http://doi.org/10.35784/iapgos.3383

received: 23.12.2022 | revised: 1.02.2023 | accepted: 12.03.2023 | available online: 31.03.2023

AUTOMATIC DETECTION OF ALZHEIMER'S DISEASE BASED ON ARTIFICIAL INTELLIGENCE

Achraf Benba, Kerchaoui Abdelilah

Mohammed V University in Rabat, Ecole Nationale Supérieure d'Arts et Métiers, Electronic Systems Sensors and Nanobiotechnologies, Rabat, Morocco

Abstract. *Alzheimer's disease is a neurodegenerative disease that progressively destroys neurons through the formation of platelets that prevent communication between neurons. The study carried out in this project aims to find a precise and relevant diagnostic solution based on artificial intelligence and which helps in the early detection of Alzheimer's disease in order to stop its progression. The study went through a process of processing MRI images followed by training of three deep learning algorithms (VGG-19, Xception and DenseNet121) and finally by a step of testing and predicting the results. The results of the accuracy metric obtained for the three algorithms were respectively 98%, 95%, 91%.*

Keywords: Alzheimer's disorder, artificial intelligence, deep learning, signal processing

AUTOMATYCZNE WYKRYWANIE CHOROBY ALZHEIMERA W OPARCIU O SZTUCZNĄ INTELIGENCJĘ

Streszczenie. *Choroba Alzheimera jest chorobą neurodegeneracyjną, która stopniowo niszczy neurony poprzez tworzenie płytek krwi, które uniemożliwiają komunikację między neuronami. Badania prowadzone w ramach tego projektu mają na celu znalezienie precyzyjnego i trafnego rozwiązania diagnostycznego opartego na sztucznej inteligencji, które pomoże we wczesnym wykryciu choroby Alzheimera w celu zatrzymania jej postępu. Badanie przeszło przez proces przetwarzania obrazów MRI, po którym następowało szkolenie trzech algorytmów głębokiego uczenia (VGG-19, Xception i DenseNet121), a na koniec etap testowania i przewidywania wyników. Wyniki metryki dokładności otrzymane dla trzech algorytmów wyniosły odpowiednio 98%, 95%, 91%.*

Slowa kluczowe: choroba Alzheimera, sztuczna inteligencja, głębokie uczenie się, przetwarzanie sygnału

Introduction

Dementia is a condition that is often chronic or progressive in nature and causes cognitive ability to deteriorate beyond what may be anticipated from the typical effects of biological aging [9]. Memory, thought, orientation, understanding, calculation, learning capacity, language, and judgment are all impacted. Currently, dementia is the seventh most common cause of mortality globally among all diseases and one of the main reasons of reliance and impairment among elderly people [9]. The major goal of this project is to develop a diagnostic method based on artificial intelligence algorithms for the early diagnosis of dementia and, more particularly, Alzheimer's disease. Artificial intelligence presents a set of theories and techniques applied to create machines capable of simulating human intelligence. The study was based on an MRI database, exploited in order to automatically detect the different stages of severity of Alzheimer's disease. To successfully find this solution, we must focus on improving the performance of these algorithms and more specifically on the accuracy metric. The more value of this metric increases the more the detection becomes precise. There are several studies in the literature which amin to classify Alzheimer disease.

In [4], the researchers proposed a method of classification of Alzheimer's disease based on the close learning by transfer, they started with the step of preprocessing the MRI images acquired in the database. ADNI data, using the Freesurfer software. The pretreatment techniques that were used are the skull stripping technique, which is used to extract the brain from the skull, a technique for normalizing intensities and correcting movement. To extract the features from the processed images, they opted to use the VGG 16 algorithm through which they performed two types of classification, a binary classification in particular between two different classes of the three proposed classes (AD, CN MCI) and a normal classification that is made between all ADNI classes. The final test accuracy result was 95.73%. In [7], the researchers proposed a classification study of Alzheimer's disease by three different CNN algorithms including GoogleNet, AlexNet and ResNet-18 based on the Transfer Learning approach. Was carried out on an age margin of 55 years to 100 years with a sample of 7800 MRI ADNI images, the database was divided into 5 classes presenting the degrees of the disease AD, CN, EMCI, LMCI MCI. The test accuracy values obtained by the study were 96.39%, 94.08% and 97.51% respectively. In [6], the researchers proposed a classification study based on the characteristics of the hippocampus (volume and morphology), they adopted the methodology of combining a multi-model based on the CNN algorithm to ensure at the same time the segmentation of the hippocampus and the classification of the disease, this method was evaluated on the T1 weighting of a database to constitute the following classes 97 AD, 233 MCI and 119 NC. of 88.9% for the binary classification of AD versus NC and the accuracy of 76.2% for the classification of MCI versus NC.

1. Database aquisition

The database used for this study was used in [5]. It has a total of 6,400 MRI images resized to 128 × 128 and divided into 4 classes, Mild_Demented (896 images), Moderate_Demented (64 images), Non_Demented (3,200 images) and Very_Mild_Demented (2,240 images). In addition, this database contains images that have been taken in two sections: a longitudinal section and a cross section. These images were taken using T1 weighting, this sequence is obtained by setting IRM to shorts TE defined as the time between delivery of the RF pulse and reception of the echo signal, and a longer TR which presents the duration between successive pulse sequences applied to the same slice.

2. Methodology

The pre-processing of the data presents the second stage of the study through which the processing of the acquired data is carried out. In our case, we have radiological images (MRI) whose pre-processing is done by applying digital algorithms to them. The choice of its processing algorithms may depend on the properties and type of the chosen radiological image.

The determination of the relative intensity (brightness) of the tissues in an MRI image is made by the following factors:

- Gradient waveforms and radio frequency pulse to use.
- The intrinsic characteristics T1 and T2 of the different tissues that means the duration of TR and RF set.
- The proton density of different tissues.

To prepare our MRI database for the training step of the chosen CNNs algorithms we used a nonlinear enhancement algorithm known as histogram equalization which allows contrast enhancement of samples with low luminosity. Its principle is that it makes it possible to separate the pixels into distinct groups if there are few output values over a wide range. It is only effective when the original image has low contrast to begin with; otherwise, histogram equalization can degrade image quality. In this case, the adaptive histogram equalization is improved. This procedure facilitates the detection of image features by CNN algorithms.

Then, we proceeded to the normalization of the intensities of the MRI images acquired, because these images are not comparable from one MRI, from one subject and from one visit to another, even if the same protocol is used. Which can influence the performance and prediction of the CNN algorithms to be trained. This approach brings the intensities to a common scale [0, 1] of all people diagnosed.

Then we proceeded to convert the IRM image, from a single-channel image (grayscale) to a three-channel image (RGB) to train the CNN transfer learning algorithms implemented in this study.

After performing the pre-processing of the acquired MRI images, we proceeded to the training stage of the applied CNN algorithms [1] based on the transfer learning technique. This study targeted three large deep learning algorithms (VGG19, DenseNet and Xception) already pre-trained on an ImageNet-type database. ImageNet is an image database organized according to the hierarchy of WordNet [10]. It is designed for use in visual object recognition software research.

Before the start of the training step of the three algorithms mentioned on the paragraph above, we ensured the division of the database into two sets with a parameter of test_size equal to 0.3, in other words, 70% of the images in the database were used to train these algorithms and 30% to test them.

To avoid performance degradation issues, a database class balancing method has been applied known as SMOTE. It is an oversampling technique that seeks to increase the size of minority classes. In our case, this technique was set to a value of random state equal to 0.

The VGG-19 algorithm is a CNN algorithm consisting of 16 layers of Conv2D and 5 layers of MaxPooling. The layers of Conv2D make it possible to extract the characteristics of an image by applying filters to its various pixels, on the other hand, the operation of the MaxPooling makes it possible to reduce the information generated by the convolutional layers to store it efficiently.

Xception is the second CNN algorithm applied. It takes the form of a linear stack of depth-separable convolution layers with a very complex architecture [3]. It includes 36 convolution layers forming the basis for network feature extraction, these 36 layers are structured into 14 modules, all of which have residual linear connections around them except for the first and last module.

The role of the residual connections added in this algorithm is to reduce the use of resources during the matrix calculation without modifying the number of parameters and to improve the performance of the algorithm [8].

In this study, this algorithm was applied with a neural network consisting of two fully connected layers identical to those of VGG-19 with the same techniques for reducing the overfitting phenomenon already used. DenseNet presents the last algorithm applied in this study, it consists of a single 7×7 Conv2D layer, 58 3×3 Conv2D, 61 1×1 Conv2D and 4 Average Pooling layers. In our case we added 3 fully connected layers with the use of three Regularizer of 0.0001, three dropout layers respectively set to 0.7, 0.5, 0.2 and two layers of BatchNormalization to minimize the overfitting problem.

As long as the number of layers of the CNN network increases, the algorithm becomes very deep, which generates the problem of the evanescence of the gradient (a very rapid decrease in the values of the gradients during the back propagation causing the cancellation of the gradient and the stop learning) which degrades the training process of the CNN algorithm. This is where this algorithm comes in to solve this problem through the simplification of the connectivity scheme between the layers. We have exceptionally resized the dimensions of the image for this algorithm because it requires 224×224×3 to have good performance.

The three algorithms (VGG-19, Xception and DenseNet121) were trained on tensor flow which is an open-source library for training machine learning algorithms. The training process was carried out based on the parameters as follow: Learning-rate of 0.00001, Adam optimizer, Epochs number of 500, Batch-size of 32 and Categorical cross entropy as function cost.

Table 1. Confusion matrix corresponding to VGG-19 Algorithm, Xception Algorithm and DenseNet121 Algorithm

		Mild_Demented	**Moderate_Demented**	**Non_Demented**	**Very_Mild_Demented**
VGG-19	**TP**	980	980	970	960
	TN	2959	2972	2954	2937
	FP	13	0	24	39
	FN	14	14	18	30
Xception	**TP**	920	940	960	950
	TN	2962	2969	2897	2858
	FP	4	0	74	110
	FN	72	49	27	40
DenseNet121	**TP**	880	870	910	860
	TN	2844	2915	2789	2750
	FP	76	0	132	161
	FN	89	104	58	118

Table 2. Performance of VGG-19 Algorithm, Xception Algorithm and DenseNet121 Algorithm

		Mild_Demented	**Moderate_Demented**	**Non_Demented**	**Very_Mild_Demented**
VGG-19	**Exactness**	**0.99**	**1**	0.99	**0.98**
	Precision	0.99	**1**	**0.98**	0.96
	Recall	**0.99**	0.99	**0.98**	**0.97**
	F1_mesure	**0.99**	0.99	**0.98**	**0.97**
Xception	**Exactness**	0.98	0.98	0.98	0.98
	Precision	**1**	**1**	**1**	**1**
	Recall	0.93	0.93	0.93	0.93
	F1_mesure	0.96	0.96	0.96	0.96
DenseNet121	**Exactness**	0.96	0.96	0.96	0.96
	Precision	0.92	0.92	0.92	0.92
	Recall	0.91	0.91	0.91	0.91
	F1_mesure	0.91	0.91	0.91	0.91

3. Results and discission

After finishing the training step of the algorithms (VGG-19, Xception, DenseNet121) on the parameters noted in table 1. They gave the accuracy results mentioned in the figures 1–3.

To ensure this training step of the three algorithms, we applied a categorical cross entropy type cost function which measures the classification performance of the CNN algorithms. It gives results between 0 and 1, the more its value tends towards 0, the more error between the prediction and the test image is minimal (figure 4–6).

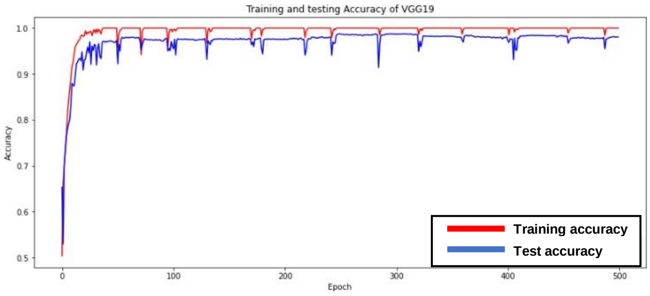

Fig. 1. Accuracy results corresponding to VGG-19

Fig. 2. Accuracy results corresponding to Xception

We need a number of epochs of 500 because a minimum value of this parameter can generate an underfitting problem that arises from the inability of the algorithms to scale their performance during the training phase. And a batch size of 32 to avoid the problem of resource consumption and acceleration of the training speed, because a very large value of this second parameter can accelerate the training speed which sometimes influences the phase learning of the CNN algorithms. However, a very small value of this parameter can slow down the training speed of the CNN algorithms for each epoch which can require more calculation and therefore more memory resources.

A large number of epochs can also cause an overfitting problem. In other words, it can influence the training process of the algorithms by diverging the training accuracy from that of the test. To avoid the appearance of this problem, we applied an optimization algorithm called Adam's algorithm, which updates the variables (Weights) of the neural network, with a value of learning rate equal to 0.00001 which allows to ensure a good adjustment of the rate of changes of these variables (Weights) in order to fix them on values where the errors between the cost function of training and test are greatly minimized.

The values of the overall accuracy of the three applied CNN algorithms were respectively 98%, 95%, 91%. To evaluate the performance of a CNN algorithm, it must first go through the confusion matrix which gives the values of the parameters TP, TN, FP, FN. In our case, we had the following confusion matrix results.

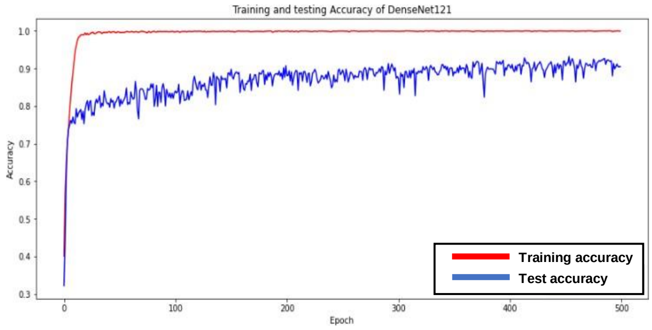

Fig. 3. Accuracy results corresponding to DenseNet121

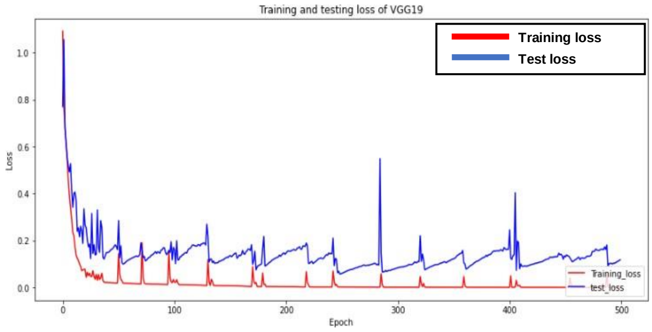

Fig. 4. Results of the cost function corresponding to VGG-19

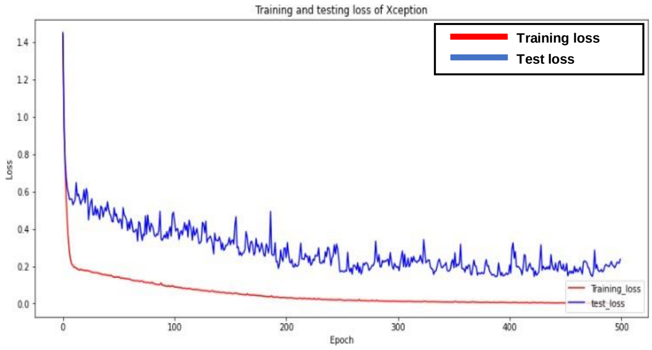

Fig. 5. Results of the cost function corresponding to Xception

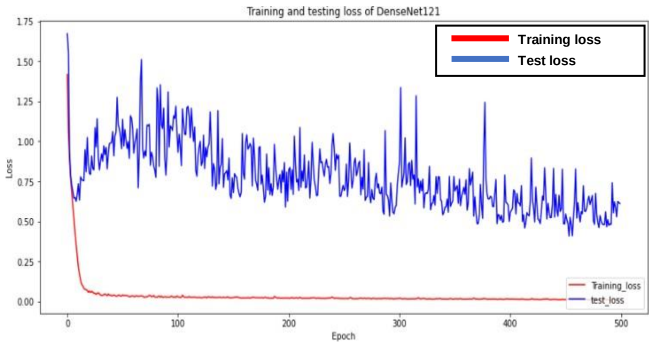

Fig. 6. Results of the cost function corresponding to DenseNet121

After calculating the values of the parameters TP, TN, FP, FN mentioned in the table 1, an evaluation of the performance of the three algorithms (VGG-19, Xception, DenseNet121) was made based respectively on the values of the four metrics accuracy, precision, recall and F1-measurements. Table 2 presents the final results of the performance analysis of the three algorithms used.

If we compare the results of the values of accuracy and precision (table 2) of the VGG-19 algorithm, we will notice that these values for the 4 classes are very close to each other with error values noted respectively 0.01, 0, 0.01, 0.02, and a maximum accuracy of 100% at the class level (Moderate_Demented) which means that this algorithm is very efficient at the level of this class compared to other classes which have completed 99% for Mild_Demented and Non_Demented and 98 respectively % for the Very_Mild_Demented class. For the results of the Xception algorithm, note that the accuracy and precision values for the 4 classes are slightly different from each other with error values noted respectively 0.02, 0.02, 0.03, 0.04, and a maximum accuracy of 98% at the level of the two classes Mild_Demented and Moderate_Demented which means that this algorithm is very efficient at the level of these two classes compared to the two other classes which have respectively achieved accuracy of 97% for Non_Demented and 96% for Very_Mild_Demented. For the results of the last DenseNet121 algorithm, note that the accuracy and precision values for the 4 classes are also different from each other with error values noted respectively 0.04, 0.03, 0.05, 0.07, and a maximum accuracy of 97% at the level of the Moderate_Demented class which means that this algorithm is efficient at the level of this class compared to the other classes which have respectively achieved accuracy of 96% for Mild_Demented, 95% for Non_Demented and 93% for Very_Mild_Demented. So from these comparison results we can conclude that the most efficient algorithm in this study is the VGG-19.

4. Conclusion

Dementia is a chronic or progressive illness that causes cognitive capacity to degrade beyond what may be expected from the normal consequences of biological aging. Memory, cognition, orientation, understanding, computation, learning capacity, language, and judgment are all influenced. Dementia is currently the sixth leading cause of mortality globally and one of the leading causes of reliance and impairment among the elderly. The primary purpose of this research is to create a diagnostic tool based on artificial intelligence algorithms for the early detection of dementia, especially Alzheimer's disease. This was accomplished by utilizing a database including numerous MRI brain scans from various classes of patients. Also we performed an evaluation of three pre-trained CNN algorithms named VGG-19, Xception and DenseNet121. This study used the transfer learning technique through which the three CNN algorithms gave the accuracy values 98%, 95% and 91% respectively.

References

[1] Albawi S. et al.: Understanding of a convolutional neural network. International conference on engineering and technology (ICET), IEEE, 2017, 1–6.
[2] Brownlee J.: Gentle introduction to the Adam optimization algorithm for deep learning. Machine Learning Mastery 3, 2017.
[3] Chollet F.: Xception: Deep learning with depthwise separable convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, 1251–1258.
[4] Jain R. et al.: Convolutional neural network based Alzheimer’s disease classification from magnetic resonance brain images. Cognitive Systems Research 57, 2019, 147–159.
[5] Kaplan E. et al.: Feed-forward LPQNet based automatic alzheimer's disease detection model. Computers in Biology and Medicine 137, 2021, 104828.
[6] Liu M. et al.: A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer’s disease. Neuroimage 208, 2020, 116459.
[7] Shanmugam J. V. et al.: Alzheimer’s disease classification using pre-trained deep networks. Biomedical Signal Processing and Control 71, 2022, 103217.
[8] Szegedy C. et al.: Inception-v4, inception-resnet and the impact of residual connections on learning. Thirty-first AAAI conference on artificial intelligence, 2017.
[9] World Health Organization, 2012. Dementia: a public health priority. World Health Organization.
[10] Yang K. et al.: Towards fairer datasets: Filtering and balancing the distribution of the people subtree in the imagenet hierarchy. Proceedings of the 2020 conference on fairness, accountability, and transparency, 547–558.

Prof. Achraf Benba
e-mail: achraf.benba@ensam.um5.ac.ma

Was born in El Jadida, Morocco, on 1990. He received the Ph.D. degree in electrical engineering from ENSAM, Mohammed V University, Rabat, Morocco, in 2017. He is a member of E2SN teams at ENSAM, Mohammed V University. His interests are in speech processing for detecting people with neurological disorders and Cardiac abnormalities.

http://orcid.org/0000-0001-7939-0790

Eng. Abdelilah Kerchaoui
e-mail: kerchaoui.abdelilah1999@gmail.com

Engineer Kerchaoui Abdelilah graduated from the National School of Arts and Crafts of Rabat, spatialized in the biomedical field and more precisely in hospital instrumentation. He began his higher education with a DUT in electrical engineering for a period of 2 years from 2017 to 2019 followed by an engineering course in biomedical field for a period of 3 years from 2019 to 2022. He realized development & research projects in medical signal processing and artificial intelligence.

http://orcid.org/0009-0004-9561-5960