Deep Learning of Brain Asymmetry Images and Transfer Learning for Early Diagnosis of Dementia

Nitsa J Herzog[1[0000-0002-4503-615X]] and George D Magoulas[2[0000-0003-1884-0772]]

[1] Department of Computer Science, Birkbeck College, University of London, WC1E 7HZ, UK
[2] Birkbeck Knowledge Lab, University of London, WC1E 7HZ, UK
{nitsa, gmagoulas}@dcs.bbk.ac.uk

Abstract. Advances in neural networks and deep learning have opened a new era in medical imaging technology, health care data analysis and clinical diagnosis. This paper focuses on the classification of MRI for diagnosis of early and progressive dementia using transfer learning architectures that employ Convolutional Neural Networks-CNNs, as a base model, and fully connected layers of Softmax functions or Support Vector Machines-SVMs. The diagnostic process is based on the analysis of the neurodegenerative changes in the brain using segmented images of brain asymmetry, which has been identified as a predictive imaging source of early dementia. Results from 300 independent simulation runs on a set of four binary and one multiclass MRI classification tasks illustrate that transfer learning of CNN-based models equipped with SVM output layer is capable to produce better performing models within a few training epochs compared to commonly used transfer learning architectures that combine CNN pretrained models with fully connected Softmax layers. However, experimental findings also confirm that longer training sessions appear to compensate for the shortcomings of the fully connected Softmax layers in the long term. Diagnosis of early dementia on unseen patients' brain asymmetry MRI data reached an average accuracy of 90.25% with both transfer learning architectures, while progressive dementia was promptly diagnosed with an accuracy that reached 95.90% using a transfer learning architecture that has the SVM layer.

Keywords: Convolutional Neural Network, Support Vector Machines, Brain Asymmetry, Neurodegenerative diseases, Transfer Learning, Dementia.

1 Introduction

Computerized analysis of the clinical data has become an important part of medical diagnostics [1], and with the introduction of Artificial Intelligence methods, a new era in medical and biomedical sciences has started. Computer vision and machine learning algorithms brought revolutionary changes in the diagnosis and treatment of many medical conditions. In particular, the use of artificial neural networks and deep learning has opened new directions to the discovery of complicated patterns in large da-

tasets and has equipped diagnostic imaging tools with additional power to analyze complex imaging data.

An area where diagnostic imaging tools, particularly Magnetic Resonance Imaging (MRI), have received considerable attention is that of diagnosis of brain disorders. MRI scan is the most popular imaging methods as it is non-invasive and does not use radiation during the sessions [2]. This work uses MRI data for the diagnosis of dementia, which has been identified as a rapidly growing problem in public health. Early dementia, or amnestic mild cognitive impairment (aMCI), belongs to the group of neurocognitive disorders and is characterized by some sort of short-time memory loss, language difficulties, lack of reasoning and judgment, hardship coping with daily routines [3]. Approximately 10% of the world population aged between 70 and 79, and 25% of the population older than 80 are diagnosed with MCI and 80% of the patients with aMCI is expected to develop severe dementia, e.g. Alzheimer's Disease (AD), within seven years. The initial stages of the disease are identified by structural markers in the images, and one of the early symptoms of dementia is a change in the level of symmetry between the left and right hemispheres of the brain [4,5]. Brain symmetry changes and their progression can be detected by computer vision algorithms, and the potential and robustness of brain asymmetry features and asymmetry images for machine learning-supported diagnosis was validated in previous work [6].

This paper focuses on the diagnostic potential of deep network architectures when transfer learning is employed. It exploits pretrained CNNs- a neural network architecture introduced in the 1980s – that became more noticeable in MRI processing only during the last decade [7]. A CNN does not require a separate feature engineering process, it allows the processing of MRI data with high-performance and has become the method of choice in multiple medical applications [8]. The transfer learning architectures investigated in this paper build on the so-called AlexNet model [9]. This CNN architecture, together with other deep networks like the VGG, the ResNet variants and architectures based on the Inception models, has been established as a standard pretrained model for deep learning applications in image processing and classification. Pretrained models often perform better than models trained from scratch and their use significantly reduces training, validation and testing time on new tasks [10].

Transfer learning enables transferring the pretrained model's knowledge to a newly established task and then creates a new architecture by adding layers to the pretrained/base model. Training of the extra layers, retraining of the base model or even training of the whole architecture are some of the strategies used to fine-tuning the transfer learning architecture using task-related data [11].

In this vein, the paper investigates transfer learning architectures that are equipped with extra fully connected Softmax layers or SVMs, for the diagnosis of early dementia in binary and multiclass classification setups.

The rest of the paper is organized as follows. Section 2 provides a short overview of related work in medicine using artificial neural networks for clinical diagnosis. Section 3 presents the MRI data sources used in the study. Brain asymmetry image processing considerations and transfer learning aspects are described in Section 4. Experiments and results are presented in Section 5. The paper ends with a discussion and future work.

2 Related work

The classification of medical pathologies using Artificial Neural Networks (ANN) and Deep Learning (DL) has attracted a lot of attention in recent years. In this context, the CNN is one of the models successfully adapted to classify imaging data in numerous applications [12]. This section presents a brief overview of recent CNN applications and transfer learning architectures, which employ CNN, relevant to the context of this paper.

In [13] a CNN-based approach that predicts AD, progressive cognitive mild impairment, and stable cognitive impairment was built and evaluated. Authors used transfer learning, which exploits pretrained initial weights, and were able to reduce the training time and increase the network's performance. For binary classification of the converted (to Alzheimer's Disease) form of mild cognitive impairment (c-MCI) vs healthy cognitively subjects (HC), and stable form of MCI (s-MCI) vs HC the obtained accuracy was 87% and 76% respectively.

Multi-Layer Perceptron and a Convolutional Bidirectional Long Short-Term Memory (ConvBLSTM) model were proposed for the diagnosis of dementia in [14]. Different clinical sources and protocols of 1851 participants of the ADNI database were combined for the classification process. The predictive results were obtained using Monte Carlo simulations [15] and demonstrated the highest accuracy of 86%.

Another study [16] proposed an unsupervised DL method for the classification of AD, MCI, and NC (Normal Cognitively). This approach extracts features with PCA (Principal Component Analysis) [17] and processes them with a Regularized Extreme Learning Machine (RELM) [18]. The investigators chose high-level features using the Softmax function (a function that takes a vector of real numbers as input and normalizes it into a probability distribution). The results of the RELM are compared with multiple kernel SVM and Import Vector Machine (IVM) (an IVM classifier based on Kernel Logic Regression uses a few data points to define the decision hyperplane and has a probabilistic output). The study confirmed that RELM improves the classification accuracy of AD and MCI from 75.33% to 80.32% for binary classification and 76.61% for multiclass classification.

Hybrid architectures based on CNN models were successfully implemented in other medical areas. For example, a combination of the Fast Region-Based Convolutional Neural Networks (R-CNN) and SVM for the detection and categorization of brain tumor improved the performance of the CNN by 1.3% on average and accelerated the speed of the training [19]. A decision fusion approach was proposed for the diagnosis of electrocardiogram (ECG) abnormalities using features extracted by CNN, and a final decision-making stage based on SVM [20]. Transfer learning methods applied to two CNN models, the Visual Geometry Group model 16 (VGG16) and the InceptionV3 network, demonstrated performance of 94% for classification of pneumonia, which improved the baseline test by 16% [21].

3 MRI Data Repositories

MRI data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu), which was launched in 2003 as a public-private partnership led by Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of Mild Cognitive Impairment and early Alzheimer's disease for up-to-date information, (see www.adni-info.org).

The created datasets include T1-waited images of MRI data of 150 subjects at the age between 55 and 65 for the Early Mild Cognitive Impairment (EMCI) and Normal Cognitively (NC), and at the age between 65 and 90 for Alzheimer's Disease (AD).

The second MRI data repository is the Open Access Series of Imaging Studies (OASIS), www.oasis-brains.org. The OASIS brain project was created by Washington University in 2007. The OASIS-2 longitudinal collection of T1-waited MRIs of 150 very mildly demented (VM-D) and non-demented (non-D) right-handed subjects, aged between 60 and 80, was used for the tests.

4 Proposed Approach

The diagnostic pipeline includes MRI image processing stages and classification based on transfer learning architecture (Fig. 1) and was implemented in Matlab using commodity hardware (Windows10 Enterprise, Intel (R) Core (TM), i7-7700 CPU@ 3.60 GHz, 16 GB RAM).

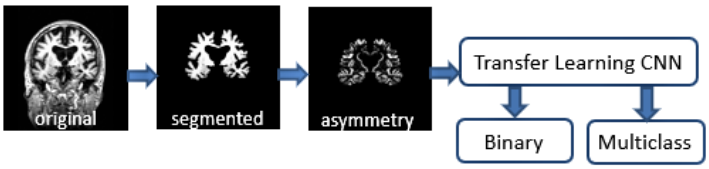

Fig. 1. The MRI data processing pipeline includes image processing stages for generating brain asymmetry images followed by deep learning classification.

4.1 MRI Image Analysis and Segmentation of Brain Asymmetry

Image processing starts from image normalization and brain segmentation from the skull and is followed by segmentation of brain asymmetries. Asymmetry segmentation was proposed in [6] for the detection of early dementia. The method is based on the theory of changes of brain symmetry in the process of loss of the gray and white matter due to initialization and development of the neurodegenerative disorder. The early changes in the brain structure contribute to increasing the symmetry between the

left and right hemisphere of the brain. Vice versa, the progression of the disease increases the degree of asymmetry. In other words, the left-sided hemispheric lateralization inherent to healthy people become gradually right-sided in the process of the development of severe dementia. One of the most common types of this sort of dementia is Alzheimer's Disease. Fig. 2 illustrates the anatomical differences between the left and the right hemisphere of the brain of a cognitively normal person.

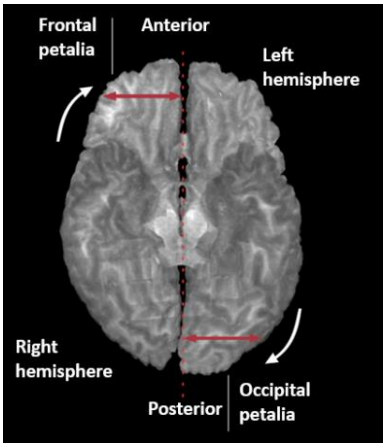

Fig. 2. Normal brain lateralisation (the Yakovlevian anticlockwise torque). (Source: CDI, Oswaldo Cruz German Hospital - Fleury Group / São Paulo 2015.)

Segmentation finds differences in the symmetry between the left and right hemispheres of the brain. The main part of the method is focused on the detection of the vertical line of brain symmetry. This line is assigned for mirroring the left hemisphere to the right and repeating the same process in the opposite direction. The next step is the subtraction of the original image from its reflected version. As a result of matrix transformation asymmetrical parts of the brain become segmented from the image. Asymmetrical areas have a different level of pixel intensity values according to the level of asymmetry. Details of the image analysis and asymmetry detection were introduced in [6]. The visual changes of an image are shown in Fig. 3.

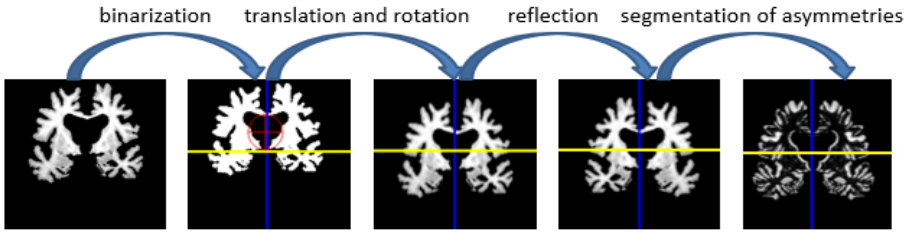

Fig. 3. Transformation stages leading to detection and segmentation of image asymmetries.

4.2 Transfer Learning Architecture

Labelled images of segmented asymmetries were equally divided into three binary classification tasks, AD vs NC, EMCI vs NC and AD vs EMCI, and a multiclass classification task, namely AD, EMCI and NC.

A pretrained CNN has been chosen as the base model of the classification system. The so-called AlexNet architecture has shown in the literature very good performance in image classification tasks (see Section 1). AlexNet has 60 million parameters, 650,000 neurons [22]. The architecture has a total of eight deep layers: five convolutional layers that are used for feature engineering and three fully connected layers. Two new concepts for CNN, which are now considered as a standard, have been introduced in this network: Local Response Normalization (LRN) and dropout. LRN is based on a non-linear function, known as Rectified Nonlinearity Units (ReLUs). ReLUs normalize the feature maps by removing the negative values. This approach significantly increases the speed of training. The dropout function reduces the model overfitting by setting to zero output of those hidden neurons, which have a probability of 0.5 and below. The 1-st convolutional layer requires an input image of size 227-by-227-by-3, where 3 is the number of RGB color channels. This layer includes 96 filters (convolutions) of size 11-by-11-by-3. The 2-nd layer has 256 filters of size 5-by-5-by-48. The 3-rd layer has 384 filters of size 3-by-3-by-256. The 4-th and 5-th layers have 384 and 256 convolutions of similar sizes of 3-by-3-by-192. The 1-st, 2-nd, and 5-th convolutions are followed by normalized max-pooling layers, which perform a down-sampled operation on the input layers. It leaves the feature maps unchanged but significantly reduces the image dimensionality. Fully connected layers compute a score of each class collected from convolution features. The last layer of AlexNet was originally configured to 1000 classes as it was trained to solve a different classification problem for the ImageNet Large-Scale Visual Recognition Challenge 2012.

Two diagnostic pathways were investigated using transfer learning. The first integrates Softmax layers to calculate the output, while the second makes a class prediction based on SVM. To this end, the last 1000-output classification layer of the AlexNet is replaced by a fully connected Softmax layer, or a fully connected SVM to fit the needs of the binary/multiclass diagnosis task. The fine-tuning strategy is based on retraining the whole architecture. The SVM layer was assembled as a separate classification module. Fig. 4 shows the architectures in compact form.

Fig. 4. Transfer learning architecture used in the experiment.

5 Experimental Study

A set of five MRI classification problems, 4 binary and 1 multiclass, were studied using the two transfer learning architectures described in Section 4.2. Binary classification datasets include 400 images of segmented asymmetries of 100 subjects of the ADNI database and 200 images of 150 subjects of the OASIS database. The multiclass dataset consists of 600 images of 150 subjects equally divided into 3 classes of the ADNI database. In all experiments, MRI data were balanced across classes: 600 images of brain asymmetries with a balanced number of AD, EMCI, and NC subjects are combined into the multiclass set and three binary datasets, EMCI vs. NC, AD vs. NC, and AD vs. EMCI. 200 images of the OASIS database compose the balanced binary set of non-D vs VM-D subjects. For each task, five independent runs of each architecture were performed increasing the number of training epochs from 5 to 40 (a total of 300 independent simulation runs). In the rest of the section, the first transfer learning architecture is denoted by M1-Softmax (Softmax layers) while the second one is denoted by M2-SVM (SVM layer). The training parameters were set as: n (n = 5, 10, 15, 20, 30, 40) is the number of epochs each architecture was trained (five independent runs conducted in each case), mini-batch size= 128, validation data frequency= 50, initial learning rate= 0.0001, and the stochastic gradient descent with momentum (SGDM) was used. The two architectures were retrained using segmented asymmetry images, whilst the final SVM layer was fine-tuned after the rest of the architecture was trained. The images of segmented asymmetry of size 256-by-256-by-3 were resized to 227×227×3 and fed into the model with 80% of the images used for training, 10% for validation, and 10% for testing. All the test results presented in this section are for unseen MRI data (unknown patients).

5.1 Experiment 1: ADNI repository

Three binary problems (AD vs EMCI, EMCI vs NC, AD vs NC) and one multiclass problem were used in this experiment. 360 images (for binary) and 540 images (for the multiclassification task) were used for training, and 40 and 60 images of unseen data were used for the testing respectively. The best available classification results on unseen data for the two architectures are presented in Table 1 in terms of accuracy (%), AUC (Area under the ROC Curve) and F1-score.

Table 1. The best available performance (%) in testing for the two transfer learning architectures and the number of epochs each architecture was trained (independent runs) in the classification tasks.

Evaluation metrics	AD vs NC		EMCI vs NC		AD vs EMCI		Multiclass (AD, EMCI, NC)	
	M1-Soft-max	M2-SVM	M1-Soft-max	M2-SVM	M1-Soft-max	M2-SVM	M1-Softmax	M2-SVM
Epoch 5								
accuracy	76.50	91.50	80.25	81.90	79.00	81.25	62.33	73.00
AUC	0.89	0.97	0.87	0.87	0.87	0.91	0.76	0.86

F1-score	0.77	0.91	0.80	0.82	0.78	0.80	0.62	0.73
Epoch 10								
accuracy	94.00	95.25	81.68	82.75	79.00	81.25	72.22	73.75
AUC	0.97	0.98	0.90	0.89	0.87	0.89	0.83	0.87
F1-score	0.94	0.95	0.83	0.82	0.79	0.86	0.73	0.74
Epoch 15								
accuracy	94.25	95.90	81.00	85.25	81.25	81.00	72.22	74.22
AUC	0.95	0.97	0.92	0.93	0.95	0.95	0.87	0.88
F1-score	0.94	0.96	0.81	0.85	0.82	0.82	0.73	0.74
Epoch 20								
accuracy	93.00	93.58	85.50	86.94	81.50	81.25	79.78	79.78
AUC	0.95	0.95	0.92	0.93	0.89	0.88	0.89	0.87
F1-score	0.93	0.94	0.86	0.89	0.82	0.81	0.79	0.79
Epoch 30								
accuracy	92.75	93.33	90.25	90.25	82.75	83.00	82.33	81.75
AUC	0.98	0.97	0.96	0.94	0.90	0.92	0.91	0.89
F1-score	0.93	0.93	0.90	0.90	0.83	0.83	0.82	0.81
Epoch 40								
accuracy	92.70	92.25	88.98	88.98	83.00	83.00	84.75	84.50
AUC	0.98	0.97	0.95	0.93	0.92	0.91	0.92	0.92
F1-score	0.92	0.92	0.89	0.89	0.83	0.83	0.85	0.84

Table 1 demonstrates that the architecture with the SVM layer (M2-SVM) is capable of good performance within a few training epochs. M1-Softmax appears to take advantage of longer training sessions to compensate for its initial shortcomings. The best performance, for example, for the AD vs NC task is 95.90% in accuracy, 0.97 in the AUC, and 0.96 in the F1-score obtained by an M2-SVM that was trained for 15 epochs. In the EMCI vs NC task, the best performance achieved by both classifiers was when they were trained for 30 epochs. In the AD vs EMCI task, an M2-SVM shows the best efficiency when trained for 30 epochs. As the number of epochs increases to 40, M2-SVM and M1-Softmax reach the same accuracy and F1-score while they are very closed results in the AUC metric. In the multiclass task, the best performance is obtained with an M1-Softmax trained for 40 epochs.

Looking at the average accuracy in testing across all tasks, M2-SVM outperforms M1-Softmax by 8.10% (on average), when trained for 5 epochs, by 1.93% (on average) when trained for 10 epochs, and by 1.31% (on average) trained for 15 epochs. The differences in average test performance per task are analyzed in Table 2.

The two architectures show similar or very close results in testing when they are trained for more than 20 epochs. For example,

Fig. **5** illustrates the average classification performance (mean classification error rate on unseen patients' MRI data) for the two architectures depending on how long they were trained (epochs) to solve the EMCI vs NC binary classification task.

Table 2. The difference in the average classification accuracy (%) between M2-SVM and M1-Softmax depending on the length of the training session (epochs).

Number of epochs trained	AD vs NC	EMCI vs NC	AD vs EMCI	Multiclass

5	10.83	5.81	5.42	10.28
10	4.17	1.25	1.91	0.28
15	1.34	1.41	2.50	~ 0

The AUC across all runs in the ADNI experiments (binary and multiclass) is in the range of 0.8338 and 0.9763. The best performance received with both transfer learning architectures for the diagnosis of early dementia (EMCI vs NC) in terms of accuracy is 90.25% with an AUC of 0.9256 when trained for 30 epochs. Progressive dementia (Alzheimer's Disease)-(AD vs NC) is diagnosed with an accuracy of 95.90% and an AUC of 0.9763 with M2-SVM trained for 15 epochs.

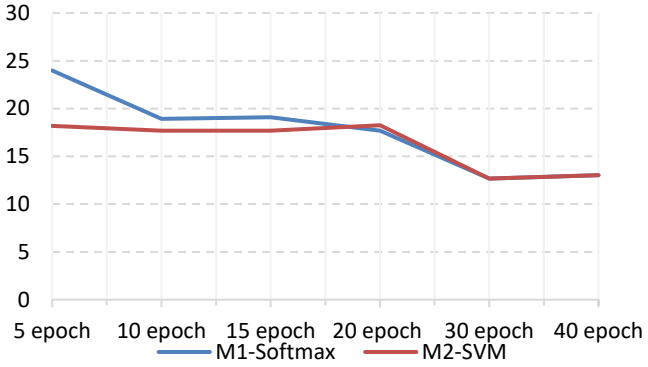

Fig. 5. Average test error rate (vertical axis) in the EMCI vs NC binary classification problem with respect to the number of epochs the two architectures were trained for.

5.2 Experiment 2: OASIS database

This database was used as a second source of MRI data to validate the behaviour of the two transfer learning architectures and the findings of Experiment 1. The OASIS-2 MRI collection is not organized like the ADNI database, so there are no EMCI and AD images to perform EMCI vs NC or AD vs NC binary classification. The dataset is separated into classes of non-Demented (non-D) patients, and patients with Very Mild dementia (VM-D) manually, according to the Clinical Dementia Rating (CDR) score equal to 0.5. The best classification results on unknown patient data are in Table 3.

Table 3. The best available classification accuracy (%) and AUC for the two transfer learning architectures in binary classification (non-D vs VM-D).

Epoch number	5	10	15	20	30	40
M1-Softmax						
Accuracy	61.50	78.25	71.50	76.00	80.00	79.00
AUC	0.80	0.86	0.84	0.86	0.88	0.87
M2-SVM						
Accuracy	76.50	80.00	82.50	75.00	80.00	78.50
AUC	0.85	0.89	0.90	0.86	0.92	0.87

Table 3 confirms the fast response of the M2-SVM architecture producing better classification results within a few training epochs. M2-SVM outperform M1-Softmax by 15% when trained for 5 epochs, 1.75% when trained for 10 epochs, and 11% when trained for 15 epochs, but as the training sessions duration gets longer the difference diminishes as happened in Experiment 1. AUC results support these findings. In particular, the M2-SVM result in epoch 30 represents the best available performance in the experiments with an accuracy of 80% and an AUC of 0.92. Fig. 6 exhibits the average classification performance (mean classification error rate on unseen patients MRI data) of the two transfer learning architectures with respect to the number of epochs were trained for. It shows that on average there are no performance differences between the two models when they are trained more than 20 epochs. At the same time, Fig. 6 shows that on average M2-SVM models are able to produce good results within a few training epochs. The best performance on average is achieved by M2-SVM models trained for 15 epochs: an averaging accuracy of 78.21% and an AUC= 0.8525 for VM-D vs non-D.

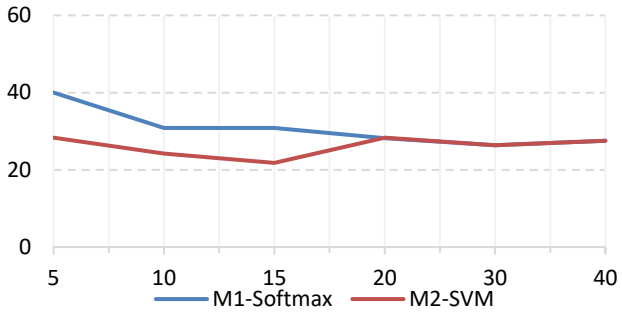

Fig. 6. Average test error rate (vertical axis) and the number of epochs (horizontal axis) of two transfer learning architectures for the VM-D vs non-D binary classification problem.

6 Discussion and Conclusions

The study investigated the effectiveness of transfer learning integration into an MRI processing pipeline for analysis of brain asymmetry and diagnosis of early dementia. Transfer learning appears to give an advantage in terms of computational time spent on training a classifier, although it is not possible to make a direct comparison with other approaches for diagnosis of dementia in the literature as MRI feature engineering and hardware differ [6]. Experiments using segmented asymmetry images from the ADNI and OASIS MRI repositories revealed that the M2-SVM architecture outperforms the M1-Softmax architecture in short training cycles (less than 15 epochs). In longer training cycles, the two architectures show no significant differences in models' efficiency. The approximate time for each training, validation and testing round on commodity hardware (Windows10 Enterprise, processor—Intel (R) Core

(TM), i7-7700 CPU@ 3.60 GHz, 16 GB RAM) is illustrated in Fig.7. Also, in the experiments, the M2-SVM architecture showed an ability to perform well using relatively small training datasets, compared to the M1-Softman architecture. This may be useful in situations where large clinical datasets are not readily available. In the literature, several studies reported limited access to clinical data [23]. Hence, when the imaging datasets are limited in size and time for testing is restricted the M2-SVM gets the advantage. Future work will investigate the potential of transfer learning architectures as the size of the MRI dataset grows beyond the 600 images used in this study.

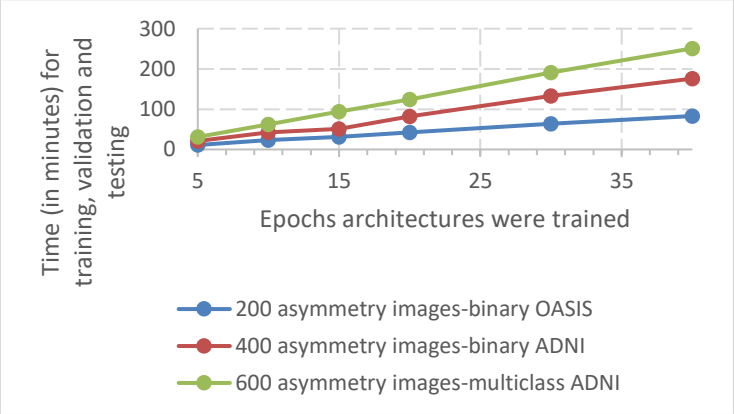

Fig. 7. Approximated time (in minutes on average) spent on training, validation and testing across the classification tasks

References

1. Zhang, Z., Xie, Y., Xing, F., McGough, M. and Yang, L.: Mdnet: A semantically and visually interpretable medical image diagnosis network. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 6428-6436. IEEE, Honolulu, Hawaii, USA (2017).
2. Segato, A., Marzullo, A., Calimeri, F., De Momi, E.: Artificial intelligence for brain diseases: A systematic review. APL bioengineering, 4(4), p.041503 (2020).
3. Janelidze, M., Botchorishvili, N.: Mild Cognitive Impairment. Alzheimer's Disease: The 21st Century Challenge, 91, IntechOpen (2018).
4. Yang, C., Zhong, S., Zhou, X., Wei, L., Wang, L., Nie, S.: The abnormality of topo-logical asymmetry between hemispheric brain white matter networks in Alzheimer's disease and mild cognitive impairment. Frontiers in aging neuroscience, 9, p.261 (2017).
5. Liu, H., Zhang, L., Xi, Q., Zhao, X., Wang, F., Wang, X., Men, W., Lin, Q.: Changes in brain lateralization in patients with mild cognitive impairment and Alzheimer's disease: A resting-state functional magnetic resonance study from Alzheimer's disease neuroimaging initiative. Frontiers in neurology, 9, p.3 (2018).
6. Herzog, N.J., Magoulas, G.D.: Brain Asymmetry Detection and Machine Learning Classification for Diagnosis of Early Dementia. Sensors, 21(3), p.778 (2021).
7. Lundervold, A.S., Lundervold, A.: An overview of deep learning in medical imaging focusing on MRI. Zeitschrift für Medizinische Physik, 29(2), pp.102-127 (2019).

8. Yamanakkanavar, N., Choi, J.Y., Lee, B.: MRI segmentation and classification of human brain using deep learning for diagnosis of Alzheimer's disease: a survey. Sensors, 20(11), p.3243 (2020).
9. Alom, M.Z., Taha, T.M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M.S., Van Esesn, B.C., Awwal, A.A.S., Asari, V.K.: The history began from alexnet: A comprehensive survey on deep learning approaches. arXiv preprint arXiv:1803.01164 (2018).
10. Tajbakhsh, N., Shin, J.Y., Gurudu, S.R., Hurst, R.T., Kendall, C.B., Gotway, M.B., Liang, J.: Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE transactions on medical imaging, 35(5), pp.1299-1312 (2016).
11. Tang, Y.: Deep learning using linear support vector machines. arXiv preprint arXiv:1306.0239 (2013).
12. Yamashita, R., Nishio, M., Do, R.K.G., Togashi, K.: Convolutional neural networks: an overview and application in radiology. Insights into imaging, 9(4), pp.611-629 (2018).
13. Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., Filippi, M., Alzheimer's Disease Neuroimaging Initiative.: Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. NeuroImage: Clinical, 21, p.101645 (2019).
14. Stamate, D., Smith, R., Tsygancov, R., Vorobev, R., Langham, J., Stahl, D., Reeves, D.: June. Applying Deep Learning to Predicting Dementia and Mild Cognitive Impairment. In IFIP International Conference on Artificial Intelligence Applications and Innovations, pp. 308-319. Springer, Cham (2020).
15. Johansen, A.M., Evers, L., Whiteley, N.: Monte carlo methods. Lecture notes, 200 (2010).
16. Lama, R.K., Gwak, J., Park, J.S., Lee, S.W.: Diagnosis of Alzheimer's disease based on structural MRI images using a regularized extreme learning machine and PCA features. Journal of healthcare engineering 2017(1), 1-11 (2017).
17. Abdi, H., Williams, L.J.: Principal component analysis. Wiley interdisciplinary re-views: computational statistics, 2(4), pp.433-459 (2010).
18. Ding, S., Xu, X., Nie, R.: Extreme learning machine and its applications. Neural Computing and Applications, 25(3), pp.549-556 (2014).
19. Khairandish M.O., Gurta R., Sharma M.: A hybrid model of faster R-CNN and SVM for tumor detection and classification of MRI brain images. International Journal of Mechanical and Production Engineering Research and Development, 10(3), 6863–6876 (2020).
20. Singh, S.A., Majumder, S.: Short and noisy electrocardiogram classification based on deep learning. In Deep Learning for Data Analytics, pp. 1-19. Academic Press (2020).
21. Yadav, S.S., Jadhav, S.M.: Deep convolutional neural network based medical image classification for disease diagnosis. Journal of Big Data, 6(1), pp.1-18 (2019).
22. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6), pp.84-90 (2017).
Strongman, H., Williams, R., Meeraus, W., Murray-Thomas, T., Campbell, J., Carty, L., Dedman, D., Gallagher, A.M., Oyinlola, J., Kousoulis, A., Valentine, J.: Limitations for health research with restricted data collection from UK primary care. Pharmacoepidemiology and drug safety, 28(6), pp.777-787 (2019).
23. Strongman, H., Williams, R., Meeraus, W., Murray-Thomas, T., Campbell, J., Carty, L., Dedman, D., Gallagher, A.M., Oyinlola, J., Kousoulis, A., Valentine, J.: Limitations for health research with restricted data collection from UK primary care. Pharmacoepidemiology and drug safety, 28(6), pp.777-787 (2019).