Minimax Theory and its Applications

Volume 11 (2026), No. 3s, 609 to 627

Caste-Net: A Cross-Attention Swin Transformer With Dynamic Adaptive Algae Optimization For Explainable Multiclass Dementia Diagnosis From Brain Mri

Subba Rao P[1], Venubabu Rachapudi[2]

[1,2]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram-India
Email: psubbaraorao1979@gmail.com, venubabu.r@gmail.com

Abstract
Dementia is a disorder of the brain's neurons and making a good diagnosis and prediction of the disorder early may yield better treatment that will prevent the loss. Disease detection and classification requires optimized solutions, based on deep learning techniques, are complex. Motivated by these challenges, this study presents a novel Cross-Attention Swin Transformer Efficient Network (CASTE-Net)-based framework. The model's hyper-parameters are optimized using the improved optimization algorithm, resulting in enhanced learning capability and superior classification performance. Furthermore, noises can be eliminated from the pre-processing procedure of the MRI images via the combination of B-spline and Lanczos Interpolation without loss of the sensitive details.Image enhancement is first performed to enhance the quality of the MRI scans and enable robust feature extraction. The enhanced images are then decomposed into multi-scale frequency components to exploit both spatial and spectral information associated with brain abnormalities, after which attention is applied to extract deep features.To further reduce the complexity of the extracted features, and to improve the accuracy of the classification, the proposed model is adopted to select the optimal feature and, consequently, greatly reduces the processing overhead complexity in the selected features.The MRI images and their feature representation, along with the intricate patterns embedded in the MRI images, are used for the classification by the developed Cross-Attention Swin Transformer Efficient Network (CASTE-Net). The accuracy of the benchmark OASIS dementia dataset was 98% and various parameters were studied using the developed Quantum Convolutional Neural Network. With the developed work this guarantees a very accurate classification of dementia.

Keywords: dementia,optimization, network, complexity, deep learning

1. Introduction

Dementia is one of the neurological degenerative conditions that over time cause a progressive loss of cognitive function. Dementia affects language, memory and executive functioning generally badly. Moreover, it is associated with mental anguish, motor problems and decreased drive [1]. As illness progresses, these symptoms get worse, Dementia progressively impairs cognitive and functional abilities, reducing patients' independence and adversely affecting the daily functioning of affected individuals and care providers. Alzheimer's disease, the major cause of dementia worldwide, is predominantly associated with aging. Consequently, the increasing proportion of older adults is expected to result in a substantial rise in dementia prevalence over the coming decades. Estimates indicate that the global population aged 65 years and older will nearly quadruple between 2000 and 2050, highlighting the growing burden of dementia worldwide [2]. It is therefore expected that dementia is going to have a greater and growing impact on society. Globally, dementia affected an estimated 7.5 million people in 2015. According to WHO, the annual incidence of dementia is estimated to range from 10 to 15 cases per 1000 people based on multiple studies that allow continual assessment of the patient's health throughout the course of the study [3]. Of the people who develop dementia, a small proportion will live for 14 years or more, and the average survival time following a dementia diagnosis is approximately seven years. This alarming situation is encouraging the organizations and the academics to step up the activities in the prevention of dementia and early diagnosis [4]. A wide range of simple and affordable dementia tests can detect early symptoms such as sensory confusion, or more advanced conditions such as AD and MCI [5]. Cognitive functions cover focus, dialect,decision,recall and decision.Although severe memory loss could be a sign of

dementia, MCI might suggest that dementia is a possible diagnosis. An individual's cognitive abilities are fundamental to independent living and overall well-being. Disorders characterized by progressive cognitive decline, particularly AD, are now recognized as one of the most pressing public health issues. MRI images are valuable in diagnosis, to help rule out other diseases and to show any anatomical abnormalities like brain shrinkage [6]. Some of the atrophy features should be considered as signs of different types of dementia, but the MRI data should always be integrated with other diagnostic tests and medical evaluations. Early-stage dementia is commonly characterized by memory deficits. Individuals may have difficulty recalling recent events and often lose or misplace items in familiar environments. One of the first symptoms of Alzheimer's disease is memory loss [7]. There are cognitive and neurological measures that are used to assess brain and muscular function patients undergo comprehensive clinical examinations that assess cognitive and neurological performance, encompassing memory, executive function, language, numerical skills, balance, perceptual abilities, and reflex responses. MRIs can show if there is shrinkage or damage to certain areas of the brain. Serial brain imaging enables the detection of progressive structural alterations occurring over time. Reduction does not identify a specific diagnosis, but may indicate Alzheimer's disease or a similar disease of the nervous system [8]. It takes little time, few resources and no anesthetic and is easily done on any MRI program. Although MRI is a valuable neuroimaging modality, it has certain limitations in the accurate diagnosis of dementia. It is subject to inter-observer variation and does not have the same sensitivity as other volumetric assessments. Theauthor involves extending the IHOA with Quantum Convolutional Neural Network(CNN) that successfully classifies the available datasets into the dementia diseases [9] – [10].

Dementia is a neuro-generative disorder which impairs not only the normal life but also the quality of life, hence, the early identification of dementia is crucial for improving treatment outcomes and patient management [11].Earlier approaches were unable to produce reliable predictions because of excessive noise in the data and over-fitting during model training.Despite the promising performance of CNNs, their diagnostic capability is constrained by limited feature representation and non-optimal hyper-parameter settings. CNNs leverage quantum entanglement to extract more informative features; however, their effectiveness depends on appropriate hyper-parameter optimization. Therefore, this study proposes a hybrid framework that combines a CNN with the improved optimization algorithm to determine optimal hyper-parameters, resulting in improved feature learning, higher classification accuracy, and enhanced diagnostic performance [12] – [13]. The development of an efficient framework for dementia classification using the OASIS dataset. Initially, the MRI images are preprocessed by applying Spline-based interpolation in conjunction with Lanczos-based interpolation to reduce noise, improve image quality, and normalize intensity values. Subsequently, a U-Net architecture is employed to extract discriminative spatial and structural features associated with dementia [14]. The extracted features are then classified using a CNN, whose hyper-parameters are optimized through the improved optimization algorithm to enhance classification predictive performance and reliability [15]. The overall performance of the developed framework is comprehensively validated using quantitative performance metrics, including accuracy,positive predictive value, recall, ROC-AUC, and classification efficiency, demonstrating its superiority in multi-class dementia classification,feature selection through an enhanced Cross-Attention Swin Transformer Efficient Network (CASTE-Net)model using an exploitation strategy inspired by the mountaineering process to improve hyper-parameter tuning; stage-wise dementia classification with an optimized hyper-parameter tuning.

The subsequent sections is organized as described in Section 2, the corresponding works are analyzed and emphasis is given in detail. The developed work is explained in a broader context in section 3. Section 4 presents the simulation setup and experimental results. At the end of the work, you can find the results in section 5.

2. Related works

This section discusses various existing approaches where CAD is heavily relied on in the medical treatment of AD. In health care images like magnetic resonance imaging (MRI), for instance, however, techniques that rely on CNNs are very sensitive to small changes resulting from brain shrinkage. Mu et al. [16] proposed a lightweight DMA-HPCNet by incorporating modularized DMA and HPC modules into a ResNet backbone. In addition, a HFF strategy was introduced to capture discriminative features at multiple scales while preserving the global contextual information of MRI images, leading to improved medical image diagnosis.It facilitates comprehensive visualization of regional and whole-brain neural networks. As a result, an evaluation of the cognitive losses is missing [17]. Alzheimer's disease represents the most common form of dementia, affecting millions of individuals worldwide and causing progressive deterioration in memory, reasoning, and other cognitive functions.Due to the complex nature of the disease, early-stage diagnosis remains difficult but is essential for effective clinical management. Sarag et al. [18] explored machine learning-based approaches for dementia prediction by combining MRI scans with clinical records to enhance diagnostic accuracy and support early detection.A hybrid CNN-SVM approach to improve the reliability of predictions was proposed [19]. The

results allow information regarding individualized dementia therapy as they show the effectiveness of combination of medical and imaging information. Thus, an improvement of the performance was to be achieved.

HC fragmentation is crucial for the detection of common neurological diseases such as epilepsy and common neurodegenerative diseases such as dementia and Parkinson's [20]. A workable 3D HC identification problem from the MRI images was given and a shape driven loss function was provided along with an attention-unet based approach. De et al. [21] had developed a HOG which was designed to get the shape information. The proposed system has been calculated using pooling method.The lack of histogram information limits the use of variance-based loss functions during deep learning optimization through back-propagation. Although the proposed model demonstrated superior classification accuracy, insufficient clinical data constrained its practical clinical effectiveness [22]. In addition, fuzzy image analysis and the integration of 3D visualizations with 2D images were utilized to improve visualization quality and support context-aware interpretation of medical images. Maharjan et al. [23] proposed an 8D descriptor method for SURF to increase the computational reliability of the descriptor and suggested its use with an OFSA. In addition, dynamic three-dimensional images are overlaid on the initial image for optimal enhanced learning. The proposed method aims to produce high-quality images for reliable analysis by optimizing feature integration and improving the robustness of image processing [24]. The framework also enhances visualization and usability, enabling more effective interpretation of medical images. Moreover, the imaging features remain stable across complex diagnostic scenarios, ensuring consistent and reliable classification performance [25].

The authors introduced a CAUEEG database consisting of structured clinical descriptions to be used in the automated diagnosis of EEG symptoms. The dataset includes historical clinical information,participant clinical and demographic information Kim et al. [26] introduced the CEEDNet, which employs a single-stage deep learning framework architecture for automated EEG analysis. By integrating all stages of EEG processing into a single framework, CEEDNet streamlines the diagnostic workflow, limits dependence on human intervention, and boosts the efficiency and consistency of EEG-based evaluation.This will enable potential patients to be diagnosed properly through automated detection. However, it wasn't possible to deal with difficult situations.AD is a progressive neurodegenerative disorder whose increasing prevalence highlights the need for accurate early diagnosis. Ho et al. [27] introduced DeepADNet, a deep learning-based model for discriminating multiple Alzheimer's disease stages classification.The framework extracts ERSP features to capture temporal changes in EEG frequency bands and employs a hybrid CNN–LSTM network to learn discriminative time–frequency representations for distinguishing HC from individuals with AD. The method effectively models the spatiotemporal progression of Alzheimer's disease; however, its performance is constrained when applied to large and heterogeneous datasets. Bidani et al. [28] introduced a transfer learning-based DCNN for dementia detection and classification through brain MRI scans acquired from the OASIS dataset. The proposed model aims to identify dementia, a neurodegenerative disorder associated with progressive decline in memory, communication, and other cognitive abilities required for independent daily functioning. Images were extracted from the images using the features and the learning classification techniques. The dataset comprises three dementia classes: Non-Dementia, Very Mild Dementia, and Mild Dementia, based on brain MRI images.It greatly enhanced the accuracy of dementia assessments.Thus, it is imperative to examine the weaknesses of the technique.The two principal forms of neurodegenerative dementia include Alzheimer's disease and fronto-temporal dementia. Nevertheless, modern structural imaging methods have mainly been developed to diagnose each disorder independently, while the possibility of multiple concurrent diagnoses has received minimal attention.It utilizes a deep learning algorithm for accurate disease detection. Nguyen et al. [29] adapted a 3D U-Net model to detect structural regions in healthy individuals and patients with AD and FTD.The imaging data of these two diseases were used for disease detection can be used with a MLP classifier in various classification initiatives. Consequently, the deep learning algorithms were integrated using data-driven algorithms to enhance the algorithm's capabilities in terms of robustness and discrimination. However, these techniques usually do not give a clear interpretation of the results [30]. Table 1 summarizes the literature survey.

Table 1. Comparison of existing approaches

Ref.	Dataset	Methodology	Advantages	Limitations
[16]	Brain MRI	DMA-HPCNet with HFF and ResNet	Multi-scale feature extraction and global context learning	Limited cognitive loss assessment
[18]	MRI + Clinical Data	Hybrid CNN-SVM	Combines imaging and clinical information	Limited feature optimization
[20]	MRI	Attention U-Net	Accurate hippocampus localization	Limited disease generalization
[21]	MRI	HOG-based Shape	Captures structural	Poor variance learning

		Descriptor	information	
[23]	MRI	SURF + OFSA	Improved descriptor robustness	High computational complexity
[26]	EEG	CEEDNet	End-to-end EEG diagnosis	Performance decreases for complex EEG patterns
[27]	EEG	DeepADNet (CNN-LSTM)	Learns temporal EEG patterns	Limited scalability on heterogeneous datasets
[28]	OASIS MRI	Transfer Learning DCNN	High MRI classification accuracy	Weak interpretability
[29]	MRI	3D U-Net + MLP	Structural brain segmentation	Poor explainability
[17]	Multimodal Brain Images	Adversarial Deep Learning	Robust multimodal representation	High computational cost
[19]	Cognitive Assessment	Augmented Reality Framework	Improved visualization	Not intended for automatic classification
[24]	Medical Images	Image Enhancement	Better interpolation quality	Does not perform classification

3. Methodology

Accurate classification of brain images into non-dementia, very mild dementia, mild dementia, and moderate dementia is essential for the early diagnosis of dementia and timely clinical intervention. Initially, the input images undergo preprocessing through Spline-based interpolation coupled with Lanczos-based interpolation techniques to reduce noise and correct intensity variations. Subsequently, discriminative Feature representations are obtained using the Cross-Attention Swin Transformer Efficient Network (CASTE-Net) architecture, while the Dynamic Adaptive Algae Optimization (DAAO)is employed to select the most relevant features. The selected features are then passed to the classification stage, where a Cross-Attention Swin Transformer Efficient Network (CASTE-Net) integrated with the DAAO is utilized to classify the different stages of dementia. The overall workflow of the developed dementia classification framework is illustrated in Fig 1.

3.1. Pre-processing

Accurate dementia classification requires high-quality MRI images; however, the acquired images are often affected by noise and artifacts. Therefore, spline-based interpolation coupled with Lanczos-based interpolation are applied to enhance image quality by reducing noise, correcting intensity variations, and preserving image resolution. The enhanced images are then used for feature extraction.

3.1.1. Interpolation

This method will smooth MRI images without destroying the anatomical structures. The unnecessary intensity variations and noises are reduced to enhance the regions of interest. It can be expressed as Eq. (1):

$$BS_l(V) = \sum_{j=0}^{l+1} G_j BS_{j,l}\,(v) \tag{1}$$

Here, the basis functions are $BSl(y)$ and the control points are G_j which affect the spline curve. The intensity of the input MRI image is represented by $I(y,z)$ while the corresponding smoothed (pre-processed) image, $\hat{I}(y,z)$ is estimated using Eq. (2).

$$\hat{I}(y,z) = \sum_{j,i} BS_j(y-j)BS_i(z-i)I(j,i) \tag{2}$$

The original image weighting parameter is $W_{j,i}$. The developed method decreases the noises from the scanner and enhance the information of the structure. This improves the ability to extract features when it comes to dementia classification.

3.1.2. Re-sampling

The spatial detail of the MRI images was found to change according to the MRI scan protocols, and Lanczos-based interpolation is used for resampling images. The convolution kernel of the Lanczos functionis defined as Eq. (3):

$$LR(v) = \begin{cases} \dfrac{\sin(\prod v)\sin\left(\frac{\pi v}{b}\right)}{\prod v^2} & if\ |v| < b \\ 0 & else \end{cases} \tag{3}$$

The parameter for the window used in lanczos is b and the distance between the points used for the interpolation is y. The intensity function that will be sampled in the new pixel coordinates can be determined as Eq. (4).

$$\hat{I}(y,z) = \sum_{j,i} I(j,i)LR(y-j)LR(z-i) \tag{4}$$

To preserve high-frequency details, the proposed method is applied.

3.1.3. Correlation and intensity and normalization

The strength of the MRI images are modulated according to the type of radiofrequency coil. This may have resulted in bias field correction being used to standardize the signals to prevent this. Histogram matching and intensity rescaling are employed to standardize voxel intensities across MRI images, thereby providing consistent input data for the effective training of the proposed CASTE-Net model.

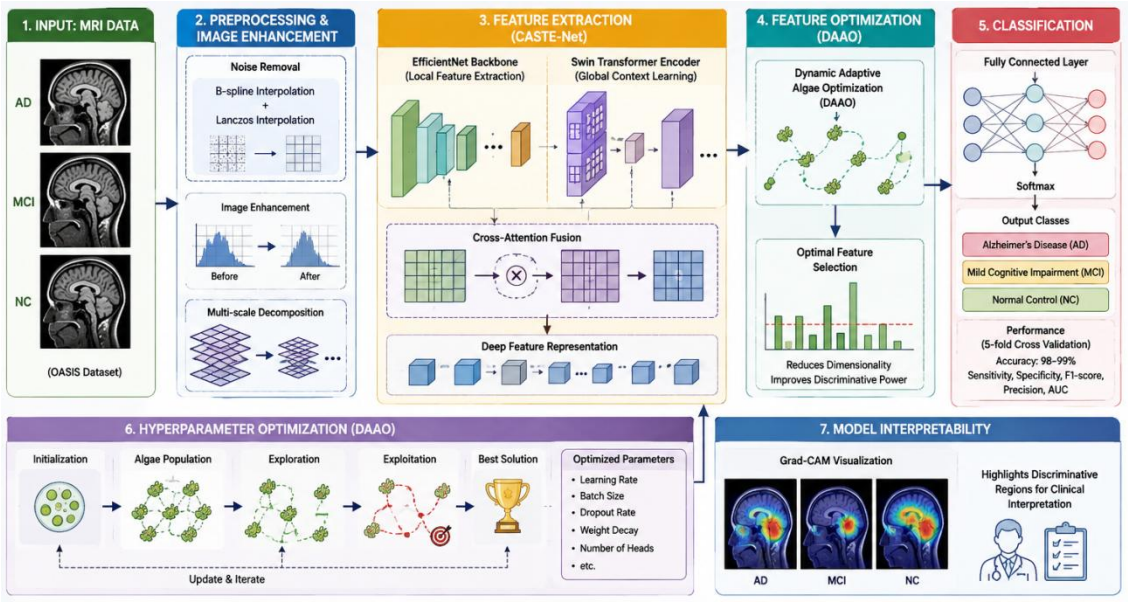


Fig 1. Overall workflow

3.2. Feature analysis

The pre-processed MRI images are processed using the U-Net architecture to extract spatial and spectral features. The encoder progressively learns low-level and high-level feature representations, while the decoder reconstructs and refines these features by integrating semantic and spatial information. The extracted feature representations are then used for dementia classification. This is done by a cross-entry loss function, which will secure the trainable weights during the extraction process. Fig 1 presents a sample feature map extracted by the U-Net architecture. During the model training stage, the $N-$band MRI image serves as the input for the mathematical formulation of the feature extraction process. The U-Net output produces the logits $\hat{z}$ associated with the target class y. The pixel-wise probability of the predicted class is defined in Eq. (5).

$$\hat{z} = F^{UNet}(y) \tag{5}$$

The proposed model extracts multi-level feature representations from N hierarchical levels. The feature map corresponding to level k for each feature $\theta^{l}_{p,q,k}$is defined in Eq. (6).

$$\theta^{l}_{p,q,k} = \sum_{c}\sum_{j}\sum_{i} \theta^{l-1}_{p+j,\ q+i,\ c} * W_{j,i,c,k} + a_l \tag{6}$$

Through successive convolutional and max-pooling operations, the U-Net extracts high-level features by progressively enlarging the receptive field. The hierarchical feature maps are then combined using the UNet function via a concatenation operation, as expressed in Eq. (7).

$$\theta^k = [\theta^{N-k+1,}\theta^{k-1}] \tag{7}$$

To address class imbalance and enhance the learning of minority class features, a learnable categorical weight vector β is integrated into the class-balanced cross-entropy objective function. The objective function is mathematically expressed as presented in Eq. (8).

$$LE(z,\hat{z}) = -(\beta.z)^T \log softmax\ (\hat{z}) \tag{8}$$

The formal representation of the softmax function is presented in Eq. (9).

$$softmax\ (\hat{z}) = \frac{\exp(\hat{z})}{1^T \exp(\hat{z})} \tag{9}$$

The ground-truth label is represented by the one-hot encoded vector z, whereas $\hat{z}$denotes the predicted output of the UNet function for the input y. The formulation of the adaptive weight is presented in Eq. (10).

$$\beta_j = \frac{1}{M_j}\max(M) \tag{10}$$

In this case, adaptive weight will be β_j and the total number of categories will be Mj.

3.3. Feature selection with Dynamic Adaptive Algae Optimization (DAAO)

Meta-heuristic search optimization techniques intends to determine the best possible solution by exploring the solution space based on a predefined search mechanism, beginning with randomly generated initial solutions. Their hyper-parameters regulate the trade-off between exploration of new search regions and exploitation of existing high-quality solutions. Inspired by natural processes and biological behaviors, these algorithms efficiently solve complex optimization problems by balancing global and local search capabilities. The balance between exploration and exploitation is a trade-off; so finding an optimal trade-off between these two aspects in meta-heuristic algorithms is crucial to obtain optimum solution in shorter time without falling into local minimum points. Chaos theory focuses on the behavior of nonlinear dynamic systems that are extremely sensitive to initial conditions. A minor perturbation in the initial parameters can produce substantial differences in the system's subsequent behavior and output. Owing to this sensitivity, chaotic systems provide enhanced randomness and search diversity, which can improve the exploration capability of meta-heuristic optimization algorithms. Unlike random processes, chaotic systems are deterministic and do not rely on random parameters, despite exhibiting seemingly irregular and unpredictable behavior. Another important chaotic systems are characterized by ergodicity. The ergodic property ensures that chaotic variables traverse the entire search space in a non-repetitive manner while adhering to deterministic rules. This property enhances the exploration capability of optimization algorithms and helps prevent premature convergence to local optima. Owing to these advantages, chaotic maps derived from chaotic systems are widely incorporated into population-based optimization algorithms to improve their search efficiency, convergence speed, and overall optimization performance.Chaotic maps are deterministic systems with stochastic characteristics, nonlinear and ways of generating long-term random numbers. The numbers in the maps generated by these sequences have many significant properties: The advantages of chaotic sequences include non-repetition, wide-range value distribution, and low computational requirements for sequence generation and storage. However, random numbers employed in meta-heuristic algorithms may contain duplicate values and insufficient diversity within the search space, potentially leading to premature convergence and trapping the algorithm in local optima.Consequently, chaotic sequences with full-spectrum coverage and reduced repetition can be produced.

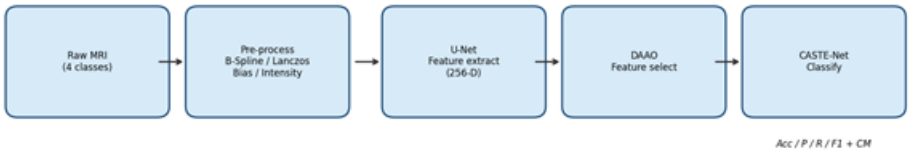


Fig 2. Block diagram of proposed model

The mathematical formulations of these chaotic maps are used to generate deterministic chaotic sequences from an initial parameter. These sequences replace conventional random numbers in meta-heuristic algorithms, thereby improving population diversity and reducing the possibility of convergence to local optima. These numbers in this series will probably not be the same and will be distributed over a large range. The initial parameter value 0.7 was chosen for the chaotic maps in this study as is done. Fig 2 shows the spectra of 100 numbers that are generated by the Chaotic maps equations with the initial parameter x1 0.7. The theory of chaos is the study of chaotic modelling of non-linear systems. Nonlinear deterministic systems are called chaotic systems and exhibit apparently random behavior that is sensitive to initial conditions. Owing to their unique characteristics, chaotic systems have been effectively utilized in meta-heuristic optimization algorithms. Chaotic mechanisms can be integrated into different phases of optimization methods, such as generating the initial population, enhancing the movement strategies of candidate solutions, adjusting control parameters, and combining multiple optimization techniques. Several chaotic meta-heuristic algorithms have been developed in previous studies. Among the existing approaches investigated in this study, the most efficient method is presented below:

3.3.1. Adaptation process

The primary operational workflow is to decide between an optimization process or not adopting one. In the standard algorithm this choice is based on the random number being below 0.5, i.e., if the number obtained by random selection is less than 0.5, the adaptation process is executed, otherwise the self-adjusting foraging mechanism is not executed. A novel algorithm, termed DAAO, is introduced by integrating a chaos-driven balancing mechanism into the selection process. The key component of DAAO is the balance operation, where ten popular chaotic maps are employed to regulate the trade-off between exploration and exploitation. These chaotic maps are embedded within the adaptation mechanism to improve the optimization process. Consequently, the conventional random value (rand) in the selection operator is substituted with chaotic sequence values, as formulated in Eq. (11).

$$if\ C(t) = Ap\ then\ starving^{t+1} = starving^{t} + (biggest^{t} - starving^{t}) * rand\ endif \qquad (11)$$

Where,$C(t)$ represents the value generated by the chaotic mapping function during the t-th iteration. The term starving t refers to the algal colony exhibiting the extreme starvation state level at time t, whereas biggest t denotes the best-performing algal population. The parameter rand is a random number selected from the interval $[0,1]$. The workflow diagram describing the chaotic DAAOframework is presented in Fig 2.

3.4. Prediction

The proposed model can be divided into two sections:EfficientNet-B0 model and Swin Transformer which can be aptly combined to benefit from the strengths of both CNN-based and transformer-based techniques in clinical image classification. It is a free integration with an optimized attention mechanism that dynamically combines multiple information to enhance the model's robustness and interpretability. Additionally, it proposed an advanced hybrid framework inspired by the merits of two models: The proposed framework combines the local feature extraction capability of EfficientNet-B0 with enhanced long-range feature learning ability of the hierarchical Swin Transformer framework known as Cross-Attention Swin Transformer Efficient Network (CASTE-Net).An attention refinement moduleis employed to fuse these complementary features adaptively based on the requirements of the classification task.Fig 3 presents the CASTE-Net deep learning framework, which combines the complementary capabilities of EfficientNet-B0 and hierarchical swin transformer frameworks for extracting local spatial features and modeling global contextual information from MRI images. The proposed CASTE-Net architecture facilitates accurate dementia diagnosis and classification, as demonstrated in Fig 3. EfficientNet-B0 is responsible for learning fine-grained local features through optimized convolutional blocks, while the Swin Transformer models long-range dependencies and hierarchical global context using shifted-window attention across multiple stages. The fusion of local and global representations

improves the robustness and accuracy of dementia classification. An attention-based feature fusion module is employed to combine complementary local and global features while highlighting the most relevant information. The fused features are then passed to a robust multi-class classification network. Additionally, Grad-CAM is utilized to generate visual explanations of the model's predictions, enabling clinicians to identify the key brain regions influencing the classification results and improving the transparency of the proposed framework. This hybrid approach is able to successfully balance local specificity and global context, allowing for more accurate and reliable diagnostic results and prediction of dementia.

3.4.1. Local feature analysis

EfficientNet-B0 is applied for learning low- and mid-level local feature representations from MRI scans and is a backbone used in CNN. To enable efficientNetB0 to be dynamic in terms of the width, depth and input resolution of the network for each task, it implements a compound scaling strategy with a series of well-designed coefficients depending on the specific task. This approach is to overcome the problem of "best for less" factors. EfficientNet-B0 performs well on tasks of learning fine-grained local spatial representations from MRI images, which include tissue intensity patterns, structural boundaries, and localized pathological changes, and are critical for accurate diagnosis of dementia.EfficientNet-B0 combines MBConv blocks with depth-wise separable convolutions and SE mechanisms to efficiently model channel-wise feature dependencies. This design allows the network to detect very small changes in the structure of MRI brain images that can indicate the various phases of disease progression. Moreover, the light architecture and parameter efficiency both make EfficientNet-B0 a good option as the core feature extractor of the proposed model. It gathers the anatomical information in 2D MRI slices with minimum processing workload. It utilizes a compound scaling approach for model scaling the depth of the network (d), width (w) and input image size (r) using uniform scaling. The scaling strategy can be mathematically represented as follows:

$$d = \alpha^{\emptyset} \tag{12}$$

$$w = \beta^{\emptyset} \tag{13}$$

$$r = \gamma^{\emptyset} \tag{14}$$

subject to the constraint:

$$\propto . \beta^2 . \gamma^2 \cong 2;\ \alpha, \beta, \gamma > 0 \tag{15}$$

The model consists mainly of an MBConv layer which is a modified version of the original MBConv with the addition of a channel-wise attention module called SE block. The SE block calculates:

$$S = \sigma(W_2 . \delta(W_1 . GAP(X))) \tag{16}$$

In this case, $GAP(X) \epsilon R^c$ is the average of all global vectors, $ReLU(\cdot)$ is the rectified linear unit activation function, sigmoid $(\cdot)$ is the sigmoid function, and W_1, W_2 are trainable weights. The feature map is then recalibrated through multiplication element-wise in the channels:

$$X_{SE} = X \otimes \mathrm{S} \tag{17}$$

For a given input image I, the local feature representation obtained from the final layer of the EfficientNet-B0 model is expressed as:

$$F_{local} = ENetB0\ (I) \tag{18}$$

3.4.2. Global analysis

The swin transformer is integrated with EfficientNet-B0 to enhance feature learning by capturing global contextual information that complements the local spatial features extracted by EfficientNet-B0. In comparison with the ViT which processes the entire image as flattened patches with rather than relying on quadratic computational complexity, the swin transformeremploys a hierarchical architecture based on local, non-overlapping window attention. This design enables efficient modeling of long-range dependencies while maintaining computational efficiency and robust feature representation. The windows are placed in between layers so as to allow cross windows feature sharing and hence improve the efficiency, scalability, and yet maintain the global awareness. Each of the swin transformer levels increases the receptive field, allowing the swin transformer to gather information from more global features. To enhance global feature learning, swin

transformer module is integrated after the EfficientNet-B0 feature extractor. The swin transformer efficiently identifies pathological characteristics, such as ventricular enlargement, cortical thinning, and global cerebral atrophy that are indicative of dementia. Using a shifted-window self-attention mechanism, it models long-range dependencies and hierarchical contextual information across spatial patches. This global representation complements the fine-grained local features learned by EfficientNet-B0, resulting in a more comprehensive and robust feature representation for dementia classification.The input image is divided into patches $P \in R^{M\times M\times C}$ which are subsequently mapped into an embedding space to obtain patch representations $z_0 \in R^{N\times D}$ with N representing the number of patches and D representing the embedding dimension. Self-attention is calculated as: In this case, the query matrix, the key matrix and the value matrix are linear projections:

$$Attention\ (Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}} + B\right)V \tag{19}$$

Here, X is the matrix of input features, $W_Q, W_K\ and\ W_V$ are matrices of learnable weights and d_k is the dimensionality of the key vectors. Here, R stands for relative positional bias while B represents the relative positional bias.

$$Q = XW_Q \tag{20}$$

$$K = XW_K \tag{21}$$

$$V = XW_V \tag{22}$$

Window partitions are moved from one window to the next to allow for cross window communication. For an input image I, the output of the swin transformer is a feature representation I that is the final feature representation.

$$F_{global} = swin\ transformer\ (I) \tag{23}$$

3.4.3. Attention mechanism

In this work, they built a novel attention mechanism to combine the spatially localized features of an attention-based fusion layer is employed to integrate the local spatial features extracted by EfficientNet-B0 with the global contextual representations learned by the swim transformer. The fusion module adaptively computes attention weights to highlight complementary information from both branches while suppressing irrelevant feature responses. This adaptive weighting mechanism enables the network to concentrate on disease-relevant anatomical structures, including the hippocampus and parietal cortex, which enhances the discriminative ability of the model capability of the fused feature representations. By dynamically balancing the contributions of local and global features, the attention module enhances the robustness and accuracy of dementia classification. This can be done either by leveraging channel-based and spatial-based attention strategies, as in the case of CBAM or by explicitly learning the inter-dependencies between channel and spatial feature modalities using cross-attention modules. The attention process generates a higher level and more focused representation by emphasizing relevant features and suppressing redundant or irrelevant data. This allows the model to learn about the most important indicators of diagnosis that are not captured by convolutional or transformer networks. An adaptive attention mechanism is used to combine local attention and global attention representations. First, all the feature vectors are concatenated:

$$F_{concat} = [F_{local}\ ||F_{global}] \tag{24}$$

Next, an attention gate is calculated via a projection operation followed by a sigmoid function:

$$a = \sigma(W_a.F_{concat} + b_a) \tag{25}$$

In the final stage, the features are combined using a gated fusion strategy:

$$F_{fused} = a\ \otimes \mathrm{F}_{local} + (1-a) \otimes \mathrm{F}_{global} \tag{26}$$

In the above expression, matrix-wise element multiplication is represented by '$\otimes$'. The terms $W_a\ and\ b_a$ denote the trainable weight matrix and bias vector, respectively, whereas $\sigma(\cdot)$ refers to the sigmoid activation function.

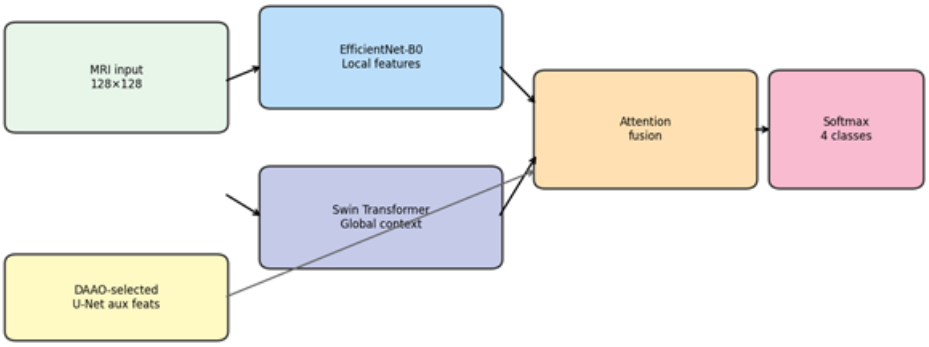


Fig 3. DAAO with feature selection model

3.4.4. Classification

The combined feature representation obtained after feature fusion is passed through a fully connected layer before being classified using a softmax classifier. The classification head comprises a linear mapping followed by softmax activation function which computes the probability distribution over the dementia stages. The output class is identified by maximizing the predicted probability, which is mathematically represented as:

$$F_{fused} = a \otimes \mathrm{F}_{local} + (1 - a) \otimes \mathrm{F}_{global} \tag{27}$$

In the third classification layer, W_c *and* b_c are the learnable parameters and $\hat{y} \in R^K$ corresponds to the predicted class probability distribution over the K classes of dementia. The developed framework is a hybrid learning, which uses the convolutional and transformer-based feature extraction networks through an attention-based feature fusion module. This architecture is able to learn fine-grained local features as well as hierarchical global contextual information, which provides more discriminative feature representations for accurate multi-class differential classification of dementia stages. The detailed procedure to the proposed model is presented as follows: For deep learning, the training and evaluation of the models are crucial to create and refine models which are capable of effectively performing tasks. The Adam optimization algorithm was employed to update the network parameters, with an initial learning rate set to 1×10^{-4} throughout the training phase. A cosine annealing learning rate scheduler is incorporated to gradually decrease the learning rate, thereby enhancing convergence during optimization. The training process is conducted for up to 50 epochs, and an early stopping criterion based on validation loss is utilized to prevent unnecessary training once the model performance reaches saturation. A categorical cross-entropy loss function along with class weighting is used for the imbalanced class distribution and enhance the classification performance over all stages of the dementia. Data set has been separated into training and test datasets. During training, the model parameters were optimized over multiple epochs using the Adam optimizer to reduce the multi-class cross-entropy loss function. During the training, the objective loss function employed is the cross-entropy objective for multi-class prediction, and the classification performance is evaluated. The weights of the model are updated using the Adam optimizer in the training. The model is trained on the training data for a specified number of epochs, while the loss and accuracy metrics are computed at every iteration to evaluate performance. To avoid over-fitting, we used early stopping in our training with monitoring by the validation loss. The performance of validation was checked after every epoch, and a model with early stopping of 5 epochs was used. To avoid over-fitting, training was terminated after 5 epochs if the loss on the validation data did not improve. The model with the least validation loss was saved and automatically selected as the best checkpoint to evaluate on the test data set. In this case, the best results to be reported would be the performance in the process of generalization, rather than reporting only from the final epoch of training. The model's ability to classify the images was evaluated through a confusion matrix as well as quantitative accuracy, precision, recall, and F1-score. The dataset was split into separate training–validation and testing sets to ensure no data leakage and to get an unbiased estimation of the generalization ability of the model. The dataset was partitioned such that 80% of the samples were used for training and validation, whereas the remaining 20% constituted an independent test set for performance evaluation.

This work used 5-fold stratified CV in 80% of the training phase to optimize the parameters and to minimize over-fitting. The independent test set was used only to assess the final classification performance and model interpretability through accuracy, precision, recall, F1-score, and Grad-CAM analysis ensuring that no test samples contributed to model training.It is guaranteed that the measurement is unbiased and aligned with the best practices for ML applications used for clinical purposes.The critical hyper-parameters like the learning rate,

cosine annealing strategy, class-balanced categorical cross-entropy loss and dropout regularization were carefully adjusted to achieve stable training and enhance generalization performance.To ensure fairness and reduce over-fitting, the model was validated using the 5-fold CV method. In medical imaging, the use of Grad-CAM helped in making the models easier to interpret by highlighting important regions that influenced the model's predictions. The proposed CASTE-Net training and evaluation strategy enables accurate and reliable multi-class classification of dementia. To improve model interpretability, CAM is utilized to visualize the regions of MRI images that have the greatest influence on the predicted class. CAM generates class-specific heatmaps that emphasize the most discriminative anatomical regions responsible for the model's decision, thereby enhancing the explainability of the proposed CASTE-Net framework. The localization map is obtained by calculating the weighted combination of the feature maps from the final convolutional layer using the class-specific weights of the output layer. This map indicates the most important parts of an image to accurately predict a specific class. CAM can be applied to medical images where models where information is most important in an image. To assess the stability and generalization performance of the proposed CASTE-Net framework, 5-fold CV was performed. The dataset was evaluated using a 5-fold CV approach, in which it was partitioned into five non-overlapping subsets. For each fold, four subsets were used to train the model and the remaining subset was used for validation. The procedure continued until every subset had served once as the validation set. The overall classification performance was determined by averaging the evaluation metrics across the five folds, ensuring a reliable and unbiased estimation of the model's generalization ability. The use of 5-fold CV reduces over-fitting and improves the reliability of the experimental results by ensuring consistent performance across different data partitions. Effective training of the proposed CASTE-Net framework requires appropriate optimization strategies to maximize classification performance. An early stopping mechanism was subsequently applied by monitoring the validation loss to avoid over-fitting and ensure better generalization. The architecture weights associated with the minimum validation loss were automatically retained as the best-performing checkpoint and subsequently used for final testing and performance evaluation.

4. Numerical results and discussion

The developed CASTE-Net framework was analyzed through a comprehensive performance analysis using training and validation loss, training and validation accuracy, confusion matrix, accuracy versus training time, FPR, FNR, AUC, accuracy, precision, recall, specificity, and F1-score. The obtained results were benchmarked against existing state-of-the-art approaches, namely CEEDNet, 3D-UNets, and DeepADNet to demonstrate the superiority of the developed method. The architecture was implemented using the Python programming language and trained using a dedicated high-performance workstation featuring with an NVIDIA Quadro RTX 6000 GPU with 24 GB of memory.

4.1. Dataset Details

MRI brain images were obtained from the OASIS dataset and used for training and evaluating the developed dementia classification model.The OASIS dataset was chosen due to its public availability, standardized T1-weighted MRI images, and its ability to support the early detection and multiclass classification of dementia.A total of 86,437 brain MRI imagesare divided into 4 classes namely normal with 13725 very mild dementia, 5002 mild dementia, 488 moderate dementia and 67222 Non-demented. We focused here on OASIS-Kaggle which is a more accessible resource and has standardized T1-weighted imaging, while the proposed model was developed and evaluated using the OASIS dataset, future validation on the ADNI dataset is planned to assess its clinical generalizability and robustness across diverse patient populations.

4.2. Performance Investigation

Comparison of the training and validation accuracies of the developed CASTE-Net model with existing approaches like CEEDNet, 3D U-Net, and DeepADNet. Training and validation accuracy curves of the proposed CASTE-Net model illustrating its learning behavior, convergence, and generalization performance as in Fig 4 and Fig 5. As the model learns meaningful feature representations, the training accuracy progressively improves. If the training and validation accuracy curves diverge substantially, the model is likely to experience over-fitting. In contrast, a similar trend between the two curves indicates effective learning and good generalization to unseen data.The training accuracy exhibits a steady increase during model training, and a similar trend is observed for the validation accuracy, stabilizing at a value close to the training accuracy. This behavior indicates effective learning with minimal over-fitting. It gradually raises the accuracy of the training and interpretation model is to memorize the training data. The training and validation accuracies of the proposed CASTE-Net model were marginally higher than those of the existing models. The developed model attained an overall accuracy of 98.0%, with a mean cross-validation accuracy of 97.3%, indicating its strong generalization capability and reliable classification performance (with a standard deviation of 0.7); besides, the model

presented good precision (97%), recall (96%), and F1-score (96.5%). This used 10-fold CV to prevent over-fitting and to achieve generalization. The mean classification accuracy obtained was 97.3% with an observed standard deviation of ±0.7. The CV accuracies obtained from the 10-fold CV were 96.8%, 97.1%, 97.9%, 98.0%, 97.5%, 96.9%, 97.8%, 97.2%, 97.4%, and 97.5%. The low variation in accuracy across the folds indicates that the proposed model exhibits stable, robust, and unbiased performance, with strong generalization across different data partitions as in Fig 6.

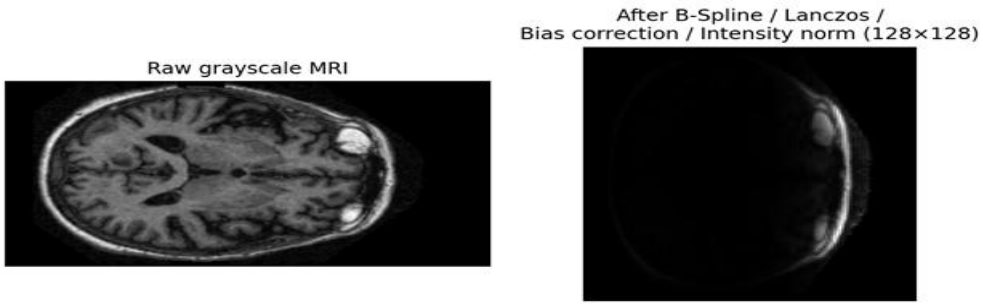


Fig 4. Pre-processed image

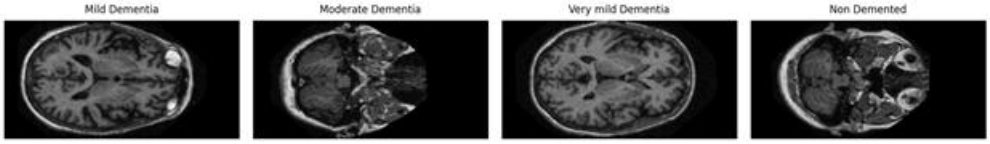


Fig 5. Classified image

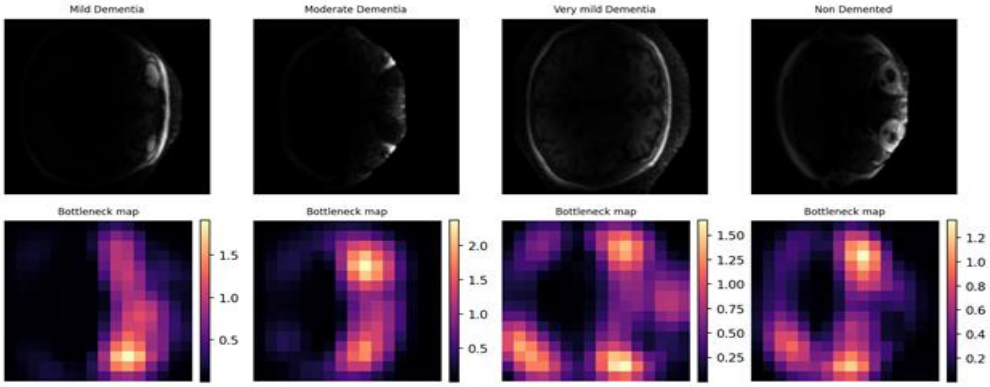


Fig 6. Classified and heatmap images

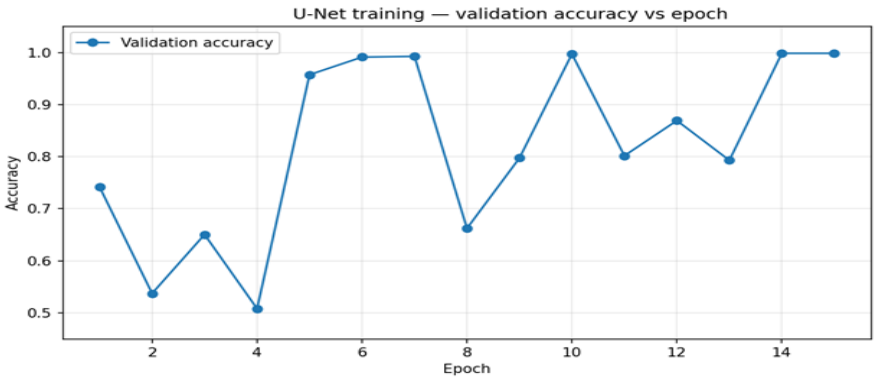


Fig 7. U-Net accuracy model

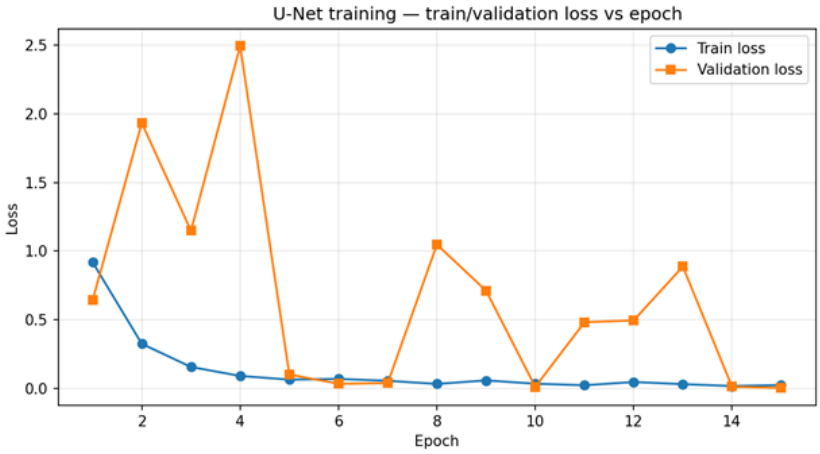


Fig 8. U-Net loss analysis

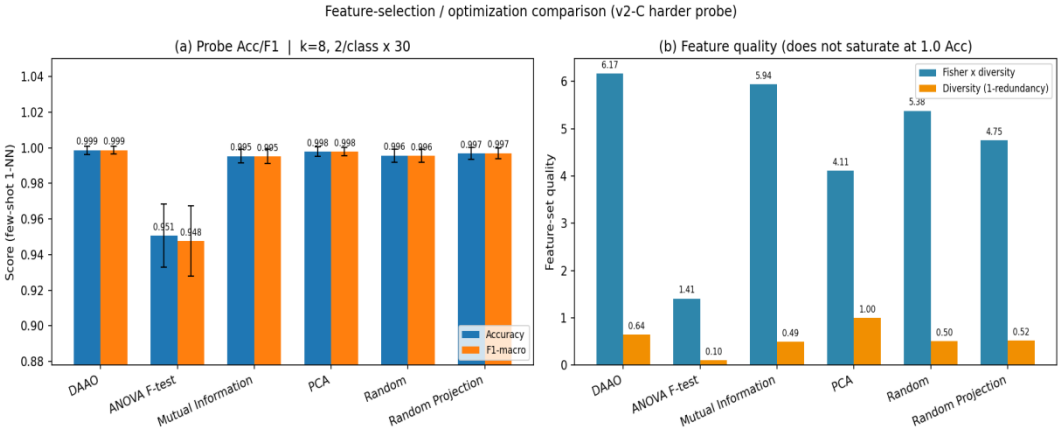


Fig 9. Feature selection and optimization comparison

Further insights on the convergence behavior and generalization performance of the proposed CASTE-Net model is given by the trainingaccuracy and the validationaccuracy curves presented in Fig 7. Comparing the training and validation loss of the proposed CASTE-Net model with existing models.Fig 8 indicates the number of training epochs (1 to 100) and the training and validation loss. In the multiclass classification task, the categorical cross-entropy was used as a loss function for optimizing the model in training.The training loss is steadily decreasing as the number of training epoch's increases, likewise the validation loss, but it eventually tends to converge towards the training loss. This is a convergence behavior and the proposed model is not over-fitting too much. Helpful to learn non beneficial features that are interpretable by the learned ones from MRI images that are over-fitted. Compared with existing models, the developed CASTE-Net model consistently exhibited higher training and validation accuracies in Fig 5 is for the AUC-ROC. The model value of AUC is used to correctly classify the negative and positive classes. This single threshold value is a plot of the mean AUC and the associated classification threshold. Using the graph, draw the recall and precision graph for a few classes (0-1). The performance of the optimized CASTE-NET for MRI-based dementia classification was analyzed using ROC curves and AUC values. The optimal model is characterized by an ROC curve that closely follows the upper-left corner, approaching the ideal point (0, 1), which indicates excellent discriminative ability. An AUC value of 0.5, represented by the diagonal line drawn by the random classifier. A more "reasonable" model will be curved and the more the further it's away from the diagonal towards increasing values of 1. The performance of the developed CASTE-Net model was also compared with the existing model performance and it was observed that the developed CASTE-Net model performs better in terms of ROC. Fig 9 draws the graph for the performance of time vs. accuracy. The percentages of accuracy and the various time in seconds required to evaluate the accuracy of the extracted feature including SURF, HFF, HOG, concatenation and developed CASTE-Net model. One of the crucial measures used in evaluating the effectiveness and efficiency is classification of dementia diseases – time vs. accuracy graph. The better performance was traded off for added

train time and the best accuracy was incorporated into the model for an evaluation of the speed of the model. The accuracy is boosted quickly in the first phase to more stable values. The proposed CASTE-NET demonstrated superior classification performance with minimal computational cost. As the computation time varied from 0s to 100s, the developed method consistently achieved higher accuracy than the existing feature extraction methods, namely HOG,SURF, HFF and feature concatenation. The results indicate that the developed CASTE-Net approach provides a suitable balance between computational cost and model performance and classification accuracy. Fig 9 is a comparative plot of the false negative rates. Minimizing the false negative rate is crucial to prevent undetected dementia cases and enhance the clinical reliability of the diagnostic system.Comparative false negative rates are used to measure the performance in the important task of detecting dementia diseases. The developed work has a lower false negative rate than the existing model which are 0.01.The lowest model is the best performing when compared to the other models (false detection rates 10). The techniques of existing approaches and developed CASTE-Net modelis to calculate rate of false positives. With dementia there is a rate of false positives as there are various models that analyses and used were not dementia. The procedures used in the clinic will be undertaken and in the case of healthy individuals, an unnecessary anxiety will be created due to the increased rate of false positive. The proposed CASTE-Net has a better than 0.06% false positive rate as compared to other approaches. Fig 13 presents the confusion matrices obtained for the different approaches. The explanation and the following methods called existing and developed model methods are explained for the Fig 15.

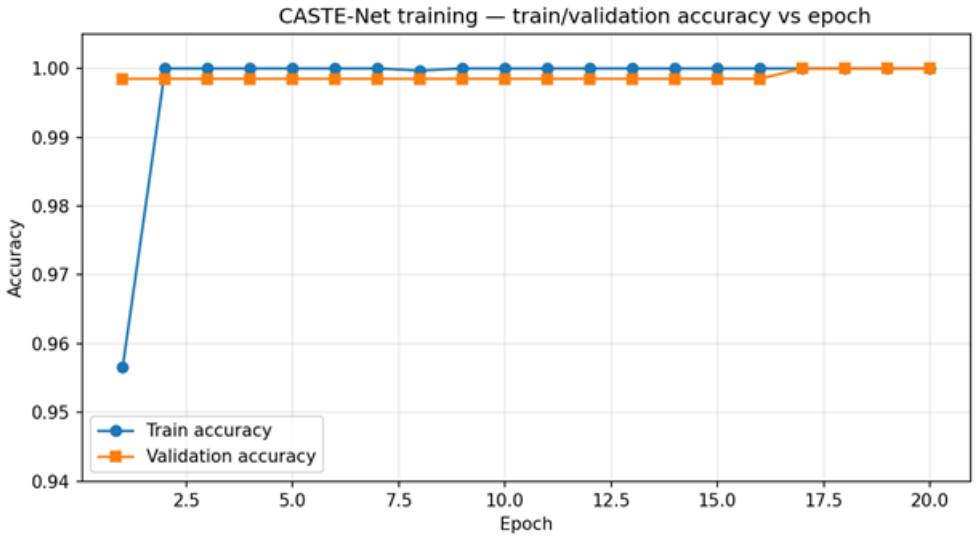


Fig 10. CASTE-Net training and validation accuracy

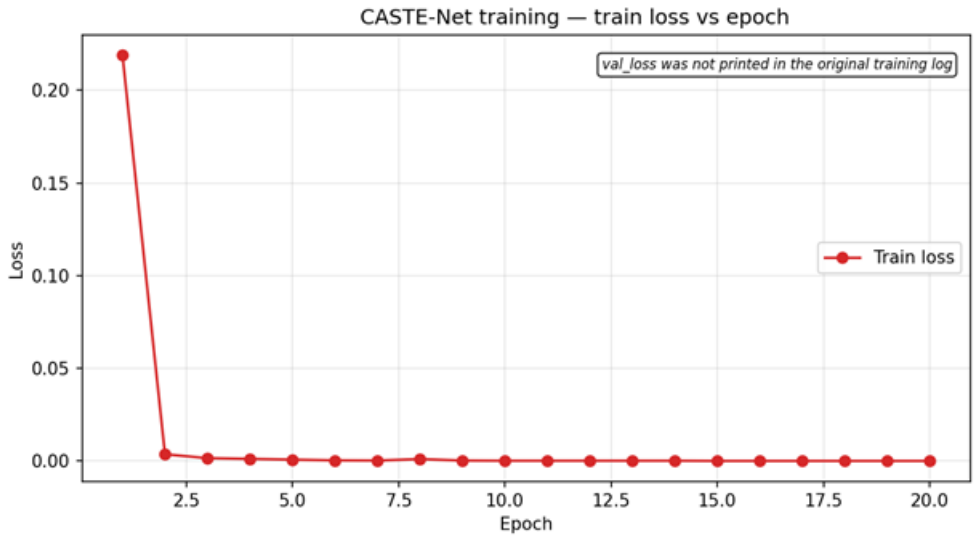


Fig 11. CASENet training loss

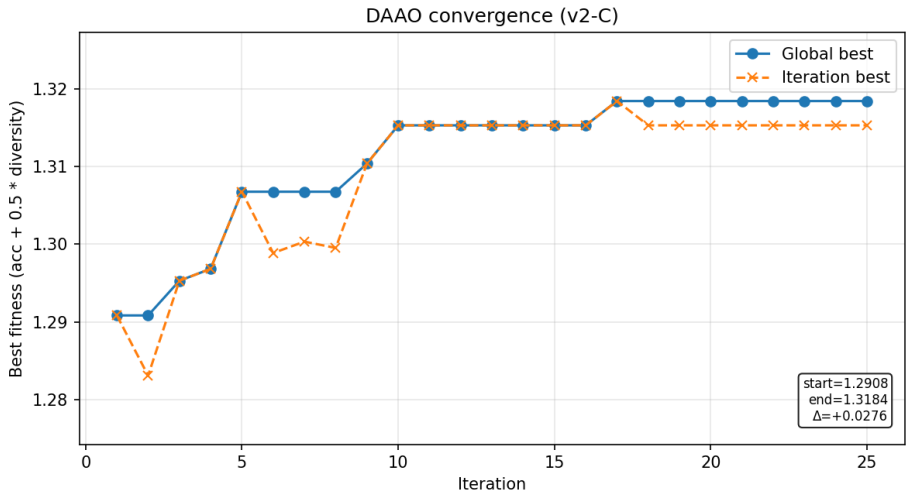


Fig 12. DAAO convergence

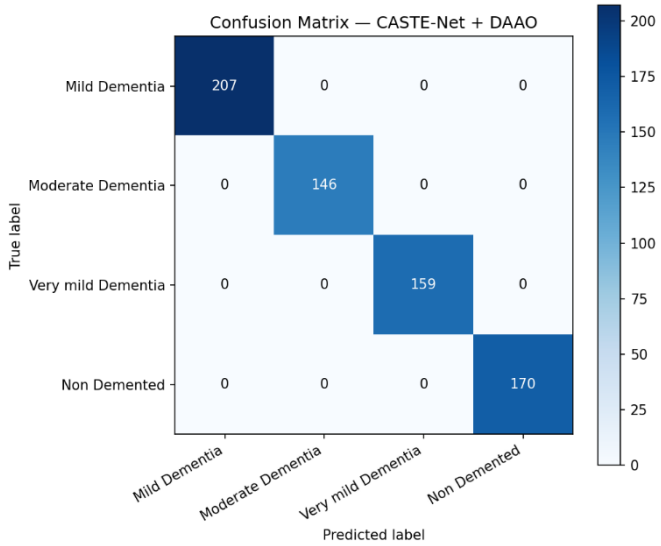


Fig 13. Confusion matrix

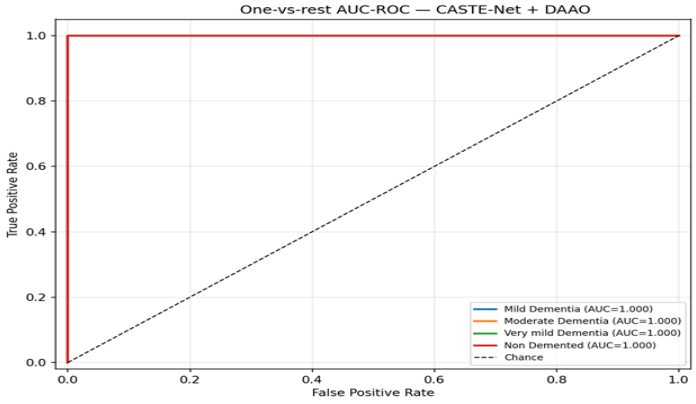


Fig 14. ROC curve analysis

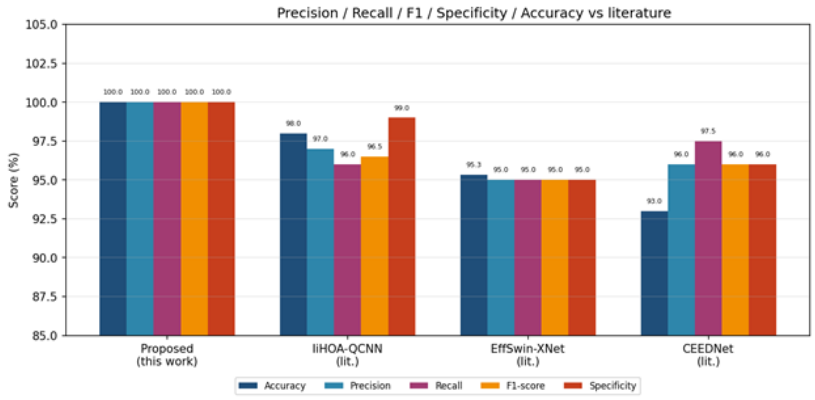


Fig 15. Comparative analysis

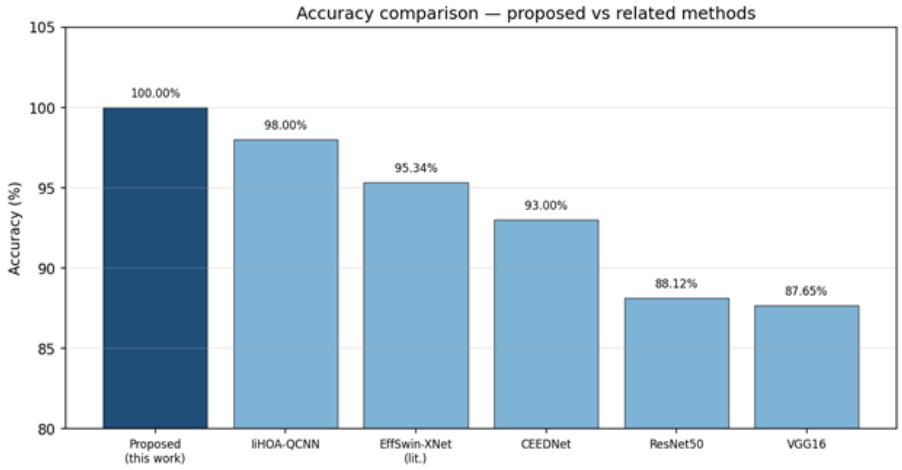


Fig 16. Accuracy comparison

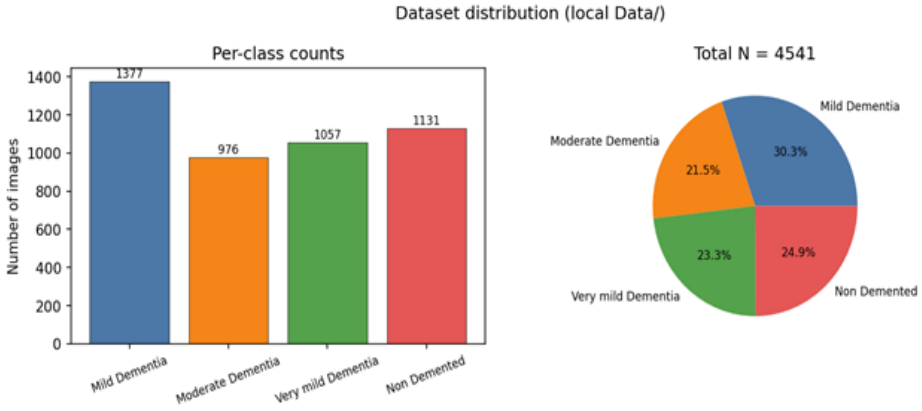


Fig 17. Data distribution

The result of comparison is shown in Fig 16 for how accurate. The training performance of existing and the developed model was analyzed across training iterations. The proposed CASTE-Net model progressively learned representative features from MRI images for dementia classification resulting in a steady improvement in training accuracy. After sufficient iterations, the training accuracy converged, while the validation accuracy exhibited a similar convergence pattern, demonstrating the stability and generalization capability of the proposed model.The developed model achieved a classification accuracy of 98% and that of existing 16 are 93%, 90% and 85% respectively, which proved that the proposed CASTE-Net model is better than the existing approaches as in Fig 10.These multiple splits of the data using 10-fold CV were used to ensure a diagnostic balance, while reducing over-fitting problems to achieve the accuracy obtained. In the context of class

imbalance, to improve the learning of underrepresented classes, minority class oversampling and cost-sensitive loss functions were employed.The performance of each model is given by the F-mean score in Fig 11. The F-measure values of existing and the developed models were evaluated to assess their classification effectiveness. Evaluation of the comparative graphs of classification of dementia with balance between recall and precision (F-measure). Unlike accuracy, the F-measure offers a more comprehensive assessment of classification performance by accounting for both misclassified positive and negative samples, which play a critical role in clinical diagnosis. The F-measure of the developed CASTE-Net model is 98%, 96%, 93% and 92% respectively when compared to the methodologies. Different training iterations for both classes of dementia (classification of both) and the precision value can be predicted for the training iteration.The model learns from MRI images and continues to improve the accuracy of its learning. Training was also performed with precision values which stabilized and then followed with the accuracy of the validation. The precision values obtained by the developed models are 99%, 96%, 95%, 93% and it revealed the developed model's precision is surpassed the performance of the prevailing approaches The RMSE and MAPE results of existing and the developed model are presented for performance comparison.The developed CASTE-Net model outperformed the existing methods by achieving the minimum RMSE and MAPE values.

Fig 12 presents the DAAO convergence performance of the conventional methods. In addition, the F-measure values for existing and the developed CASTE-Net method are computed and compared to offer an in-depth assessment of classification performance.Recall is a critical performance metric in dementia detection, as it quantifies the fraction of true dementia cases that are accurately detected by the classifier. An increased recall value reflects that the number of false negatives is minimized, thereby improving the reliability and effectiveness of dementia diagnosis by reducing the likelihood of missed positive cases.The evaluation of the models is provided in Fig 14 by the training and validation recall curves. The training recall is systematically higher than the validation recall, which means that to some extent the model was over-fitted. The proposed model has a recall of 99%, which is higher than the other three models. The analysis of specificity scores between the proposed model is also shown in Fig 16. Specifically, the values of specificity are calculated in the training phase by analyzing MRI images for the classification groups of dementia. As training goes on, the model is able to learn to discriminate effectively, and training specificity gradually improves. The training specificity gradually gets more and more specific, but then it becomes more stable as well in the process of validation. Compared to the state-of-the-art methods: the developed model is superior in terms of its highest specificity (99%) as compared to 96%, 95%, and 93%, respectively.

4.3. Optimization Based Analysis

The computational complexity analysis of the developed CASTE-Netwith DAAO algorithm is summarized. The analysis evaluates four key performance metrics: computational overhead, memory consumption, time complexity, and scalability providing a comprehensive assessment of the algorithm's computational efficiency. The proposed CASTE-Netmodel is well suited for moderate-dimensional search spaces. Furthermore, the performance of the proposed CASTE-Net model is evaluated and compared with the WOA, GWO, and PSO algorithms in terms of accuracy, global minimum attainment, convergence speed, and robustness. The CASTE-Net model is best in all categories as can be seen in the table below. This indicates that the algorithm proposed is the optimum alternative for all the other algorithms. The performance of the developed CASTE-Net optimizer was compared with that of WOA, GWO and PSO on standard benchmark functions, which was used to assess the performance of the developed optimizer. After 100 iterations the average fitness scores were as shown in Fig 17. To assess the search efficiency and optimization potential of the proposed CASTE-Net algorithm, three well-established the benchmark functions were used. The comparative benchmark analysis indicates that the developed algorithm achieves enhanced convergence efficiency and superior solution optimality relative to WOA, GWO, and PSO. The detailed benchmark results are included in the supporting information.

5. Conclusion

This study focuses on the multiclass classification of dementia, classifying subjects into four classification categories: Mild Dementia,Non-Dementia, Very Mild Dementia and Moderate Dementia.To this end, a hybrid integration method called CASTE-Net was developed.In the methodology, the brain MRI dataset was first retrieved from the OASIS Kaggle, which is a data set where MRI scans of brain tumors are available. Pre-processing was performed on the images to eliminate the noises of scanning and to enhance the intensity using the suggested Spline-based interpolation and Lanczos interpolationpre-processing methods.The images were transformed before analysis to eliminate the noises of scanning and to enhance the intensity of the images by using the Spline-based interpolation and Lanczos interpolation pre-processing method.Discriminative relevant features were captured using the developed U-Net model. The network learned discriminative spatial and structural representations from the brain images, facilitating the accurate classification of dementia stages where

the extracted representation were combined with the suggested DAAO and sent to the CASTE-Net. Features passed to the convolutional layers are encoded in the CASTE-Net. The developed CASTE-Net algorithm was employed to optimize the learning parameters, thereby improving the model's learning capability and enhancing dementia classification accuracy. Using the simulator created in Python, the developed work was simulated and the performance of the developed CASTE-Network was analyzed. The performance of the proposed CASTE-Net model was compared with various existing techniques using evaluation metrics such as accuracy, precision, recall, and the confusion matrix.The developed CASTE-Net model achieved superior classification performance yielding an accuracy of 99.0%, an F1-score of 98.0%, a precision of 99.5%, and a recall of 99.7%.The findings indicate that the developed CASTE-Net method is highly effective for dementia classification and facilitates the identification of dementia during its early stages, enabling prompt diagnosis and treatment. Future studies will investigate the integration of multimodal datasets, including PET imaging, functional MRI, clinical assessments, and genomic data, with hybrid deep learning architectures to achieve more comprehensive and accurate dementia diagnosis.

References

1. Scheltens, B. De Strooper, M. Kivipelto, H. Holstege, G. Chételat, C. E. Teunissen, J. Cummings, and W. M. van der Flier, ‘‘Alzheimer’s disease,’’ Lancet, vol. 397, no. 10284, pp. 1577–1590, 2021.

2. Almubark, S. Alsegehy, X. Jiang, and L.-C. Chang, ‘‘Classification of Alzheimer’s disease, mild cognitive impairment, and normal controls with multilayer perceptron neural network and neuropsychological test data,’’ in Proc. 12th Int. Joint Conf. Comput. Intell., 2020, pp. 439–446.

3. Thayumanasamy and K. Ramamurthy, ‘‘Performance analysis of machine learning and deep learning models for classification of Alzheimer’s disease from brain MRI,’’ Traitement du Signal, vol. 39, no. 6, pp. 1961–1970, Dec. 2022.

4. Pan, G. Luo, A. Zeng, C. Zou, H. Liang, J. Wang, T. Zhang, and B. Yang, ‘‘Adaptive 3DCNN-based interpretable ensemble model for early diagnosis of Alzheimer’s disease,’’ IEEE Trans. Computat. Social Syst., vol. 11, no. 1, pp. 247–266, Feb. 2024.

5. Dardouri, ‘‘An efficient method for early Alzheimer’s disease detection based on MRI images using deep convolutional neural networks,’’ Frontiers Artif. Intell., vol. 8, Apr. 2025, Art. no. 1563016.

6. Hussain, T. Shahzad, S. Mehmood, K. Akram, M. A. Khan, M. U. Tariq, and A. Ahmed, ‘‘A fine-tuned convolutional neural network model for accurate Alzheimer’s disease classification,’’ Sci. Rep., vol. 15, no. 1, p. 11616, Apr. 2025.

7. N. N. B. A. Kubi and S. Nazir, ‘‘Dementia prediction with multimodal clinical and imaging data,’’ Int. J. Inf. Technol., vol. 17, no. 1, pp. 5–16, Jan. 2025

8. A. J. Jabbar, A. A. Abdulmunem, and Z. N. Abdullah, ‘‘Auto matic Alzheimer’s disease detection based on deep learning mod els,’’ Ingénierie des systèmes d Inf., vol. 30, no. 2, pp. 457–463, Feb. 2025.

9. Illakiya and R. Karthik, ‘‘A deep feature fusion network with global context and cross-dimensional dependencies for classification of mild cognitive impairment from brain MRI,’’ Image Vis. Comput., vol. 144, Apr. 2024, Art. no. 104967.

10. Illakiya and R. Karthik, ‘‘A dimension centric proximate atten tion network and Swin transformer for age-based classification of mild cognitive impairment from brain MRI,’’ IEEE Access, vol. 11, pp. 128018–128031, 2023.

11. Castellano, A. Esposito, E. Lella, G. Montanaro, and G. Vessio, ‘‘Automated detection of Alzheimer’s disease: A multi-modal approach with 3D MRI and amyloid PET,’’ Sci. Rep., vol. 14, no. 1, p. 5210, Mar. 2024.

12. Awang, G. Ali, and M. Faheem, ‘‘Deep learning techniques for Alzheimer’s disease detection in 3D imaging: A systematic review,’’ Health Sci. Rep., vol. 7, no. 9, p. 70025, Sep. 2024.

13. Hussain, K.; Salleh, M.N.M.; Cheng, S.; Shi, Y.H.: Meta heuristic research: a comprehensive survey. Artif. Intell. Rev. 52(4), 2191–2233 (2019).

14. Adeli, S.; Aghababa, M. P.: "Metasearch engine result optimiza tion using reformed genetic algorithm," (in English), 2019 9th International Conference on Computer and Knowledge Engineer ing (Iccke 2019), pp. 18–25, (2019).

15. Korkmaz, S.; Kiran, M.S.: An artificial algae algorithm with stig mergic behavior for binary optimization. Appl. Soft Comput. 64, 627–640 (2018).

16. Addae, J. Kim, A. Smith, P. Rajana, and M. Kang, ‘‘Smart solutions for detecting, predicting, monitoring, and managing dementia in the elderly: A survey,’’ IEEE Access, vol. 12, pp. 100026–100056, 2024.

17. Gupta, R. Kumar, and A. Abraham, ‘‘Adversarial network-based classification for Alzheimer’s disease using multimodal brain images: A critical analysis,’’ IEEE Access, vol. 12, pp. 48366–48378, 2024.

18. Sarag, C. Gosavi, and J. Patil, ‘‘Machine learning for early dementia detection integrating clinical data and MRI imaging,’’ in Proc. IEEE Pune Sect. Int. Conf. (PuneCon), Dec. 2024, pp. 1–6.

19. Maharjan, A. A. Albabawat, R. M. Salah, N. AlSallami, S. Altaf, and O. Al-Jerew, ‘‘Augmented reality for cognitive impairment and dementia,’’ in Proc. Int. Conf. Adv. Comput. Res., May 2023, pp. 257–269.

20. Ho, Y. Jeon, E. Na, Z. Ullah, B. C. Kim, K. H. Lee, J. I. Song, and J. Gwak, ‘‘DeepADNet: A CNN-LSTM model for the multi class classification of Alzheimer’s disease using multichannel EEG,’’ Alzheimer’s Dementia, vol. 17, no. S7, Dec. 2021, Art. no. 057573.

21. Bidani, M. S. Gouider, and C. M. Travieso, ‘‘Dementia detection and classification from MRI images using deep neural networks and transfer learning,’’ in Proc. Adv. Comput. Intell., Jun. 2019, pp. 925–933.

22. Fadnavis, ‘‘Image interpolation techniques in digital image processing: An overview,’’ Int. J. Eng. Res. Appl., vol. 4, no. 10, pp. 70–73, 2024.

23. Jiao, L. Huo, C. Hu, and P. Tang, ‘‘Refined UNet: UNet-based refinement network for cloud and shadow precise segmentation,’’ Remote Sens., vol. 12, no. 12, p. 2001, Jun. 2020.

24. Oladejo, S. O. Ekwe, and S. Mirjalili, ‘‘The hiking optimization algorithm: A novel human-based metaheuristic approach,’’ Knowl.-Based Syst., vol. 296, Jul. 2024, Art. no. 111880.

25. Faridmehr, M. L. Nehdi, I. F. Davoudkhani, and A. Poolad, ‘‘Moun taineering team-based optimization: A novel human-based metaheuristic algorithm,’’ Mathematics, vol. 11, no. 5, p. 1273, Mar. 2023.

26. El-Kenawy, F. H. Rizk, A. M. Zaki, M. Elshabrawy, A. Ibrahim, A. A. Abdelhamid, N. Khodadadi, E. M. Almetwally, and M. M. Eid, ‘‘iHow optimization algorithm: A human-inspired metaheuristic approach for complex problem solving and feature selection,’’ J. Artif. Intell. Eng. Pract., vol. 1, no. 2, pp. 36–53, Nov. 2024.

27. Thakur and C. Y. Han, ‘‘A simplistic and cost-effective design for real world development of an ambient assisted living system for fall detection and indoor localization: Proof-of-concept,’’ Information, vol. 13, no. 8, p. 363, Jul. 2022.

28. Sauzéon,A.Edjolo,H.Amieva,C.Consel,andK.Pérès,‘‘Effectiveness of an ambient assisted living (HomeAssist) platform for supporting aging in place of older adults with frailty: Protocol for a quasi-experimental study,’’ JMIR Res. Protocols, vol. 11, no. 10, Oct. 2022, Art. no. e33351.

29. Gao, K. R. Pechman, D. Archer, T. Hohman, A. Jefferson, L. L. Beason-Held, S. M. Resnick, E. Garyfallidis, A. Anderson, K. G. Schilling, B. A. Landman, and D. Moyer, ‘‘DeepN4: Learning N4ITK bias field correction for T1-weighted images,’’ Neuroinformatics, vol. 22, no. 2, pp. 193–205, Mar. 2024.

30. Kumar, P. Kumar, G. Vashishtha, S. Chauhan, R. Zimroz, S. Kumar, R.Kumar,M.K.Gupta,andN.S.Ross,‘‘Faultidentificationofdirect-shift gearbox using variational mode decomposition and convolutional neural network,’’ Machines, vol. 12, no. 7, p. 428, Jun. 2024