A Transfer Learning Approach to Classify the Brain Age from MRI Images

Animesh Kumar, Pramod Pathak, and Paul Stynes(✉)

National College of Ireland, Dublin, Ireland
x18184731@student.ncirl.ie, {Pramod.Pathak,Paul.Stynes}@ncirl.ie

Abstract. Predicting brain age from Magnetic Resonance Imaging (MRI) can be used to identify neurological disorders at an early stage. The brain contour is a biomarker for the onset of brain-related problems. Artificial Intelligence (AI) based Convolutional Neural Networks (CNN) is used to detect brain-related problems in MRI images. However, conventional CNN is a complex architecture and the time to process the image, large data requirement and overfitting are some of its challenges. This study proposes a transfer learning approach using InceptionV3 to classify brain age from the MRI images in order to improve the brain age classification model. Models are trained on an augmented OASIS (Open Access Series of Imaging Studies) dataset which contains 411 raw and 411 masked MRI images of different people. The models are evaluated using testing accuracy, precision, recall, and F1-Scores. Results demonstrate that InceptionV3 has a testing accuracy of 85%. This result demonstrates the potential for InceptionV3 to be used by medical practitioners to detect brain age and the potential onset of neurological disorders from MRI images.

Keywords: Transfer learning · InceptionV3 · Neurological disorder · Brain age · MRI images

1 Introduction

The World Health Organization (WHO) indicates that 50 million people are suffering from neurological disorders such as Alzheimer Disease (AD). Medical practitioners are able to deduce the physiological or biological age of a person as the brain structure changes over time [8] which can assist with the early detection of neurological disorders. Age estimation can be performed based on either cortical anatomy or MRI images of grey matter, white matter, and cerebrospinal fluid present inside the brain as shown in Fig. 1.

Recently, AI-based CNN is revolutionizing the way medical data such as MRI Images are analyzed [1, 12, 16, 17]. However, conventional CNN is a complex architecture and the time to process the image, large data requirement and overfitting are some of its challenges [3, 13]. Transfer learning (TL) is a research problem in machine learning (ML) that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem where training data could be partially or completely

L. Bellatreche et al. (Eds.): BDA 2020, LNCS 12581, pp. 103–112, 2020.
https://doi.org/10.1007/978-3-030-66665-1_8

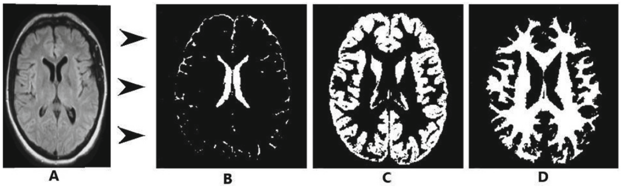

Fig. 1. Brain MRI segmentation - A. T1-weighted brain MRI, B. Cerebrospinal fluid, C. Grey matter, D. White matter

different from the testing data [14]. There are several types of Transfer Learning models such as InceptionV3 and DenseNet. Inceptionv3 is a convolutional neural network for assisting in image analysis and object detection. DenseNet is a type of convolutional neural network that utilizes dense connections between layers, through Dense Blocks, which connect each layer to every other layer in a feed-forward fashion. A DenseNet architecture is a logical extension of ResNet. Transfer learning (TL) based models such as InceptionV3 and DenseNet have been shown to decrease the processing time and model overfitting respectively [14]. It is generally used for smaller datasets for more accurate predictions.

The aim of this research is to investigate a novel transfer learning approach to classify brain age from the MRI images in order to improve the brain age classification model. The transfer learning approach uses InceptionV3. The model performance is evaluated through accuracy, precision, recall, F1-Scores, and processing time.

The contribution is a novel transfer learning approach that classifies age groups into six different categories such as 10 to 20, 21 to 30 years of age and so on where the age categories are known as bins.

This paper discusses related work in Sect. 2 with a focus on machine learning approaches to medical image classification. Section 3 discusses the research methodology used in this research. Section 4 discusses the results and discussion. Section 5 concludes the research and discusses future work.

2 Related Work

This section discusses various approaches to medical image classification such as machine learning, deep learning, and transfer learning.

Originally, analysis of the brain was mostly dependent on regression-based algorithms such as Relevance Vector Regression (RVR) [5, 6] in order to classify MRI images.

Franke et al. [5] used the RVR method to study the impact of diabetes mellitus on brain age using MRI images whilst Gaser et al. [6] used the RVR method to detect Alzheimer Disease (AD). The RVR method has a self-learning process to decide the parameters for the best model fit [6]. The RVR method had a limitation of skipping white matter lesions which is a biomarker for brain age prediction. Besteher et al. [2] included depression parameters to analyze changes in brain age. Although, the result demonstrated no major deviation in brain age. Nakano et al. [11] proposed a comparison between normal and abnormal development in a newborn baby. The model used two architecture Principal

Component Analysis Regression (PCAR) and Manifold Learning-PCAR (ML-PCAR), where ML-PCAR proved to be more accurate. However, the model failed to differentiate between 0 to 3-month-old new-borns due to the different diameters of the brain MRI.

Recently, AI has enhanced the analysis of MRI images using deep learning with techniques such as enhanced dimensionality reduction and feature extraction incorporated into the image classification models such as CNN. The CNN model has been used in the analysis of 2D [8] and 3D [12] MRI images.

Huang et al. [8] proposed a deep learning model VGG Net based on the CNN model to estimate the age of a person based on the brain MRI image. The model was applied to the IXI dataset. The results were comparable to recent research. However, the model was limited to healthy brain images and did not consider unhealthy MRI images. The research would not be able to identify if a 10-year-old child had a brain MRI of a 70-year-old person, which would indicate that the child was unhealthy and may have a neurological disorder.

Qi et al. [12] proposed an enhanced 3D CNN model with an added dense block (sub-DenseNet model) to estimate the age from MRI images. Their model showed that it helped to minimize the gradient vanishing problem and increases the fitting ability of the model. The use of a sub DenseNet demonstrates the potential to solve the overfitting problem.

Ueda et al. [16] proposed a 3D CNN model to estimate the age from brain MRI images. They used an Aoba medical center collected dataset of 1000 MRI images. The study reported improved accuracy. However, the 3D CNN architecture extracts high dimensionality features from the images at the cost of higher processing time.

Bermudez et al. [1] proposed a novel deep learning approach of using conventional CNN and volumetric feature processors to predict the brain age. They used the OASIS and IXI based datasets. The OASIS dataset images are present in GIF format with segmentation in grey matter, white matter, and cerebrospinal fluid features. The OASIS MRI images are smaller in size and consume less time to process in comparison to the NIFTY format used in the IXI dataset. The MRI images in the OASIS dataset are normalized, bias field corrected, and well documented in advance for research purposes [9] however the IXI dataset is not.

Wang et al. [17] investigated whether gray matter in brain atrophy is an established biomarker for dementia prediction in unhealthy people. This approach demonstrates the need to look at unhealthy people in order to identify Brain MRI segmentation such as grey matter.

Siar and Teshnehlab [15] proposed an AlexNet based CNN model. The age categories were divided into 5 bins ranging from 10 years to 70 years. The model is implemented with three different classification layers (SVM, Decision Tree, SoftMax). The SoftMax demonstrated the highest accuracy of 79% on 1290 images that were self-collected. The accuracy with unequal age categories bins provides an aid to medical practitioners to narrow down the patient with possible neurological disorders as older people tend to have more neurological disorders. The proposed classification of age groups in bins is of interest in this research.

CNN based models were proved to be highly efficient on large datasets using complex architectures like ResNet and AlexNet. However, it takes a longer processing time. CNN models also require a large dataset in order to minimize the overfitting problem. This would indicate a need to look at alternative models to deep learning.

Transfer learning is an approach used to address the problem of multiple domain training. The transfer learning-based models used to extract features from one domain and apply it to another similar domain for model training [14].

Ren et al. [13] proposed a transfer learning-based 3D CNN model that was trained on a UK Biobank[1] dataset that contains 9850 MRI images. The trained weights are used for model training on an NKI[2] dataset having 395 MRI images. The approach of having the transfer learning model trained on a large dataset then the transfer learning model being applied to a smaller dataset for training worked well for solving the overfitting problem.

Transfer learning-based models were also implemented for brain tumor detection [3]. The research compared 9 different transfer learning models for predicting tumor classes. The result demonstrated that the models with fewer layers performed better than models with a higher number of layers like ResNet101 with 3064 images.

Ding et al. [4] proposed a model to diagnose Alzheimer Disease from MRI images in early stages using InceptionV3. The result demonstrated that the model has a lower precision of 55% for mild cognitive impaired (MCI) on an ADNI dataset. The result also demonstrated a precision of 18% with an independent dataset that is a self-collected dataset. The result demonstrated that the model has a dependency on the clinal distribution of MRI images per class for the training dataset.

The literature review indicates that conventional CNN models face challenges such as model overfitting, and processing time with small datasets. The literature also indicates that there is a need to include MRI images from both healthy and unhealthy people in order to classify age estimation based on grey matter, white matter, and cerebrospinal fluid. Alternative approaches such as transfer learning show promise for addressing the large processing time and overfitting. For overfitting, the transfer learning models can be trained on a large dataset such as ImageNet. Then the transfer learning model can be applied to a smaller dataset for training. The literature demonstrates that transfer learning approaches like DenseNet is a useful approach for solving the overfitting problem. The review also demonstrates that InceptionV3 has better performance on a smaller dataset and reduces the overall model processing time. This review demonstrates the need for a classification model that can handle a full Brain MRI segmentation on grey matter, white matter, and cerebrospinal fluid in both healthy and unhealthy people.

3 Research Methodology

The research methodology of this study discusses the step-by-step process as shown in Fig. 2.

The first step involves the collection of data from the OASIS neuro-imaging dataset [9]. The OASIS neuro-imaging dataset consists of 436 different brain MRI images in various masking forms such as increasing the contrast of the MRI images. The data set

1 https://www.ukbiobank.ac.uk/data-showcase/.

2 https://fcon_1000.projects.nitrc.org/indi/enhanced/.

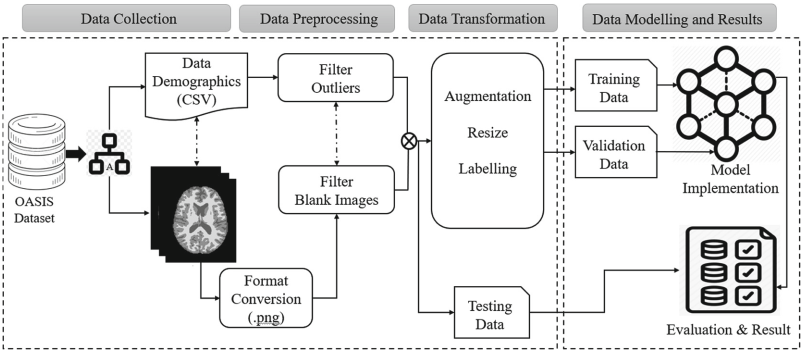

Fig. 2. Process flow diagram

is used for detecting brain age and neurological disorders such as Alzheimer disease. The mean age of the dataset is 51.35 years with age ranging from 18 years to 96 years. The changes in brain contour with ae are shown below as shown in Fig. 3.

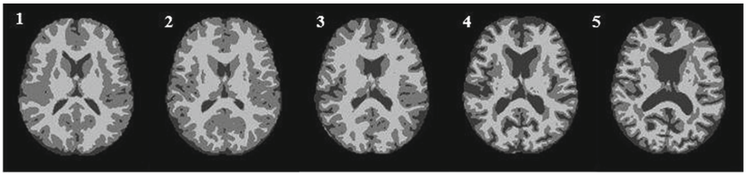

Fig. 3. Change in brain orientation with aging – 1. 20yr 2. 40yr 3. 60yr 4. 80yr 5. 96yr

The second step involves data preprocessing in which the dataset was checked for blank images. If there were blank images, then the corresponding demographic data was removed. The dataset is converted from GIPHY to (Portable Graphic Format) PNG format as the GIPHY short moving images take up 4 times the memory and increase the processing time. The PNG MRI images consisted of 436 raw and 436 masked images. The pre-processed MRI images were resized into 175X175 pixels and 3 channels using OpenCV2 library and saved with their unique ids in different class folders. Outliers were removed leaving an augmented OASIS dataset with 411 raw and 411 masked images as shown in Table 1. Images and labels were categorized into 6 different age group bins as shown in Table 1. Data was split into training and test sets. The test set was taken as 20% of the dataset. The training set was augmented, resized, and labeled as per the model requirement and split into training and validation set with a ratio of 80:20 respectively. Categorized test, train, and validation data were fed for model training and testing. Model weights and information per epochs were stored for evaluation and analysis. Model Evaluation was performed using metrics like accuracy, precision, recall, and F1-Score.

The third step involves data transformation in which the MRI images were augmented with 12 different filters for data augmentation [10] as shown in Fig. 4.

Table 1. Class distribution according to age

Age Groups	10–20	21–30	41–60	61–70	71–80	81–90	Total
Bin	1	2	3	4	5	6	–
Data	50	108	64	45	89	55	411

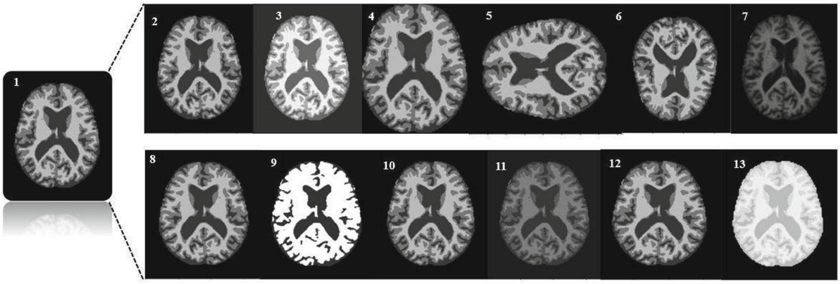

Fig. 4. Data augmentation – 1. Original data 2. Left-right flip 3. Brightness (0.2) 4. Center cropping (0.8) 5. Rotation 90 6. Upside down 7. Random contrast 8. Saturation (10) 9. Adjust contrast (8), 10. Random Hue 11. Segmented 12. Random gamma 13. Random saturation

The fourth step involves data modeling and results. Data modeling involves the implementation of InceptionV3 and DenseNet. The InceptionV3 model has a parallel processing mechanism with 11 stacked inception modules shown in different color coding in Fig. 5. Each block uses the same color inception sub-block having convolutional filters, pooling layers, and activation functions as rectified linear units. The concatenation layer is added with fully connected layers of size 1024 and 9 with SoftMax classifier [4].

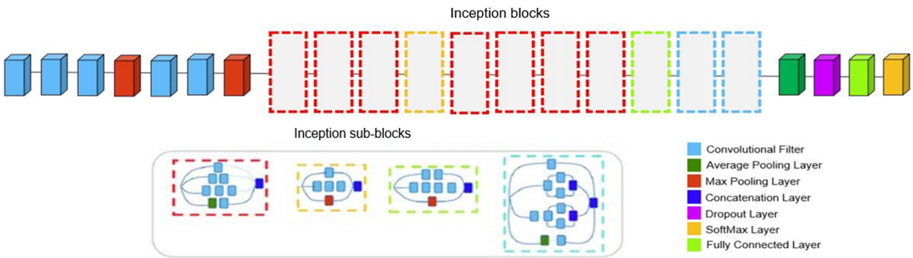

Fig. 5. InceptionV3 network architecture [4]

DenseNet is a deeper model with five dense blocks and feature reuse mechanism using a concatenation network [7]. It contains all the similar feature maps connected to each other to preserve feed-forward nature [7]. Each dense block comprises batch normalization, ReLU activation function, CNN and max-pooling layers. Pretrained models like InceptionV3 and DenseNet have a definite set of layers with varied fully connected layers as per requirement.

The training and validation dataset were passed through the image generator which fetched labels and performed real-time augmentation. The models were implemented in python language using Keras neural network library, integrated on top of TensorFlow framework. The epoch count (100) and image size 175X175 were agreed based on several model iterations and kept constant throughout the experiments for comparison. The total MRI images after augmentation for training are 7896. Model execution logging like accuracy, validation, and losses were stored in CSV file using CSV logger and improved weights were stored in google drive for future reference. An early stopping mechanism was also integrated for efficient model run. A generalized step per epoch and validation per epoch were assigned having values equals to the training or validation count upon batch size.

The InceptionV3 and DenseNet model was supported by pre-trained weights from ImageNet imported using Keras library. The hyperparameters used are batch size of 32, learning rate of 0.0001, categorical cross-entropy as loss function, and Adam as optimizer.

4 Results and Discussion

The aim of this experiment was to compare the accuracy of transfer learning models DenseNet and InceptionV3 in order to improve the brain age classification model. Model performance was evaluated using accuracy, precision, recall, F1-Scores metrics, and processing times as shown in Table 2. In order to make a useful comparison to the DenseNet and InceptionV3 experiments, this research compared the results with the deep learning model proposed by Huang et al. [8].

Table 2. InceptionV3, DenseNet and 2D CNN model comparison.

Method	Testing accuracy	Precision accuracy	Recall	F1-scores	Time (Sec)
2D-CNN	47%	11%	28%	15%	19124
DenseNet	60%	18%	17%	17%	89982
InceptionV3	85%	86%	85%	84%	7418

Huang et al. [8] proposed a 2D-CNN model based on VGG Net to estimate the age of a person based on the brain MRI image using the IXI dataset. This research replicated this experiment and extended the experiment to the augmented OASIS dataset instead of the IXI dataset. The 2D-CNN model was simulated using 5 convolutional layers. Results demonstrate that the testing accuracy with 411 raw MRI images and 411 masked MRI images. The testing accuracy was 47% as shown in Table 2. This indicates a problem of model overfitting due to unequal samples in the validation dataset. These results demonstrate that the CNN model is not suitable for analyzing MRI images on a dataset of 411 raw and 411 masked images.

The DenseNet model accuracy and model loss for the training and validation dataset is shown in Fig. 6. The model achieved a testing accuracy of 60% with precision as low as 18% for the fourth bin (61 to 70 years age) containing 45 MRI. The model accuracy and loss curves are inconsistent as shown in Fig. 6. This indicates a problem of model overfitting due to mis-representation of images using a validation dataset. These results demonstrate that the DenseNet model is not suitable for analyzing MRI images on a dataset of 411 raw and 411 masked images.

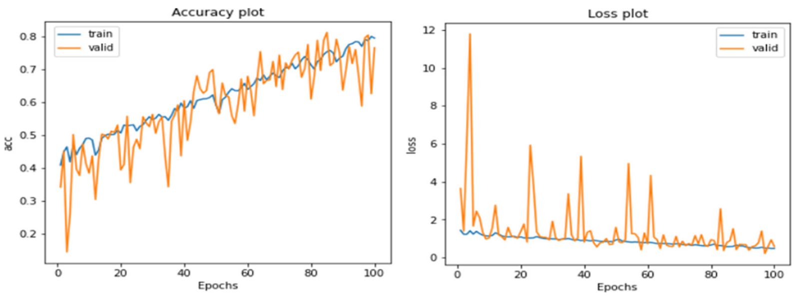

Fig. 6. DenseNet model accuracy and loss curve plots

The InceptionV3 model accuracy and model loss curves for the training and validation dataset is shown in Fig. 7. Both model accuracy and loss are consistent with a constant increase in training and validation accuracy in the accuracy plot as shown in Fig. 7.

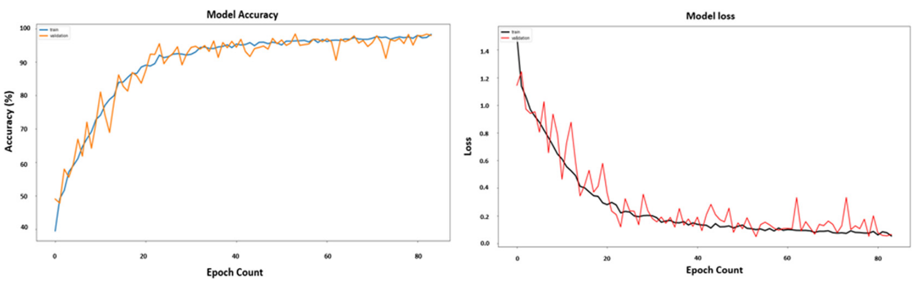

Fig. 7. InceptionV3 accuracy and loss curve

The validation accuracy and training accuracy are comparable which demonstrates that the model is not overfitting. Results show that the InceptionV3 model achieved a testing accuracy of 85% as shown in Table 2. The precision of 86% demonstrates that the model correctly classified MRI images of people with neurological disorders. The precision shows promise for using transfer learning models such as InceptionV3 for the analysis of MRI images. The result demonstrates that the InceptionV3 model's processing time is $1/12th$ the time taken by DenseNet and $2\frac{1}{2}$ times faster than 2D CNN models.

The processing time of the models has a correlation to the number of training parameters. The InceptionV3 model has a lesser number of training parameters than DenseNet. Due to which InceptionV3 processing time is $1/12th$ of DenseNet processing time on GPU. The results demonstrate that the testing accuracy for 2D-CNN is nearly half of the InceptionV3 model. It shows the inefficiency of 2D CNN on small dataset like the augmented OASIS with 411 raw and 411 masked images. The results show promise that InceptionV3 outperforms 2D-CNN and DenseNet on a smaller dataset like the augment OASIS dataset.

5 Conclusion and Future Work

This research proposed a transfer learning approach to classify brain age from MRI images in order to improve the brain age classification model. Two transfer learning models were compared namely, InceptionV3 and DenseNet. Results demonstrate that the InceptionV3 model outperformed DenseNet by 40% with a testing accuracy of 85% using an augmented OASIS dataset with 411 raw and 411 masked images after preprocessing. The results show promise for assisting medical practitioners in the early detection of neurological disorders with small datasets. Future work includes extending this research to investigate the application of transfer learning on larger datasets with 1290 to 3064 images.

References

1. Bermudez, C., et al.: Anatomical context improves deep learning on the brain age estimation task. Magn. Reson. Imaging **62**, 70–77 (2019)
2. Besteher, B., Gaser, C. and Nenadić, I.: Machine-learning based brain age estimation in major depression showing no evidence of accelerated aging. Psychiatry Res. Neuroimaging **290**, 1–4 (2019)
3. Chelghoum, R., Ikhlef, A., Hameurlaine, A., Jacquir, S.: Transfer learning using convolutional neural network architectures for brain tumor classification from MRI images. In: Maglogiannis, I., Iliadis, L., Pimenidis, E. (eds.) AIAI 2020. IAICT, vol. 583, pp. 189–200. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-49161-1_17
4. Ding, Y., et al.: A deep learning model to predict a diagnosis of Alzheimer disease by using ^{18}F-FDG PET of the brain. Radiology **290**(2), 456–464 (2019)
5. Franke, K., Gaser, C., Manor, B., Novak, V.: Advanced brainAGE in older adults with type 2 diabetes mellitus. Front. Aging Neurosci. **5**, 90 (2013)
6. Gaser, C., Franke, K., Klöppel, S., Koutsouleris, N., Sauer, H.: BrainAGE in mild cognitive impaired patients: Predicting the conversion to Alzheimer's disease, PLoS ONE **8**(6), e67346 (2013)
7. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4700–4708 (2017)
8. Huang, T.-W.: Age estimation from brain MRI images using deep learning. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp. 849–852. IEEE (2017)

9. Marcus, D.S., Fotenos, A.F., Csernansky, J.G., Morris, J.C., Buckner, R.L.: Open access series of imaging studies: longitudinal MRI data in nondemented and demented older adults. J. Cogn. Neurosci. **22**(12), 2677–2684 (2010)
10. Mikolajczyk, A., Grochowski, M.: Data augmentation for improving deep learning in image classification problem. In: 2018 International Interdisciplinary Ph.D. Workshop (IIPhDW), pp. 117–122. IEEE (2018)
11. Nakano, R., et al.: Neonatal brain age estimation using manifold learning regression analysis. In: 2015 IEEE International Conference on Systems, Man, and Cybernetics, pp. 2273–2276. IEEE (2015)
12. Qi, Q., Du, B., Zhuang, M., Huang, Y., Ding, X.: Age estimation from MR images via 3D convolutional neural network and densely connect. In: Cheng, L., Leung, A.C.S., Ozawa, S. (eds.) ICONIP 2018. LNCS, vol. 11307, pp. 410–419. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-04239-4_37
13. Ren, Y., Luo, Q., Gong, W., Lu, W.: Transfer learning models on brain age prediction. In: Proceedings of the Third International Symposium on Image Computing and Digital Medicine, pp. 278–282 (2019)
14. Shao, L., Zhu, F., Li, X.: Transfer learning for visual categorization: a survey. IEEE Trans. Neural Netw. Learn. Syst. **26**(5), 1019–1034 (2014)
15. Siar, M., Teshnehlab, M.: Age detection from brain MRI images using the deep learning. In: 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE), pp. 369–374. IEEE (2019)
16. Ueda, M., et al.: An age estimation method using 3D-CNN from brain MRI images. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), pp. 380–383. IEEE (2019)
17. Wang, J.: Gray matter age prediction as a biomarker for risk of dementia. Proc. Natl. Acad. Sci. **116**(42), 21213–21218 (2019)