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

Alzheimer's disease Classification Using Histogram of Oriented Gradient, Transfer Learning, and Capsules Network

Iliass Zine-dine*[,1,a], Jamal Riffi[1,b], Khalid El Fazazy[1,c], Mohamed Adnane Mahraz[1,d], Hamid Tairi[1,e]

Submitted:10/03/2024 **Revised**: 25/04/2024 **Accepted**: 02/05/2024

Abstract: Alzheimer's disease (AD) is a widespread neurodegenerative condition that affects the brain and causes cognitive impairment. This disease falls within the fields of medicine and health care; recently, it has been considered one of the most studied diseases of the nervous system. AD has no cure or any way to slow or stop its progression. Machine learning (ML) techniques, specifically pattern recognition in biomedical sciences from disease diagnosis to treatment, have become one of the important ways that researchers better understand the whole situation and solve complex medical problems. Deep learning (DL) is a powerful machine-learning model for classification while extracting high-level features. The importance of classifying this type of medical data is to develop a predictive classifier system to know the type of disease from the characteristics extracted from the images or to estimate the stage of the disease. The proper choice and the best application of appropriate computer vision methods and techniques to extract important features from the medical image lead to efficient classification. This paper presents a methodology comprising three main phases. Firstly, we preprocess the data using standard computer vision techniques and graphic functions, along with the Histogram of Oriented Gradient (HOG) descriptor, to extract relevant features, accelerate model training, and prevent overfitting. Secondly, we feed these extracted features into two convolutional neural network models, VGG-16 and ResNet50, to extract deep features. We then concatenate these deep features into a single vector. Subsequently, these concatenated features are evaluated using a RandomForestRegressor to select the most relevant ones, reduce the dimensionality of our dataset, and enhance the interpretability and reliability of our model's decisions. Finally, the selected features are utilized as inputs for a CapsNet model to extract spatial features and perform classification. This comprehensive approach leverages the strengths of each technique to achieve robust and accurate classification outcomes. The proposed method provided effective results (94.27%), surpassing several recent experiments on the topic of AD in terms of accuracy, and demonstrated significant value in the early prediction of AD through the evolution of computer vision method and machine learning and its applications in the medical domain.

Keywords: Alzheimer's disease, Histogram of Oriented Gradient, VGG-16, ResNet50, Capsule Network.

1. Introduction

Alzheimer's disease is a neurological, irreversible, and progressive brain condition that affects several brain functions. It kills brain cells, impairing memory, cognitive abilities, and eventually the capacity to do the most basic tasks. The cognitive decline caused by this disorder ultimately leads to dementia [1]. A thorough medical examination, including a patient history, a mental state examination, a physical exam, and a neurobiological exam, are all necessary for the diagnosis of AD. In addition to those evaluations, resting-state functional magnetic resonance imaging offers a non-invasive way to assess functional brain activity and changes in the brain. For the aforementioned reasons, AD is a key public health concern and it considered one of the most widespread neurodegenerative disorders throughout the world [2].

According to recent statistics, the incidence of AD is increasing in later life, with people over 65 having the highest risk [3]. Based on the World Health Organization (WHO) statistics, it is expected that AD will affect one out of every 85 people in the world by the year 2050 [4]. During the stages of AD, degeneration of the brain progresses with time. Therefore, it is crucial to divide AD patients into various categories depending on the phase of the disease [5]. This division is critical because the patients at different stages of the disease are required to be treated differently, and the same medication cannot be used for all of them. For that purpose, the classification of different stages of the disease can be very helpful in the treatment of symptoms of the disease to improve the patient's quality of life.

Early diagnosis and treatment of AD are effective treatments. Especially at an early stage, this technique is a challenging task. Usually, a neuropsychological examination is used for the early diagnosis of AD [6]. The accuracy of psychological and cognitive tests is very dependent on the ability and experience of the clinician. Therefore, medical experts are responsible for analyzing the interpretation of medical data; this is quite difficult and

[1] Laboratory of Informatics, Signals, Automatics and Cognitivism (LISAC), Faculty of Sciences Dhar El Mehraz (FSDM), Sidi Mohamed Ben Abdellah University (USMBA), Fès, Morocco
[a] ORCID ID : 0000-0001-7134-4888
[b] ORCID ID : 0000-0003-0818-7706
[c] ORCID ID : 0000-0002-6333-2295
[d] ORCID ID : 0000-0002-0966-9654
[e] ORCID ID : 0000-0002-5445-0037
** Corresponding Author Email: zinedine.iliass@gmail.com*

limited for a medical expert to do because of the subjectivity and high complexity of the images [7]. Thus, it is important to develop automatic detection and classification techniques.

It could be as simple as a computer vision processing problem, but it could also be a non-trivial problem if the compositions, dimensions, and pixels of the images are taken into account. At present, international researchers concentrate on the early diagnosis and prediction of AD from its earliest stages. Intervention of artificial intelligence and computer vision is the best approach to accomplishing this task [8].

Statistical machine learning techniques have demonstrated early success in the automated detection of AD. Deep learning techniques such as convolutional neural networks (CNN) and sparse auto encoders have surpassed statistical methods. However, deep learning techniques recently taught profound structures from scratch, which has two types of drawbacks: First, the correct training of a deep learning network necessitates an enormous amount of annotated training data, which can be problematic, particularly in the medical imaging domain. Second, deep network training necessitates the meticulous tuning of numerous parameters, and suboptimal tuning can lead to overfitting or underfitting, whichever can negatively impact performance [9].

Transfer learning is one of the techniques for training a deep network (particularly CNN) from scratch. This method has proven to be robust for several applications in various fields, such as networks trained on natural images used with medical images. The trained CNNs are carefully constructed using large-scale datasets like ImageNet in the various computer vision domains, such as object recognition, medical image detection, document classification, etc [10].

The classification of AD is a difficult job because different criteria are dependent on the construction of the nervous system's tissue and cells. Nowadays, the evolution of technology and the digitalization of medical instruments aid doctors in correctly detecting and classifying AD in the early phase [11]. The field of ML is concerned with this task, and it is considered a productive space for innovation for many researchers. DL is the latest machine learning approach that has shown superior performance over traditional ML in identifying complex structures in complex, high-dimensional data, particularly in the field of computer vision [12]. Recently, there has been a lot of interest in the intervention of DL for automated categorization and the early diagnosis of AD. However, DL has several restrictions. Deep learning has several restrictions. The extensive use of neural network techniques in task classification, the absence of a dataset with vowel labels, and rigorous computational efficiency prevent researchers from discovering more accurate performance measurements [13]. For these reasons, we considered a technique that would provide us with an elevated level of accuracy in identifying objects and give us improved prediction scores in the experimental results.

Lately, new research has been based on the extraction of features from the image, which will be fed to the classifiers. Iago R. et al. [14] suggested an approach using CNN layers to extract more features that will feed ML classifiers in order to classify AD. In this regard, Ruhul et al. [15] offered in a survey different feature extraction (FE) techniques like voxel morphometry, texture, wavelet transform, graph/network, eigenvector, harmonic function, scale-invariant, and Artificial Neural Network (ANN).

Recently, it has become evident that features extraction methods are frequently used in medical image classification. We also cannot ignore the effective role of convolutional neural (CNN) architectures in solving this task. In this study, we propose an approach aimed at achieving higher prediction scores through the utilization of a capsule network (CapsNet) architecture, comprising three primary phases: data preprocessing, feature extraction, and classification. Initially, standard computer vision and infographic functions are employed in the preprocessing phase, alongside the Histogram of Oriented Gradient (HOG) as an image descriptor. Subsequently, the features extracted by HOG are fed into VGG-16 and ResNet50 models during the feature extraction phase to extract deep features, thereby accelerating model training, preventing overfitting, and enhancing the overall effectiveness and robustness of the classification framework. Finally, the features selected by the CNNs models are used as inputs for a Capsule Network model to extract spatial features and perform classification. The proposed system recorded very good classification performance (94.27%); based on these methods above; it appears that the functionalities extracted in this classification task have an important and effective role. Therefore, our approach demonstrate significant value in the early prediction of AD through the advancement of computer vision methods and machine learning, applied within the medical domain.

The rest of the article is structured as follows: Section 2 offers related works. In section 3, the proposed approach are illustrated. In section 4, the experiments, results, and discussion are demonstrated. Finally, conclusion and plan for future studies are presented in Section 5.

2. RELATED WORK

The main objective of this section is to review current research that has used feature extraction, machine learning, and deep learning methods to detect and classify AD. Lately, traditional ML techniques have been used to accurately detect AD. Mohammad et al. [16] used the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset to classify the five different stages of the disease using six different ML and data mining algorithms, including k-nearest neighbors (k-NN), decision trees (DT), rule induction, Naive Bayes, generalized linear models (GLM), and deep learning algorithms, but their results fell short of expectations and are still not applicable. In 111 papers from 2006 to 2016, Pellegrini et al. [17] examined the ML methods used to treat dementia and cognitive impairment. It placed emphasis on the creation of original ML models using a multidisciplinary approach. In this regard, Jose et al. [18] gave an overview of brain diseases' neuroimaging methods. According to some research, using ML techniques can help identify the underlying neurological reasons for various types of brain illnesses.

Another kind of method based on deep learning proves to be very effective while managing vast amounts of data. In 2017, Litjens et al. [19] offered a review of DL techniques for medical image analysis. Although DL models are referred to as "black boxes" it is mentioned that some statistical methods can be utilized to evaluate network uncertainty. Shen et al. [20] presented a survey on deep learning for Alzheimer’s disease. It also supported this point of uncertainty in prediction by deep learning models. Many methods are proposed for better feature selection (FS) from neuroimaging data. Ahmadlou et al. [21] noted that non-linear features may show significant differences in certain electroencephalogram (EEG) sub-bands but may not help in separating the two groups (CN and AD) in band-limited EEG. To manage this, visibility graphs (VG) were used for extracting features from EEG signals. Rodrigues et al. [22] employed short-time Fourier transform (STFT) and WT features for classification utilizing ANN.

A different technique for extracting characteristics from MR images is to extract cortical characteristics employing surface-based morphometry. Cho et al. [23] suggested a powerful incremental classification approach employing cortical thickness data, which was translated into the spatial frequency domain via a manifold harmonic transform. In addition to MRI scans, a variety of other forms of data can be analyzed to learn more about a subject's health, including cognitive and physiological testing, genetic makeup, and demographics.

Data from multiple modalities provides complementary information that can improve the efficiency of machine learning algorithms. To achieve this purpose, Ye et al. [24] suggested a multiple kernel learning (MKL) architecture that combined ROI and tensor features into a single feature space. Zhang et al. [25] presented a novel multimodal Laplacian regularized least squares (mLapRLS) approach to leverage unlabeled samples in aiding the classification. Liu et al. [26] discussed that selecting features separately from each modality neglects the powerful inter-modal correlation within each subject, which results in suboptimal efficiency. As a solution to this issue, a novel multi-task learning-based feature selection method was offered that jointly chose sparse features from all the modalities. The MKL techniques previously addressed integrate numerous kernels from various modalities linearly, and they are very dependent on the weights given to each modality. To solve these problems, Tong et al. [27] suggested non-linearly combining data from multiple modalities by employing the non-linear graph fusion (NGF) method. Kim et al. [28] suggested a multimodal sparse hierarchical extreme learning machine (MSH-ELM) to gather high-level characteristics from many modalities and combine them for classification.

Deep learning is a multi-layered neural network adept at learning intricate data structures to produce a high capacity for abstraction. Suk et al. [29] used SAE and demonstrated that non-linear correlations within the features can improve the diagnosis accuracy of AD, MCI (Mild Cognitive Impairment), and MCIc. Payan et al. [30] combined SAE and CNN and discovered that 3D-CNN performs better in classifying AD and MCI from CN (Control Normal). The best accuracy was obtained by multi-kernel SVM (MK-SVM) with low-level features (LLF) and SAE-learned features (SAEF). Liu et al. [31] employed SAE-based DL architecture for the diagnosis of AD in four stages by acquiring high-level features and a soft-max logistic regression (LR). Hosseini et al. [32] suggested a deeply supervised adaptive 3D-CNN (DSA-3DCNN) to classify AD and MCI from CN, which was pre-trained by a 3D convolutional autoencoder (3D-CAE) utilizing features from MRI. In this regard, the combination of models In another paper, Suk et al. [33] combined sparse regression models with a deep CNN known as the deep ensemble sparse regression network (DeepESRNet) and were able to clinically diagnose AD. Suk et al. [34] combined DL and state-space modeling and functional dynamics with the time of rs-fMRI to form a biomarker for the diagnosis of MCI and the classification of AD, another technique based on voxels. Davatzikos et al. [35] used voxel-based analysis for the classification of AD and frontotemporal dementia (FTD) from CN using the high-dimensional PC method.

High-dimensional multi-variate discriminant analysis provides more diagnostic accuracy than traditional measuring. Lopez et al. [36] employed principal component analysis (PCA) to minimize the dimensions and used Bayesian classifiers to compare CN with AD. To develop a more reliable classifier, Termenon et al. [37] employed a two-stage sequential ensemble of classifiers using a relevance vector machine (RVM) to create the discriminant function. Chincarini et al. [38] chose informative features from intensity- and textural-based

MRI features using the Random Forest classifier (RF). To take advantage of the complementary knowledge that many modalities supply, Hor et al. [39] introduced MRI and PET scans with tree-based feature transformations and used RF for classification.

The majority of existing medical MR imaging research focuses on the automatic classification of AD [40]. Recently, several researchers have presented various strategies for detecting AD in MR images; the following table represents previous studies carried out on different datasets.

Table 1. Related work approaches to classification methods, feature extraction, and accuracy of Alzheimer's disease classification.

Literature	***Classification Methods***	***Modality***	***Accuracy***
Liu et al. [41]	Stacked auto-encoder +MKSVM	MRI,PET	91%
Li et al. [42]	PCA+ Restricted Boltzmann Machine	MRI,PET, CSF	91.4%
T. Brosch and R. C. Tam [43]	SAE+CNN	MRI	93.8%
Srivastava et al. [44]	SAE+3DCNN	MRI	95.39%
Aderghal et al. [45]	2DCNN	MRI	91%
Suk et al. [46]	Deep Boltzmann machine	MRI,PET	95%
Shi et al. [47]	Deep polynomial network	MRI	95%
Lian et al. [48]	Hierarchical-CNN	MRI	90%
Xin et al. [49]	3DCNN	MRI	91%

Now, the current studies have shown that deep learning methods are widely employed in expert and intelligent techniques as well as in medical image analysis. The limitations of the earlier-outlined approaches should be considered when using classification for AD. The primary flaw in the previously suggested systems is that they only consider the different categorizations of the MRI image dataset, ignoring the features to be extracted. In addition, these models are data-hungry and have complex computing efficiency; thus, they have dissatisfactory performance. For these reasons, we thought about a method that gives us better prediction scores.

Real-time performance is a critical factor in medical diagnosis, particularly in emergency situations where timely and accurate decisions are essential for patient treatment and prognosis. Evaluating a model's inference time and computational resource requirements ensures its suitability for real-world applications. Models designed for such scenarios must balance speed and accuracy to provide reliable diagnostics without compromising computational efficiency [50].

The primary contribution of this study can be summarized as follows: during the preprocessing stage, we employed common computer vision and infographic techniques to facilitate subsequent tasks, alongside HOG as an image descriptor, to extract relevant and significant features. These features extracted by HOG accelerate model training, prevent overfitting, and thereby enhance the overall effectiveness and robustness of the classification framework, were then fed into VGG-16 and ResNet50 during the subsequent "feature extraction" phase to derive deep features. Additionally, we utilized a capsule network model to extract spatial features and ultimately, for classification. To validate the efficacy of our approach, experiments were conducted on a famous AD dataset, and the results were compared with existing methodologies. Our approach has shown considerable value in the early prediction of AD through the advancement of computer vision method and machine learning and its applications in the medical field.

3. Material and methods

In this section, we outline the overarching framework of our proposed approach. Figure 1 provides a visual representation detailing the essential components of our method designed for Alzheimer's disease classification. This approach comprises three primary modules: data preprocessing, feature extraction, and classification. We describe the initial preprocessing steps applied to the input MRI images, including normalization, resizing, cropping, and augmentation, to prepare them for classification (Section 3.2). To accelerate model training, prevent overfitting, and enhance the overall effectiveness and robustness of the classification framework, we incorporate Histogram of Oriented Gradients as a feature extractor. The features extracted by this descriptor are then fed into both VGG-16 and ResNet50 CNN models to extract deep features, which well concatenated into a single vector. Subsequently, these concatenated features are evaluated using a RandomForestRegressor to select the most relevant ones, reduce the dimensionality of our dataset, dimensionality, and bolster the interpretability and dependability of our model's decisions (Section 3.3). Finally, the selected features are utilized as inputs for a CapsNet model to extract spatial features and perform classification (Section 3.4). The following subsections discuss the three important elements in more detail.

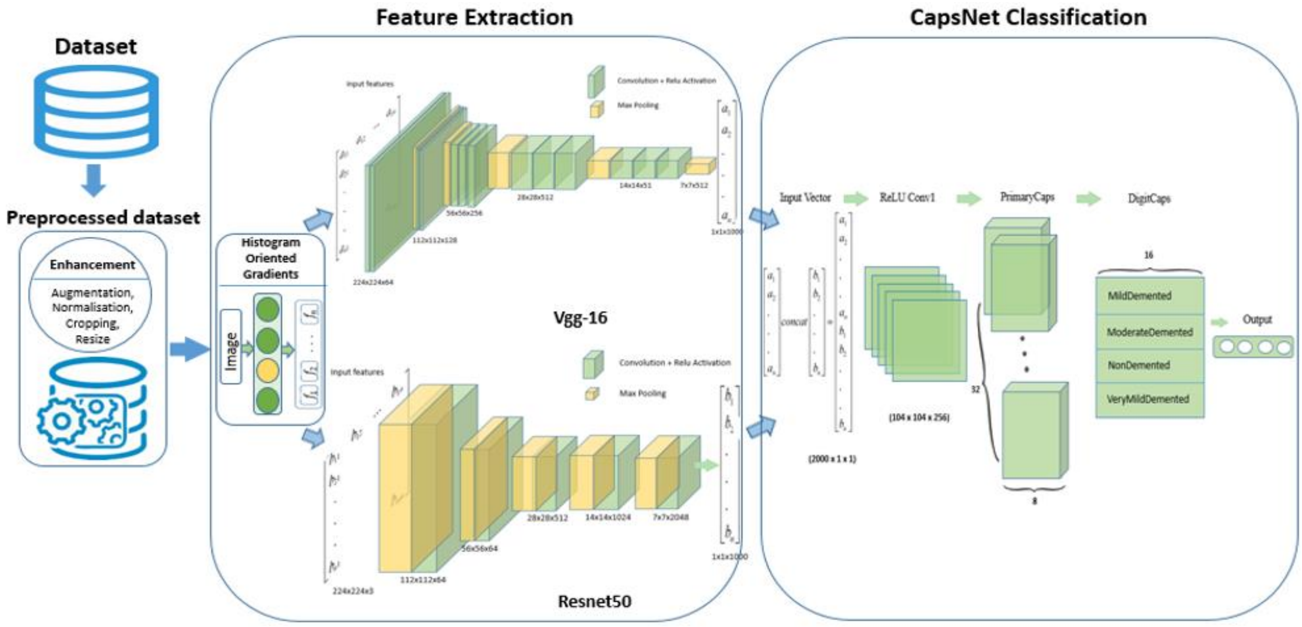

Fig. 1. Proposed architecture of oriented gradient histogram, Vgg-16, ResNet50 and CapsNet for AD classification.

3.1. Dataset

The Alzheimer's disease utilized in our research is crucial for classification, offering real-world data reflecting clinical complexities. It enables algorithm development and evaluation, facilitating supervised learning and serving as a benchmark for advancing medical image analysis.

3.1.1. Data description

In this study, we used the free Kaggle Alzheimer's dataset, which consists of 6400 2D slices, these MR images are divided as follows: 896 mildly mentally impaired, 64 moderately mentally impaired, 3200 non-demented, and 2240 very mildly mentally impaired subjects. We started to preprocess the dataset by applying famous imaging methods like normalization, resizing, cropping, and augmentation to facilitate the use of next functions, these techniques applied in the data processing phase allowed us to have (33984 images). To allow us to have a model that predicts with a higher score during training, we convert these images into tensors to simplify the tasks of the feature extraction functions that will be applied later, such as VGG-16, and ResNet50. Finally, we created our CapsNet model after gathering the embeddings derived from these two neural network models to identify AD.

3.1.2. Data preprocessing

The application of techniques and methods of preprocessing the dataset in order to have more relevant information that will be exploited in the input of a classifier to predict results has become a trend in the fields of computer graphics and computer vision. In this paper, we will use well-known image-processing methods; the key strategies employed in this section of the treatment are covered below.

Data crop: Nearly all of the images in our brain MRI datasets have undesirable spaces, which results in subpar classification performance. Therefore, it is vital to crop the photographs in order to eliminate unnecessary portions and use only the pertinent information [50]. In this study, we employ the cropping approach, which computes extreme points and returns a geographic subset of an object as specified by an extent object. The application of this method consists of five steps: First, we load the original MR pictures. Secondly, we apply thresholding in order to create binary images; thirdly, we undertake dilation and erosion processes to reduce image noise; and fourthly, we use the threshold images' largest contour to determine the images' four extreme points (extreme top, extreme bottom, extreme right, and extreme left). Finally, we crop the image based on the contour and extreme point data. Bicubic interpolation is used to enlarge the cropped images.

Data augmentation: Data augmentation is a technique that involves transforming the original dataset to produce a synthetic dataset. It is a procedure that generates additional training data by applying transformations to existing data to obtain new data [51]. This method involves creating numerous duplicates of the original image with various scales, orientations, locations, brightness, and other characteristics. It has been shown that augmenting existing data can increase the classification accuracy of the model, rather than collecting new data. As our MRI dataset is not particularly large in this research, we performed image augmentation to increase the size of the dataset.

3.2. Features extraction

The technique of turning raw data into numerical features that can be handled while keeping the information in the original dataset is known as feature extraction [52]. Compared to using ML on raw data directly, it produces better outcomes. In this study, we applied standard computer vision and infographic techniques to preprocess our dataset. Additionally, we extract pertinent features utilizing the HOG descriptor to speed up the models training; avoid overfitting, and thereby augmenting the overall effectiveness and robustness of the classification framework, these characteristics gleaned by HOG then serve as inputs for the CNN models VGG-16 and ResNet50 to extract deep caracteristics. CNN models, particularly pre-trained ones such as VGG-16 or ResNet50, have the capacity to generate numerous features from our image data. However, not all of these features are equally important for effective classification. By employing RandomForestRegressor, we were able to identify the most significant features, thereby reducing the dimensionality of our dataset. This reduction can potentially improve our model's performance by mitigating overfitting and expediting the training process. Additionally, it enhances the interpretability and reliability of our model's decisions, contributing to a more robust classification framework; finally, we concatenated them and fed them into the capsule network model to extract spatial features and ultimately, for classification.

3.2.1. Histogram of Oriented Gradients

Histogram of Oriented Gradients (HOG) is a feature descriptor used in computer vision and image processing for object detection, similar to the Canny Edge Detector and Scale-Invariant Feature Transform (SIFT). This approach counts occurrences of gradient orientation in localized parts of an image. This method is comparable to edge orientation histograms, scale-invariant feature transformation descriptors, and shape contexts, but it is more accurate because it is computed on a dense grid of equally spaced cells and employs overlapping local contrast normalization. Four types of normalization are explored. The unnormalized vector containing all the histograms of a single block is denoted by v, its k-norm by ‖v‖k, and e is a low-value constant. The normalization factor is then defined by:

- L2-norm: $f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}}$
- L2-hys: L2-norm followed by clipping (limiting the maximum values of v to 0.2) and renormalizing
- L1-norm: $f = \frac{v}{(\|v\|_1 + e^2)}$
- L1-sqrt: $f = \sqrt{\frac{v}{(\|v\|_1 + e^2)}}$

The L2-Hys, L2-norm, and L1-sqrt norms achieve similar performance, while L1-norm performs worse, but still significantly outperforms no normalization. In our approach we applied the first L2-norm normalization, the figure below illustrates the application of the hog function on our dataset.

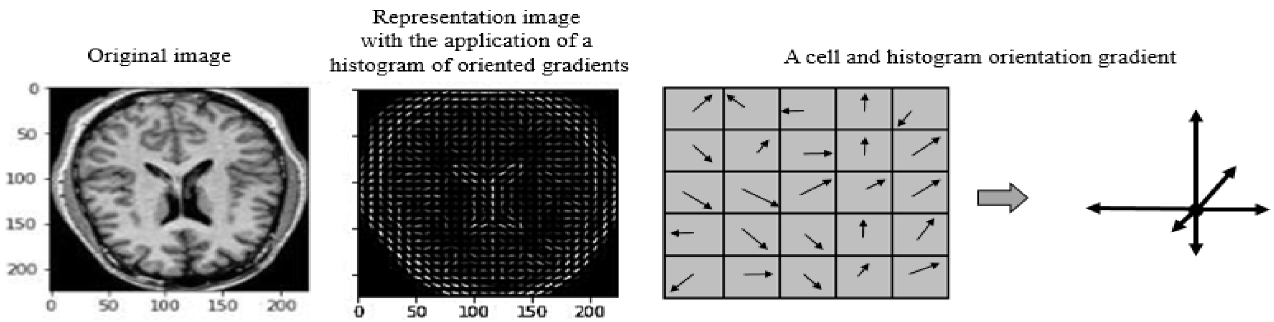

Fig. 2. Representation of features extracted from Alzheimer's disease image by the HOG algorithm.

The integration of the HOG function into our approach holds significant importance for several reasons. Firstly, HOG is widely acknowledged for its ability to capture texture and shape information within an image, making it a powerful tool for feature extraction. By leveraging HOG as an image descriptor in our methodology, we can extract relevant and discriminative features, which are crucial for the task of medical image classification. Furthermore, HOG provides a compact representation of the extracted features, aiding in reducing the dimensionality of the data and enhancing the efficiency of subsequent classification algorithms. By incorporating HOG into our approach, we can enhance the quality of the extracted features and consequently, the accuracy and robustness of our classification model.

3.2.2. Convolutional neural network models

Convolutional Neural Networks (CNNs) are a class of deep learning models specifically designed for tasks involving image processing and pattern recognition. CNNs excel in feature extraction from images due to their unique architecture, which includes convolutional layers, pooling layers, and fully connected layers. These networks are adept at capturing hierarchical patterns and spatial relationships within images, making them ideal for tasks such as object recognition, image classification, and segmentation [53].

One of the key advantages of CNNs is their ability to automatically learn relevant features. This feature extraction capability accelerates model training by reducing the dimensionality of the input space and focusing on the most discriminative features. Moreover, CNNs can avoid overfitting by incorporating techniques such as dropout and regularization, which help generalize learned patterns to unseen data [54].

In this study, to accelerate model training and prevent overfitting, we incorporated the crucial and relevant features extracted by HOG as inputs to both VGG-16 and ResNet50, two well-known methods for extracting deep features. We utilized the convolutional, pooling, and normalization layers, while excluding the fully connected classification layer. Classification phase will be conducted later in CapNet model. In the following two paragraphs, we define and present VGG-16 and ResNet50, two CNN models utilized in our approach.

Vgg-16 is a convolutional neural network model trained on millions of images from different categories. It deals with the classification tasks of image processing. It belongs to the family of models called VGGNet (Visual Geometry Group Network). Deep layers and uniform architecture characterize this model. The model has 16 trainable weight layers, hence the name "Vgg-16". Specifically, the architecture of VGG-16 is composed of 13 convolution layers, followed by fully connected layers and output layers [55]. For our task, We utilized the convolutional, pooling, and normalization layers, while excluding the fully connected classification layer to extract deep features, Convolution layers use small filters (3x3) with a stride of 1 and pooling layers with a stride of 2. In the final layer, we aggregated the features extracted by VGG-16 into a vector of dimension (1.1.1000) to concatenate it with the other vector extracted by ResNet50, which has the same dimension.

Resnet50 is a deep learning model used for computer vision applications where weighting layers learn residual functions with reference to the input layer. ResNet architectures were developed to solve the problem of gradients disappearing or exploding when networks become very deep [56]. The ResNet50 architecture is made up of 50 layers composed of residual blocks; these layers are expressed by stages (Input Layer, Initial Convolutional Block, Four Main Stages, Residual Block, Global Average Pooling (GAP), Fully Connected (FC) Layer). For our task, we employed all these layers except for the fully connected classification layer to extract deep features. Moreover, in the final layer, we aggregated the features extracted by Resnet50 into a vector of dimension (1.1.1000) to concatenate it with the other vector extracted by VGG-16, which has the same dimension.

The vector concatenation technique utilized in this study enhances model performance by incorporating crucial information, facilitating a more comprehensive data representation. This can potentially enhance generalization and foster more dependable classification decisions. Furthermore, by reducing data dimensionality, concatenation aids in mitigating overfitting and enhancing model interpretability.

Hyperparameters: In this study, we utilized the PyTorch library for our experiments. Notably, the VGG-16 and ResNet50 architectures come with predefined kernel size, padding, and stride parameters, which are not explicitly specified during the training phase. Throughout training, the parameters, including weights and biases, of these models are updated using the optimizer. We employed Stochastic Gradient Descent (SGD) as our optimizer with a learning rate of 0.001 to facilitate model convergence and optimization. To measure the model's performance and guide the training process, we adopted cross-entropy loss as the loss function. Our training protocol involved 10 epochs, each comprising a batch size of 32 instances, ensuring thorough model optimization and robustness evaluation.

Discussion: The incorporation of HOG alongside CNN models such as VGG-16 and ResNet50 holds significant value in our approach for medical image classification. HOG serves as an effective initial feature extraction technique, capturing important texture and shape information from the input images. By utilizing HOG in conjunction with VGG-16 and ResNet50, the extracted features are fed into these deep learning architectures, enriching the feature representation with both local and global context. The combination of HOG with CNN models allows for a more comprehensive characterization of image content, speed up the models training; avoid overfitting, and enhancing the model's ability to discriminate between different classes. Moreover, by concatenating the outputs of VGG-16 and ResNet50 and considering them as input to a capsule network, we exploit the complementary strengths of these models on extract deep and spatial features, further improving the overall classification performance. This strategic integration of HOG and CNN models optimizes feature representation and facilitates more accurate and robust classification of our dataset.

3.2.3. Concatenation

Vectors concatenation involve combining several vectors together to form a longer vector and then processing it according to the chosen method [57]. Specifically, we concatenated feature vectors extracted by VGG-16 and Resnet50, each of size (1.1.1000), resulting in a combined vector of (1.1.2000). Subsequently, we utilized the RandomForestRegressor algorithm to evaluate and identify the most pertinent features from this concatenated vector. This process not only helped in reducing the dimensionality of our dataset but also enhanced the interpretability and reliability of our model's decisions. Finally, the resulting features were employed as inputs for the classifier. In this research, we employed a 10-fold cross-validation approach to ensure the reliability and statistical significance of the results, as well as to rigorously evaluate the model's performance. The dataset was divided into 10 equal subsets (folds). In each iteration, one fold was designated as the test set, while the remaining nine folds were used for training the model. This process was repeated 10 times, with each fold serving as the test set exactly once. By averaging the performance metrics across all iterations, we achieved a comprehensive and unbiased assessment of the model's effectiveness.

3.3. Classification

Automatic classification, or supervised classification, is the algorithmic categorization of objects based on statistical data. In our study, our goal is to classify AD. We started by processing our images, applying the usual methods such as normalizing, resizing, cropping, and augmenting. Then, we applied the HOG descriptor and two convolutional neural network models (Vgg-16 and Resnet50) to extract more information; these extracted features will be inputs to our final classification model (Capsnet).

3.3.1. Capsule network overview

Convolutional neural networks are very powerful at extracting useful features from images, but they can also encounter difficulties related to orientation and relative spatial relationships between features. A new type of neural network based on so-called capsules has solved this problem. Similar to how our brains can, capsules are a stacked collection of neuronal layers that can recover visual information. Each capsule in a CapsNet model is capable of obtaining the probability of a specific object being present, and it is made up of many neurons that display various instantiation properties related to the underlying object, such as rotation and size. It is a collection of neurons whose activity vectors describe a variety of characteristics. In other words, a capsule is a vector with a variety of attributes, and the length of the vector denotes the probability that the object it represents will exist. A vector's length is often decreased to less than one using a squashing function [58]. The input layer, hidden layers (dense reshape, capsule convolution, etc.), and output layer compose the three levels of the CapsNet architecture. The final layer of the capsnet is a fully connected layer. In this research, we applied the softmax activation function:

$$c_{ij} = \frac{e^{x_i}}{\sum_{i=1}^{n} e^{x_i}}$$
(1)

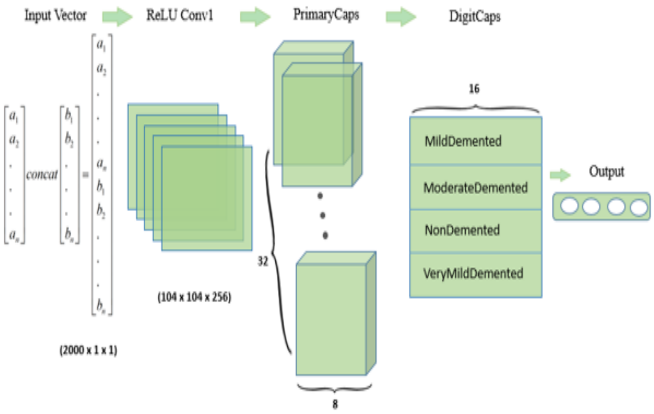

Fig. 4. The stages of the capsule network architecture applied to our features extracted to classify AD.

which generates the probability distribution, where x_i denote the i^{th} element of the vector, n is the number of class.

Considering u_i output of capsule i, its prediction for parent capsule j is computed as:

$$\hat{u}_{\frac{j}{i}} = w_{ij} u_i$$
(2)

where $\hat{u}_j$ is the prediction vector of the output of the j^{th} Capsule in a higher level computed by Capsule i in the layer below, and w_{ij} is the weighting matrix that needs to be learned in the backward pass, thus, for parent capsule input vector j is calculated as:

$$s_j = \sum_i c_{ij} \hat{u}_{\frac{j}{i}}$$
(3)

Lastly, to prevent the output vectors of Capsules from exceeding one, the following non-linear squashing function is employed to generate the final output of each Capsule based on its initial vector value defined in the previous equation.

$$v_j = \frac{\|s_j\|^2}{1+\|s_j\|^2}\frac{s_j}{\|s_j\|} \quad (4)$$

where s_i is the input vector to Capsule j and v_i is the output, therefore, we calculate the agreement aij for updating log probabilities and coupling coefficients as follows:

$$a_{ij} = v_j . \hat{u}_{j/i} \quad (5)$$

We summarized and compared all the steps and their associated equations for capsules and neurons in the figure below.

		Capsule	Neuron
Input		Vector(u_i)	Scalar(x_i)
Operation	Affine Transform	$\hat{u}_{j/i} = w_{ij} u_i$	-
	Weighting/Sum	$s_j = \sum_i c_{ij} \hat{u}_{j/i}$	$a_j = \sum_i w_i x_i +$
	Nonlinear Activation	$v_j = \frac{\|s_j\|^2}{1+\|s_j\|^2}\frac{s_j}{\|s_j\|}$	$h_j = f(a_j)$
Output		Vector(u_j)	Scalar(h_j)

Fig. 3. The steps and their associated equations for Capsule and

3.3.2. Capsule Network framework

The architecture of a capsule network has several stages that work together to capture hierarchical features and spatial relationships in the input data. It applies to the task of object recognition, particularly in situations where the spatial relationships between parts of an object are crucial. The figure below illustrates the different steps of the capsule network architecture applied to our extracted features to classify AD.

The main stages of the architecture of a capsule network are: 1) the input layer that receives the input data; each input capsule represents a primary feature or part of an object. 2) Convolutional layer with primary capsules to extract low-level features. 3) Primary Capsule Layer to capture the presence of specific features and their instantiation parameters, such as orientation and scale. 4) Digit capsule layer to represent higher-level features or parts of an object. 5) Routing by agreement to determine the connection strengths between capsules in different layers. 6) Squashing Function to ensure that the length of the output vector is between 0 and 1. 7) Final Classification Layer (softmax layer) to produce the output probabilities for different classes.

3.3.3. Capsule Network parameters

In our study, we employed a Capsule network to classify AD into 4 distinct classes. We leveraged the PyTorch library for our experiments, utilizing its versatile functionalities for neural network development. The hyperparameters chosen for our study were carefully selected to ensure optimal model performance. Specifically, we set the number of primary capsules to 32, the number of output capsules to 4, and the dimensionality of the output capsules to 16. Additionally, we performed routing iterations of 3 to refine the routing weights during the network's forward pass. For training, we employed a learning rate of 0.001 and trained the model for 40 epochs. Throughout the training process, we utilized the cross-entropy loss function to compute the discrepancy between predicted and actual class labels, and we optimized the model parameters using the Adam optimizer. This comprehensive approach enabled us to effectively train the Capsule network and achieve robust classification performance in identifying different stages of AD.

3.4. Discussion

Our approach to solving the complex task at hand is underpinned by a meticulously crafted three-phase framework designed to extract meaningful insights from our data. Firstly, we embark on the crucial phase of processing and augmenting our data, where we meticulously preprocess and enhance the raw input to ensure its suitability for subsequent analysis. This initial step is pivotal in ensuring the robustness and quality of our dataset, thereby laying a solid foundation for the subsequent phases. Following this, we adopt a novel approach by harnessing the power neural network models, VGG-16 and ResNet50, in tandem. By concatenating these models, we aim to leverage their complementary strengths to extract the most relevant and discriminative features from our data. This fusion not only enhances the richness and depth of our feature representations but also facilitates a more comprehensive understanding of the underlying patterns within the dataset. Lastly, we employ a cutting-edge capsule network model for the final classification stage. Unlike traditional convolutional architectures, capsule networks excel at capturing hierarchical spatial relationships within the data through dynamic routing between capsules. This unique mechanism enables the model to extract spatial characteristics in a more nuanced and sophisticated manner, leading to superior classification performance. Through the seamless integration of these three phases, our approach represents a holistic and innovative methodology designed to unlock new insights and drive advancements in the field. Real-time performance is critical in medical diagnosis, particularly in emergencies where fast and accurate decisions are essential. In our study, we evaluated the model's inference time and achieved classification results within one minute, demonstrating a balance between speed and accuracy for practical applications.

4. Experiments and results

In this section, we present the experimental setup and results of our study on Alzheimer's disease Classification Using Histogram of Oriented Gradient, Transfer Learning, and Capsules Network. Initially, we provide an overview of our approach and the frameworks employed for implementing the code. Following this, we detail the evaluation metrics used to assess the performance of our classification models. Subsequently, we conduct a comprehensive comparison between our approach and related works in the field. Finally, we delve into a discussion of our findings and offer insights into future research directions.

4.1. Experimental Setting

In this experiment, after the image pre-processing phase (normalization, resizing, cropping, augmentation, etc.), we employ an oriented histogram gradient descriptor and two convolutional neural network models (Vgg-16 and Resnet50) as feature extractors to extract key characteristics. Then we concatenated these features gleaned evaluated them by a RandomForestRegressor and used them as inputs for the Capsnet model. This novel strategy improved our classification prediction score. All trials were carried out on a computer with an NVIDIA GeForce GTX 1070 Ti GPU.

4.2. Results

In evaluating the performance of our classification models, we employed several key metrics, including precision, recall, accuracy, F1-score, loss and accuracy, and confusion matrix. In the following two paragraphs, we present the performance evaluation, the loss and accuracy curves, and confusion matrix.

4.2.1. Performances Evaluation

Evaluating the performance of a machine learning model involves employing a range of metrics and techniques to gauge its effectiveness, accuracy, and ability to generalize to new data. These metrics help assess how well the model might perform on unseen data and identify issues like overfitting or underfitting. Our experiment's effectiveness was determined using specific performance metrics tailored to the classification task, including precision, recall, accuracy, and f1-score.

Precision: the percentage of results that are relevant and it is defined as:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (6)$$

Recall: the percentage of total relevant results correctly classified by the proposed algorithm which is defined as:

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (7)$$

Accuracy: formally, accuracy has the following definition:

$$Accuracy = \frac{True\ Positive + True\ Negative}{Total} \quad (8)$$

F1-score: is a machine learning metric that can be used in classification models; f1-score has the following definition:

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (9)$$

In this study, we used the Alzheimer's dataset from a Kaggle competition for an Alzheimer's disease classification task to obtain the empirical results of this research. The primary aim of this experiment is to utilize common computer vision and infographic functions for processing our dataset. We also aim to extract relevant features using the HOG descriptor and use them as input for two CNN models, VGG-16 and ResNet50. Subsequently, we intend to feed these extracted features into a Capsule network classifier model to achieve high-performance classification. In the following paragraph, we present another two famous metrics (loss and accuracy) to evaluate the performance of our model.

4.2.2. Accuracy and loss

Accuracy: Accuracy is a metric of a classification model's efficacy; it is the ratio of correctly predicted instances to the total number of instances. In another sense, precision refers to the model's accuracy rate for predictions. The following figures illustrate the prediction results.

Loss: is a measure of the performance of the model; it quantifies the difference between the predicted values and the actual values; the total of the errors produced for each example in the training or validation sets constitutes the loss. Therefore, we presume that "the lower the loss, the better the model" [59].

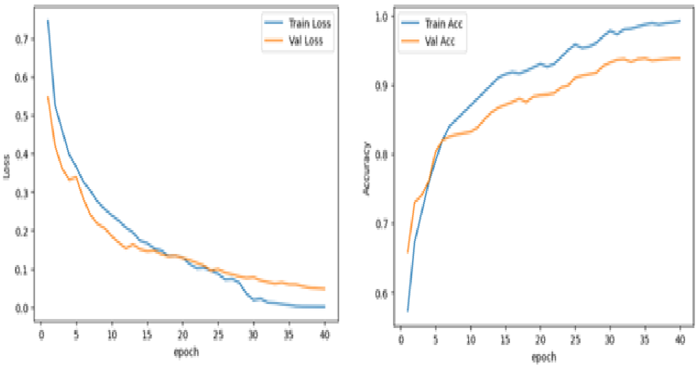

Fig. 5. The CapsNet architecture's loss and accuracy.

4.2.3. Confusion Matrix

A confusion matrix is a N*N matrix used to assess the effectiveness of a classification model, where N is the number of target classes. In the matrix, actual target values compare with those that the ML model anticipated.

The fundamental terms that will assist us in identifying the measures we're after are as follows:

- ✓ True Positives (TP): when the actual value is Positive and predicted is also Positive.
- ✓ True negatives (TN): when the actual value is Negative and prediction is also Negative.
- ✓ False positives (FP): When the actual is negative but prediction is Positive. Also known as the Type 1 error.
- ✓ False negatives (FN): When the actual is Positive but the prediction is Negative. Also known as the Type 2 error.

The figure below illustrates the confusion matrix of the capsule network architecture applied to our extracted features to classify AD.

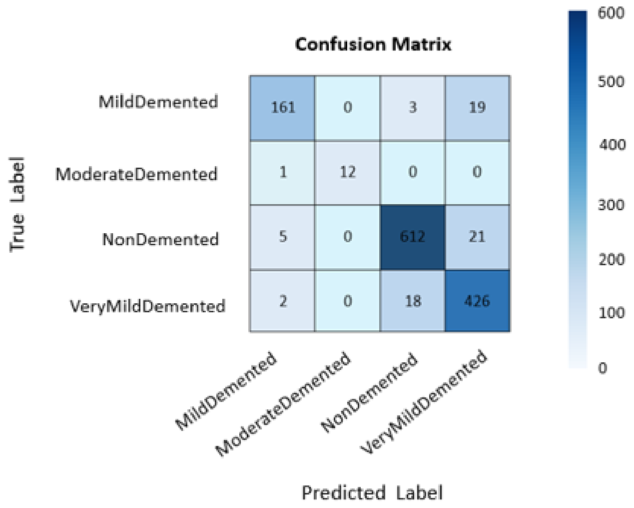

Fig. 6. Our CapsNet model's confusion matrix.

4.2.4. Discussion

In assessing the efficacy of our proposed method for AD classification, the utilization of key evaluation metrics such as accuracy and loss curves alongside a confusion matrix is paramount. The accuracy curve provides valuable insights into the overall performance of our model across epochs, illustrating its ability to correctly classify instances. Concurrently, the loss curve offers a visualization of the model's convergence during the training process, allowing us to gauge its optimization trajectory. Additionally, the confusion matrix provides insights into classification accuracy and potential misclassifications, and delivers a detailed breakdown of the model's classification performance, offering a comprehensive view of its predictive capabilities across different classes. In this study, by leveraging these evaluation tools, we can effectively validate the robustness and effectiveness of our approach in accurately classifying Alzheimer's disease cases, thereby affirming its potential clinical utility and contribution to medical diagnostics.

4.3. Disscussion

Previous work on early prediction and classification of AD achieves different levels of accuracy; this task is a crucial challenge. The techniques applied in the various stages of each approach: data import, preprocessing, feature extraction, and classification are the factors that contribute to adding value to our method proposed. In our study, the application of preprocessing functions (normalization, resizing, cropping, and augmentation) and feature extraction (Hog, Vgg-16, and Resnet50), thus the application of the capsule network algorithm allowed us to have a better classification of AD.

In the following table, we cite methods and their precision for classification prediction of Alzheimer's disease. The benefit of our study is that our approach applied to the same dataset contains powerful imaging functions (hog, vgg-16, resnet50). These methods allow us to extract crucial information, which will feed our classifier and obtain scores of higher accuracy. The table below shows the related work with this study.

Table 2. Performance comparison between our proposed method and different approaches.

Literature	Dataset	Classification Methods	Feature Extraction Method	Accuracy
Cosimo et al. (2018) [60]	Binary Class	2D CNN	Deep learning	89.8%
Zhang et al. (2021) [61]	ADNI (2 class)	73D Residual Attention Deep Neural Network	CNN	91%
Khedher et al. (2015) [62]	ADNI (2 class)	Partial least squares, principal component analysis	SVM	88.41
Ge et al. (2019) [63]	ADNI (2 class)	Multiscale	networks	93.53%

		convolutional neural	Softmax	
Payan and Montana (2015) [64]	ADNI (2 class)	3D convolutional neural network	networks Softmax	95.39%
Rachna Jain et al. (2018) [65]	ADNI (3 class)	Vgg-16	CNN	95.73%
Daniel Stamate et al. (2020) [66]	ADNI (3 class)	MLP ConvBLSTM (SMOTE) ConvBLSTM (GAIN)	Relief	89% 82% 82%
Acharya et al. (2021) [67]	Kaggle (4 class)	Transfer learning	Softmax	95.7%
Murugan et al. (2021) [68]	Kaggle (4 class)	DEMentia NETwork based cnn (SMOTE) DEMentia NETwork based cnn (without SMOTE)	Softmax	95.23% 85%
Marcia Hon and Naimul Mefraz Khan. (2017) [69]	Kaggle (4 class)	Transfer learning	Inception V4	96.25%
Proposed	Kaggle (4 class)	CapsNet	Vgg-16 Concat Resnet50 Hog + vgg-16 Hog + Resnet50 Hog + (vgg-16 Concat Resnet50)	89% 92% 91% **94.27%**

Our research brings a novel approach to Alzheimer's disease classification by leveraging advanced preprocessing techniques coupled with state-of-the-art deep learning architectures. Unlike previous works that often overlook the importance of preprocessing, we prioritize this stage by integrating common computer vision and infographic methods. Additionally, we employ the Histogram of Oriented Gradients (HOG) as an image descriptor to extract relevant and significant features. By doing so, we ensure that our model receives high-quality input data, speed up the models training, avoid overfitting, and setting a strong foundation for subsequent processing. During the feature extraction phase, we capitalize on the power of VGG-16 and ResNet50, renowned for their ability to extract deep features from images. This integration allows us to capture intricate patterns and nuances crucial for accurate disease classification. Finally, we employ a capsule network model for classification, a cutting-edge technique known for its ability to capture spatial relationships within data. By combining these elements, our approach aims to outperform previous methods in Alzheimer's disease classification by enhancing feature extraction and classification accuracy. Furthermore, we evaluated the model's real-time performance, achieving classification results within a one minute time frame and a classification accuracy of 94.27%, demonstrating its efficiency and suitability for practical applications.

5. Conclusion and perspective

In this study, we propose a hybrid model for Alzheimer's disease classification based on neural network architectures. Our research is structured into three key phases: Firstly, the dataset processing phase involves standard computer vision and infographic techniques like normalization, resizing, cropping, and augmentation. Secondly, in the feature extraction phase, we leverage HOG, VGG-16, and ResNet50 models to extract pertinent fea27tures, thereby expediting the models training, preventing overfitting, and fortifying the overall efficacy and resilience of the classification framework. Finally, in the classification phase, we concatenate these extracted features into a single vector, evaluate, and select them using a RandomForestRegressor, then utilize them as input in the CapsNet model to extract spatial features and predict outcomes.

Our approach's strength lies in leveraging feature extractors such as HOG, VGG-16, and Resnet50, and combining their extracted characteristic vectors for classification using the Capsnet model. This strategy optimizes the predictive capabilities of individual classifiers, thereby improving the overall accuracy and efficiency of our classification framework.

Finally, our experimental findings underscore the efficacy of the diverse feature extraction methods employed (HOG, VGG-16, and Resnet50) in capturing the profound attributes of MR images, along with the synergy achieved through their concatenation. Furthermore, the capability of the Capsnet model to extract spatial features and perform classification adds another layer of efficiency to our approach. These results demonstrate promising advancements in medical imaging analysis, particularly in computer-assisted diagnosis in digital pathology.

Several studies in disease classification often fall short of meeting medical experts' expectations due to issues such as poor performance, data dependency, or reliance on computationally complex deep learning models. In our future endeavors, we aim to investigate alternative large-scale datasets and devise methodologies to overcome these limitations, striving to advance the field of medical image analysis and provide more reliable tools for disease diagnosis and prognosis.

References

[1] S. Saman et T. Ghassem, «Classification of Alzheimer's Disease Using fMRI Data and Deep Learning Convolutional Neural Networks,» n° %1https://doi.org/10.48550/arXiv.1603.08631, 2016.

[2] S. Saman et T. Ghassem, «DeepAD: Alzheimer's Disease Classification via Deep Convolutional Neural Networks using MRI and fMRI,» n° %1https://doi.org/10.1101/070441, 2017.

[3] R. Farheen, G. K. Muhammad Usman, R. Asim, I. Sajid, S. Tanzila, R. Amjad et M. Zahid, «A Deep Learning Approach for Automated Diagnosis and Multi-Class Classification of Alzheimer's Disease Stages Using Resting-State fMRI and Residual Neural Networks,» n° %1https://doi.org/10.1007/s10916-019-1475-2, p. 2019.

[4] M. Tanveer, B. Richhariya, R. U. Khan, A. H. Rashid, P. Khanna, M. Prasad, C. T. LinAuthors Info et Claims, «Machine Learning Techniques for the Diagnosis of Alzheimer's Disease: A Review,» n° %1https://doi.org/10.1145/3344998, 2020.

[5] F. Zhao, X. Fanyu, Q. Xuedan, L. Cai et Y. Lili, «Classification of Alzheimer's disease based on brain MRI and machine learning,» n° %1https://doi.org/10.1007/s00521-019-04495-0, 2019.

[6] Y. Nagaraj, C. Jae Young et L. Bumshik, «MRI Segmentation and Classification of Human Brain Using Deep Learning for Diagnosis of Alzheimer's Disease: A Survey,» n° %1https://doi.org/10.3390/s20113243, 2020.

[7] L. Siqi, L. Sidong, C. Weidong, P. Sonia, K. Ron et F. Dagan, «Early diagnosis of Alzheimer's disease with deep learning,» n° %110.1109/ISBI.2014.6868045, 2019.

[8] K. Bijen, K. Goo-Rak et L. Ramesh, «Comparative analysis of Alzheimer's disease classification by CDR level using CNN, feature selection, and machine-learning techniques,» n° %1 https://doi.org/10.1002/ima.2231, 2019.

[9] Z.-d. Iliass, F. Anass, K. E. F. Jamal Riffi, E. B. Ismail, M. Mohamed Adnane et T. Hamid, «Brain tumor classification using features extraction and ensemble learning,» 2025.

[10] G. Ahmed, M. Ghulam et H. M Shamim, «Cervical cancer classification using convolutional neural networks and extreme learning machines,» n° %1https://doi.org/10.1016/j.future.2019.09.015, 2020.

[11] Zine-dine, J. Riffi, E. F. Khalid, M. Mohamed Adnane et T. Hamid, «Brain Tumor Classification using Machine and Transfer Learning,» n° %1DOI: 10.5220/0010762800003101, 2022.

[12] J. Taeho, N. Kwangsik et J. S. Andrew, «Deep Learning in Alzheimer's Disease: Diagnostic Classification and Prognostic Prediction Using Neuroimaging Data,» n° %1https://doi.org/10.3389/fnagi.2019.00220, 2019.

[13] W. Ahmad, P. B. Salehi et G. Gaurav, «Alzheimer's Disease Diagnosis using Deep Learning Techniques,» n° %1DOI: 10.35940/ijeat.C5345.029320, 2020.

[14] R. S. Iago R., L. S. Gabriela S., d. S. Rodrigo G., d. S. Wellington P. et d. A. F. Roberta A., «Model Based on Deep Feature Extraction for Diagnosis of Alzheimer's Disease,» n° %110.1109/IJCNN.2019.8852138, 2019.

[15] H. Ruhul Amin, M. Arnab Kumar, N. S. Samarendra, S. P. Babu et K. Debdatta, «A Survey on Classification Algorithms of Brain Images in Alzheimer's Disease Based on Feature Extraction Techniques,» n° %110.1109/ACCESS.2021.3072559, 2021.

[16] S. Muhammad, A. Shahzad, G. Aziz, N. Aneeta et U. Amina, «Classification of Alzheimer's Disease using Machine Learning Techniques,» n° %110.5220/0007949902960303, 2019.

[17] P. Enrico, B. Lucia, D. C. Maria, H. Valdes, C. Francesca M., G.-C. Victor, A. Devasuda, D. Samuel, M.-M. Susana, J. Dominic et P. Cyril, «Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: A systematic review. Alzheimer’s,» n° %1https://doi.org/10.1016/j.dadm.2018.07.004, 2018.
[18] M.-P. José María, D. Mahsa, L.-A. María, I.-M. Yasser, Z. Yashar et C. and Alan, «Structural neuroimaging as clinical predictor: A review of machine learning applications,» n° %1https://doi.org/10.1016/j.nicl.2018.08.019, 2018.
[19] L. Geert, K. Thijs, E. B. Babak, A. S. Arnaud Arindra, C. Francesco, G. Mohsen, V. D. L. Jeroen Awm, V. G. Bram et S. and Clara I., «A survey on deep learning in medical image analysis,» n° %1https://doi.org/10.1016/j.media.2017.07.005, 2017.
[20] S. Dinggang, W. Guorong et a. H.-I. Suk, «Deep learning in medical image analysis,» n° %1https://doi.org/10.1146/annurev-bioeng-071516-044442, 2017.
[21] Mehran, A. Hojjat et A. Anahita, «New diagnostic EEG markers of the Alzheimer’s disease using visibility graph,» n° %1https://doi.org/10.1007/s00702-010-0450-3, 2010.
[22] R. Pedro et T. João Paulo, «Artifcial neural networks in the discrimination of Alzheimer’s disease,» n° %1https://doi.org/10.1007/978-3-642-24352-3_29, 2011.
[23] C. Youngsang, S. Joon-Kyung, J. Yong et S. Sung Yong, «Alzheimer’s Disease Neuroimaging Initiative,» 2012.
[24] Y. Jieping, C. Kewei, W. Teresa, L. Jing, Z. Zheng, P. Rinkal, B. Min, J. Ravi, L. Huan et A. Gene, «Heterogeneous data fusion for alzheimer’s disease study,» n° %110.1145/1401890.1402012, 2008.
[25] Z. Daoqiang et S. Dinggang, «Semi-supervised multimodal classifcation of alzheimer’s disease,» n° %1DOI: 10.1109/ISBI.2011.5872715, 2011.
[26] L. Feng, W. Chong-Yaw, C. Huafu et S. and Dinggang, «Inter-modality relationship constrained multimodality multi-task feature selection for Alzheimer’s disease and mild cognitive impairment identifcation,» n° %1https://doi.org/10.1007/978-3-642-40811-3_39, 2014.
[27] T. Tong, G. Katherine, G. Qinquan, C. Liang et R. Daniel, «Multi-modal classifcation of Alzheimer’s disease using nonlinear graph fusion. Pattern Recogn,» n° %1https://doi.org/10.1016/j.patcog.2016.10.009, 2017.
[28] K. Jongin et L. Boreom, «Identifcation of Alzheimer’s disease and mild cognitive impairment using multimodal sparse hierarchical extreme learning machine.,» n° %1https://doi.org/10.1002/hbm.24207, 2018.
[29] S. Heung-Il et S. Dinggang, «Deep learning-based feature representation for AD/MCI classifcation,» n° %1https://doi.org/10.1007/978-3-642-40763-5_72, 2013.
[30] P. Adrien et M. Giovanni, «Predicting Alzheimer’s disease: A neuroimaging study with 3D convolutional neural networks,» n° %1https://doi.org/10.48550/arXiv.1502.02506, 2015.
[31] L. Siqi, L. Sidong, C. Weidong, C. Hangyu, P. Sonia, K. Ron, F. M. Dagan et F. J., «Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer’s disease,» n° %110.1109/TBME.2014.2372011, 2015.
[32] H.-A. Ehsan, G. Georgy et E.-B. Ayman, «Alzheimer’s disease diagnostics by a deeply supervised adaptable 3D convolutional network,» n° %1https://doi.org/10.48550/arXiv.1607.00556, 2016.
[33] S. Heung-Il et S. Dinggang, «Deep ensemble sparse regression network for Alzheimer’s disease diagnosis,» n° %1https://doi.org/10.1007/978-3-319-47157-0_14, 2016.
[34] S. Heung-Il, W. Chong-Yaw, L. Seong-Whan et S. Dinggang, «State-space model with deep learning,» n° %1https://doi.org/10.1016/j.neuroimage.2016.01.005, 2016.
[35] D. Christos, M. R. Susan, W. X, P. P, Christopher et M. Clark, «Individual patient diagnosis of AD and FTD via high-dimensional pattern classifcation of MRI,» n° %1https://doi.org/10.1016/j.neuroimage.2008.03.050, 2008.
[36] L. M., R. J., G. J. M., S.-G. D., A. I., S. F. et P. C. G., «Automatic tool for Alzheimer’s disease diagnosis using PCA and bayesian classifcation rules,» n° %1https://doi.org/10.1049/el.2009.0176, 2009.
[37] T. M. et G. Manuel, «A two stage sequential ensemble applied to the classifcation of Alzheimer’s disease based on mri features,» n° %1https://doi.org/10.1007/s11063-011-9200-2, 2012.
[38] C. Andrea, B. Paolo, C. Piero, G. Gianluca, E. Mario, O. Chiara, R. Luca, S. Sandro, R. Guido et B. Roberto, «Local MRI analysis approach in the diagnosis of early and prodromal Alzheimer’s disease,» n° %1https://doi.org/10.1016/j.neuroimage.2011.05.083, 2011.
[39] H. Soheil et M. Mehdi, «Learning in data-limited multimodal scenarios: Scandent decision forests and tree-based features,» n° %1https://doi.org/10.1016/j.media.2016.07.012, 2016.

[40] G. Neha, C. Mahipal Singh et M. B. Rajesh, «A review on Alzheimer's disease classification from normal controls and mild cognitive impairment using structural MR images,» n° %1https://doi.org/10.1016/j.jneumeth.2022.109745, 2023.
[41] L. S., L. S., C. W., C. H., P. S., K. R., F. D. et F. M. J., «Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease,» n° %110.1109/TBME.2014.2372011, 2015.
[42] L. F., T. L, H. T. K, J. S., S. D. et L. and J., «A robust deep model for improved classification of AD/MCI patients,» n° %1DOI: 10.1109/JBHI.2015.2429556, 2015.
[43] B. T. et R. Tam, «Manifold learning of brain MRIs by deep learning,» n° %1https://doi.org/10.1007/978-3-642-40763-5_78, 2013.
[44] S. N., H. G., S. I. et S. R., «Dropout: A simple way to prevent neural networks from overfitting,» 2014.
[45] Karim, B.-P. Jenny et A. Karim, «Classification of sMRI for Alzheimer's disease Diagnosis with CNN: Single Siamese Networks with 2D+ε Approach and Fusion on ADNI,» n° %1https://doi.org/10.1145/3078971.3079010, pp. 494 - 498, 2017.
[46] S. Heung-Il, L. Seong-Whan et S. Dinggang, «Subclass-based multi-task learning for Alzheimer's disease diagnosis,» n° %1 https://doi.org/10.3389/fnagi.2014.00168, 2014.
[47] Y. S. Holtzman et D. M., «Interplay between innate immunity and Alzheimer disease: APOE and TREM2 in the spotlight,» n° %1https://doi.org/10.1038/s41577-018-0051-1, 2018.
[48] L. Chunfeng, L. Mingxia, Z. Jun et S. Dinggang, «Hierarchical Fully Convolutional Network for Joint Atrophy Localization and Alzheimer's Disease Diagnosis using Structural MRI,» n° %1doi: 10.1109/TPAMI.2018.2889096, 2018.
[49] Z. Xin, H. Liangxiu, Z. Wenyong, S. Liang et Z. Daoqiang, «An Explainable 3D Residual Self-Attention Deep Neural Network For Joint Atrophy Localization and Alzheimer's Disease Diagnosis using Structural MRI,» n° %110.1109/JBHI.2021.3066832, 2015.
[50] Y. Jianzhou, L. Stephen, K. Sing Bing et T. Xiaoou, «Learning the Change for Automatic Image Cropping.,» Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 971-978, 2013.
[51] M. Kiran, M. Surajit et N. Bhushan kumar, «A review: Data pre-processing and data augmentation techniques.,» n° %1https://doi.org/10.1016/j.gltp.2022.04.020, 2022.
[52] D. ping Tian, «A Review on Image Feature Extraction and Representation Techniques.,» International Journal of Multimedia and Ubiquitous Engineering, Vols. %1 sur %2Vol. 8, No. 4, 2013.
[53] S. Deepak et P. Amer, «Brain tumor classification using deep CNN features via transfer learning.,» n° %1https://doi.org/10.1016/j.compbiomed.2019.103345, 2019.
[54] H. Acharya, R. Mehta et D. Kumar Singh, «Alzheimer Disease Classification Using Transfer Learning,» n° %1doi: 10.1109/ICCMC51019.2021.9418294, 2021.
[55] N. Deepa et S. Chokkalingam, «Optimization of VGG16 utilizing the Arithmetic Optimization Algorithm for early detection of Alzheimer's disease,» n° %1doi.org/10.1016/j.bspc.2021.103455, 2022.
[56] L. V. Fulton, D. Dolezel, J. Harrop, Y. Yan et C. P. Fulton, «Classification of Alzheimer's Disease with and without Imagery Using Gradient Boosted Machines and ResNet-50,» n° %1doi.org/10.3390/brainsci9090212, 2019.
[57] R. Venkatesan, J. R. Alex Noel, P. T. Krishnan et N. Ganesh R., «A Customized VGG19 Network with Concatenation of Deep and Handcrafted Features for Brain Tumor Detection.,» n° %1https://doi.org/10.3390/app10103429, 2020.
[58] S. Sara, F. Nicholas et H. Geoffrey E., «Dynamic Routing Between Capsules,» Advances in Neural Information Processing Systems, 2017.
[59] P. R., R. Sumantra Dutta, M. Pravat K. et G. Shantanu, «High-Accuracy Detection of Early Parkinson's Disease through Multimodal Features and Machine Learning.,» n° %1https://doi.org/10.1016/j.ijmedinf.2016.03.001, 2016.
[60] Cosimo, M. Nadia, B. Alessia, H. Amir et M. Francesco C., «A Convolutional Neural Network approach for Classification of Dementia Stages based on 2D-Spectral Representation of EEG recordings,» n° %1https://doi.org/10.1016/j.neucom.2018.09.071, 2018.
[61] Z. Xin, H. Liangxiu, Z. Wenyong, S. Liang et Z. Daoqiang, «An explainable 3D residual self-Attention deep neural network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI,» n° %110.1109/JBHI.2021.3066832, 2021.
[62] K. L., R. J., G. J.M. et F. S. A. Brahim, «Initiative AsDN, Early diagnosis of Alzheimer ′s disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images,» n° %1https://doi.org/10.1016/j.neucom.2014.09.072, 2015.
[63] G. C, Q. Q, I, Y-H, A. Gu et Jakola, «Multiscale Deep Convolutional Networks for Characterization and Detection of Alzheimer's Disease Using MR

Images,» n° %1doi: 10.1109/ICIP.2019.8803731, 2019.

[64] Payan et G. Montana, «Predicting Alzheimer's Disease: a Neuroimaging Study with 3D Convolutional Neural Networks,» n° %1https://doi.org/10.48550/arXiv.1502.02506, 2015.

[65] J. Rachna, J. Nikita, A. Akshay et J. H. D., «Convolutional Neural Network based Alzheimer's Disease Classification from Magnetic Resonance Brain Images,» n° %1https://doi.org/10.1016/j.cogsys.2018.12.015, 2018.

[66] S. Daniel, S. Richard, T. Ruslan, V. Rostislav, L. John, S. Daniel et R. & David, «Applying Deep Learning to Predicting Dementia and Mild Cognitive Impairment,» https://doi.org/10.1007/978-3-030-49186-4_26, 2020.

[67] Heta, M. Rutvik et S. Dheeraj Kumar, «Alzheimer Disease Classification Using Transfer Learning.,» n° %110.1109/ICCMC51019.2021.9418294, 2021.

[68] S. Murugan, V. Chandran, S. M. G., G. Xiao-Zhi, E. B. et A. M., «DEMNET: A Deep Learning Model for Early Diagnosis of Alzheimer Diseases and Dementia From MR Images,» n° %110.1109/ACCESS.2021.3090474, 2021.

[69] M. H. Khan et M. and Naimul, «Towards Alzheimer's Disease Classification through Transfer Learning,» n° %110.1109/BIBM.2017.8217822, 2017.