

Classification of Alzheimer's disease stages from magnetic resonance images using deep learning

Alejandro Mora-Rubio¹, Mario Alejandro Bravo-Ortíz¹, Sebastián Quiñones Arredondo¹, Jose Manuel Saborit Torres², Gonzalo A. Ruz^{3,4,5} and Reinel Tabares-Soto^{1,5,6}

¹ Department of Electronics and Automation, Universidad Autónoma de Manizales, Manizales, Caldas, Colombia

² Unidad Mixta de Imagen Biomédica FISABIO-CIPF, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana, Valencia, Spain

³ Center of Applied Ecology and Sustainability (CAPES), Santiago, Chile

⁴ Data Observatory Foundation, Santiago, Chile

⁵ Facultad de Ingeniería y Ciencias, Universidad Asdolfo Ibáñez, Santiago, Chile

⁶ Department of Systems and Informatics, Universidad de Caldas, Manizales, Caldas, Colombia

ABSTRACT

Alzheimer's disease (AD) is a progressive type of dementia characterized by loss of memory and other cognitive abilities, including speech. Since AD is a progressive disease, detection in the early stages is essential for the appropriate care of the patient throughout its development, going from asymptomatic to a stage known as mild cognitive impairment (MCI), and then progressing to dementia and severe dementia; is worth mentioning that everyone suffers from cognitive impairment to some degree as we age, but the relevant task here is to identify which people are most likely to develop AD. Along with cognitive tests, evaluation of the brain morphology is the primary tool for AD diagnosis, where atrophy and loss of volume of the frontotemporal lobe are common features in patients who suffer from the disease. Regarding medical imaging techniques, magnetic resonance imaging (MRI) scans are one of the methods used by specialists to assess brain morphology. Recently, with the rise of deep learning (DL) and its successful implementation in medical imaging applications, it is of growing interest in the research community to develop computer-aided diagnosis systems that can help physicians to detect this disease, especially in the early stages where macroscopic changes are not so easily identified. This article presents a DL-based approach to classifying MRI scans in the different stages of AD, using a curated set of images from Alzheimer's Disease Neuroimaging Initiative and Open Access Series of Imaging Studies databases. Our methodology involves image pre-processing using FreeSurfer, spatial data-augmentation operations, such as rotation, flip, and random zoom during training, and state-of-the-art 3D convolutional neural networks such as EfficientNet, DenseNet, and a custom siamese network, as well as the relatively new approach of vision transformer architecture. With this approach, the best detection percentage among all four architectures was around 89% for AD vs. Control, 80% for Late MCI vs. Control, 66% for MCI vs. Control, and 67% for Early MCI vs. Control.

Submitted 20 February 2023

Accepted 21 June 2023

Published 24 August 2023

Corresponding author

Alejandro Mora-Rubio,
alejandro.morar@autonoma.edu.co

Academic editor

Ahmed Elazab

Additional Information and
Declarations can be found on
page 12

DOI 10.7717/peerj-cs.1490

© Copyright

2023 Mora-Rubio et al.

Distributed under

Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Bioinformatics, Artificial Intelligence, Computer Vision, Data Mining and Machine Learning, Neural Networks

Keywords Computer-aided diagnosis, Convolutional neural networks, Digital image processing, Supervised learning

INTRODUCTION

Alzheimer's disease (AD) is a progressive neurological disorder that primarily affects memory and cognition. Other symptoms, such as language disorders, hallucinations, and in some cases, seizures or parkinsonian features, may also be present (*Kumar et al., 2022*). One theory regarding the causes of AD links the β -amyloid peptide and mutations in its precursor proteins to weakened brain structures, which in turn impact various cognitive processes in the patient's brain. This theory is supported and assessed through two key markers: amyloid plaques and neurofibrillary tangles (*Ballard et al., 2011*).

Macroscopically, changes in brain morphology play a crucial role in the diagnosis of AD, with frontotemporal lobe atrophy being a common feature. This loss of brain volume in the region results in diminished function of the amygdala and hippocampus, which are responsible for memory processes. Medical imaging techniques, such as magnetic resonance imaging (MRI) and positron emission tomography (PET) scans, are utilized by specialists to evaluate the morphological changes in the brain caused by AD (*van Oostveen & de Lange, 2021*). In addition to brain scans, using cognitive tests, like the Mini-Mental State Examination (MMSE), that consider specific physical, psychological, and social skills is essential for a reliable diagnosis of the disease.

Regarding AD progression, as the disease advances the symptoms change, and a series of consecutive stages can be identified, following the convention used by Alzheimer's Disease Neuroimaging Initiative (ADNI, <https://adni.loni.usc.edu>). A brief description of each stage is presented below.

- Early mild cognitive impairment (EMCI): the affected person starts experiencing episodes of memory loss with words or the location of household items, nevertheless he can function independently and participate in social activities.
- Mild cognitive impairment (MCI): the affected people are susceptible to forgetting recent occurrences, becoming disoriented in their homes, and having difficulties with communication. This is often the longest stage of AD, lasting up to 4 years.
- Late mild cognitive impairment (LMCI): at this stage of the disease, patients may need help with daily tasks, facing increasing difficulty communicating and controlling their movements. Their memory and cognitive skills worsen, and changes in behavior and personality may occur.
- Alzheimer's disease (AD): as the disease progresses, the affected person requires increasing levels of attention and aid with daily tasks. This stage is characterized by growing unawareness of time and space, problems recognizing family and close friends, difficulty walking, and behavioral disturbances that may even lead to aggression.

Furthermore, with the rise of deep learning (DL) and its successful implementation in medical imaging applications, researchers have been working on automatic systems able to learn from MRI and PET scans to support the diagnosis of AD (*LaMontagne et al., 2019; Petersen et al., 2010; Zeng et al., 2018*). This kind of decision support system, created with artificial intelligence at its core, has been used successfully in different situations in the healthcare field improving diagnostic performance, reducing medical errors, and providing a better quality of service in developing countries (*Currie & Rohren, 2021; Porumb et al., 2020*). In general, DL techniques are developed to find patterns in the data and extract the relevant features for optimal classification (*Sharma, Sharma & Sharma, 2016; Chan et al., 2020; Bravo-Ortiz et al., 2021*), making them suitable for tasks with large volumes of labeled data based on human experiences, such as classifying the stages of AD using brain imaging (*Liu et al., 2014; Chitradevi & Prabha, 2020; Zeng, Li & Peng, 2021*). In particular, for AD diagnosis, most of the research papers consider a 3D convolutional neural network (CNN) approach to take advantage of the multidimensional features in brain scans (*Payan & Montana, 2015; Huang et al., 2019*). Alternatively, the work of *Cheng & Liu (2017)* proposes a 2D CNN approach, combined with recurrent neural networks to link the features of the three-dimensional scans. In the development of such systems, image pre-processing is key to achieving the best possible results, and in that regard, the main tool is FreeSurfer, which allows operations like skull suppression, bias field correction, anatomical registration, segmentation, reconstruction, and parcellation of the cortical surface (*Fischl, 2012*).

Acknowledging the growing number of patients with AD, and the importance of improving their quality of life, it is essential to develop more effective and efficient tools to diagnose the disease. This article presents a DL-based approach to the classification of MRI scans in the different stages of AD, using the EfficientNet, DenseNet, a custom Siamese architecture, and a Vision Transformer, the detection percentage among all four architectures was around 89% for AD vs. Control, 80% for Late MCI vs. Control, 66% for MCI vs. Control, and 67% for Early MCI vs. Control, and it is organized as follows: the Methods section describes the dataset, CNN architectures, and training set-up; the Results section presents the experiments and classification results obtained; the Discussion section reviews the results and their impact on the field; and finally the Conclusion of this work.

METHODS

Dataset

For this task, open-access imaging databases ADNI (<https://adni.loni.usc.edu>) and OASIS (*LaMontagne et al., 2019*) are the most commonly used sources of MRI scans to train and evaluate DL models. The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD.

For the purpose of this article, a dataset comprising images from both databases was used. There are a total of 2,559 images from 1,126 subjects, where the gender distribution is

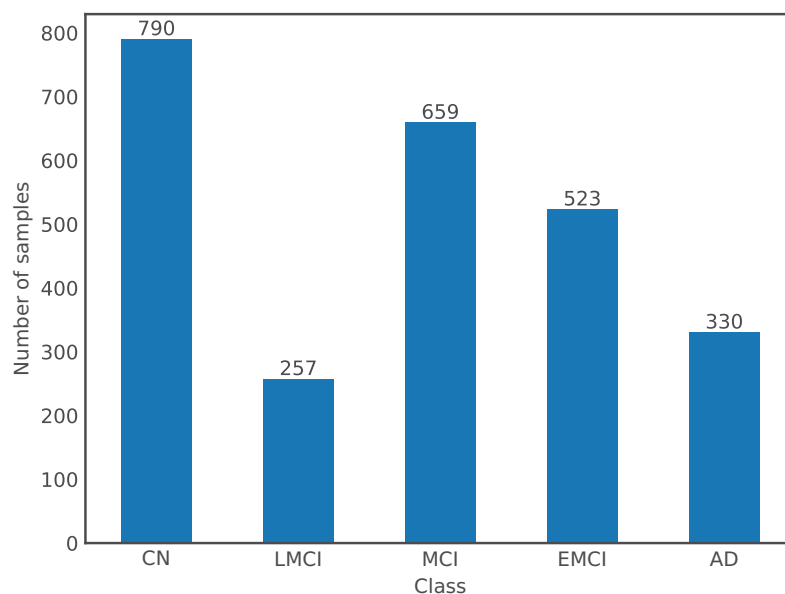


Figure 1 Class distribution in the dataset.

Full-size DOI: 10.7717/peerj-cs.1490/fig-1

relatively even with 54% of images from males and the remaining 46% from females; and ages ranging from 55 to 96 years old with a mean of 75 years old. Following the labels used by ADNI, there are a total of five different classes in the dataset: Cognitively Normal or Control (CN), EMCI, MCI, LMCI, and AD; OASIS labels, which are expressed numerically according to the Clinical Dementia Rating, were changed to ADNI format: label 0 corresponds to CN, (0,1) to MCI, and [1,3] to AD. [Figure 1](#) shows the number of samples per class in the dataset.

In order to reduce data-induced bias in the model, the images were partitioned into the training, validation, and test sets on a subject basis, obtaining a training set with 683 patients (1,569 images), a validation set with 222 patients (513 images), and a test set with 221 patients (477 images), where every session from a patient is contained in a single set. Additionally, a similar age distribution within each partition was also maintained (see [Fig. 2](#)).

Deep learning architectures

Siamese 3D

In this type of architecture, the data flows through parallel routes at the same time, which are then combined to make a final prediction. In this case, each path is composed of sets of 3D convolutional layers, 3D batch normalization, and average pooling operations, in the end, the feature maps are flattened, concatenated, and then passed to the fully connected layers for the final prediction. It is worth mentioning, that this is a custom architecture, designed for the classification of AD stages using 3D MRI scans ([Saborit Torres, 2019](#)). In the proposed configuration, this model has a total of 392,822 trainable parameters for the binary classification tasks, and 392,923 for the 3-way classification tasks. [Figure 3](#) shows the model architecture.

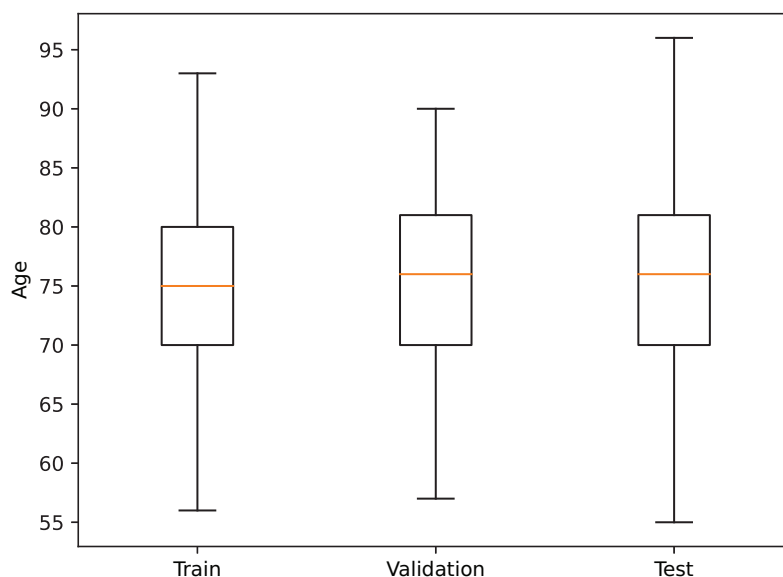


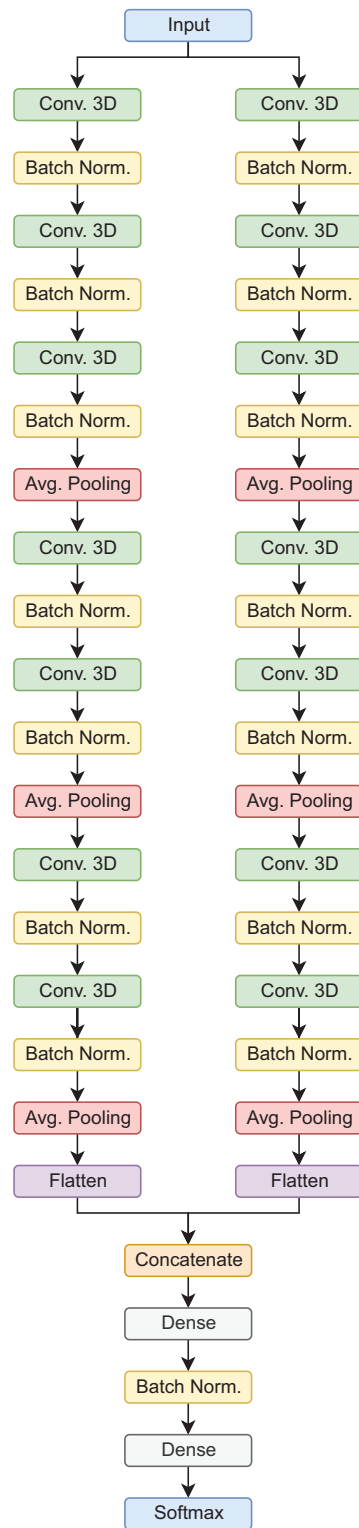
Figure 2 Age distribution in each dataset partition. Full-size  DOI: 10.7717/peerj-cs.1490/fig-2

Vision transformer

Vision transformers (ViT) ([Dosovitskiy et al., 2020](#)) are a groundbreaking approach to computer vision that has recently gained popularity due to their impressive performance on various tasks. Unlike traditional CNNs, which rely on hand-engineered feature maps, vision transformers use self-attention mechanisms to learn features directly from the input data dynamically. This approach has several advantages, including increased flexibility and scalability and improved performance on object detection, segmentation, and classification tasks. Additionally, ViT ([Khan et al., 2022](#)) has shown promise in domains beyond computer vision, including natural language processing and speech recognition. One of the critical innovations of ViT is self-attention, which allows the network to attend to different parts of the input data and dynamically adjust the importance of each feature. ViT ([Khan et al., 2022](#)) contrasts traditional CNNs, which rely on fixed, pre-defined feature maps that may only be optimal for some tasks. Another advantage of vision transformers is their ability to learn from large-scale datasets. Pre-training on massive datasets such as ImageNet has been shown to significantly improve the performance of ViT on downstream tasks, demonstrating the importance of unsupervised learning in deep learning models ([Dosovitskiy et al., 2020](#)). Despite their impressive performance, ViT is still a relatively new approach to computer vision, and there is ongoing research into how best to design and optimize these models. In the proposed configuration, this model has a total of 1,853,698 trainable parameters.

DenseNet

DenseNet was originally proposed as a general-purpose image classification architecture, where the dense convolutional blocks improved significantly the vanishing gradient problem experienced by very deep neural networks. Each of these dense convolutional blocks refers to a set of convolutional layers that are all connected between them, this

**Figure 3** Siamese 3D model architecture.Full-size  DOI: 10.7717/peerj-cs.1490/fig-3

means that “the feature-maps of all preceding layers are used as inputs, and its own feature-maps are used as inputs into all subsequent layers” (Huang et al., 2016). By concatenating several of these blocks, the DenseNet architecture and all its variants are created. This network has been successfully used in different biomedical imaging classification problems, such as the detection of COVID-19 in computed tomography scans avoiding manual delineation of the lesions (Xiao et al., 2022), the detection of AD from brain scans based on features extracted from the hippocampal region (Cui & Liu, 2019), and the detection of skin cancer by classifying the images of skin lesions (Villa-Pulgarin et al., 2022), which motivate its inclusion in this work. In the proposed configuration, this model has a total of 11,244,674 trainable parameters for the binary classification tasks, and 11,245,699 for the 3-way classification tasks.

EfficientNet

EfficientNet refers to a series of convolutional neural networks that scale the three dimensions of the network at the same time, achieving better classification results with fewer trainable parameters compared to other state-of-the-art architectures. The dimensions affected are depth (number of layers), width (number of channels or filters per layer), and resolution (size of the input images and feature maps). In particular, the B7 variant was used for this article, which is the biggest one as presented in the original paper (Tan & Le, 2019). As for DenseNet, this architecture has been implemented in various tasks of biomedical image processing such as the detection of diabetic retinopathy from digital fundus images (Jaiswal et al., 2021; Wang et al., 2020), the detection of tuberculosis in chest X-ray images (Munadi et al., 2020), and aging classification based on skin microstructure characteristics (Moon & Lee, 2022). In the proposed configuration, this model has a total of 68,764,114 trainable parameters for the binary classification tasks, and 68,766,675 for the 3-way classification tasks.

Training procedure

Regarding the model training set-up, the MONAI (MONAI Consortium, 2022) was used, it is based on PyTorch and allows easy implementation of input data pipelines with image intensity scaling to the [0,1] range, image resizing to $91 \times 91 \times 91$ pixels, and spatial data-augmentation operations, such as random 90° rotation, random flip, and random zoom. The latter operations aim to improve the models’ generalization ability by adding variability to the images during training. In terms of training hyperparameters, all models were trained for 250 epochs, using a weighted cross entropy loss to account for data imbalance, and the Adadelta optimizer with a learning rate of 1.0, $\rho = 0.95$, and $\varepsilon = 1 \times 10^{-7}$. All the experiments were performed using NVIDIA Quadro RTX 8000 graphic processor units.

RESULTS

This section presents the experiments performed over the previously described dataset, and the classification results achieved for each one of them. The first experiment consists of the binary classification between the Control cases and AD, which is the benchmark task in

Table 1 Classification results for two classes—control and Alzheimer’s disease. The best results for each test metric is indicated in bold italics.

Model	Train accuracy	Validation accuracy	Test accuracy	Test sensitivity	Test specificity
Siamese 3D	0.9844	0.8646	0.7448	0.6776	0.5943
ViT	0.7804	0.9203	0.8902	0.8902	0.7401
DenseNet	0.9656	0.8854	0.7448	0.6322	0.5905
EfficientNet B7	0.9672	0.8750	0.8542	0.8085	0.6835

AD detection. [Table 1](#) presents the training, validation, and test results achieved for this task, where the ViT architecture performed the best.

[Figure 4A](#) shows the receiver operator curve for the CN vs. AD task using the predictions from the EfficientNet model.

The second experiment involves the binary classification between Control cases against LMCI. [Table 2](#) presents the results for this experiment.

[Figure 4B](#) shows the receiver operator curve for the CN vs. LMCI task using the predictions from the DenseNet model.

Similarly, the third experiment consists of the classification of Control cases against MCI. [Table 3](#) presents the results for this experiment.

[Figure 4C](#) shows the receiver operator curve for the CN vs. MCI task using the predictions from the DenseNet model.

The fourth experiment consists of the classification of Control and EMCI, which is significantly more challenging since the physical effect of the disease is not as aggressive as for AD, but ultimately it is more relevant for physicians since it would help them identify potential AD patients in the early stages. [Table 4](#) presents the results for this experiment.

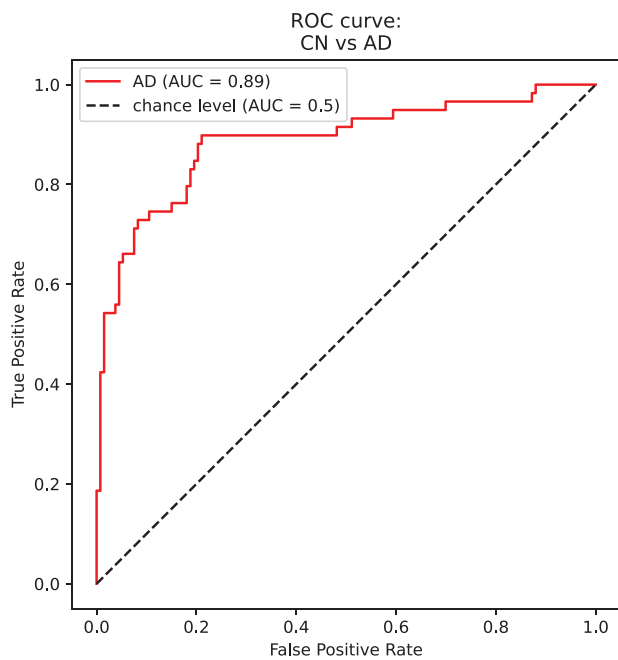
[Figure 4D](#) shows the receiver operator curve for the CN vs. EMCI task using the predictions from the Siamese 3D model.

The fifth experiment aimed to classify the early and late stages of the disease progression; the early stages were the Control and EMCI cases, and MCI, LMCI, and AD were considered as late stages. The main goal was to take advantage of the complete dataset aiming to improve classification results. [Table 5](#) presents the results for the fourth experiment.

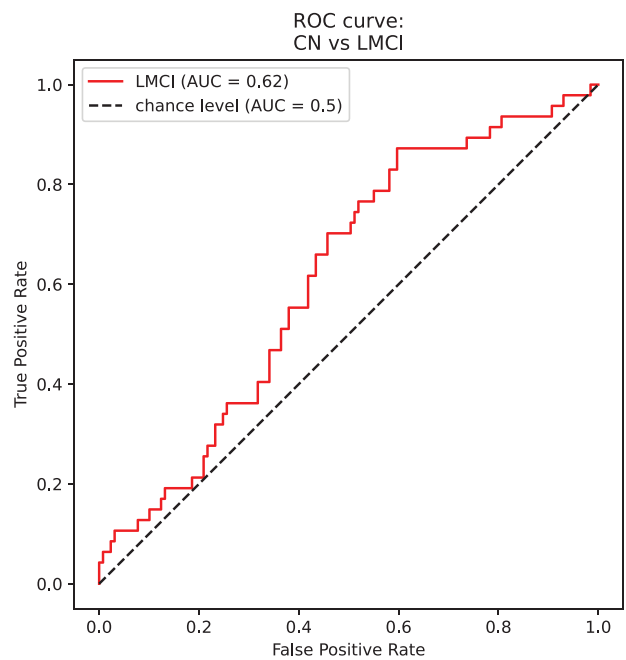
The sixth and final experiment consists of a multi-class classification problem involving Control, MCI (including early and late MCI), and AD cases. [Table 6](#) presents the results of this experiment.

DISCUSSION

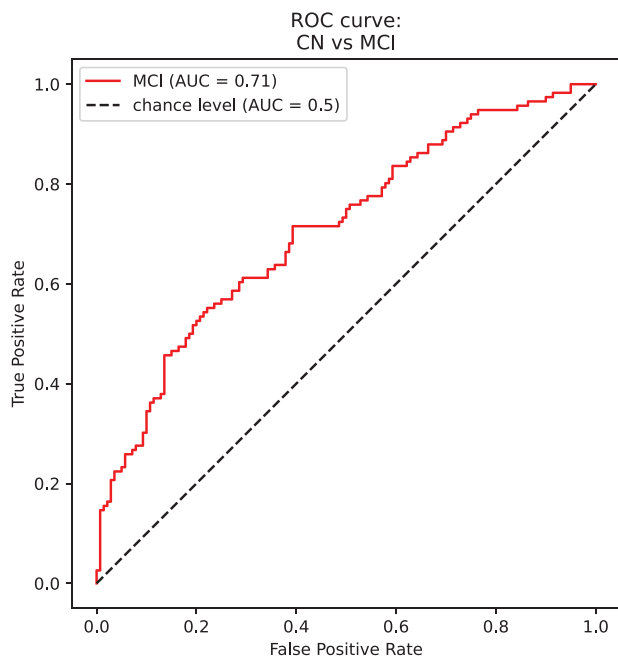
This article presents a DL-based approach to the classification of MRI scans in the different stages of AD, using a curated set of images from both ADNI and OASIS databases. There are studies from different countries around the world that suggest an increasing prevalence of dementia in older patients, and how AD can be considered one of the leading causes of it ([Rasmussen & Langerman, 2019](#)), it is important for the scientific and healthcare community to develop and improve current diagnostic techniques, in particular during its



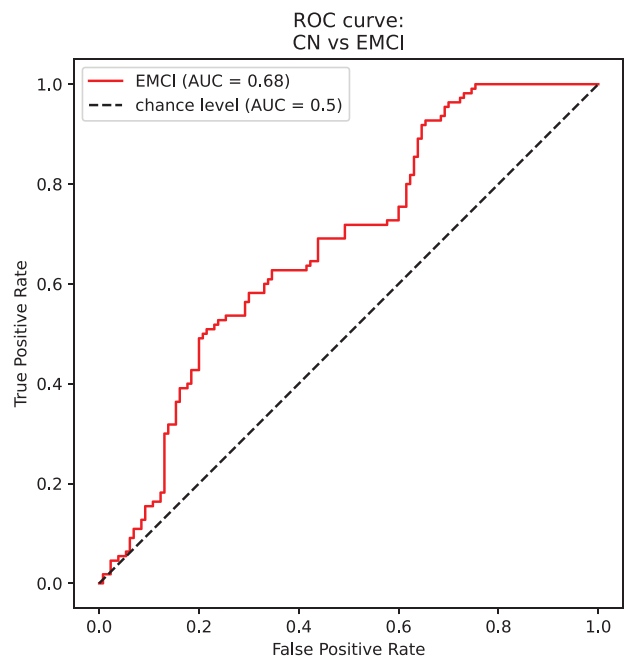
(a)



(b)



(c)



(d)

Figure 4 ROC curves. (A) Control and Alzheimer's disease. (B) Control and late mild cognitive impairment. (C) Control and mild cognitive impairment. (D) Control and early mild cognitive impairment.

Full-size  DOI: [10.7717/peerj-cs.1490/fig-4](https://doi.org/10.7717/peerj-cs.1490/fig-4)

Table 2 Classification results for two classes—control and late mild cognitive impairment. The best results for each test metric is indicated in bold italics.

Model	Train accuracy	Validation accuracy	Test accuracy	Test sensitivity	Test specificity
Siamese 3D	0.7656	0.8021	0.7578	0.5156	0.5156
ViT	0.7656	0.9070	0.8056	0.6870	0.6530
DenseNet	0.9984	0.8385	0.7031	0.6991	0.6633
EfficientNet B7	0.7375	0.8177	0.7734	0.5833	0.5833

Table 3 Classification results for two classes—control and mild cognitive impairment. The best results for each test metric is indicated in bold italics.

Model	Train accuracy	Validation accuracy	Test accuracy	Test sensitivity	Test specificity
Siamese 3D	0.8966	0.7305	0.6250	0.6017	0.6017
ViT	0.8506	0.7103	0.6432	0.5970	0.6420
DenseNet	0.9976	0.7188	0.6641	0.6449	0.6633
EfficientNet B7	0.9976	0.7383	0.6641	0.6530	0.6591

Table 4 Classification results for two classes—control and early mild cognitive impairment. The best results for each test metric is indicated in bold italics.

Model	Train accuracy	Validation accuracy	Test accuracy	Test sensitivity	Test specificity
Siamese 3D	0.7331	0.7344	0.6719	0.6749	0.6511
ViT	0.7932	0.7020	0.6320	0.6720	0.6311
DenseNet	0.9596	0.7383	0.5938	0.6094	0.5913
EfficientNet B7	0.5807	0.6523	0.5042	0.4915	0.4640

Table 5 Classification results for two classes—early and late stages. The best results for each test metric is indicated in bold italics.

Model	Train accuracy	Validation accuracy	Test accuracy	Test sensitivity	Test specificity
Siamese 3D	0.8542	0.6875	0.6473	0.6292	0.6292
ViT	0.7920	0.6990	0.6540	0.6560	0.6520
DenseNet	1.0000	0.7051	0.6540	0.6505	0.6505
EfficientNet B7	0.7760	0.7246	0.6652	0.6550	0.6550

Table 6 Classification results for three classes—control, mild cognitive impairment, and Alzheimer's disease. The best results for each test metric is indicated in bold italics.

Model	Train accuracy	Validation accuracy	Test accuracy	Test sensitivity	Test specificity
Siamese 3D	0.6250	0.5645	0.5625	0.3585	0.6664
ViT	0.6120	0.5780	0.5510	0.4260	0.6650
DenseNet	0.5964	0.5820	0.5446	0.4199	0.6861
EfficientNet B7	0.5443	0.5703	0.5893	0.4163	0.6765

early stages. Since there is yet to be a treatment that can revert the physiological effects of severe dementia, the early diagnosis of the disease is imperative for appropriate care of the patient, and a realistic projection of the following years and stages. In this regard, brain imaging and CAD systems are of high interest to researchers, where the high capacity of DL algorithms to extract relevant features from images can be leveraged to detect the subtle changes present in the early stages of the disease.

Our methodology encompasses image pre-processing using FreeSurfer, spatial data augmentation techniques such as rotation, flip, and random zoom during training, and state-of-the-art 3D CNNs, including EfficientNet, DenseNet, and a custom siamese network, as well as the relatively new approach of Vision Transformer architecture. Using this approach, the best detection percentages among all three architectures were approximately 89% for AD vs. Control, 80% for Late MCI vs. Control, 66% for MCI vs. Control, and 67% for Early MCI vs. Control. It is noteworthy that all three models involved in the experiments achieve competitive results and outperform one another in specific cases, with the Siamese 3D model's performance on the CN vs. Early MCI task and the ViT model's performance on the CN vs. AD task being particularly remarkable. Regarding these results, it is essential to acknowledge the difficulty in comparing them with similar research papers due to the numerous factors affecting model performance. If these factors are not adequately reported, reproducing the results and directly comparing them with ours becomes nearly impossible. Consequently, the dataset and selected partition are crucial for obtaining reliable results that accurately represent how the model would perform in real-life scenarios. As reported in medical tasks, a significant portion of predictive models presented in the scientific literature offer little to no value to the healthcare community due to the presence of bias in the models, which affects the accuracy of predictions (Roberts et al., 2021, Wynants et al., 2020). Bearing this in mind, the code and references to the data used during the experimentation stage of our work are available in a GitHub repository (<https://github.com/MoraRubio/alzheimer-stages-dl>), and this article describes, to the best of our ability, the materials and methods used throughout the process.

In particular, compared to the referenced works on classification using 3D CNNs, The work by Huang et al. (2019) also includes FDG-PET images, which offer relevant information for the diagnosis, but at a greater cost for the patient or healthcare system, resulting in less viable option for developing countries with lower resources, as well as, fewer images for model training and evaluation. On the other hand, the work by Payan & Montana (2015), uses an autoencoder for feature extraction instead of feeding the images directly to the CNN, which adds extra complexity for model development and deployment. In this work, only MRI images are considered, since they are more common and accessible than PET images, and the images are directly fed to the CNN architectures. As mentioned above, a direct comparison of the detection percentages is not easily performed.

Acknowledging the limited size of the involved dataset, and the number of DL methods; future work should be aimed at exploring different network architectures, such as Convolutional Vision Transformers, to enhance the detection accuracy of the different stages. Initially, this will focus on binary classification to differentiate intermediate stages

from Control and AD patients. However, the ultimate goal is to identify those MCI patients who are most likely to develop AD, particularly in the early stages, where prediction is crucial for healthcare professionals in providing optimal patient care. Moreover, assembling a larger image dataset while adhering to ethical principles (Tabares-Soto *et al.*, 2022) and aiming to minimize data-induced bias in the models can positively impact the detection capabilities and, more importantly, the generalization ability of the models, which is essential for developing viable CAD systems. Furthermore, it is vital to conduct in-depth analysis of the results using visualization or interpretation techniques, such as activation maps, occlusion sensitivity, or gradient-based heat maps. These methods could enable comparison with physical markers assessed by medical specialists or even identify new regions of interest for AD detection.

CONCLUSIONS

Early diagnosis of AD is key to providing adequate care to the patient, but existing detection techniques are not definitive enough to provide a certain diagnosis. The evaluation of brain morphology and its changes through time is one of the tools used by specialists to diagnose the disease, however, changes in the early stages are difficult to identify in plain sight, and the use of CAD systems is of great interest for this task. The results presented in this article regarding the classification of different stages of AD using brain MRI indicate that DL models do have the ability to correctly identify control/healthy patients from the ones suffering from dementia, nevertheless, there is still a lot of work to be done in order to create reliable models and test them appropriately in real-life scenarios.

ACKNOWLEDGEMENTS

Data were provided in part by OASIS-3: Longitudinal Multimodal Neuroimaging: Principal Investigators: T. Benzinger, D. Marcus, J. Morris; NIH P30 AG066444, P50 AG00561, P30 NS09857781, P01 AG026276, P01 AG003991, R01 AG043434, UL1 TR000448, R01 EB009352. AV-45 doses were provided by Avid Radiopharmaceuticals, a wholly owned subsidiary of Eli Lilly. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the Universidad Autonoma de Manizales as part of the project “Detección de COVID-19 en imágenes de rayos X usando redes neuronales convolucionales” with code 699-106, and also to the projects “CH-T1246: Oportunidades de Mercado para las Empresas de Tecnología—Compras Públicas de Algoritmos Responsables, Éticos y Transparentes”, ANID PIA/BASAL FB0002, and ANID/PIA/ANILLOS ACT210096. Data collection and sharing for this project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of

Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. There was no additional external funding received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

Universidad Autonoma de Manizales: 699-106, ANID PIA/BASAL FB0002 and ANID/PIA/ANILLOS ACT210096.

National Institutes of Health Grant: U01 AG024904.

DOD ANID PIA/BASAL FB0002.

ANID/PIA/ANILLOS ACT210096.

National Institutes of Health Grant: U01 AG024904.

DOD ADNI (Department of Defense): W81XWH-12-2-0012.

National Institute on Aging.

National Institute of Biomedical Imaging and Bioengineering.

AbbVie, Alzheimer's Association.

Alzheimer's Drug Discovery Foundation.

Araclon Biotech.

BioClinica, Inc.

Biogen.

Bristol-Myers Squibb Company.

CereSpir, Inc.

Cogstate.

Eisai Inc.

Elan Pharmaceuticals, Inc.

Eli Lilly and Company.

EuroImmun.

F. Hoffmann-La Roche Ltd.

Genentech, Inc.

Fujirebio.
GE Healthcare.
IXICO Ltd.
Janssen Alzheimer Immunotherapy Research & Development, LLC.
Johnson & Johnson Pharmaceutical Research & Development LLC.
Lumosity.
Lundbeck.
Merck & Co., Inc.
Meso Scale Diagnostics, LLC.
NeuroRx Research.
Neurotrack Technologies.
Novartis Pharmaceuticals Corporation.
Pfizer Inc.
Piramal Imaging.
Servier.
Takeda Pharmaceutical Company.
Transition Therapeutics.
The Canadian Institutes of Health Research.
Foundation for the National Institutes of Health.
Northern California Institute for Research and Education.
Alzheimer's Therapeutic Research Institute at the University of Southern California.
University of Southern California.
Avid Radiopharmaceuticals.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Alejandro Mora-Rubio conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Mario Alejandro Bravo-Ortíz performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Sebastián Quiñones Arredondo performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Jose Manuel Saborit Torres analyzed the data, prepared figures and/or tables, data collection and preparation, and approved the final draft.
- Gonzalo A. Ruz conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Reinel Tabares-Soto conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The data is available from ADNI (<https://adni.loni.usc.edu>) and OASIS (<https://www.oasis-brains.org>).

The search parameters for ADNI database are available at Zenodo: Alejandro Mora-Rubio. (2023). MoraRubio/alzheimer-stages-dl: First Release (v1.0). Zenodo. <https://doi.org/10.5281/zenodo.7855386>.

REFERENCES

- Ballard C, Gauthier S, Corbett A, Brayne C, Aarsland D, Jones E. 2011.** Alzheimer's disease. *The Lancet* **377(9770)**:1019–1031 DOI [10.1016/S0140-6736\(10\)61349-9](https://doi.org/10.1016/S0140-6736(10)61349-9).
- Bravo-Ortiz MA, Arteaga-Arteaga HB, Tabares-Soto R, Padilla-Buritica JI, Orozco-Arias S. 2021.** Cervical cancer classification using convolutional neural networks, transfer learning and data augmentation. *Revista EIA* **18(35)**:100–111 DOI [10.24050/reia.v18i35.1462](https://doi.org/10.24050/reia.v18i35.1462).
- Chan H-P, Samala RK, Hadjiiski LM, Zhou C. 2020.** Deep learning in medical image analysis. In: Lee G, Fujita H, eds. *Deep Learning in Medical Image Analysis. Advances in Experimental Medicine and Biology*. Vol. 1213. Cham: Springer DOI [10.1007/978-3-030-33128-3_1](https://doi.org/10.1007/978-3-030-33128-3_1).
- Cheng D, Liu M. 2017.** Combining convolutional and recurrent neural networks for alzheimer's disease diagnosis using pet images. In: *2017 IEEE International Conference on Imaging Systems and Techniques (IST)*. Piscataway: IEEE, 1–5.
- Chitradevi D, Prabha S. 2020.** Analysis of brain sub regions using optimization techniques and deep learning method in Alzheimer disease. *Applied Soft Computing* **86(4)**:105857 DOI [10.1016/j.asoc.2019.105857](https://doi.org/10.1016/j.asoc.2019.105857).
- Cui R, Liu M. 2019.** Hippocampus analysis by combination of 3-D densenet and shapes for alzheimer's disease diagnosis. *IEEE Journal of Biomedical and Health Informatics* **23(5)**:2099–2107 DOI [10.1109/JBHI.2018.2882392](https://doi.org/10.1109/JBHI.2018.2882392).
- Currie G, Rohren E. 2021.** Intelligent imaging in nuclear medicine: the principles of artificial intelligence, machine learning and deep learning. In: *Seminars in Nuclear Medicine*. vol. 51: Amsterdam: Elsevier, 102–111.
- Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J, Houlsby N. 2020.** An image is worth 16×16 words: transformers for image recognition at scale. *ArXiv preprint*. DOI [10.48550/arXiv.2010.11929](https://doi.org/10.48550/arXiv.2010.11929).
- Fischl B. 2012.** Freesurfer. *Neuroimage* **62(2)**:774–781 DOI [10.1016/j.neuroimage.2012.01.021](https://doi.org/10.1016/j.neuroimage.2012.01.021).
- Huang G, Liu Z, van der Maaten L, Weinberger KQ. 2016.** Densely connected convolutional networks. *arXiv* DOI [10.48550/arXiv.1608.06993](https://doi.org/10.48550/arXiv.1608.06993).
- Huang Y, Xu J, Zhou Y, Tong T, Zhuang X, ADNI. 2019.** Diagnosis of Alzheimer's disease via multi-modality 3D convolutional neural network. *Frontiers in Neuroscience* **13**:509 DOI [10.3389/fnins.2019.00509](https://doi.org/10.3389/fnins.2019.00509).
- Jaiswal AK, Tiwari P, Kumar S, Al-Rakhami MS, Alrashoud M, Ghoneim A. 2021.** Deep learning-based smart IoT health system for blindness detection using retina images. *IEEE Access* **9**:70606–70615 DOI [10.1109/ACCESS.2021.3078241](https://doi.org/10.1109/ACCESS.2021.3078241).
- Khan S, Naseer M, Hayat M, Zamir SW, Khan FS, Shah M. 2022.** Transformers in vision: a survey. *ACM Computing Surveys (CSUR)* **54(10s)**:1–41 DOI [10.1145/3505244](https://doi.org/10.1145/3505244).
- Kumar A, Sidhu J, Goyal A, Tsao JW. 2022.** Alzheimer disease. In: *StatPearls [Internet]*. Treasure Island (FL): StatPearls Publishing.

- LaMontagne PJ, Benzinger TL, Morris JC, Keefe S, Hornbeck R, Xiong C, Grant E, Hassenstab J, Moulder K, Vlassenko AG, Raichle ME, Cruchaga C, Marcus D. 2019. OASIS-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and alzheimer disease. *MedRxiv* 7:44 DOI 10.1101/2019.12.13.19014902.
- Liu S, Liu S, Cai W, Pujol S, Kikinis R, Feng D. 2014. Early diagnosis of alzheimer's disease with deep learning. In: *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*. Piscataway: IEEE, 1015–1018.
- MONAI Consortium. 2022. MONAI: medical open network for AI. Available at <https://monai.io/>.
- Moon CI, Lee O. 2022. Skin microstructure segmentation and aging classification using CNN-based models. *IEEE Access* 10:4948–4956 DOI 10.1109/ACCESS.2021.3140031.
- Munadi K, Muchtar K, Maulina N, Pradhan B. 2020. Image enhancement for tuberculosis detection using deep learning. *IEEE Access* 8:217897–217907 DOI 10.1109/ACCESS.2020.3041867.
- Payan A, Montana G. 2015. Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks. *ArXiv preprint*. DOI 10.48550/arXiv.1502.02506.
- Petersen RC, Aisen P, Beckett LA, Donohue M, Gamst A, Harvey DJ, Jack C, Jagust W, Shaw L, Toga A, Trojanowski J, Weiner M. 2010. Alzheimer's disease neuroimaging initiative (ADNI): clinical characterization. *Neurology* 74(3):201–209 DOI 10.1212/WNL.0b013e3181cb3e25.
- Porumb M, Stranges S, Pescapè A, Pecchia L. 2020. Precision medicine and artificial intelligence: a pilot study on deep learning for hypoglycemic events detection based on ecg. *Scientific Reports* 10(1):1–16 DOI 10.1038/s41598-019-56927-5.
- Rasmussen J, Langerman H. 2019. Alzheimer's disease—why we need early diagnosis. *Degenerative Neurological and Neuromuscular Disease* 9:123–130 DOI 10.2147/DNND.S228939.
- Roberts M, Driggs D, Thorpe M, Gilbey J, Yeung M, Ursprung S, Aviles-Rivero AI, Etmann C, McCague C, Beer L, Weir-McCall JR, Teng Z, Gkrania-Klotsas E, Rudd JHF, Sala E, Schönlieb C-B. 2021. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nature Machine Intelligence* 3(3):199–217 DOI 10.1038/s42256-021-00307-0.
- Saborit Torres JM. 2019. Análisis poblacional de patología de Alzheimer a partir de imágenes de resonancia magnética. Available at <http://hdl.handle.net/10251/130025>.
- Sharma S, Sharma V, Sharma A. 2016. Performance based evaluation of various machine learning classification techniques for chronic kidney disease diagnosis. *ArXiv preprint*. DOI 10.48550/arXiv.1606.09581.
- Tabares-Soto R, Bernal-Salcedo J, García-Arias ZN, Ortega-Bolaños R, Hermosilla MP, Arteaga-Arteaga HB, Ruz GA. 2022. *Analysis of ethical development for public policies in the acquisition of AI-based systems*. Pennsylvania: IGI Global, 184–212.
- Tan M, Le QV. 2019. Efficientnet: rethinking model scaling for convolutional neural networks. *ArXiv* DOI 10.48550/arXiv.1905.11946.
- van Oostveen WM, de Lange EC. 2021. Imaging techniques in Alzheimer's disease: a review of applications in early diagnosis and longitudinal monitoring. *International Journal of Molecular Sciences* 22(4):1–34 DOI 10.3390/ijms22042110.
- Villa-Pulgarin J, Ruales-Torres A, Arias-Garzón D, Bravo-Ortiz M, Arteaga-Arteaga H, Mora-Rubio A, Alzate-Grisales J, Mercado-Ruiz E, Hassaballah M, Orozco-Arias S, Cardona-Morales O, Tabares-Soto R. 2022. Optimized convolutional neural network models for skin lesion classification. *Computers, Materials and Continua* 70(2):2131–2148 DOI 10.32604/cmc.2022.019529.

- Wang J, Yang L, Huo Z, He W, Luo J. 2020. Multi-label classification of fundus images with efficientnet. *IEEE Access* 8:212499–212508 DOI [10.1109/ACCESS.2020.3040275](https://doi.org/10.1109/ACCESS.2020.3040275).
- Wynants L, Calster BV, Collins GS, Riley RD, Heinze G, Schuit E, Bonten MM, Damen JA, Debray TP, Vos MD, Dhiman P, Haller MC, Harhay MO, Henckaerts L, Kreuzberger N, Lohmann A, Luijken K, Ma J, Navarro CLA, Reitsma JB, Sergeant JC, Shi C, Skoetz N, Smits LJ, Snell KI, Sperrin M, Spijker R, Steyerberg EW, Takada T, Kuijk SMV, Royen FSV, Wallisch C, Hooft L, Moons KG, Smeden MV. 2020. Prediction models for diagnosis and prognosis of COVID-19: systematic review and critical appraisal. *British Medical Journal* 6:m1328 DOI [10.1136/bmj.m1328](https://doi.org/10.1136/bmj.m1328).
- Xiao B, Yang Z, Qiu X, Xiao J, Wang G, Zeng W, Li W, Nian Y, Chen W. 2022. PAM-DenseNet: a deep convolutional neural network for computer-aided COVID-19 diagnosis. *IEEE Transactions on Cybernetics* 52(11):12163–12174 DOI [10.1109/TCYB.2020.3042837](https://doi.org/10.1109/TCYB.2020.3042837).
- Zeng N, Li H, Peng Y. 2021. A new deep belief network-based multi-task learning for diagnosis of Alzheimer's disease. *Neural Computing and Applications* 35:1–12 DOI [10.1007/s00521-021-06149-6](https://doi.org/10.1007/s00521-021-06149-6).
- Zeng N, Qiu H, Wang Z, Liu W, Zhang H, Li Y. 2018. A new switching-delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer's disease. *Neurocomputing* 320(5):195–202 DOI [10.1016/j.neucom.2018.09.001](https://doi.org/10.1016/j.neucom.2018.09.001).