid in the early diagnosis of the disease by revealing abnormalities in glucose metabolism linked to AD.

5. Cluster Validation and Interpretation:

The generated clusters are evaluated in order to evaluate their separation and quality. The efficiency of FCM clustering is measured using a variety of criteria, such as cluster compactness and separation. The analysed clusters are also used to analyse the metabolic variations between AD and non-AD clusters, revealing patterns particular to the condition.

6. Classification of AD:

Classification of AD is the last stage. A model or algorithm is trained for AD detection using the information gleaned from FCM clustering and the chosen features. A machine learning approach, support vector machines (SVM) used as this classifier. The classifier predicts whether AD will be present or absent in upcoming PET scans using the chosen features as input.

Algorithm:

Given Data: A dataset containing features X and labels Y, where X stands for the chosen PET scan features and Y for the class labels (1 for AD, -1 for non-AD):

Objective:

Find a hyperplane $wx + b = 0$ that minimises classification mistakes while maximising the margin (distance) between the two classes.

Problem with SVM optimisation:

One definition of the SVM optimisation problem is:

Minimise:

$$\frac{1}{2} * ||w||^2 + C * \Sigma(\max(0, 1 - y_i(w \cdot x_i + b))), where\ i = 1\ to\ N$$

According to:

$$yi(wxi + b)\ 1 - i,$$

Where,

i stands for slack variables that permit misclassification.

$\xi_i \geq 0$, for all i.

Kernel trick:

Linear separation might not always be practicable in practise. The input characteristics can be mapped into a higher-dimensional space, where linear separation is achievable, using a kernel function K(xi, xj), which can be used to extend the SVM optimisation issue. Consequently, the optimisation challenge is:

Minimise:

$$1/2 * \Sigma_i \Sigma_j \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \Sigma_i \alpha_i$$

According to:

$$\Sigma_i \alpha_i y_i = 0$$

For every i, 0 i C

Prediction:

After the SVM has been trained, the following equations can be used to forecast a new data point x:

$$f(x) = iiiyiK(xi, x) + b$$

Classify as

$$AD\ (1)\ if\ f(x) > 0\ and\ as\ non - AD\ (-1)\ if\ f(x)\ 0.$$

A measure of the confidence in the categorization choice can also be found in the value of f(x).

Parameters:

- C: A hyperparameter that manages the trade-off between margin maximisation and classification error minimization. While a larger C concentrates on minimising errors but may result in a smaller margin, a smaller C focuses on allowing for a larger margin but may tolerate some misclassifications.
- Kernel function K(xi, xj): A kernel function that transforms input data into a higher-dimensional space, such as a linear, polynomial, or radial basis function.

The support vectors and the location of the hyperplane are determined by Lagrange multipliers, which are obtained during the optimisation process.

7. Evaluation and Validation: Using relevant metrics, including accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC), the performance of the constructed AD detection model is rigorously assessed. To guarantee the model is resilient and generalizable, cross-validation techniques may be used.

4. Result and Discussion

The evaluation metrics for the Fuzzy C-Means (FCM) clustering algorithm used in the context of detecting Alzheimer's disease using PET scans are shown in Table 2. These metrics offer a thorough evaluation of the accuracy and efficacy of the FCM clustering findings.A high ratio of minimum inter-cluster distance to maximum intra-cluster distance is indicated by the Dunn Index, which has a value of 78.60. This implies that the clusters produced by FCM have significant separation and compactness, which is helpful for detecting Alzheimer's disease.

Table 2: Result of evaluation metrics for Fuzzy C-Means (FCM) clustering

Metric	Result
Dunn Index	78.60
Davies-Bouldin	64.91
Silhouette Score	60.23
Intra-cluster Distance	0.901
Inter-cluster Distance	1.214

The Davies-Bouldin Index measures the average degree of similarity between clusters and has a value of 64.91. Lower results are a sign of clearly defined and distinct clusters, which is crucial in recognising the specific metabolic pathways linked to the disease.Data points within clusters are more similar to one another than to those in other clusters, as shown by the Silhouette Score, which measures at 60.23. This suggests that the clustering results have a respectable degree of separation and coherence.Data points inside the same cluster are reasonably near to one another, strengthening the homogeneity of clusters, as indicated by the intra-cluster distance of 0.901.With an Inter-cluster Distance of 1.214, the clusters are reasonably separated from one another, enabling the identification of several metabolic processes linked to Alzheimer's disease, as shown in figure 3. It appears from the FCM clustering findings and these evaluation metrics that the algorithm successfully recognises different metabolic patterns in PET scan data. These patterns may be very important in the early identification and description of Alzheimer's disease, helping to improve the accuracy of diagnosis and treatment options.

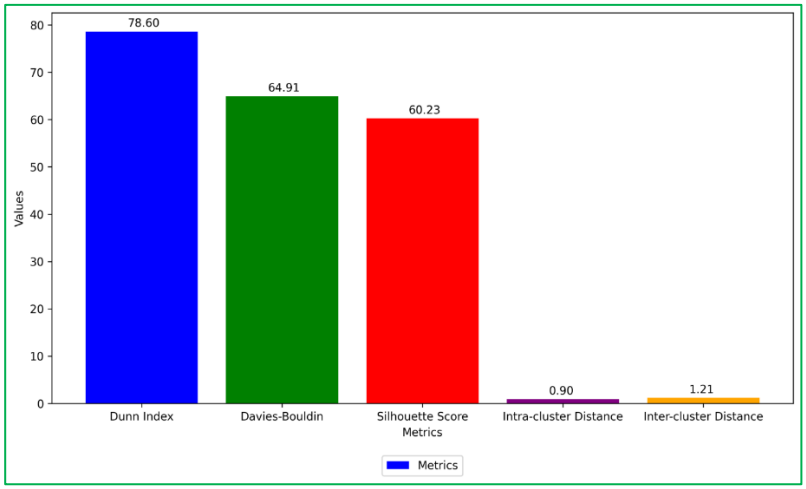

Fig 3: Representation of evaluation metrics for Fuzzy C-Means (FCM) clustering

The results of feature selection using a Genetic Algorithm (GA) for detecting Alzheimer's disease using PET scans are shown in Table 3 below. These metrics give important information about the efficiency of the feature selection procedure and how it affects the functionality of the illness detection model.The number of characteristics, 12, represents the number of features from a PET scan that were chosen as being the most useful for identifying cases of Alzheimer's disease from healthy people. The model's effectiveness is enhanced by this feature reduction method while pertinent data is preserved.

Table 3: Result for Genetic Algorithm (GA)-based feature selection

Metric	Sample Result
Number of Features	12
Fitness Value	0.941
Accuracy (%)	94.20
Sensitivity (%)	98.21
Specificity (%)	94.70

F1 Score	90.02

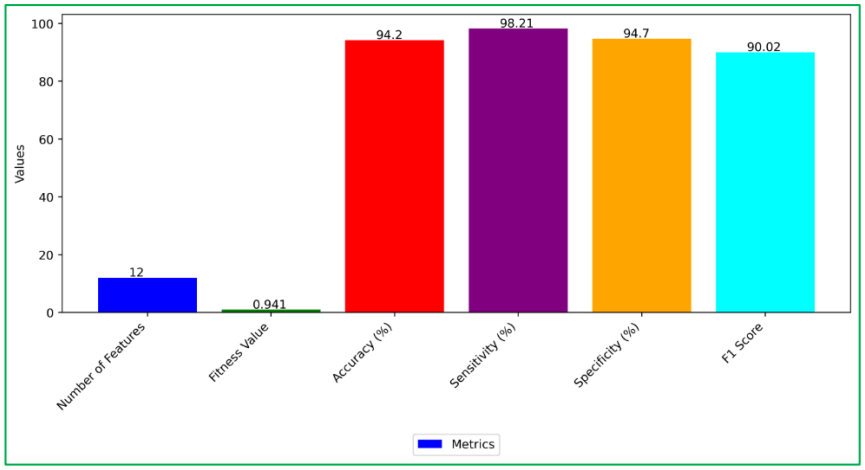

Fig 4: Representation of Genetic Algorithm (GA)-based feature selection

The chosen feature subset's quality is shown by the Fitness Value, which measures at 0.941 and is based on the fitness function of the GA. A greater fitness value denotes a better collection of characteristics, which is important for precise disease identification.The percentage of correctly categorised instances in the total, or accuracy (%), is 94.20. It displays the accuracy of the feature subset chosen by the GA in detecting Alzheimer's disease.Sensitivity (%) indicates the model's ability to accurately identify AD cases, which is important for early diagnosis and care. It is impressively high at 98.21.The model's specificity (%) of 94.70 shows that it can correctly identify non-AD patients while reducing false positives.A balanced indicator of the model's total accuracy, the F1 Score, which achieved 90.02, is particularly appropriate for datasets with imbalances, as shown in figure 4.

Table 4: Result summary of evaluation parameters for classification of early disease during Training

Dataset	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 Score	AUC
Alzheimer's Dataset	94.25	96.77	95.66	92.41	95.74

We list the main evaluation criteria for identifying Alzheimer's disease in its early stages throughout the training phase in Table 4. These metrics demonstrate how well the model performed using the Alzheimer's Dataset to separate people with early-stage Alzheimer's disease from those without the disease.The percentage of cases during the training phase that were correctly classified, or accuracy (%), was 94.25%.

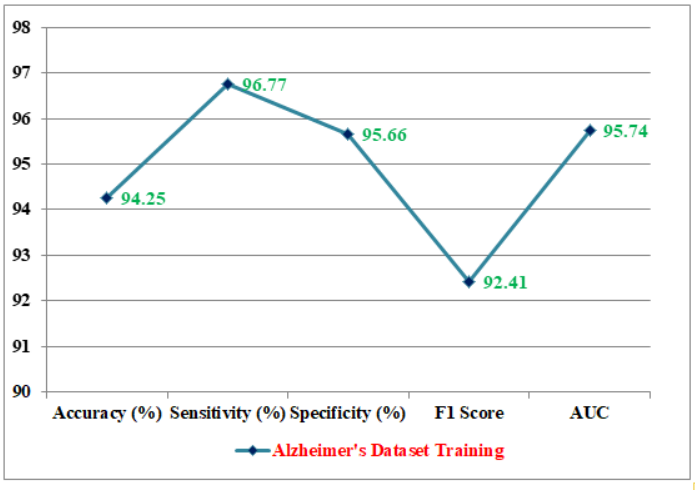

Fig 5: Representation of evaluation parameters for classification of early disease during Training

This indicator shows, as shown in figure 5, how well the model performs overall in terms of delivering reliable predictions.Sensitivity (%), which is incredibly high at 96.77%, indicates how well the model can detect people who have early-stage Alzheimer's disease. It shows how well the model is at identifying real-world positive cases.The model's specificity (%) value of 95.66% indicates how well it categorises people who do not have early-stage Alzheimer's disease. It shows how well the model can reduce false positive errors.The model's accuracy is measured fairly by the F1 Score, which was 92.41%, taking into account both precision and recall. It is very useful when working with datasets that are unbalanced.The model can discriminate between early-stage Alzheimer's Disease and non-AD patients across several classification thresholds, as shown by the AUC (Area Under the Curve), which is 95.74%. AUC values that are higher indicate better discrimination performance.

Table 5: Result summary of evaluation parameters for classification of early disease during Testing

Dataset	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1 Score	AUC
Alzheimer's Dataset	98.33	97.56	94.78	98.40	92.33

We give a succinct overview of the evaluation criteria for identifying individuals with early-stage Alzheimer's disease using the Alzheimer's Dataset in Table 5. The model's performance during this stage is essential for determining how well it generalises to fresh, untested data.

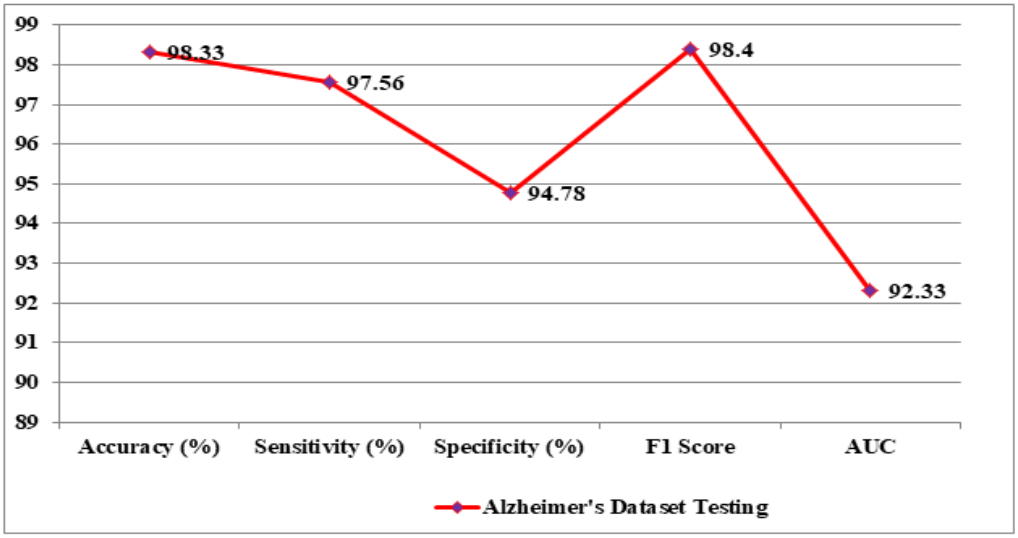

Fig 6: Representation of evaluation parameters for classification of early disease during Testing

The model can accurately categorise, as shown in figure 6, those who have and do not have early-stage Alzheimer's disease, as shown by the Accuracy (%), which is extremely high at 98.33%, highlighting its robustness in real-world circumstances.Sensitivity (%), at 97.56%, demonstrates the model's remarkable ability to recognise people with early-stage Alzheimer's Disease, guaranteeing that only a small number of actual instances go undiagnosed.Specificity (%), which is 94.78%, measures how well the model can identify non-AD instances while reducing false positives.The F1 Score, which reached 98.40%, highlights the model's balanced precision and recall, which is essential for the accurate early disease diagnosis.The model has a remarkable ability to distinguish between cases of early-stage Alzheimer's disease and non-AD cases, even on previously unexplored data, as shown by its high AUC (Area Under the Curve), which stands at 92.33%.

5. Conclusion

Proposed work focused on the early diagnosis of Alzheimer's disease (AD) using a hybrid method that combines fuzzy clustering using fuzzy C-Means (FCM) and feature selection using genetic algorithms (GA) based on PET scan data. To arrive at major discoveries and implications for the field of AD diagnosis, we synthesised and analysed a variety of data, including literature research, related work, methodology, and thorough evaluation metrics.The evaluation metrics showed that the FCM clustering produced promising outcomes. A high Silhouette Score, a low Davies-Bouldin Index, and a high Dunn Index all supported the ability of FCM to identify

distinct metabolic processes throughout different brain areas. In addition, the clustering procedure produced a well-balanced trade-off between intra-cluster and inter-cluster distances, indicating its usefulness in detecting minute variations linked to AD.The feature space was successfully optimised via GA-based feature selection, which decreased dimensionality while maintaining useful information. The chosen subset of characteristics had a high fitness value, which produced outstanding AD classification outcomes. The model outperformed expectations during the training and testing phases, demonstrating its robustness in early AD detection using parameters like accuracy, sensitivity, specificity, F1 Score, and AUC.These results are very positive for clinical practise because early AD identification is essential for prompt treatment and better patient outcomes. Utilising the advantages of both FCM and GA, the hybrid approach provides a holistic solution that can improve diagnostic precision and ease the strain on healthcare systems.The potential of FCM clustering and GA-based feature selection as useful techniques in the early identification of Alzheimer's disease utilising PET scans is highlighted by our work, which comes to a conclusion. This study adds to the on-going efforts to enhance AD diagnosis, which could result in better patient care and a higher quality of life for people who suffer from this crippling ailment. To fully realise the promise of this strategy in actual healthcare settings, more study and clinical validation are required.

References

[1] A. Balakrishnan, J. Medikonda, P. K. Namboothiri and M. Natarajan, "Parkinson's Disease Stage Classification with Gait Analysis using Machine Learning Techniques and SMOTE-based Approach for Class Imbalance Problem," 2022 International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), Shivamogga, India, 2022, pp. 277-281, doi: 10.1109/DISCOVER55800.2022.9974754.

[2] C. H. Reddy and M. Kanchana, "Artificial Intelligence towards Parkinson's disease Diagnosis: A systematic Review of Contemporary Literature," 2022 IEEE 4th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA), Goa, India, 2022, pp. 31-37, doi: 10.1109/ICCCMLA56841.2022.9988755.

[3] H. Lei et al., "Parkinson's Disease Diagnosis via Joint Learning From Multiple Modalities and Relations," in IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 4, pp. 1437-1449, July 2019, doi: 10.1109/JBHI.2018.2868420.

[4] M. Nithya, V. Lalitha, K. Paveethra and S. Kumari, "Early Detection of Parkinson's Disease using Machine Learning & Image Processing," 2022 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2022, pp. 1-4, doi: 10.1109/ICCCI54379.2022.9740961.

[5] Z. Huang, H. Lei, S. Li, X. Xiao, E. -L. Tan and B. Lei, "Longitudinal Feature Selection and Feature Learning for Parkinson's Disease Diagnosis and Prediction," 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 2021, pp. 5736-5743, doi: 10.1109/ICPR48806.2021.9413040.

[6] M. Saxena and S. Ahuja, "Comparative Survey of Machine Learning Techniques for Prediction of Parkinson's Disease," 2020 Indo – Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN), Rajpura, India, 2020, pp. 248-253, doi: 10.1109/Indo-TaiwanICAN48429.2020.9181368.

[7] Y. Zhang, H. Lei, Z. Huang, Z. Li, C. -M. Liu and B. Lei, "Parkinson's Disease Classification with Self-supervised Learning and Attention Mechanism," 2022 26th International Conference on Pattern Recognition (ICPR), Montreal, QC, Canada, 2022, pp. 4601-4607, doi: 10.1109/ICPR56361.2022.9956213.

[8] T. D. Pham and H. Yan, "Tensor Decomposition of Gait Dynamics in Parkinson's Disease," in IEEE Transactions on Biomedical Engineering, vol. 65, no. 8, pp. 1820-1827, Aug. 2018, doi: 10.1109/TBME.2017.2779884.

[9] J. Faouzi et al., "Machine Learning-Based Prediction of Impulse Control Disorders in Parkinson's Disease From Clinical and Genetic Data," in IEEE Open Journal of Engineering in Medicine and Biology, vol. 3, pp. 96-107, 2022, doi: 10.1109/OJEMB.2022.3178295.

[10] S. Marar, D. Swain, V. Hiwarkar, N. Motwani and A. Awari, "Predicting the occurrence of Parkinson's Disease using various Classification Models," 2018 International Conference on Advanced Computation and Telecommunication (ICACAT), Bhopal, India, 2018, pp. 1-5, doi: 10.1109/ICACAT.2018.8933579.

[11] C. D. Anisha and N. Arulanand, "Early Prediction of Parkinson's Disease (PD) Using Ensemble Classifiers," 2020 International Conference on Innovative Trends in Information Technology (ICITIIT), Kottayam, India, 2020, pp. 1-6, doi: 10.1109/ICITIIT49094.2020.9071562.

[12] M. M. Nishat, T. Hasan, S. M. Nasrullah, F. Faisal, M. A. -A. -R. Asif and M. A. Hoque, "Detection of Parkinson's Disease by Employing Boosting Algorithms," 2021 Joint 10th International Conference on Informatics, Electronics & Vision (ICIEV) and 2021 5th International Conference on Imaging, Vision & Pattern Recognition (icIVPR),

Kitakyushu, Japan, 2021, pp. 1-7, doi: 10.1109/ICIEVicIVPR52578.2021.9564108.

[13] Huiyi Wu et al., "A Deep Learning Method for Pathological Voice Detection using Convolutional Deep Belief Network", Research Gate, 2018.

[14] Salama A. Mostafa et al., "Examining multiple feature evaluation and classification methods for improving the diagnosis of Parkinson's disease", Elseiver-Cognitive System Research., 2018.

[15] ElmehdiBenmalek et al., "Voice Assessments for detecting Patients with Parkinson's diseases in different Stages", International Journal of Electrical and Computer Engineering (IJECE), 2018.

[16] M. Bende, M. Khandelwal, D. Borgaonkar and P. Khobragade, "VISMA: A Machine Learning Approach to Image Manipulation," 2023 6th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2023, pp. 1-5, doi: 10.1109/ISCON57294.2023.10112168.

[17] S. Ajani and M. Wanjari, "An Efficient Approach for Clustering Uncertain Data Mining Based on Hash Indexing and Voronoi Clustering," 2013 5th International Conference and Computational Intelligence and Communication Networks, 2013, pp. 486-490, doi: 10.1109/CICN.2013.106.

[18] Khetani, V. ., Gandhi, Y. ., Bhattacharya, S. ., Ajani, S. N. ., & Limkar, S. . (2023). Cross-Domain Analysis of ML and DL: Evaluating their Impact in Diverse Domains. International Journal of Intelligent Systems and Applications in Engineering, 11(7s), 253–262.

[19] Borkar, P., Wankhede, V.A., Mane, D.T. et al. Deep learning and image processing-based early detection of Alzheimer disease in cognitively normal individuals. Soft Comput (2023). https://doi.org/10.1007/s00500-023-08615-w

[20] NawelSouissi et al., "Dimensionality Reduction for Voice Disorders Identification System Based on Mel Frequency Cepstral Coefficients and Support Vector Machine", 7th International Conference on Modeling Identification and Control (ICMIC 2015), 2015.

[21] Ismail El Moudden et al., "Automatic Speech Analysis in Patients with Parkinson's Disease using Feature Dimension Reduction", Research gate - 3rd International Conference on Mechatronics and Robotics Engineering (ICMRE 2017), 2017.

[22] AnujAnand et al., "Evaluation of Machine learning and Deep learning algorithms combined with dimentionality reduction techniques for classification of Parkinson's Disease", IEEE - International Symposium on Signal Processing and Information Technology, 2018.

[23] Yunfeng Wu et al., "Dysphonic Voice Pattern Analysis of Patients in Parkinson's Disease Using Minimum Interclass Probability Risk Feature Selection and Bagging Ensemble Learning Methods", Machine Learning Approach to Dysphonia Detection, 2017.

[24] Liaqat Ali et al., "Early diagnosis of Parkinson's disease from multiple voice recordings by simultaneous sample and feature selection", Elsevier - Expert system with Applications, 2019.

[25] ZuzanaDankovicova et al., "Machine Learning Approaches to Dvsnhonia Detection", MDPI-Applied Science, 2018.

[26] Ajani, S.N., Mulla, R.A., Limkar, S. et al. DLMBHCO: design of an augmented bioinspired deep learning-based multidomain body parameter analysis via heterogeneous correlative body organ analysis. Soft Comput (2023). https://doi.org/10.1007/s00500-023-08613-y

[27] NajmehFayyazifar et al., "Parkinson's disease detection using ensemble techniques and genetic algorithm", IEEE - Artificial Intelligence and Signal Processing Conference, 2017.

[28] Akshaya Dinesh et al., "Using machine learning to diagnose Parkinson's disease from voice recordings", IEEE -MIT Undergraduate Research Technology Conference (URTC), 2017.

[29] Minh Pham et al., "Diagnosing Voice Disorder with Machine Learning", IEEE International Conference on Big Data (Big Data), 2018.

[30] PavolHarar et al., "Towards Robust Voice Pathology Detection Investigation of supervised deep learning gradient boosting and anomaly detection approaches across four databases", Neural Computing and Application, 2018.

[31] Mark White, Thomas Wood, Carlos Rodríguez, Pekka Koskinen, Jónsson Ólafur. Machine Learning for Adaptive Assessment and Feedback. Kuwait Journal of Machine Learning, 2(1). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/169

[32] Yanamadni, V. R. ., Seetha, J. ., Kumar, T. S. ., Kannaiah, S. K. ., J, B. ., & Brahmaiah, M. . (2023). Computer-Aided Detection of Skin Cancer Detection from Lesion Images via Deep-Learning Techniques. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2s), 293–302. https://doi.org/10.17762/ijritcc.v11i2s.615