Alzheimer's Disease Diagnosis Assistance Through the Use of Deep Learning and Multimodal Feature Fusion

Angela Díaz-Cadena, Irma Naranjo Peña, Hector Lara Gavilanez, Diana Sanchez Pazmiño, and Miguel Botto-Tobar

Abstract Patients suffering from Alzheimer's disease (AD) lose their ability to think and frequently forget what they have learned during their life. There are currently no effective therapies available for this illness. The sooner the disease is recognized, the better the therapy alternatives and the greater the possibility of eliminating Alzheimer's. Computer-assisted diagnosis, or CAD, is a method that integrates neuroimaging with deep learning algorithms trained on multimodal pictures. The CAD system is powered by deep learning algorithms that were trained to function by being exposed to a diverse spectrum of artistic outputs. Each component of the system affects the functioning of these algorithms. In recent years, several multimodal feature learning-based alternative techniques for extracting and integrating latent. We were able to achieve our aim because we devised several novel approaches. Here are some more detailed illustrations of imaging techniques: This diagnostic category includes imaging procedures such as MRI and PET scans. Given the complexities of the procedures utilized, providing a complete assessment of the immeasurable value of the data obtained is difficult. An image-based multimodal fusion approach is proposed as a result, our understanding of the brain's structure and operation has grown significantly. The technique's primary emphasis is the grey matter of the brain.

A. Díaz-Cadena · I. Naranjo Peña · H. Lara Gavilanez · D. Sanchez Pazmiño
University of Guayaquil, Guayaquil, Ecuador
e-mail: angela.diazca@ug.edu.ec

I. Naranjo Peña
e-mail: irma.naranjop@ug.edu.ec

H. Lara Gavilanez
e-mail: hector.larag@ug.edu.ec

D. Sanchez Pazmiño
e-mail: diana.sanchezp@ug.edu.ec

M. Botto-Tobar (✉)
Research Group in Artificial Intelligence and Information Technology, University of Guayaquil, Guayaquil, Ecuador
e-mail: miguel.bottot@ug.edu.ec

Eindhoven University of Technology, Eindhoven, The Netherlands

D. Koundal et al. (eds.), *Data Analysis for Neurodegenerative Disorders*,
Cognitive Technologies, https://doi.org/10.1007/978-981-99-2154-6_8

We were able to provide more accurate diagnoses to individuals suffering from neurological diseases. To accomplish our purpose, we use the registration and mask coding procedures. This had a direct impact on the creation of a well-rounded theory aimed primarily at the automobile sector. In addition, we put our image fusion approach to the test with a 3D basic convolutional neural network for binary classification and a 3D multi-scale CNN for multiple classification tasks. These two networks are linked by the fact that they are 3D convolutional neural networks. In a three-dimensional situation, both functions admirably. Using the ADNI dataset, researchers revealed that their suggested picture fusion algorithm outperformed cutting-edge approaches for detecting Alzheimer's disease. Furthermore, as compared to feature fusion and single-modal approaches, its overall performance is significantly superior.

Keywords Alzheimer's disease · Computer-assisted diagnosis · Convolution neural network · Feature fusion

1 Introduction

It is a neurodegenerative illness, which means that brain cells progressively and inexorably die over time. People with cognitive impairment are unable to execute daily tasks because their cognitive abilities have deteriorated. As the symptoms of Alzheimer's disease increase, both patients and those who care for them experience a decrease in their overall quality of life [1]. By the year 2020, global spending on dementia-related medical care, nursing home stays, and end-of-life treatments is expected to exceed $305.2 billion [2]. According to the most recent predictions, there will be 115 million Alzheimer's disease patients globally by 2050. This demonstrates the critical need for better therapies and an accurate technique for identifying Alzheimer's disease.

Alzheimer's disease has an unknown aetiology, making it difficult to determine what causes it. The scientific community recognizes that these two pathways play a significant role in both neurodegeneration and synaptic loss [3, 4]. The capacity of a patient to fulfil the diagnostic criteria for that condition determines whether they have normal control, moderate cognitive impairment (MCI), or Alzheimer's disease. Normal control, MCI, and Alzheimer's disease are all terms used to describe the same illness. One of these situations is within "normal control," making it the least perilous of the three. Alzheimer's disease is frequently identified with structural magnetic resonance imaging, also known as diagnostic MRI. The ability of structural MRI to highlight critical parts of brain architecture and its enhanced soft-tissue resolution are largely responsible for this. The cerebral cortex, temporal lobes, parietal lobes, and anterior cingulate gyrus are all part of this group of brain structures [5, 6]. Alzheimer's disease is characterized by an enlargement of the ventricles, which oversee creating cerebrospinal fluid. The hippocampus and other areas of the brain have been revealed to be smaller than they were previously. When brain tissue is scanned using magnetic resonance spectroscopy, sharp, three-dimensional (3D) pictures are produced. These

graphics may help you identify structural issues and have a better understanding of them. While conducting their research, some studies that employed MRI as a clinical diagnostic technique for Alzheimer's disease revealed some unexpected new facts. Trained support vector machines on voxels representing grey matter (GM) to improve their understanding of the features of MR images [7]. Until recently, white matter, grey matter, and cerebrospinal fluid were all thought to be different, independent components of the brain. The researchers could plainly determine that Alzheimer's patients were not identical to healthy controls. When the spatial normalization method was completed, we turned our focus to the GM tissue densities and gave them additional thought because of the strong link between GM and AD, no other approach could have produced this result [8, 9]. On the other hand, when I learned more about cerebral spinal fluid and grey matter, I developed a newfound excitement. Effectively recognized neuroimaging data by using GM volume as a single feature for all 93 ROIs in the annotated MR image. The only methodologies we used to reach this result were GM volume analysis and multiple-kernel learning [10]. This change was made in anticipation of the anticipated favorable impact of the migration on the efficacy with which data categorization may be performed. The findings of these studies suggest that when attempting to diagnose Alzheimer's disease, magnetic resonance imaging (MRI) should concentrate primarily on GM tissue. It is worth noting that this result is consistent with what Zhu and colleagues observed [11, 12].

Functional techniques, which rely largely on this technology, rely on PET imaging's capacity to give a rapid and accurate study of brain-related activity. Functional approaches may thus thrive in conditions such as these. Finally, it has the potential to be utilized to test individuals at high risk for Alzheimer's disease's cognitive symptoms decades before the illness manifests itself in a clinical environment. One technique that may be used to successfully attain this goal is to compare the person's intellect and memory to a standard. FDG-PET is a highly useful diagnostic tool because it can discriminate between favorable and detrimental morphological changes [13–15]. Because the number of brain structures decreases with age, it may be difficult to appropriately assess a person's mental health based just on the morphological changes shown by an MRI of the brain. This is because physical changes do not necessarily influence mental health. This is described because of a person's brain naturally weakening with age (for example, in individuals who are older than 75 years). According to the study's findings, this characteristic is more widespread among elderly people (those aged 75 and over). In cases like these, PET testing may be able to offer a more exact evaluation of the patient's present state of health than more traditional approaches.

2 Background

In contrast to functional PET imaging, which may only highlight metabolic changes, structure-based MRI may be able to illustrate how the structure of the brain has

developed over time. This would be preferable to functional PET imaging. When it comes to pinpointing the exact site of lesions, structural MRI outperforms functional PET imaging. Consider the following example: One strategy for improving the precision of the Alzheimer's disease diagnosis procedure is to combine MRIs and PET scans into a single multimodal approach. In this scenario, a reference is necessary [16–19]. The challenges raised by multimodal learning have been handled in several ways, many of which rely on the integration of many components. This is one of the most widely used methods: These techniques produce high-dimensional semantic characteristics by beginning with a diverse set of unimodal inputs [20, 21]. According to Shi and colleagues' study, two-layered deep polynomial networks may be able to recognize the ethereal properties of pictures such as MRI and PET scans [22]. The inclusion of these attributes into a second SDPN facilitates the integration of data from diverse neuroimaging modalities. We were able to merge data from photographs with different pixel sizes by including a second deep neural network, which was quite useful. A variety of data sources are integrated into a process known as "feature fusion," and it has been demonstrated that this produces more accurate experimental findings than using only one data source [23–25]. This method has been nicknamed "the black box" since it cannot explain why observed outcomes differ from those predicted. Data is collected from a variety of sources using deep learning approaches based on fusion and a multi-channel input network. As a result, the conditions for the total number of model parameters quickly become much stiffer.

The multimodal technique is one of the easiest ways to accomplish medical data fusion. Many independent input pictures are merged to generate a single composite image using this approach. This image can assist medical experts in making a more accurate diagnosis and determining the best course of therapy. If this method is used, the patient is more likely to respond positively to the treatment they are receiving. Feature fusion algorithms, on the other hand, merge information from many pictures to generate a more accurate whole. The integrated visualizations highlight the data's numerous modal features while also providing a more realistic picture of the data. GM tissue, on the other hand, is required for a correct diagnosis of Alzheimer's disease. MRI scans, which are restricted to finding just morphological abnormalities in the patient's brain due to their low resolution, cannot analyze the metabolic rate of the patient's complete brain like PET scans do. A PET scan may be used to assess the patient's whole body. Once all the scan data has been obtained, any genetic differences determined to be of low significance are removed, while any genetic differences deemed to be of high significance are saved for future research. Feature extraction is the practice of removing sections of an image that are deemed superfluous or insignificant. As a result, the viewer will be made aware of the major points of interest in the image.

Our lengthy investigation enabled us to divide the most relevant findings into two categories. Here, an unusual approach for detecting Alzheimer's disease is presented that incorporates a substantial number of photos. We hope that using this strategy, we will be able to improve the information representation capabilities of a wide range of neuroimaging modalities. This is accomplished by contrasting and comparing the capabilities of each of these CNN versions. This is demonstrated by comparing the

diagnostic effectiveness of the proposed fused modality to that of the two CNNs. In this post, we will provide evidence to back up our work. In the next chapters, we will continue to dissect the project into its constituent elements and undertake a much more in-depth investigation of each. In this section, we will go through the dataset we utilized as well as the photo fusion approach we employed in depth. Convolutional neural networks can extract and categorize a wide range of characteristics from neuroimaging data. These studies were carried out to explore if our suggested picture fusion method could be employed in the context of an AD diagnostic paradigm. The purpose of this study is to see how viable the picture fusion approach we presented is within the framework of an AD diagnosis paradigm. There are additional comparisons of AD and MCI to NC. NC and AD can also be contrasted. As a result, there are now four possible combinations in total. The similarity between NC and Alzheimer's illness is one such connection. The fifth subsection of this section contains both the judge's decision and the final ruling. Section 4 contains the argument, which is stated below.

3 The Proposed Method

Unmistakable visible evidence that the components that have been assembled successfully complement one another [26–35]. To do this, the data from the two scans will be blended. As a result of this, the conclusions will be more trustworthy. Combining photos from several sources is one approach that may be utilized to attain this aim. As a result of this, we may have higher expectations for receiving a more accurate diagnosis. The composite picture modality must be sent across a network with only one channel to complete the diagnostic process. Another citation is required in this area of the text. The Alzheimer's Association created the image. It is critical in this process to merge many photographs into a single image, then identify certain components inside the merged image, and finally categorize the resulting images. Which is the full word (PET). Following this, we'll most likely analyze the GM-PET pictures to find the semantic traits they have. All the information gathered about everyone is sent into a two-tiered classifier. These layers are also known as the FC layer and the SoftMax layer. In this instance, categorizing individuals into the appropriate category based on their characteristics will be easier.

Since its inception, multimodal image fusion has been able to integrate complementary data from photographs collected using several modalities. Before contemporary technology, things were done differently. Figure 1 depicts how we utilize the MRI as a mask in our method of image fusion. This allows you to avoid the FDG-PET scan in the whole GM area. This industry is heavily reliant on Alzheimer's disease diagnosis. The MRI scan is used as a mask to achieve this effect. This article will lead you through each stage of the multi-stage picture compositing process.

The following is a list of the MRI processing pipeline stages that we performed in the order that they were done. After removing a portion of the patient's skull, the MRI is registered using MNI152, and the tissue is segmented on the MRI. The

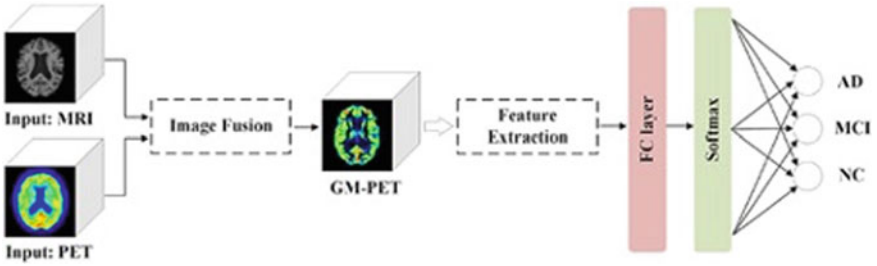

Fig. 1 An approach that incorporates a variety of imaging data sources for diagnosis of Alzheimer's disease

following protocols must be followed to successfully finish the PET technique: The steps of this procedure are as follows: The Origin PET was matched to the MNI's MRI, the MNI's PET to the GM's MRI, and the MNI-GM PET back to the Origin PET.

It has been demonstrated that the FreeSurfer 6.0 program's "watershed" module may be used to skull-strip structural MRI data [29]. This instance is depicted in Fig. 2. To remove surrounding tissues such as the skull and those that aren't part of the brain, the data can be filtered using the watershed segmentation approach. This immediately reduces the amount of background noise and irrelevant data in the research participant's brain volume. It is now known as SS-MRI, and it is expected that this will be the only aspect of brain tissue structure that is maintained. Toxins will continue to be eliminated from various bodily areas.

The affine translation from the SS-MRI space to the MNI152 space is shown in Fig. 2. This transition is a common design for global brain atlases. This strategy may be applied to a variety of additional imaging modalities. Affinity linear transformations are widely used in healthcare. To get the most exact results from the registration procedure, the motions of the participants in the scanner must be constrained as much as possible about a reference frame. This increases the precision of following treatments that segment the tissue. Now, MNI-MRI data that has already been recorded is used as the input mode for unimodal Alzheimer's disease classification tasks.

The FMRIB Automated Segmentation Tool (FAST) module, which is part of the FSL package, is used to extract and segment the GM area. This is done to achieve the planned aims. The fully automated procedure begins with an input picture and ends with probabilistic and/or partial volume tissue segmentation. A variety of criteria can be met by modifying the bias field. Unlike previous techniques that depended on finite mixture models, our approach is dependable and devoid of background noise. Figure 2c depicts the results of the GM tissue segmentation.

It may be possible to fix the grayscale divergence problem that we addressed by co-registering the MNI-GM-PET picture with the appropriate. This approach was used to generate the GM-PET picture shown in Fig. 2. Our method of registration, which simultaneously accounts for and corrects for the divergence brought on by the affine translation, maintains the spatial resolution of the initial PET image. Following

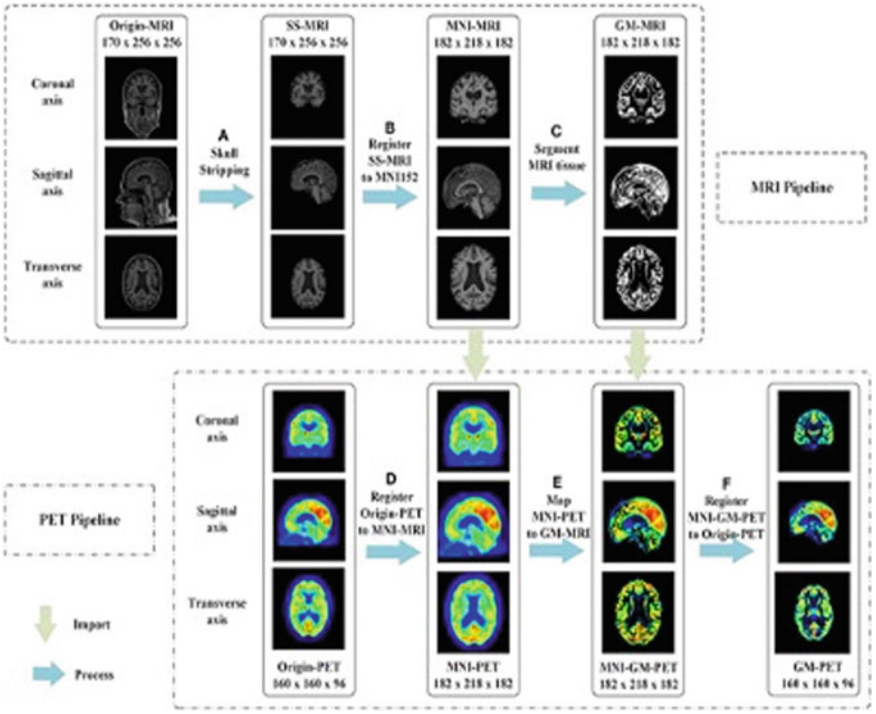

Fig. 2 The multimodal image fusion approach involves combining data from a variety of sources

the application of an affine transformation, an image that differs from the original PET scan is created. It is possible to achieve both goals if work is done on both at the same time. Switching to a resolution that uses less processing power and storage space on the device could be a good idea.

3.1 The Data Sources

The data used in this investigation were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database, which may be viewed online at the following URL: https://adni.loni.usc.edu. The use of this dataset enabled the gathering of the information required for this investigation. A multinational collaboration of researchers known as the Alzheimer's Condition Neuroimaging Initiative is currently employing imaging technologies to understand more about Alzheimer's disease. This program was created with the primary purpose of doing in-depth research on Alzheimer's disease (ADNI). The primary purpose of this research is to identify clinical, radiological, genetic, and biochemical markers for Alzheimer's disease. The two major outcomes of this endeavor are the early discovery of the illness

Table 1 Regarding demography, the data are displayed using the mean and standard deviation

Subjects	Quantity	Gender ratio	Age	MMSE	CDR
NC	128	72/56	74	28.15	0.03
MCI	162	109/53	75	25.33	1.27
AD	97	55/42	74	17.84	2.75

and the monitoring of the illness's course. To reach this aim, we will perform this research. Because the bulk of ADNI participants came from a varied range of North American nations, data gathering, and synthesis required a wide range of sources. Data from the ADNI participants was obtained and analyzed for this purpose. This book covers several centuries following the Common Era. The relationship between the two imaging modalities was investigated using FDG-PET and T1-weighted MRI images of the same people. This is merely one of the numerous factors that influenced the individuals' willingness to participate in the study. MPRAGE scans are so named because it is widely accepted that MRI scans are the most accurate imaging technology. Table 1 displays the clinical data obtained from research participants.

For an MRI picture to be properly processed, the following steps must be completed in the correct order: This phenomenon is induced by a combination of slowing time, B1 anomalies, and the resonant N3. The brightness of the picture may be changed using either a Grad warp or a B1 calibration scan. In Adobe Photoshop, you may receive the required scans for each of them. If the gradient model has distorted your geometry, you may correct it by doing a Grad warp calibration scan or a B1 calibration scan. Grad warp may also be used to rectify any irregularities in the brightness of the image formed by the gradient model with a few clicks. Using a peak-sharpening approach on the N3 histogram, it is feasible to raise the total signal intensity and get the desired result. An Example Because various manufacturers build the RF coils, they use them in very different ways, the pictures will require substantial post-processing before they can be used for anything. Before incorporating the material into our investigation, we carefully prepared it.

Many steps must be completed before the FDG-PET images from the initial baseline can be appropriately analyzed. This is the point at which the analysis may begin. These technologies are used to generate PET data, which can then be efficiently communicated via a variety of channels. If all these processes are accomplished in the correct order, the intended outcome should be realized. Following the injection, the patient will undergo six FDG-PET scans, each lasting five minutes; the first scan will begin between 30 and 60 min later. The following frames will be co-registered with the initial recovered frame to create a time-varying co-registered image. This feature was designed with the purpose of providing a dynamic and current perspective of the patient in mind. To prepare for the procedure, the various frames are co-registered before the process that will merge the distinct frames into one. The objective is to minimize the influence of patient movements on the examination's findings. Using the data from the co-registration research, we do the computation described into determining the average of the six unique photos. The image is then transformed

into a grid of voxels with the following dimensions: 160 by 160 by 96, with each voxel having a 1.5 mm side length. This grid will eventually act as the framework for the finished piece of art. This has been done to guarantee that the move goes off without a hitch. The adjustments have had an impact on the structure's front and rear commissures. To assure the accuracy of our results, we closely adhere to the guidelines indicated below. The result is then overlaid using a mask designed specifically for the photo's subject to preserve uniform brightness throughout. If you adopt this method, the individual voxels that make up the mask will have an average value of 1, adding an extra layer of protection. The third stage of the technique, which is specifically designed for the scanner, comprises applying the image in its normalized and filtered form. At half the maximum, this approach generates a picture with an isotropic resolution of up to 8 mm full width. This is done to offer the reader a consistent visual experience from paragraph to paragraph.

CNN's current success may be attributed to the network's wide range of services in the field of medical image classification. Convolutional neural networks in two dimensions (2D) are utilized in techniques. However, in directions perpendicular to the plane in which they function, they neglect the anatomical environment. This happens as we analyze the 3D medical picture slice by slice. According to research, a 3D CNN that employs 3D data as a full input has the potential to outperform a 2D CNN. This is true even if it raises both the computational complexity and the memory requirements. This remains true despite the increasing complexity of computing and the demands placed on memory. Both designs are discussed in the chapter. In the sections that follow, we will go into further detail about how we employed each of these CNNs to effectively complete AD classification tasks. Both CNNs were constructed by our team using the TensorFlow framework. The goal of this study is to compare how well the GM-PET modality works to how well other CNNs work so that conclusions can be made about how to use it.

3.2 Simple 3D CNN

The advantages and disadvantages of each of these additional options will be discussed in greater depth later. Before proceeding, we will discuss these considerations in further detail in the phrases that follow. Only four of the 3D Simple CNN's 11 layers shown in Fig. 3 are used for convolutions. The diagram demonstrates how to tell the difference between the two. Because it has fewer parameters, the 3D Simple CNN is less prone to becoming overtrained than deeper networks.

The image below shows how the central node acts as the structural support for the Conv-main block (s). The three phases that must be performed to complete this section must be always followed. Continue reading to learn more about these shows. This approach also employs an s-dimensional convolution matrix (ReLU). To safeguard the "Feature Extraction" section of our system, we have four Conv-blocks that may be arranged in one of three ways: 3, 8, or 32. (3,64). This results in a two-fold increase in total channel count at each iteration, as well as a directly

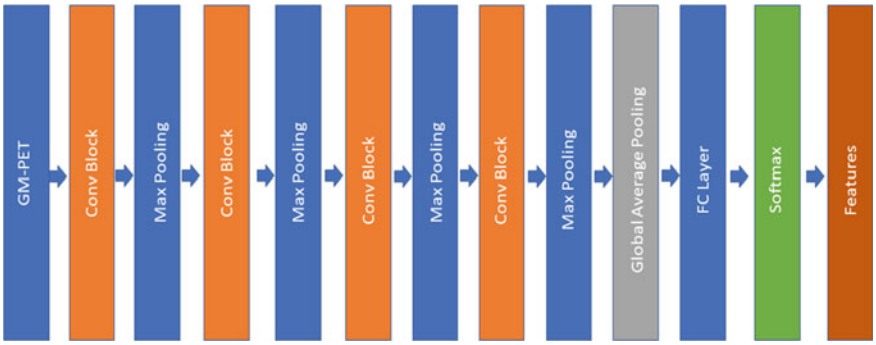

Fig. 3 A 3D basic CNN structure is used to categorize AD

proportional increase in the size of the convolution kernels (3, 3, 3). A layer called the max-pooling layer separates every pair of Conv-blocks in the third dimension. Numerous indicators indicate that this stratum will grow greatly in the future (2, 2, 2). This effect is achieved by combining the two layers. The total is larger than the sum of its parts due to the synergistic influence of these features functioning together. An FC layer and a softmax layer are then integrated into a single structure for use in AD classification. This ensures that the desired outcomes are obtained. You'll determine whether to proceed with this step based on how well the Feature Extraction phase went. It seems to reason that our approach of integrating several photos will outperform alternative methods in a head-to-head comparison. This is because to the ease with which a 3D Simple CNN may be created.

3.3 Utilizes Three-Dimensional Multi-Scale Convolutional Neural Networks

Many UNet-based networks have accomplished a variety of biological image recognition tasks effectively [36–38]. This is because a U-shaped network with skip links may be able to capture data that is both location-specific and contextually relevant more effectively. Figure 4 shows why we feel it is critical to try using a 3D Multi-Scale CNN to identify Alzheimer's disease patients. This realization was the catalyst for the initiative.

The Group Label Prediction module handles multi-scale data collection and integration, while the Feature Extraction module handles group predictions. Both aspects comprise the Multi-Scale Data Acquisition and Integration System. One of the many useful tools offered in the Feature Extraction module is the Feature Extraction subset. As shown in Fig. 4, convolutional layers in a common CNN design contain channel counts and kernel sizes of (3, 3). We were able to alleviate overfitting by making improvements such as reducing the number of channels in the convolutional layers.

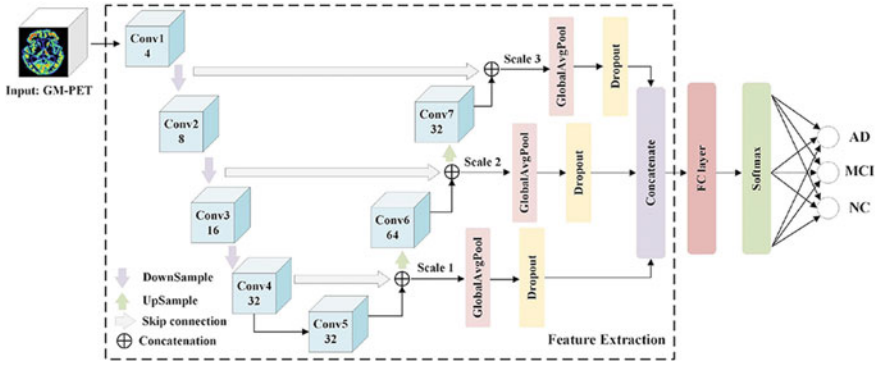

Fig. 4 The Alzheimer's disease classification using CNN

Deep layers are typically associated with pictures rich in semantic significance, whereas shallow layers are frequently associated with image characteristics rich in fine detail. The latter is generally characterized by extremely thin layers. Both types of data must be collected to make an accurate diagnosis of Alzheimer's disease, which can be accomplished through a variety of methods. There are various ways to accomplish this. After being down sampled, the outputs of the first and second convolutional layers are mixed with those of the seventh and sixth convolutional layers. This is done to guarantee that everyone understands each other. At this time, the outputs of layers 4 and 5 of the convolutional neural network are also intermingled. While processing 3D scans as inputs, this scenario employs three alternative scales to circumvent the GPU's memory access limits. The GAP layer and dropout layer results are mixed before being sent to the next classifier. The integrity of the multi-resolution components must be preserved.

Accuracy may improve because of this change. This guarantees that the multi-resolution properties of the original are retained. Finding a multi-scale feature with a high level of granularity could be very important for making a reliable test for Alzheimer's disease shown in Fig. 5.

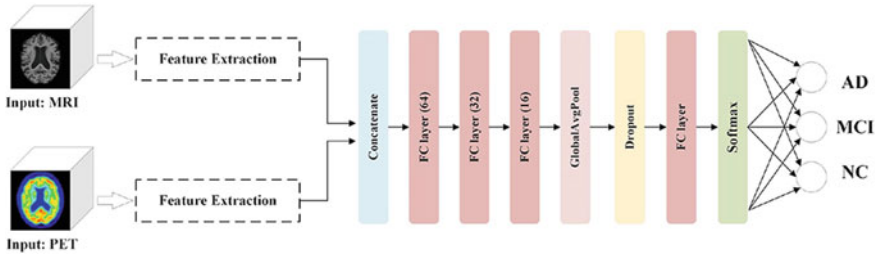

Fig. 5 Framework for Alzheimer's disease diagnosis: increased production and effectiveness

4 Results

If high-resolution 3D data were employed, the CNN training step would need more processing power than is currently accessible. To swiftly produce singleton data, it is important to first choose from the input data randomly after data purification. The Cutting Room should be your first stop before doing anything else. Figure 2 shows that, across all imaging modalities, there were significant portions of background space. We can reduce the amount of information required without negatively affecting any brain areas if we just eliminate background information that isn't beneficial in the right locations. The new MRI machine is much smaller than the previous one, measuring 176 inches by 208 inches by 176 inches as opposed to 182 inches by 218 inches by 182 inches. Both the PET and the GM-original PET's metric dimensions have been lowered from 192 by 192 by 96 to 112 by 128 by 96. To elaborate on the second point, the sampling procedure's components are as follows: Each time a slice is collected, a transverse axis built into the sample divides it in half. An MRI scan has 176 by 208 by 88 pixels, a PET scan has 112 by 128 by 48 pixels, and a GM-PET scan contains pixels that are the same size as PET pixels. These figures can be used to make estimates regarding image size. As a result, even if the resolution decreases, the sample size may increase. This adds to the network model's ongoing development and improvement.

All the networks assessed here were built with TensorFlow, a deep learning framework. These tests are intended to differentiate between normal ageing and various kinds of dementia, including Alzheimer's disease. Earlier studies separated the population into two groups due to the ease with which AD and NC could be recognized from one another. Studies must be conducted to learn more about the distinctions between the two types of information. Adam, who has a learning rate of 1e-4 at the start of the training phase, helps to keep exact weights when developing the network optimizer. Because we used tenfold cross-validation, we were able to perform extremely exact computations, making it much simpler to compare our results to those of other organizations. Comprehensive testing and analysis of a sample. Following the statistical analysis, the individuals in the sample are divided into ten distinct subgroups. After we exceed the 500-repeat barrier, we will alternate between two methods after each trial to fine-tune the learning rate. If we continue in this fashion, we will be able to accurately modify our learning rate. (1) If the rumor is confirmed, it will be a severe setback. Consider the following to be the worst-case scenario. (2) The learning rate is lowered if the validation set's accuracy does not improve after 20 repetitions of exposure to the training set. As a result, it has occurred even if the average pace of human learning has not doubled. We need to get everything organized as soon as feasible so that we can accomplish this assignment as soon as possible. If the validation loss doesn't go down after 50 training cycles, the operation will stop right away.

When assessing the overall performance of the system, the classification accuracy (ACC), sensitivity (SEN), and specificity (SP) metrics will be widely used. The standard deviation (SD) is a way, to sum up, the results of a tenfold test. The mean

standard deviation is a statistical measure of variation. Our primary goal is to assess the potential usefulness of our picture fusion approach in comparison to the current diagnostic paradigm for tasks involving AD classification. In addition to utilizing the results of previous unimodal scans such as MRI and PET as a baseline, our strategy for detecting Alzheimer's disease includes the feature fusion method. This is now a possibility because the module is compatible with both formats. This is one of the module's probable applications. The use of this data may aid in the discovery and treatment of health concerns. Overfitting can be reduced to some extent by using a GAP layer in conjunction with a dropout layer when building a model. After you achieve the correlation fusion level, it is strongly encouraged that you build three FC layers, each with a different number of nodes: 64, 32, and 16.

4.1 Identifying Differences by Comparing AD and NC

Table 2 summarizes occurrences based on whether they were categorized using a single modality, several modalities, or different network topologies, all of which contribute to their classification as AD or NC. While multi-modality systems are preferable because they can use data from both types of scans, single-modality techniques rely solely on MRI or PET scans (MRI and PET). Image fusion is one of the multi-modality strategies mentioned, although feature fusion is another option. When we compare our picture fusion strategy to the other two multimodal approaches already in use, the case for its superiority becomes stronger. Its sensitivity, at 93.33%, was only second to the gold standard. Despite having the lowest accuracy and specificity of the three approaches, the feature fusion methodology has the highest sensitivity (95.55%). When paired with a tried-and-true image fusion approach, the 3D Multi-Scale CNN achieved the greatest possible rate of classification accuracy (95.22%). Classification accuracy increased by at least 4.75 percentage points as compared to unimodal techniques, while sensitivity and specificity increased by 6.27 and 3.46% points, respectively. By combining many images into one, we beat both AD and NC in a classification task competition.

4.2 Identifying Differences Between MCI and NC Outcomes

The network topologies and classification accuracy of various MCI and NC approaches are shown in Table 3. The proposed image fusion solution beats other approaches by a substantial margin. With the help of the 3D Simple CNN, our image fusion strategy was able to obtain a best-case classification accuracy of 88.48.6.5%. This approach was the most successful overall since it was exceptionally sensitive (93.44%) and specific (82.18%). The proposed picture combining technique improved classification accuracy by 6.11%, sensitivity by 1.25%, and specificity by 11.62%, suggesting that it efficiently includes a wide range of information. When

Table 2 Outcomes of using AD and NC might change under different network installation and configuration scenarios

Network	Modalities	Accuracy	Sensitivity	Specificity
3D simple CNN	Unimodal MRI	90.90	87.42	92.08
	Unimodal PET	93.20	90.24	95.38
	Feature fusion	94.33	95.55	92.73
	Proposed method	95.22	93.33	96.15
3D multiscale CNN	Unimodal MRI	89.99	87.22	91.54
	Unimodal PET	90.47	88.17	91.92
	Feature fusion	94.77	94.44	94.61
	Proposed method	95.22	94.47	95.37

Table 3 Findings from MCI and NC study that looked at a range of different modalities and networks

Network	Modalities	Accuracy	Sensitivity	Specificity
3D simple CNN	Unimodal MRI	80.57	88.60	70.24
	Unimodal PET	73.10	73.92	71.68
	Feature fusion	99.59	93.20	70.85
	Proposed method	77.12	94.55	83.29
3D multiscale CNN	Unimodal MRI	77.12	78.61	75.38
	Unimodal PET	69.66	76.05	81.75
	Feature fusion	94.28	91.74	74.66
	Proposed method	86.11	85.72	86.71

combined with the 3D Multi-Scale CNN, our technique of fusing pictures still gave the highest rates of accuracy (85.09%), specificity (85.60%), and sensitivity (84.69%). Our technique outperformed the alternatives chosen by most experts by at least 11.33% points. On the MCI versus. NC test for classifying, the suggested method of "image fusion" was the best choice.

4.3 Variations in Observed Results Between AD and MCI

Table 4 is a summary of what was found when single and multimodal techniques, as well as supporting networks. It is designed to be used in a therapeutic environment using the information in Table 4. We were able to raise the accuracy of Alzheimer's disease diagnosis to 84.837% by combining our picture fusion approach with the usage of the 3D Simple CNN. As a result, we were able to achieve an 84% success rate. Ranked second thanks to a successful combination of high specificity (94.69%) and sensitivity (68.98%). When compared to unimodal approaches, the picture fusion

methodology improved classification accuracy by 6.53%, sensitivity by 10.83%, and specificity by 5.00%. We were able to obtain previously unattainable classification accuracy by including a 3D Multi-Scale CNN into our picture-fusing technique. Surprisingly, the overall score was 80.805.9%. This is a significant juncture in our journey. The feature fusion method was also the most precise. Based on what we found, our method was the best way to tell the difference between people with Alzheimer's and those with moderate cognitive impairment.

Classification assessments, including comparisons of AD, normal aging, and MCI in the context of normal aging, are significantly more difficult than the binary classification tasks (NC). When applied to a job requiring three classifications, the efficacy of both unimodal and feature fusion systems decreased significantly; nonetheless, our image fusion approach remained the most successful tactic across all assessment criteria. The classification accuracy of the 3D Simple CNN is 75.45%, the sensitivity is 59.518%, and the specificity is 100%. This percentage is high 85.41% to be precise.

Our photo fusion method outperformed the competition in terms of classification accuracy by at least 10.73% in terms of sensitivity and 6.2% in terms of specificity. We discovered that combining multiple photos with the 3D Multi-Scale CNN resulted in the highest possible classification accuracy of 71.52%. The specificity was 83.40%, and the sensitivity was 55.67%, according to the data. Furthermore, we observed that our image fusion approach beat its current counterparts in terms of sensitivity (4.03% points), specificity (2.37% points), and accurate classification rate (93.3%). Our strategy of employing fused pictures to address the problem involving several classes has been shown to be effective (Table 5).

Table 4 Results of multiple modalities using different networks for people with AD and MCI in the form of unit percentages

Network	Modalities	Accuracy	Sensitivity	Specificity
3D simple CNN	Unimodal MRI	83.58	57.60	88.61
	Unimodal PET	89.41	68.57	90.71
	Feature fusion	92.10	79.44	90.26
	Proposed method	95.94	79.30	95.71
3D multiscale CNN	Unimodal MRI	79.51	63.81	88.60
	Unimodal PET	84.18	72.02	80.40
	Feature fusion	91.58	64.52	96.05
	Proposed method	91.91	82.28	96.05

Table 5 The deployment of AD, MCI, or NC will have varied results depending on the approach and infrastructure employed

Network	Modalities	Accuracy	Sensitivity	Specificity
3D simple CNN	Unimodal MRI	75.11	58.21	89.19
	Unimodal PET	71.76	54.61	86.50
	Feature fusion	76.59	59.79	81.25
	Proposed method	85.65	61.42	96.52
3D multiscale CNN	Unimodal MRI	77.35	50.68	80.83
	Unimodal PET	60.48	53.94	85.09
	Feature fusion	79.26	62.75	92.14
	Proposed method	82.63	66.78	94.51

The specified ratio of one amount to another

4.4 Discussions and Evaluations of the Most Cutting-Edge Research Methodologies

The results of task-specific classification were examined using the given image fusion approach, and the findings were compared to those obtained utilizing cutting-edge multimodal methodologies (Table 6). Our method, which combines image fusion with a 3D basic CNN, surpassed every existing multimodal diagnostic tool on every test currently used to detect Alzheimer's disease. The results corroborated this. Although the preparation for our technique to fuse multimodal images is significant, the significant reduction in network parameters gained as a direct result more than compensates for the effort. In contrast to the original collection of pictures gathered in diverse ways, the classification network gets a single unified image. This photo collection has taken the place of the previous one. We can greatly cut the amount of time we need to spend using this strategy. When compared to other techniques developed for the same objective, the current method for picture fusion does not result in a considerable increase in the amount of processing complexity or memory required.

4.5 Conceptualization in Three and Four Dimensions

As shown in Fig. 6, we investigated the source images and related attributes over a wide range of modalities and subject groups to demonstrate the efficacy of our image fusion technique. This was done to compare our findings to those obtained from other studies that employed different approaches. The images to the left of each cell provide a new perspective on the subject by lighting it from various angles. When MRI and PET brain slices were examined, we discovered that the Alzheimer's patient had the lowest metabolic rate and the greatest loss of brain tissue. This was

Table 6 The following are some areas where our classifiers outperform those already in use: AD, MCI, and NC diagnoses

Method	AD versus NC	MCI versus NC	AD versus MCI	AD versus CI versus NC
Multi-modality classification	92.5	93.2	–	64.80
Robust deep model for improved classification	92.5	88.5	82.2	–
Multi-modal classification	92.9	80.6	–	71.3
Multimodal and multiscale deep neural networks	85.60	96.48	–	–
Multi-modality cascaded convolutional neural networks	94.37	85.45	–	–
Multi-modal AD classification	90.13	93.64	–	–
Hypergraph-based multi-task feature selection	93.62	91.1	–	–
Multimodal data analysis	82.10	93.64	–	–
Proposed method	95.22	90.59	–	–

evident when juxtaposing the two scenarios. Given the two options, this was an astute comment. The capacity of GM-PET to totally replace MRI and PET for measuring metabolic levels may be available soon. PET and MRI are the current gold standards for identifying brain atrophy. Because the final image only showed the GM area, GM-PET scans revealed no artefacts in the surrounding brain tissue, particularly near the skull. We believe our technique for image fusion is better than others since it takes advantage of the richness of information included in the photos.

Figure 6 depicts some of the imaging modalities available for the diagnosis of neurodegenerative diseases. These disorders include dementia, mild cognitive impairment, and other comparable difficulties. The photos on the right of each of the nine cells labeled with the letters "A" through "I" show the Grad-CAM findings for each of the nine separate slicing's across the subjects shown in the images of the cells. The contour regions of interest are shown in red in the 3D Grad-CAM output, while the metabolic characteristic zones are shown in yellow. Depending on their metabolic features, the areas of interest are indicated with a green circle on both the MRI and the gamma-metabolic positron emission tomography (GM-PET) images. The areas of interest are readily visible in both pictures. A green circle has been

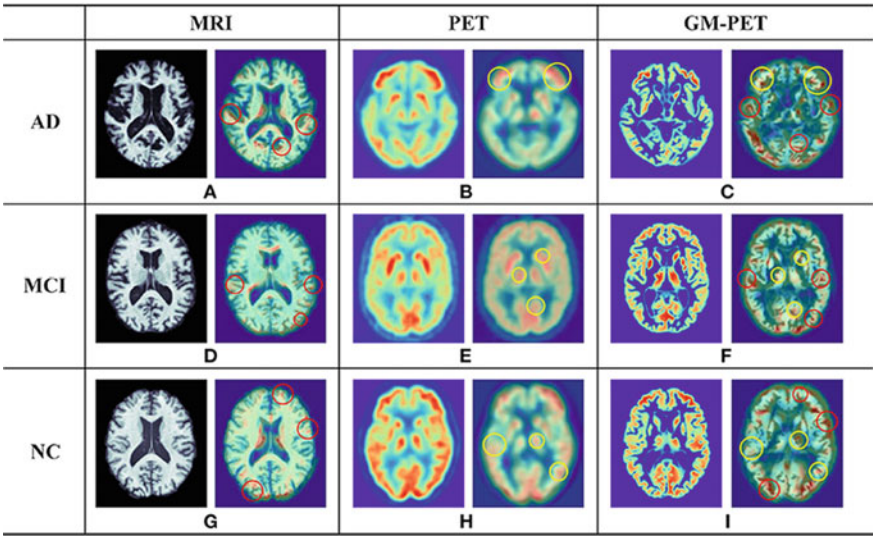

Fig. 6 Imaging modalities

drawn around these locations to draw attention to them. In theory, two circles might exist simultaneously in the same location and time.

One of the most significant and critical variables was determining whether the multimodal GM-PET data was sufficient for CNN's feature extraction module. Figure 6 shows how we were able to successfully incorporate 3D Grad-CAM technology into the 3D environment. The images in the cells on the far right depict the fundamental CNN's second convolutional layer. The highlighted image CAMs earn a higher relevance score from the convolutional layer. The red circles reflect several notable traits found in the MRI images. The shape of the item and the degree of surface roughness along its edges were the key concerns in the evaluation. According to the PET slice analysis results, there was a strong correlation between the areas of interest and those with greater metabolic rates. The issue areas were marked by yellow circles. The convolutional layer of the GM-PET model was expected to merge the contour and metabolic data to generate a single score. For example, the versatility of the GM-PET modality would allow for the use of a broader variety of Alzheimer's disease diagnostic criteria. This observation might be due to the adaptability of the medium in question.

This is done so that multimodal data can highlight more subtle illness implications. To create a fused GM-PET modality, the proposed fusion technique combines data from FDG-PET scans conducted in the same imaging region with data from brain MRI scans that identify GM tissue. This technique is carried out to make the modality accessible. Because the GM tissue region has been determined to be the most helpful indication of the existence of the illness in prior research, diagnostic techniques for Alzheimer's disease often focus on this location (10, 11, 45). (10, 11, 45). Figure 2

shows a GM-PET image demonstrating the success of this fusion process. As seen in Fig. 2, this picture combines structural information from the MRI with metabolic information from the subject's brain PET scan. View the graphic to learn more about how the combined images operate. To decide, compare the two images side by side and examine the differences. To put it another way, you will be able to independently verify these outcomes. We were able to observe that this was the case. The use of software enabled the identification of this significant advance. When this happens, it means that GM-PET may be able to provide more accurate modality information that may be used for categorizing. Furthermore, when it comes to addressing the challenge of aligning multiple properties seen in multimodal pictures, our image fusion technique may be more successful than approaches based on multimodal feature learning. This is since our method blends the two learning theories. This happens during the user's system enrolment and after registration is finished.

To complete each of these assignments (AD versus MCI versus NC), you must select if a certain item falls into the yes or no group (AD versus MCI versus NC). To avoid this difficulty, we suggested employing a three-dimensional multi-scale convolutional neural network (CNN) that considers the dimensions of both the size and placement of the features. During the development process, the following checks were made to ensure that none of these networks inappropriately fitted their training data. As an initial step, the total number of convolutional layers must be lowered. Following that, we will try a convolutional layer with fewer channels. Add GAP and dropout layers as a last step to eliminate any remaining signs of noise. A single-input network is also employed in the proposed Alzheimer's disease diagnosis paradigm. Feature fusion techniques, on the other hand, take advantage of multiple-input networks. To do this, our image fusion technology integrates data from many imaging techniques into a single, comprehensive picture. It is the precise circumstance because of the aforementioned factors. Furthermore, our image-fusing technique may be successful in drastically reducing the overall number of CNN parameters. It may be a direct result of it.

To assess the success of the approach we provided for merging images, we ran several tests and studies on it. This is since multimodal approaches use the proposed image fusion strategy for feature fusion. This is the root of the problem. This is because when the findings of different multimodal research approaches are combined, a plethora of data is created. When we were given a difficult task that included three distinct categories, our approach for integrating photographs performed considerably better than the other method for merging characteristics. This indicates not just the overall success of our picture fusion technique, but also how effectively it adapts to the many categorization networks that we subject it to. This is since independent investigations undertaken by CNN and HLN both yielded the same results. Furthermore, when compared to previous advances in the field of multimodal learning-based systems, our solution for image integration performed significantly better. However, there were times when it fell short of expectations in terms of sensitivity and specificity. Even though the proposed approach of photo fusion was frequently used to produce the best results, it did not work on a few occasions.

5 Conclusion

In this section, we'll investigate GM-PET imaging, a hybrid imaging technology that combines MRI and PET scans to diagnose Alzheimer's disease. Our strategy mainly relies on combining many photos. Aside from structural imaging, the GM-PET approach may one day be able to offer information on how the brain operates. Furthermore, the mode considerably reduces the amount of visual noise, making it much simpler for the viewer to focus on the important aspects of the image. Using cutting-edge 3D Grad-CAM technology, we were given a birds-eye perspective of the CNN broadcast area. It is hard to say whether this effort was successful. It was anticipated that by doing so, the study's findings would eventually be incorporated into routine therapy practices. According to the findings of our study, our technique of image fusion surpasses both the unimodal and feature fusion approaches, demonstrating that the recently found approach of picture fusion is superior to the other after comprehensive testing. As a result, our picture fusion technique is not only a very successful way of performing AD classification tasks, but it is also extremely easy.

References

1. Song, J. et al.: An effective multimodal image fusion method using MRI and PET for Alzheimer's disease diagnosis. Frontiers Digital Health, **3** (2021)
2. Carrion, C., Folkvord, F., Anastasiadou, D., Aymerich, M.: Cognitive therapy for dementia patients: a systematic review. Dement Geriatr Cogn Disord. **46**, 1–26 (2018). https://doi.org/10.1159/000490851
3. Association, A.: Alzheimer's disease facts and figures. Alzheimers Dement. **16**, 391–460 (2020). https://doi.org/10.1002/alz.12068
4. Theofilas, P., Ehrenberg, A.J., Nguy, A., Thackrey, J.M., Dunlop, S., Mejia, M.B., et al.: Probing the correlation of neuronal loss, neurofibrillary tangles, and cell death markers across the Alzheimer's disease Braak stages: a quantitative study in humans. Neurobiol Aging. **61**, 1–12 (2018). https://doi.org/10.1016/j.neurobiolaging.2017.09.007
5. Wang, C., Saar, V., Leung, K.L., Chen, L., Wong, G.: Human amyloid β peptide and tau co-expression impairs behavior and causes specific gene expression changes in Caenorhabditis elegans. Neurobiol Dis. **109**, 88–101 (2018). https://doi.org/10.1016/j.nbd.2017.10.003
6. Dai, Z.: Applications, opportunities and challenges of molecular probes in the diagnosis and treatment of major diseases. Chin Sci Bull. **62**, 25–35 (2017). https://doi.org/10.1360/N972016-00405
7. Sakalle, A., Tomar, P., Bhardwaj, H., Alim, M.: A modified LSTM framework for analyzing COVID-19 effect on emotion and mental health during pandemic using the EEG signals. J. Healthcare Eng., (2022)
8. Sakalle, A., Tomar, P., Bhardwaj, H., Iqbal, A., Sakalle, M., Bhardwaj, A., Ibrahim, W.: Genetic programming-based feature selection for emotion classification using EEG signal. J. Healthcare Eng., (2022)
9. Liu, M., Zhang, D., Shen, D.: Ensemble sparse classification of Alzheimer's disease. Neuroimage **60**, 1106–1116 (2012). https://doi.org/10.1016/j.neuroimage.2012.01.055
10. Suk, H.I., Lee, S.W., Shen, D.: Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. Neuroimage **101**, 569–582 (2014). https://doi.org/10.1016/j.neuroimage.2014.06.077

11. Zhu, Q., Yuan, N., Huang, J., Hao, X., Zhang, D.: Multi-modal AD classification via self-paced latent correlation analysis. Neurocomputing **355**, 143–154 (2019). https://doi.org/10.1016/j.neucom.2019.04.066
12. Farooq, A., Anwar, S., Awais, M., Rehman, S.: A deep CNN based multi-class classification of Alzheimer's disease using MRI. In: Proceedings of the International Conference on Imaging Systems and Technique, pp. 1–6. Beijing, IEEE (2017). https://doi.org/10.1109/IST.2017.8261460
13. Ge, C., Qu, Q., Gu, I.Y.H., Jakola, A.S.: Multi-stream multi-scale deep convolutional networks for Alzheimer's disease detection using MR images. Neurocomputing **350**, 60–69 (2019). https://doi.org/10.1016/j.neucom.2019.04.023
14. Kashyap, R.: Big data analytics challenges and solutions. Big Data Anal. Intell. Healthcare Manage., 19–41 (2019). https://doi.org/10.1016/b978-0-12-818146-1.00002-7 [Accessed 30 Aug 2022]
15. Tiwari, S., Gupta, R., Kashyap, R.: To enhance web response time using agglomerative clustering technique for web navigation recommendation. Adv. Intell. Syst. Comput., 659–672 (2018). https://doi.org/10.1007/978-981-10-8055-5_59 [Accessed 30 Aug 2022]
16. Kashyap, R.: Machine learning for internet of things. Adv. Wireless Technol. Telecommun., 57–83 (2019). https://doi.org/10.4018/978-1-5225-7458-3.ch003 [Accessed 30 Aug 2022]
17. Kashyap, R.: Object boundary detection through robust active contour based method with global information. Int. J. Image Mining **3**(1), 22 (2018). https://doi.org/10.1504/ijim.2018.10014063 [Accessed 30 Aug 2022]
18. Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D.: Multimodal classification of Alzheimer's disease and mild cognitive impairment. Neuroimage **55**, 856–867 (2011). https://doi.org/10.1016/j.neuroimage.2011.01.008
19. Li, Y., Meng, F., Shi, J.: Learning using privileged information improves neuroimaging-based CAD of Alzheimer's disease: a comparative study. Med Biol Eng Comput. **57**, 1605–1616 (2019). https://doi.org/10.1007/s11517-019-01974-3
20. Bi, X.A., Hu, X., Wu, H., Wang, Y.: Multimodal data analysis of Alzheimer's disease based on clustering evolutionary random forest. IEEE J. Biomed. Health Inform. **24**, 2973–2983 (2020). https://doi.org/10.1109/JBHI.2020.2973324
21. Li, F., Tran, L., Thung, K.H., Ji, S., Shen, D., Li, J.: A robust deep model for improved classification of AD/MCI patients. IEEE J. Biomed. Health Inform. **19**, 1610–1616 (2015). https://doi.org/10.1109/JBHI.2015.2429556
22. Tong, T., Gray, K., Gao, Q., Chen, L., Rueckert, D.: Multi-modal classification of Alzheimer's disease using nonlinear graph fusion. Pattern Recogn. **63**, 171–181 (2017). https://doi.org/10.1016/j.patcog.2016.10.009
23. Shi, J., Zheng, X., Li, Y., Zhang, Q., Ying, S.: Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease. IEEE J. Biomed. Health Inform. **22**, 173–183 (2018). https://doi.org/10.1109/JBHI.2017.2655720
24. Lu, D., Popuri, K., Ding, G.W., Balachandar, R., Beg, M.F., Weiner, M., et al.: Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images. Sci Rep. **8**, 1–13 (2018). https://doi.org/10.1038/s41598-018-22871-z
25. Nair, R., Bhagat, A.: An introduction to clustering algorithms in big data. Encyclopedia of Information Science and Technology, Fifth Edition, pp. 559–576, (2021). https://doi.org/10.4018/978-1-7998-3479-3.ch040 [Accessed 14 Jun 2022]
26. Sharma, A., Singh, K., Koundal, D.: A novel fusion based convolutional neural network approach for classification of COVID-19 from chest X-ray images. Biomed. Signal Process. Control, **77**, 103778.26 (2022).; Nair, R., Soni, M., Bajpai, B., Dhiman, G., Sagayam, K.: Predicting the death rate around the world due to COVID-19 using regression analysis. Int. J. Swarm Intell. Res., **13**(2), 1–13 (2022). https://doi.org/10.4018/ijsir.287545
27. Agrawal, M., Kumar Shukla, P., Nair, R., Nayyar, A., Masud, M.: Stock prediction based on technical indicators using deep learning model. Comput. Mater. Continua, **70**(1), 287–304 (2022). https://doi.org/10.32604/cmc.2022.014637

28. Liu, M., Zhang, J., Yap, P.T., Shen, D.: View-aligned hypergraph learning for Alzheimer's disease diagnosis with incomplete multi-modality data. Med Image Anal. **36**, 123–134 (2017). https://doi.org/10.1016/j.media.2016.11.002
29. Bartos, A., Gregus, D., Ibrahim, I., Tintěra, J.: Brain volumes and their ratios in Alzheimer's disease on magnetic resonance imaging segmented using Freesurfer 6.0. Psychiatry Res Neuroimaging, **287**, 70–4 (2019). https://doi.org/10.1016/j.pscychresns.2019.01.014
30. Bhat, S., Koundal, D.: Multi-focus image fusion using neutrosophic based wavelet transform. Appl. Soft Comput. **106**, 107307 (2021)
31. Jenkinson, M., Bannister, P., Brady, M., Smith, S.: Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage **17**, 825–841 (2002). https://doi.org/10.1006/nimg.2002.1132
32. Jenkinson, M., Smith, S.: A global optimisation method for robust affine registration of brain images. Med. Image Anal. **5**, 143–156 (2001). https://doi.org/10.1016/S1361-8415(01)00036-6
33. Koundal, D.: Texture-based image segmentation using neutrosophic clustering. IET Image Proc. **11**(8), 640–645 (2017)
34. Malhotra, P., Gupta, S., Koundal, D., Zaguia, A., Enbeyle, W.: Deep neural networks for medical image segmentation. J. Healthcare Eng., (2022)
35. Nair, R., Alhudhaif, A., Koundal, D., Doewes, R.I., Sharma, P.: Deep learning-based COVID-19 detection system using pulmonary CT scans. Turk. J. Electr. Eng. Comput. Sci. **29**(8), 2716–2727 (2021)
36. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 234–41. Cham, Springer (2015). https://doi.org/10.1007/978-3-319-24574-4_28
37. Li, X., Chen, H., Qi, X., Dou, Q., Fu, C.W., Heng, P.A.: H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. IEEE Trans Med Imaging. **37**, 2663–2674 (2018). https://doi.org/10.1109/TMI.2018.2845918
38. Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods. **18**, 203–211 (2020). https://doi.org/10.1038/s41592-020-01008-z