

Early Diagnosis of Mild Cognitive Impairment with 2-Dimensional Convolutional Neural Network Classification of Magnetic Resonance Images

Luca M. Heising

Tilburg University, The Netherlands
L.M.Heising@tilburguniversity.edu

Spyros Angelopoulos

Tilburg University, The Netherlands
S.Angelopoulos@tilburguniversity.edu

Abstract

We motivate and implement an Artificial Intelligence (AI) Computer Aided Diagnosis (CAD) framework, to assist clinicians in the early diagnosis of Mild Cognitive Impairment (MCI) and Alzheimer's Disease (AD). Our framework is based on a Convolutional Neural Network (CNN) trained and tested on functional Magnetic Resonance Images datasets. We contribute to the literature on AI-CAD frameworks for AD by using a 2D CNN for early diagnosis of MCI. Contrary to current efforts, we do not attempt to provide an AI-CAD framework that will replace clinicians, but one that can work in synergy with them. Our framework is cheaper and faster as it relies on small datasets without the need of high-performance computing infrastructures. Our work contributes to the literature on digital transformation of healthcare, health Information Systems, and NeuroIS, while it opens novel avenues for further research on the topic.

1. Introduction

The digital transformation of healthcare has been a prominent topic in health information systems (HIS) the last decade [1]. Artificial intelligence (AI), as part of the digital transformation toolkit, can enable a more efficient and effective healthcare provision [2]. One of the most promising areas of HIS is the application of AI in assisting medical diagnosis via medical images [3]. Advances in AI show that such applications can perform on par with medical experts on diagnosis via Magnetic Resonance Images (MRI) [4], [5], [6].

Such applications to date, oppose AI to clinicians' performance [7]. For instance, there have been more than 20,000 studies on deep learning (DL) methods for MRI analyses the last ten years, which compare the performance of AI to clinicians' performance [8]. Recent work suggests that studies should focus on the comparison of performance between clinicians using AI, and their performance without an AI aid [7].

The recent global pandemic of COVID-19, however, revealed another urgent need of early disease diagnosis: the ability to make predictions based on few cases. The AI Computer-Aided-Diagnosis (CAD) frameworks to date, are based on large amounts of data, and require high performance computing (HPC) infrastructures. To address that lacuna, we propose a synergistic approach, in which clinicians and scientists collaborate for faster, cheaper, and more accurate diagnosis, while the AI-CAD frameworks can rely on small datasets to make accurate-enough predictions.

Neuro-Information-Systems (NeuroIS), as an interdisciplinary field that combines cognitive neuroscience and IS, can address this gap and advance the theory and practice of HIS [9]. For MRI diagnosis with AI, this translates to finding which approaches can provide the most efficient diagnosis and apply them to assist clinicians with a HIS that compliments their knowledge and skills. The literature on AI in healthcare to date, focuses on cancer, the nervous system, and cardiovascular diseases because these are the leading causes of disability and mortality [10]. However, a promising frontier where AI can assist clinicians is Alzheimer's Disease (AD). Whilst there was limited progress in the search for treatment to slow down the progress of AD, last year Biogen released promising clinical studies for a new drug. To start treatment in early stages, an early diagnosis is needed. As it can take up to 20 years before the patient shows any cognitive decline, it can be difficult to diagnose AD in early stages. An AI-CAD framework could assist clinicians in AD diagnosis via MRI.

In this paper, thus, we motivate and implement an AI-CAD framework for the early diagnosis of Mild Cognitive Impairment (MCI) and AD to assist clinicians. By doing so, we contribute to the extant literature on digital transformation of healthcare [1], HIS [11], and NeuroIS [12]. The approach that we proposed in this paper can be extended for the diagnosis of other diseases, and to further enhance the digital transformation of healthcare [3].

2. Background

AD is the most common reason for dementia, caused by an accumulation of β -amyloid ($A\beta$) plaques, and abnormal amounts of *tau* proteins in the brain. This results in synapse loss, where the impulse does not reach the neurons, and in loss of structure or function of neurons, including their death, causing memory impairment and other cognitive problems [13]. More than 44 million people worldwide suffer from AD and it is projected to triple by 2050 [14]. AD has strong impact on patients' cognitive and physical functioning, resulting in death [15]. Moreover, there is also a strong societal and economic impact as global cost of AD is estimated at more than \$818 billion [14].

The recent developments in slowing AD decline has increased the relevance of its early diagnosis [16], and MCI plays an important role in the early diagnosis of AD. MCI is a syndrome where patients have greater cognitive decline than expected, but it does not affect their lives [18]. Although some MCI patients remain stable or return to cognitively normal (CN), there is a 10-15% risk per year of progression to AD [13].

The AD etiology was unknown for the last decade, and diagnosis relied on neurocognitive tests [18]. This approach was inefficient and unreliable because AD could not be distinguished from other dementias. The development of biomarkers improved the diagnosis of AD, and many studies have focused on identifying biomarkers. A common method to diagnose AD is hippocampus segmentation, which relates to memory function [19], and its small volume is an AD biomarker [20]. For a long time, AD diagnosis was done manually by looking at the brain structure and size of the hippocampus on MRI, which requires a lot of practice and precision [21]. There has been an increasing amount of studies on automated methods for hippocampus segmentation using machine learning (ML) and DL with promising results [22], [23]. Hippocampus segmentation for the diagnosis of AD and MCI, however, requires clinicians' expertise and is sensitive to interrater and intra-rater variability [22].

Convolutional Neural Networks (CNN), which is a successful method for image classification, can support clinicians in diagnosing AD and MCI. CNN can significantly improve the performance of image classification [24], and are becoming increasingly popular in MRI analysis. Recent studies have showed that CNN can achieve similar results as a specialist on classifying MRI of skin cancer patients [5]. Similar approaches with three-dimensional (3D) as well as two-dimensional (2D) CNN have also been used for AD diagnosis with promising results. The filter of a 3D CNN slides along the three dimensions of the input image, resulting in 3D feature maps, whereas in a 2D

Study	2D CNN		3D CNN	
	AD	MCI	AD	MCI
Basaia et al. [25]	-	-	.99	.87
Feng et al. [27]	-	-	.95	.86
Korolev et al. [28]	-	-	.80	-
Liu et al. [29]	-	-	.85	-
Liu et al. [30]	-	-	.91	-
Senanayake et al. [31]	-	-	.76	.75
Hon and Khan [32]	.96	-	-	-
Sarraf and Tofighi [33]	.97	-	-	-
Sarraf and Tofighi [34]	.99	-	-	-
Wang et al. [26]	.98	-	-	-

CNN the filter only slides along the height and width of the input image. Therefore, the latter case results in 2D feature maps with need for less parameters and, thus, less computational power and time. Most prior studies used 3D CNN, achieving accuracy up to 99% for AD diagnosis, and up to 87% for MCI [25], while others obtained similar results with 2D CNN [26].

A summary of 2D and 3D approaches in the literature is shown in Table 1. Prior studies, however, have not applied 2D approaches for detecting MCI. We will, thus, investigate whether a 2D CNN is capable of early diagnosing MCI as well as AD.

3. Methodology

Neural networks (NN) are inspired by the human brain to resolve problems that are simple for humans but complex for machines [35]. CNN is the most common NN architecture for image analysis. Fully connected NN take multiple inputs, and hidden layers perform calculations on them, while the neurons connect to each other. Neurons in CNN, however, connect to others close to them, and all have the same weight. CNN thus, need fewer parameters than fully connected NN, which results in small risk of overfitting, higher accuracy, and faster processing time. Moreover, in CNN there is no need to transform images to one-dimensional, a process which can result in loss of structural information, as CNN learn the relationships among pixels of input by extracting representative features with kernel convolutions [13]:

$$S(i, j) = (I \times K)(i, j) = \sum_m \sum_n I(m, n)K(i - m, j - n)$$

where I is the input and K is the kernel; the input indices are represented by i and j , and the kernel indices are represented by m and n . A 2D CNN extracts features by sliding along the height and width of images, resulting in 2D feature maps. 3D CNN also slide along the depth of images thus, the kernel does not have the same depth as the input, resulting in 3D feature maps. Although previous work established that 3D CNN perform better for patch classifications, the results between 2D and 3D approaches for whole image labeling did not differ much [36]. A 3D CNN, however, is computationally expensive, and, due to the number of parameters, require more training data [37].

The data used in the preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership. The primary goal of ADNI has been to test whether MRI, other biological markers, and clinical as well as neuropsychological assessment can be combined to measure the progression of MCI and early AD. The ADNI is separated into 3 studies of 5 years, while the first was prolonged by 2 years under the name ADNI-GO. In total, 2517 people of ages 55-90 participated in the study. The ADNI encourages the use of their standardized datasets to ensure consistency in analysis and direct comparison of various methods among studies. We, therefore, used their two standardized datasets ‘ADNI:Complete 2Yr 1.5T’, and ‘ADNI:Complete 3Yr 1.5T’, which contain MRI that have passed quality control assessment [38].

Our dataset consists of 1941 images, distributed in 658 MRI of CN subjects, 411 MRI of AD patients, and 872 MRI of MCI patients. The data come from 99 AD patients, 148 MCI patients, and a control group of 135 CN subjects. We present the demographic information of the included subjects in Table 2 to enable comparison with the samples used in other studies.

The dataset was split into one with CN and AD subjects (1063 MRI), and one with CN and MCI subjects (1524 MRI). Since the participants of the ADNI study returned for more than one check-up, any patient can have up to 12 MRI, which are not identical as they are taken at different moments, and every MRI in the standardized dataset was treated independently.

The standardized datasets are preprocessed, without noisy MRI, while three more processing steps have been performed: i) gradwarp correction to correct inherent nonlinearity of the imaging gradients in MRI scanners, ii) B1 non-uniformity for a greater acquisition accuracy of the MRI, and iii) bias field correction, which can lead to spurious intensity

	MCI	AD	CN
Images	891	412	662
Subjects	148	99	135
Gender	105 M / 43 F	52 M / 47 F	66 M / 69 F
Age	$\mu = 75.59$ $\sigma = 7.40$	$\mu = 75.85$ $\sigma = 7.61$	$\mu = 77.19$ $\sigma = 5.24$

variations. Due to the different scanners and techniques used by the ADNI, the MRI were of different sizes, and, therefore had to be resized. MRI with sizes (256, 256, 180), (256, 256, 170), (256, 256, 166), (256, 256, 184) and (192, 192, 160) were resized to be: (192, 192, 160) with the zoom function of the Scipy library. In total, 28 MRI from the MCI dataset and 11 MRI from the AD dataset had opposite dimensions (e.g. 180, 256, 256), which means reshaping them would deform them, and, thus, these MRI were removed from the dataset. Resizing the MRI results in a different range of pixel values, and, therefore, to assure that the pixel values of all MRI had the same range, min-max normalization was applied, scaling all pixel values between 0 and 1 as follows:

$$z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}$$

where x is the MRI data and z_i the i^{th} normalized MRI. The dataset was then split into train set, validation set, and test set with ratio 60:20:20. After preprocessing, our AD dataset consisted of 1063 MRI, with 652 CN, and 411 AD, while the MCI dataset consisted of 1524 MRI, with 872 MCI, and 652 CN subjects.

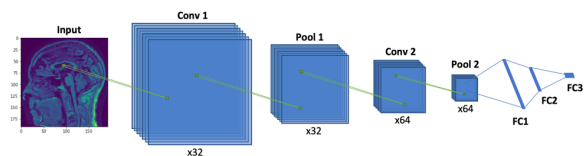


Figure 1. Convolutional network architecture

Layer	C1	P1	C2	P2	FC1	FC2	FC3
Kernel	3x3	2x2	3x3	2x2	-	-	-
Filter	32	32	64	64	128	64	2

A NN consists of an input layer, hidden layers, and an output layer. A CNN has hidden layers divided into convolution, pooling, activation, and classification layers. We based our architecture on LeNet-5, which includes two convolutional layers, two pooling layers, and two fully connected layers (Figure 1 and Table 3).

The Rectified Linear Unit (ReLU) is the most commonly used activation function in DL because of its simplicity and performance [39], [40], but when the activation values are zero, the gradient will be zero from that point on and, thus, the model will not learn anymore. This is also referred to as the dying ReLU problem. As a remedy, we employ the Leaky Rectified Linear Unit (LReLU) as activation function for all convolutional layers. LReLU addresses this problem by allowing for a small non-zero gradient [41]. The LReLU activation function in our model is defined as:

$$y(x) = \begin{cases} x, & \text{if } x < 0 \\ 0.01x, & \text{otherwise} \end{cases}$$

where x is the input. A Sigmoid activation function was applied to the dense layer, which outputs the probability of the images' class, with 0 if healthy and 1 if sick (AD or MCI). The Sigmoid function in our model, with x being the input data, is described as:

$$\sigma(x) = \frac{1}{(1 + e^{-x})}$$

The batch size was set to 32, and the optimizer of the model was Adam [42], with a learning rate of 10^{-4} .

The CNN was built in a Jupyter Notebook using Python 3.6.4, Tensorflow 2.1.0, and Keras 2.3.1. TensorFlow is an interface for expressing ML algorithms, and an implementation for executing them [43]. Keras is a NN application programming interface (API), designed to enable fast experimentation with NN [44]. To load the MRI in NIfTI format, the Nilearn library was used, which is developed especially for statistical learning on NeuroImaging data. The scikit-learn and SciPy libraries were used for data preprocessing. The development, testing, and application of the model took place on Google Cloud Console, where we used a storage bucket to store the datasets, and one compute engine instance with 64 vCPUs and 416 GB of memory to build our model.

4. Results

The model evaluation was measured in accuracy, recall, precision, and f1-score. Recall provides sensitivity information on how many patients were correctly identified. Precision expresses how many of the positives that the model returns were actually

positive. F1-score is the harmonic mean between precision and recall. A NN adjusts its weights to optimize the loss, which is calculated with the use of binary cross entropy loss:

$$CE = - \sum_{i=1}^{C'} t_i \log(s_i) = t_1 \log(s_1) - (1 - t_1) \log(1 - s)$$

where C represents the classes, s_i is the predicted probability value for class i , and t is the true probability for that class. Since the data was unevenly distributed, the accuracy baseline of random guessing was also calculated. The baseline was calculated with respect to the class distribution of the dataset. First, we trained and tested our model on the AD dataset. After passing the baseline of random guessing on the training data ($>.725$) with an accuracy of .910, we applied the same model on the MCI dataset. The random guessing baseline for the test dataset of the AD model was .519 and for the test dataset of the MCI model was .507. Models' over-epochs performance is shown in figure 2 for AD, and in figure 3 for MCI.

The graphs indicate a normal learning curve, however, as the performance keeps increasing on the train data, the validation performance flattens, which implies overfitting. This appears to be stronger on the MCI dataset than on the AD dataset. Our model achieved an accuracy of .732 on the AD test-set. Irrespective of overfitting, the achieved test accuracy surpasses the random guessing baseline of .519.

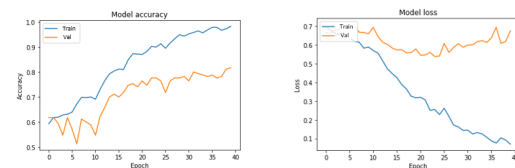


Figure 2. Model performance on AD dataset.
Left: Accuracy; Right: Loss

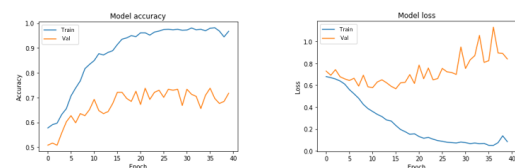


Figure 3. Model performance on MCI dataset.
Left: Accuracy; Right: Loss

Study	Subjects	Images	Dimensions	Accuracy	
				AD	MCI
Basaia et al. [25]	645	-	3D	.99	.87
Feng et al. [27]	193	-	3D	.95	.86
Korolev et al. [28]	111	111	3D	.80	-
Liu et al. [29]	193	-	3D	.85	-
Liu et al. [30]	902	-	3D	.91	-
Senanayake et al. [31]	-	322	3D	.76	.75
Hon and Khan [32] *	200	6400	2D	.96	-
Sarraf and Tofighi [33] **	43	367,200	2D	.97	-
Sarraf and Tofighi [34] **	302	62,335	2D	.99	-
Wang et al. [26] **	98	17,738	2D	.98	-
Our	234	1,063	2D	.73	.73

* accuracy before transfer learning = .74
** used fMRI slices independently

The model predicted MCI with accuracy of .734, passing the random guessing baseline of .507, and appeared to be overfitting. Table 4 presents the performance metrics of the models on the test sets.

Data	Loss	Acc.	Prec.	Recall	F1	MRI
AD	1.261	.732	.876	.640	.740	1074
MCI	1.101	.734	.822	.923	.870	1553

The model performs better than chance on both datasets, with similar loss, accuracy, and precision, and achieves a better recall on the MCI data as 92.3% of the MCI patients were correctly identified. The MCI dataset was larger than the AD one, which may explain why the performance of MCI is similar to predicting AD, and for recall even significantly higher.

5. Discussion

Comparing our study to previous ones, we find a large difference in dataset size (Table 5). Some of the

prior studies only report the number of subjects, but the number of images can differ from these since one subject can have up to 12 images in these datasets. As expected, studies with larger datasets, achieved a higher accuracy. Moreover, some studies with a 2D method treated the slices independently, thereby enlarging the size of their dataset, however, the MRI is not treated as a whole. We are able, thus, to predict AD better than chance by .213, and MCI by .227.

We expected the model to perform worse on detecting MCI than AD, which was not the case. Although the results for predicting MCI are better than the AD ones, we cannot conclude that the model performs better for MCI than for AD, as the MCI dataset was larger than the AD one. The MCI dataset had 489 more positive labels than the AD dataset. Our work, thus, resulted in an AI-CAD framework that can assist clinicians in the early diagnosis of MCI and AD with high-enough accuracy, based on a small dataset, and without the need of HPC infrastructures.

5.1. Implications

The implications of our work are threefold. First, we contribute to the line of research on using CNN for AD and MCI diagnosis, by applying a 2D approach.

Second, whilst AI-CAD frameworks have been thoroughly studied, they have not been proposed as a tool for assisting clinicians. Moreover, the existing approaches are often complicated, and make use of large amounts of data. Our proposed framework provides high-enough accuracy to assist clinicians in their diagnosis for early AD and MCI without the need of large datasets or HPC infrastructures. Our proposed approach can also be extended for other diseases, as well as for cases where time is scarce or limited data is available. Furthermore, whilst the literature on AI-CAD frameworks is mostly approached from a computer science perspective, clinicians have been shown to lack trust in them [46]. Our work addresses that lacuna by providing a synergistic approach between clinicians and scientists, thereby contributing to the literature on HIS [11] and NeuroIS [12]. The NeuroIS literature focusses on different disciplines from theory-focused to design-focused [9]. Our final contribution, thus, is on bridging these disciplines. The work we present here, thus, extends current efforts in the field, and can open avenues for further research.

5.2. Limitations

Our approach has some limitations that should be acknowledged. First, some preprocessing steps were not performed because of the required operating system (OS). An important step in MRI preprocessing is skull stripping. In this step, the brain tissue is removed from the image to reduce noise. Commonly used techniques are the brain extraction tool (BET) [45], the Brain Surface Extractor (BSE) [47], and the Robust Brain Extraction (ROBEX) [48]. The automated methods in Python for these techniques, however, are only useable under Unix. Due to the used OS, skull stripping could not be applied to all images in the dataset and was therefore omitted. As noise in the data can lead to unexpected results in classification due to false feature extraction, it is expected that skipping this step has direct consequences to the performance of the CNN. Skull stripping (BET2 in particular) leads to a better segmentation [49], it is however difficult to say how large the impact is for classification tasks. Second, our model showed overfitting, which means that it includes more terms or uses more complicated approaches than necessary [50]. Regularization can control overfitting and dropout regularization is a commonly used approach because it is computationally inexpensive, and it prevents co-adaptation among feature map units [37]. We added dropout regularization to our model, and although it performed slightly better on training and validation data, it performed worse on the test data. For that reason, we omitted dropout regularization.

5.3. Future Research

Our work offers a faster and cheaper method for classifying MRI, as resources can be saved from not having to acquire computational power. We proposed a 2D approach for MCI and AD detection because the 3D one has drawbacks related to its computational requirements. As a next step of our research project, we intend to replicate the existing 3D approaches in the literature, and compare their execution time with the one of our model's, on the same computational infrastructure. Such a comparison of execution time will further illustrate the merits of our proposed approach before proposing it for testing and implementation to the broader healthcare system.

Whilst we will further develop the accuracy of the model before proceeding, our next step is to evaluate the performance of clinicians using our AI-CAD framework, and their performance without the AI aid. For this step, we have already established collaboration with hospitals and clinics that have geriatric units specializing in AD and MCI, and we are ready to proceed with a pilot phase. Such collaboration will provide us with unique datasets to further enhance the externality of our model, as most studies using CNN on AD and MCI are based on the ADNI datasets.

Moreover, quantum computing (QC) is a new paradigm that could offer a solution to the challenge we address. The field of quantum DL provides innovative algorithms that offer significant speedups over the classical DL approaches [51]. We aim, thus, to further develop an AI-CAD framework with a Quantum CNN (QCNN). In classical computing a bit can either be $|0\rangle$ or $|1\rangle$; a qubit in QC can be $|0\rangle$, $|1\rangle$, or in a superposition state $\alpha|0\rangle + \beta|1\rangle$ with amplitudes $(\alpha, \beta) \in \mathbb{C}$ such that $|\alpha|^2 + |\beta|^2 = 1$. In classical computing, processes are executed in sequence, while in QC processes run simultaneously further optimizing, thus, the total run time. Moreover, QCNN do not suffer from the issue of causal asymmetry that can affect classic CNN [52]. Should the QC paradigm become more easily accessible in the coming years, such an approach can provide clinicians with fast access to predictive models to address urgent situations, as well as the lack of large datasets.

6. Conclusion

Prior studies have used CNN to diagnose MCI and AD, most of which applied 3D approaches, with promising results. 3D CNN, however, has drawbacks that relate to needs for HPC infrastructures. Other studies have focused on detecting AD with a 2D CNN, achieving similar results as the 3D approach. Despite

the relevance of detecting MCI, prior studies did not investigate how these methods perform on detecting MCI. Our main goal was to determine whether a 2D CNN can be used to diagnose AD and MCI. Our work resulted in an AI-CAD framework that can assist clinicians in the early diagnosis of MCI and AD with high-enough accuracy, based on a small dataset, and without the need of HPC infrastructures. Our next step will be to evaluate the performance of clinicians with and without the use of our AI-CAD framework. As a follow-up study, we intent to develop an AI-CAD framework with a QCNN, which could provide clinicians with fast access to predictive models to address urgent situations and the lack of large datasets.

7. References

- [1] R. Agarwal, G. Gao, C. DesRoches, and A.K. Jha, “The digital transformation of healthcare: Current status and the road ahead”, *Information Systems Research*, 21(4), 2010, pp. 796-809.
- [2] N. Noorbakhsh-Sabet, R. Zand, Y. Zhang, and V. Abedi, “Artificial Intelligence Transforms the Future of Health Care”, *The American Journal of Medicine*, 132(7), 2019, pp. 795-801.
- [3] I. Bardhan, H. Chen, and E. Karahanna, “Connecting systems, data, and people: a multidisciplinary research roadmap for chronic disease management”, *MIS Quarterly*, 44(1), 2020, pp. 185-200.
- [4] A.A. Cruz-Roa, J.E. Arevalo Ovalle, A. Madabhushi, and F.A. González Osorio, A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection, *Lecture Notes in Computer Science*, Springer, Berlin, Heidelberg, 2013, pp. 403-410.
- [5] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, and S. Thrun, “Dermatologist-level classification of skin cancer with deep neural networks”, *Nature*, 542(7639), 2017, pp. 115.
- [6] C.D. Lehman, A. Yala, T. Schuster, B. Dontchos, M. Bahl, K. Swanson, and R. Barzilay, “Mammographic breast density assessment using deep learning: Clinical implementation”, *Radiology*, 90(1), 2019, pp.52-58.
- [7] G. Briganti, and O, Le Moine, “Artificial Intelligence in Medicine: Today and Tomorrow”, *Frontiers in Medicine*, 7, 2020, pp. 27.
- [8] X. Liu, L. Faes, A.U. Kale, S.K. Wagner, D.J. Fu, A. Bruynseels, T. Mahendiran, G. Moraes, M. Shamdass, C. Kern, J.R. Ledsam, M.K. Schmid, K. Balaskas, E.J. Topol, L.M. Bachmann, P.A. Keane, and A.K. Denniston, “A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis”, *The Lancet Digital Health*, 1(6), 2019, pp. e271-e297.
- [9] R. Riedl, M. Hubert, and P. Kenning, “Are there neural gender differences in online trust? An fMRI study on the perceived trustworthiness of eBay offers”, *MIS quarterly*, 34(2), 2010, pp. 397-428.
- [10] V. Abedi, N. Goyal, G. Tsivgoulis, N. Hosseinichimeh, R. Hontecillas, J. Bassaganya-Riera, L. Eljovich, J.E. Metter, A.W. Alexandrov, D.S. Liebeskind, A.V. Alexandrov, and R. Zand, “Novel Screening Tool for Stroke Using Artificial Neural Network”, *Stroke*, 48(6), 2017, pp. 1678-1681.
- [11] M.W. Chiasson, and E. Davidson, “Pushing the contextual envelope: developing and diffusing IS theory for health information systems research”, *Information and Organization*, 14(3), 2004, pp. 155-188.
- [12] A. Dimoka, F.D. Davis, A. Gupta, P.A. Pavlou, R.D. Banker, A.R. Dennis, A. Ischebeck, G. Müller-Putz, I. Benbasat, D. Gefen, P.H. Kenning, R. Riedl, J. Brocke, and B. Weber, “On the use of neurophysiological tools in IS research: Developing a research agenda for NeuroIS”, *MIS Quarterly*, 36(3), 2012, pp. 679-702.
- [13] M.W. Weiner, D.P. Veitch, P.S. Aisen, L.A. Beckett, N.J. Cairns, J. Cedarbaum, R.C. Green, D. Harvey, C.R. Jack, W. Jagust, J. Luthman, J.C. Morris, R.C. Petersen, A.J. Saykin, L. Shaw, L. Shen, A. Schwarz, A.W. Toga, and J.Q. Trojanowski, “2014 Update of the Alzheimer’s Disease Neuroimaging Initiative: A review of papers published since its inception”, *Alzheimer’s & Dementia*, 11(6), 2015, pp. e1-e120.
- [14] M. Prince, A. Comas-Herrera, M. Knapp, M. Guerchet, and M. Karagiannidou, “World Alzheimer report 2016: improving healthcare for people living with dementia: coverage, quality and costs now and in the future”, *Alzheimer’s Disease International (ADI)*, 2016.
- [15] B. Winblad, P. Amouyel, S. Andrieu, C. Ballard, C. Brayne, H. Brodaty, A. Cedazo-Minguez, B. Dubois, D. Edvardsson, H. Feldman, and L. Fratiglioni, “Defeating Alzheimer’s disease and other dementias: a priority for European science and society”, *The Lancet Neurology*, 15(5), 2016, pp. 455-532.
- [16] D.J. Selkoe, “Alzheimer disease and aducanumab: Adjusting our approach”, *Nature Reviews Neurology*, 15(7), 2019, pp. 365-366.
- [17] S. Gauthier, B. Reisberg, M. Zaudig, R.C. Petersen, K. Ritchie, K. Broich, S. Belleville, H. Brodaty, D. Bennett, H. Chertkow, and J.L. Cummings, “Mild cognitive impairment”, *The Lancet*, 367(9518), 2006, pp. 1262-1270.
- [18] E. Kolibas, V. Korinkova, V. Novotny, K. Vajdickova, and D. Hunakova, “ADAS-Cog (Alzheimer’s Disease

- Assessment Scale-cognitive subscale)-validation of the Slovak version”, Bratislavske Lekarske Listy, 101(11), 2000, pp. 598-602.
- [19] W.B. Scoville, and B. Milner, “Loss of recent memory after bilateral hippocampal lesions”, *The Journal of Neuropsychiatry and Clinical Neurosciences*, 12(1), 2000, pp. 103--a.
- [20] D.P. Devanand, G. Pradhaban, X. Liu, A. Khandji, S. De Santi, S. Segal, H. Rusinek, G.H. Pelton, L.S. Honig, R. Mayeux, Y. Stern, M.H. Tabert, and M.J. de Leon, “Hippocampal and entorhinal atrophy in mild cognitive impairment”, *Neurology*, 68(11), 2007, pp. 828-836.
- [21] C.R. Jack Jr, “MRI-based hippocampal volume measurements in epilepsy”, *Epilepsia*, 35, 1994, pp. S21-S29.
- [22] D. Ataloglou, A. Dimou, D. Zarpalas, and P. Daras, “Fast and precise Hippocampus Segmentation through Deep Convolutional Neural Network Ensembles and Transfer Learning”, *Neuroinformatics*, 17(4), 2019, pp. 563-582.
- [23] L. Cao, L. Li, J. Zheng, X. Fan, F. Yin, H. Shen, and J. Zhang, “Multi-task neural networks for joint hippocampus segmentation and clinical score regression”, *Multimedia Tools and Applications*, 77(22), 2018, pp. 29669-29686.
- [24] Q. Li, W. Cai, X. Wang, Y. Zhou, D.D. Feng, and M. Chen, “Medical image classification with convolutional neural network”, *International Conference on Control Automation Robotics Vision*, 2014, pp. 844-848.
- [25] S. Basaia, F. Agosta, L. Wagner, E. Canu, G. Magnani, R. Santangelo, M. Filippi, and Initiative, A.D.N. “Automated classification of Alzheimer’s disease and mild cognitive impairment using a single MRI and deep neural networks”, *NeuroImage: Clinical*, 21, 2019, pp. 101645.
- [26] S.H. Wang, P. Phillips, Y. Sui, B. Liu, M. Yang, and H. Cheng, “Classification of Alzheimer’s disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling”, *Journal of Medical Systems*, 42(5), 2018, pp. 85.
- [27] C. Feng, A. Elazab, P. Yang, T. Wang, F. Zhou, H. Hu, X. Xiao, and B. Lei, “Deep Learning Framework for Alzheimer’s Disease Diagnosis via 3D-CNN and FSBi-LSTM”, *IEEE Access*, 7, 2019, pp. 63605-63618.
- [28] S. Korolev, A. Safiullin, M. Belyaev, and Y. Dodonova, “Residual and plain convolutional neural networks for 3D brain MRI classification”, *IEEE International Symposium on Biomedical Imaging*, 2017, pp. 835-838.
- [29] M. Liu, D. Cheng, K. Wang, Y. Wang, and Initiative, A.D.N. “Multi-Modality Cascaded Convolutional Neural Networks for Alzheimer’s Disease Diagnosis”, *Neuroinformatics*, 16(3), 2018, pp. 295-308.
- [30] M. Liu, J. Zhang, E. Adeli, and D. Shen, “Landmark-based deep multi-instance learning for brain disease diagnosis”, *Medical Image Analysis*, 43, 2018, pp. 157-168.
- [31] U. Senanayake, A. Sowmya, and L. Dawes, “Deep fusion pipeline for mild cognitive impairment diagnosis”, *IEEE International Symposium on Biomedical Imaging*, 2018, pp. 1394-1997.
- [32] M. Hon, and N.M. Khan, “Towards Alzheimer’s disease classification through transfer learning”, *IEEE International Conference on Bioinformatics and Biomedicine*, 2017, pp. 1166-1169.
- [33] A. Sarraf, and G. Tofighi, “Classification of alzheimer’s disease using fmri data and deep learning convolutional neural networks”, *ArXiv Preprint*, ArXiv:1603.08631, 2016.
- [34] A. Sarraf, and G. Tofighi, “DeepAD: Alzheimer’s Disease Classification via Deep Convolutional Neural Networks using MRI and fMRI”, *BioRxiv*, 2016.
- [35] A. Graves, A. Mohamed, and G. Hinton, “Speech recognition with deep recurrent neural networks”, *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 6645-6649.
- [36] M. Lai, “Deep learning for medical image segmentation”, *ArXiv Preprint*, ArXiv:1505.02000, 2015.
- [37] J. Bernal, K. Kushibar, D.S. Asfaw, S. Valverde, A. Oliver, R. Martí, and X. Lladó, “Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review”, *Artificial Intelligence in Medicine*, 95, 2019, pp. 64-81.
- [38] B.T. Wyman, D.J. Harvey, K. Crawford, M.A. Bernstein, O. Carmichael, P.E. Cole, P.K. Crane, C. DeCarli, N.C. Fox, N. J.L. Gunter, D. Hill, R.J. Killiany, C. Pachai, A.J. Schwarz, N. Schuff, M.L. Senjem, J. Suhy, P.M. Thompson, M. Weiner, and C.R. Jack, “Standardization of analysis sets for reporting results from ADNI MRI data”, *Alzheimer’s & Dementia*, 9(3), 2013, pp. 332-337.
- [39] J. Bernal, K. Kushibar, D.S. Asfaw, S. Valverde, A. Oliver, R. Martí, and X. Lladó, “Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review”, *Artificial Intelligence in Medicine*, 95, 2019, pp. 64-81.
- [40] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning”, *Nature* 521, 2015, pp. 436-444.
- [41] L. Lu, Y. Shin, Y. Su, and G.E. Karniadakis, “Dying ReLU and Initialization: Theory and Numerical Examples”, *ArXiv Preprint*, ArXiv:1903.06733, 2019.
- [42] D.P. Kingma, and J. Ba, “Adam: A method for stochastic optimization”, *ArXiv preprint*, arXiv:1412.6980, 2014.

- [43] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, and S. Ghemawat, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems", ArXiv Preprint ArXiv:1603.04467, 2016.
- [44] F. Chollet, Keras documentation. keras.io, 2015
- [45] S.M. Smith, "Fast robust automated brain extraction. Human brain mapping", 17(3), 2002, pp. 143-155.
- [46] G. Briganti, and O. Le Moine, "Artificial Intelligence in Medicine: Today and Tomorrow", *Frontiers in Medicine*, 7(27), 2020.
- [47] D.W. Shattuck, S.R. Sandor-Leahy, K.A. Schaper, D.A. Rottenberg, and R.M. Leahy, "Magnetic resonance image tissue classification using a partial volume model", *NeuroImage*, 13(5), 2001, pp. 856-876.
- [48] J.E. Iglesias, C.Y. Liu, P.M. Thompson, and Z. Tu, "Robust brain extraction across datasets and comparison with publicly available methods", *IEEE transactions on medical imaging*, 30(9), 2011, pp. 1617-1634.
- [49] J. Acosta-Cabronero, G.B. Williams, J.M.S Pereira, G. Pengas, and P.J. Nestor, "The impact of skull-stripping and radio-frequency bias correction on grey-matter segmentation for voxel-based morphometry", *NeuroImage*, 39(4), 2008, pp. 1654-1665.
- [50] D.M. Hawkins, "The problem of overfitting", *Journal of Chemical Information and Computer Sciences*, 44(1), 2004, pp. 1-12.
- [51] I. Kerenidis, J. Landman, and A. Prakash, "Quantum Algorithms for Deep Convolutional Neural Networks", arXiv preprint arXiv:1911.01117, 2019.
- [52] V.G. Ivancevic, and T.T. Ivancevic, *Quantum Neural Computation (Vol. 40)*. Springer, 2010.