

Brain Age Estimation from MRI Images using 2D-CNN instead of 3D-CNN

MR Görüntülerinden 3D-ESA Yerine 2D-ESA Kullanılarak Beyin Yaşı Tahmini

Muazzez Buket Darıcı¹ , Şüheda Yıldırım² , Murat Gezer³ 



ABSTRACT

Human Brain Age has become a popular aging biomarker and is used to detect differences among healthy individuals. Because of the specific changes in the human brain with aging, it is possible to estimate patients' brain ages from their brain images. Due to developments of the ability of CNN in classification and regression from images, in this study, one of the most popular state of the art models, the DenseNet model, is utilized to estimate human brain ages using transfer learning. Since this process requires high memory load with 3D-CNN, 2D-CNN is preferred for the task of Brain Age Estimation (BAE). In this study, some experiments are carried out to reduce the number of computations while preserving the total performance. With this aim, center slices of each three brain planes are used as the inputs of the DenseNet model, and different optimizers such as Adam, Adamax and Adagrad are used for each model. The dataset is selected from the IXI (Information Extraction from Images) MRI data repository. The MAE evaluation metric is used for each model with different input set to evaluate performance. The best achieved Mean Absolute Error (MAE) is 6.3 with the input set which consisted of center slices of the sagittal plane of brain scan and the Adamax parameter.

Keywords: Brain Age Estimation, Brain MRI, 2D-CNN

ÖZ

İnsan Beyin Yaşı, son zamanlarda popüler bir yaşlanma biyobelirteci haline geldi ve sağlıklı kişiler arasındaki farklılıkları tespit etmek için kullanıldı. Yaşlanmayla birlikte insan beynindeki spesifik değişiklikler nedeniyle, hastaların beyin yaşlarını beyin görüntülerinden tahmin etmek mümkündür. Evrimsel Sinir Ağlarının (ESA) gelişen görüntü sınıflama ve regresyon yeteneğinden yola çıkılarak, bu çalışmada en popüler ESA modellerinden biri olan DenseNet modeli öğrenme aktarımı yöntemiyle kullanılarak insan beyni yaşı tahmini gerçekleştirilmiştir. 3D-ESA yüksek bellek yükü gerektirdiğinden Beyin Yaşı Tahmin (BAE) görevi için 2D-CNN tercih edilmiştir. Bu deneyde, toplam performans korunurken hesaplama yükünü azaltmak için bazı deneyler yapılmıştır. Bu amaçla, her üç beyin düzleminin merkez dilimleri DenseNet modelinin girdileri olarak kullanılmıştır ve her model için Adam, Adamax ve Adagrad gibi farklı optimizasyonlar kullanılmıştır. Veri kümesi, IXI MRI veri havuzundan seçilmiştir. Performansı değerlendirmek için ortalama mutlak hata (MAE) metriği her model için kullanılmıştır. Bu çalışmada en düşük Ortalama Mutlak Hata (MAE), beynin sagittal düzleminin merkez dilimlerini içeren giriş kümesiyle ve Adamax parametresiyle 6.3 olarak elde edilmiştir.

Anahtar kelimeler: Beyin Yaşı Tahmini, Beyin MR, 2D-ESA

¹(Res. Asst.), Kadir Has University, Faculty of Engineering and Natural Sciences, Electrical - Electronics Engineering, Istanbul, Turkey

²(Master Student), Kadir Has University, Faculty of Economics, Administrative and Social Sciences, Management Information Systems, Istanbul, Turkey

³(Lect. PhD.), Istanbul University, Informatics Department, Istanbul, Turkey

ORCID: M.B.D. 0000-0002-0943-9381;
Ş.Y. 0000-0002-5577-1586;
M.G. 0000-0002-7286-3943

Corresponding author:

Muazzez Buket DARICI,
Kadir Has University, Faculty of Engineering and Natural Sciences, Electrical - Electronics Engineering, Istanbul, Turkey
E-mail address: buket.darici@khas.edu.tr

Submitted: 07.04.2021

Revision Requested: 20.06.2021

Last Revision Received: 08.07.2021

Accepted: 25.07.2021

Published Online: 16.09.2021

Citation: Darıcı, M B., Yıldırım, S., & Gezer, M. (2021). Brain age estimation from MRI images using 2D-CNN instead of 3D-CNN. *Acta Infologica*, 5(2), 373-385.
<https://doi.org/10.26650/acin.911202>

1. INTRODUCTION

Brain tissue has a complex structure, and trying to understand it will provide important advantages in the fight against many diseases and in developing new tools related to this structure. One particular benefit of understanding the brain is that we can gain an insight into the development of the brain. According to studies, the brain experiences some changes throughout the years (Giedd et al., 1999; Lenroot & Giedd, 2006). For that reason, predicting brain age has various benefits for the diagnostic processes of brain-related diseases. Because of this, there have been numerous studies conducted and many methods adopted to grasp various dynamics related to brain age. It has been observed that there is a certain pattern in morphological changes in the brain during brain development and healthy aging. Volumetric changes in brain tissues such as grey matter, white matter (WM) and Cerebrospinal fluid (CSF) have been examined (Jiang et al., 2020; Ito et al., 2018; Li et al., 2018). Studies have shown that the volume of the grey matter decreases with age, while CSF's volume increases, and the volume of WM shows small changes. T1-weighted images, which are a type of MR image, have been used to analyze volumetric changes in brain sciences.

Some brains change outside of ordinary developmental patterns. Particular diseases such as Alzheimer's, dementia and cancer can change the morphological structure of the brain. In such cases, knowledge of ordinary aging provides great benefits. Some studies benefit from the ordinary aging processes while examining cognitive impairment, Alzheimer's and neurological disorders (Liem et al., 2017; Franke & Gaser, 2012; Cole et al., 2017a; Rokicki et al., 2020).

Biomedical data widen with every passing day. Medical doctors have made their decisions manually based on the images or records taken to collect these data. However, these partially subjective decision-making resources and processes lead scientists to look for more standardized approaches than that. In time, scientists from other departments have supported clinical decision-making processes by contributing their expertise. Computer scientists are among the ones who have made substantial contributions to the medical literature. Researchers have advanced some means for computer-aided diagnosis through some machine learning and deep learning techniques (Doi, 2007; Tang et al., 2009). It is a widely supported view that these tools have improved radiologists' performance and diagnostic processes (Yin & Chiu, 2004). Today, scientists and medical experts attach great importance to decision-making processes based on these data.

Studying the brain has its difficulties. It is a three-dimensional entity with multiple parts. Minor changes in these parts could indicate many alterations and transformations in the mind and some other parts of the body. For this reason, observing these changes necessitates a complex and interdisciplinary approach. These changes are aimed to be observed using different methods and tools. T1-weighted images, which are a type of MR image, have been used to analyze volumetric changes in brain sciences. The convolutional neural network is used in many studies, and 3D-CNN and 2D-CNN are some methods that are highly benefited from in relevant studies (T. Huang et al., 2017; Levakov et al., 2020; Symposium & Imaging, 2019; Rossi et al., 2019; Pardakhti & Sajedi, 2020; Peng et al., 2021).

There are some studies in the literature on age estimation from T1-weighted images. In one of these studies, the relationship among age, sex and the volume of the brain was observed (Luders et al., 2009). The results suggested that the water content of the grey matter decreases more harshly in men than women. Additionally, the water content of the female grey matter is on average 1.2% higher than the respective male grey matter water content. In another study, the researchers designed a cortical surface pattern (CSP) combining cortical thickness with curvatures, which constructed an accurate human age estimation model (Wang et al., 2014). Their results suggested that CSP is sensitive to brain development. It is much more powerful than voxel-based morphometry used in previous methods for age estimation.

There are some disadvantages of using 3D-CNN to predict brain age. The initial obstacle is that this method requires high memory load as all brain slices are examined. Even though we have enough memory capacity, the computers that are utilized should be strong enough to compute many units of data simultaneously. Taking these conditions into consideration, it is not always possible to adapt the 3D-CNN model. Two studies in the literature used 2D convolutions instead of 3D convolutions in their model to reduce computational time and memory requirements (Rossi et al., 2019; Hong et al., 2020). Both studies trained their complex models on all sections (slices) of 3D brain planes to predict brain age.

The primary purpose of this study is to achieve a three-dimensional insight into brain age. At first, it was planned to adopt the 3D-CNN model. However, this idea was abandoned because of the difficulties that come with it and not enough memory capacity. Therefore, we tried to change our direction to make our method usable for everybody else (Rossi et al., 2019). Since other methods such as 3D-CNN need high-level hardware configurations and much time to provide a result, we planned to feed our model with MR images of some slices of the brain to overcome this hardship.

In this study, our aim is to create a model using 2D-CNN, which is trained in a shorter time than 3D-CNN and other manually created methods and has close success rates. Additionally, thanks to the method used in this study, it is possible to provide small-scale health institutions a chance to apply this method for the early diagnosis of neurodegenerative diseases without huge investments and long waiting times.

To address the disadvantages of the 3D-CNN approach, in this study, 3D-CNN was replaced by 2D-CNN. Thanks to this approach, it was aimed to reduce memory and computational requirements. Many experiments were conducted to reach the best results in a short time:

- All 3D brain scans were divided into their three planes which are axial, sagittal and coronal. In order to further reduce the computational requirements, instead of using all slices of the axial, sagittal and coronal views, we used only the center slices that represented 2D images that were processed independently in the DenseNet121 model, which consisted of 2D convolutions.
- To decide on which center slice of the brain planes provided better prediction on brain age estimation, the training process of the DenseNet121 model, which consisted of three different input sets, was repeated using different optimizers such as Adam, Adamax and Adagrad.

Our best score was compared to a 3D-CNN (Cole et al., 2017a) and some complex models proposed by Rossi et al. (2019) and Jonsson et al. (2019). The results showed that the approach of using only the 2D center images of each brain plane allows comparable results using significantly less computation memory in much less time to be achieved.

2. DATA

2.1 Dataset

The dataset used in this study consists of brain MR images. T1-weighted MR images were obtained from the IXI dataset. This data repository has about 600 healthy individual MRI images. The acquisition protocol for each sample contains T1, T2, PD, MRA and DTI images. The entire dataset has been gathered at three different hospitals in London, the United Kingdom. The dataset is publicly available and can be accessed at <https://brain-development.org/ixi-dataset/>. This dataset consists of samples from different age groups and different sexes. The numbers of samples for age ranges based on sex may be seen in Figure 1.

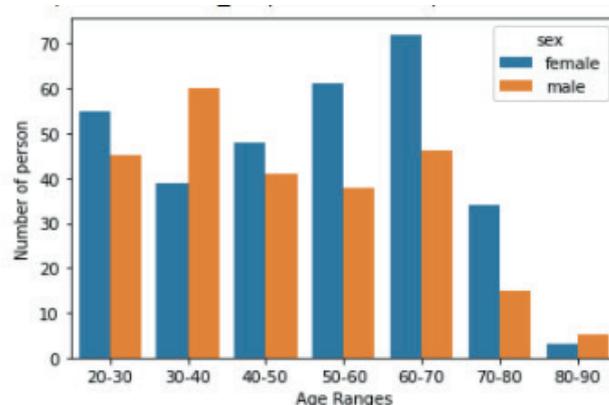
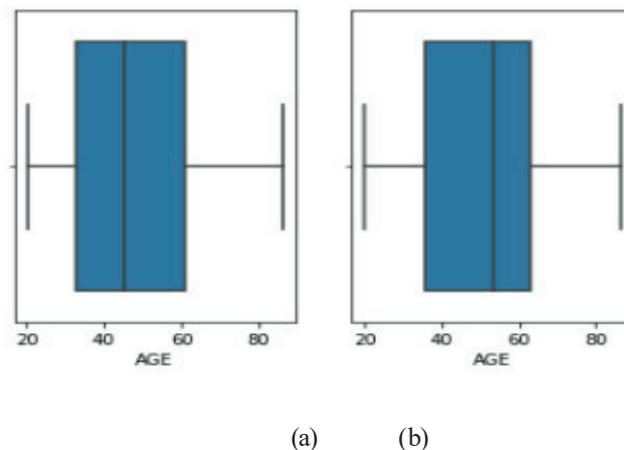


Figure 1. The numbers of samples for each age range in the IXI dataset.

The boxplot shown in Figure 2 provides an idea of the symmetry of the data, how tightly the data are grouped and whether there is skewness in the data. It is shown that the mean age is 47 for males, and it is 50 for females, and the maximum and minimum ages for each sex are the same, which are 86 and 20, respectively.



In this study, we used 563 T1-weighted healthy samples of the IXI dataset, 400 samples for the training set and validation set, and 163 samples as a test set for the analysis.

2.2 Preprocessing

All images in the IXI dataset are in the NIFTI file format. These 3D images should be preprocessed before using them in models. This preprocessing involves the realignment and normalization methods. These preprocessing techniques were applied to the dataset by SPM8. SPM8 is a version of the SPM software. SPM is used to organize and interpret functional neuroimaging data. In order to compare data from several scans or research participants, all brain images have to be in the same 3D space. In SPM, this is accomplished by normalizing the images into the space defined by the MNI template. Realignment is a motion correction process. Realignment adjusts movement between slices. After these two preprocesses, the normalized image dimensions are $79 \times 95 \times 68$.

An MRI head scan provides three MRI planes: coronal, sagittal and axial, as seen in Figure 3 (Padmanaban et al., 2020).

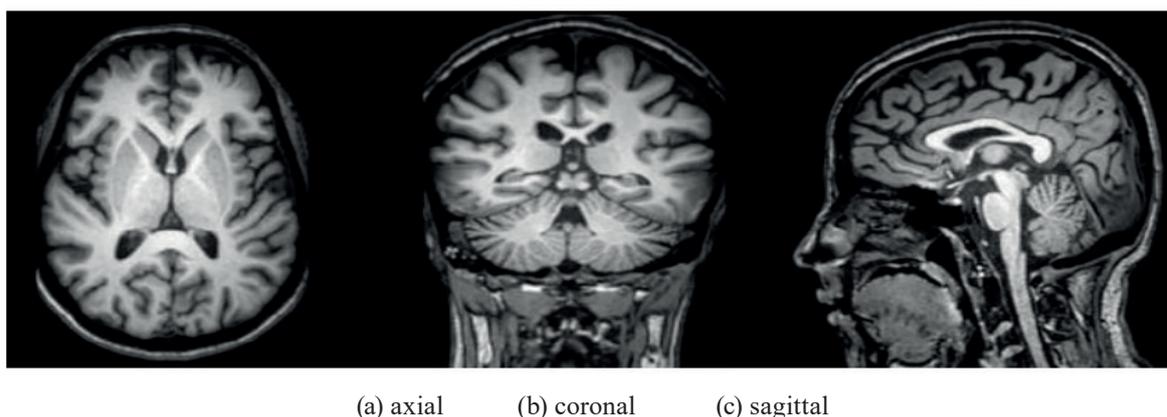


Figure 3. Planes of brain scan.

These planes generate orientations of the head scan. For example, the axial plane generates brain scans from top to down, the coronal plane generates brain scans from front to back, and the sagittal plane generates brain scans from one side to the other laterally. In this study, since the available 3D image dimensions were $79 \times 95 \times 68$, for one brain MRI result, we had 79 95×68 2D sagittal images, 95 79×68 2D coronal images and 68 79×95 2D axial images.

3.METHODOLOGY

In this study, DenseNet121, a type of densely connected convolutional network, was used as the transfer learning approach. Each plane of the brain scan was fed to this model as an input set. The effects of some different optimizers were investigated in terms of the mean absolute error.

3.1 DenseNet121 using Transfer Learning

The Convolutional Neural Network (CNN) is one of the most prevalently preferred deep learning architectures. CNN has been used for different purposes such as image classification, voice recognition, pattern recognition and segmentation in various fields. A simple CNN consists of layers as the convolution layer, pooling layer and fully connected layer (Darici et al., 2020). It is possible to obtain different CNNs by changing the size of the filters used in these layers and the order of these layers. Computer vision, thanks to CNN, has had some notable successes. However, recently, because of improvements in imaging techniques and the increasing size of images, using simple CNNs may not be adequate to obtain successful results. In such cases, the need for more robust and deeper structures increases. For this purpose, some robust architectures such as DenseNet (G. Huang et al., 2017), ResNet (He et al., 2016), MobileNet (Howard et al., 2017), Inception (Szegedy et al., 2015), Xception (Chollet, 2017) and VGG have been proposed. The main idea behind DenseNet is to connect all layers directly to each other, as seen in Figure 4 (G. Huang et al., 2017).

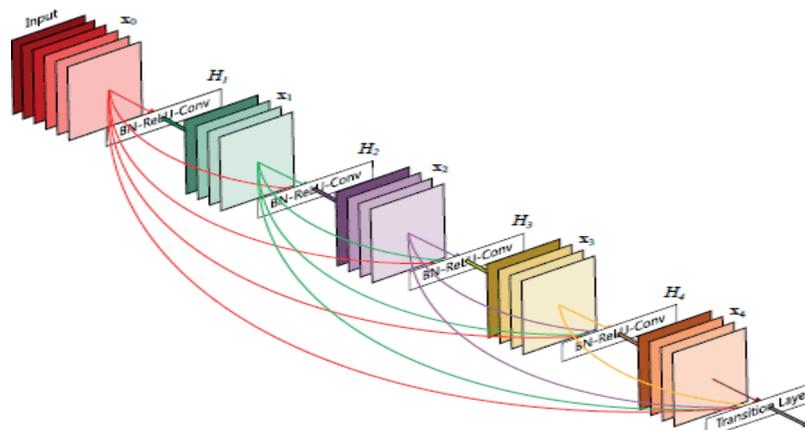


Figure 4. All layers are connected to each other in DenseNet.

This architecture has been trained on two CIFAR datasets (Krizhevsky, 2009) that have colored images with dimensions of 32x32 as the Street View House Numbers (SVHN) dataset (Netzer et al., 1952) that contains 32x32 colored digit images and ImageNet dataset, which is the ILSVRC 2012 classification dataset (Jia Deng et al., 2009) which consists of more than 1 million images. Within these datasets, Huang et al. have achieved state-of-the-art results with DenseNet architecture. Additionally, thanks to its fewer parameters, requirements and lower degree of needing computational resources, DenseNet may be more preferable to other state-of-the-art structures. To be able to reach successful results with these models, it is needed to have many images in the dataset and use powerful machines with high memory resources. In this study, since we could not meet both needs, we decided to use DenseNet with the transfer learning approach.

Transfer learning, which is a deep learning methodology, involves the reuse of pre-learned information for some problems to solve another problem. In the Transfer Learning approach, the features and weights that are obtained in some pre-trained models are reused for different tasks. This approach is beneficial for achieving high success levels quickly with small datasets (Darici, 2020).

Since the initial layers of all models extract some basic features of datasets such as edges, the reuse of these weights representing the dataset features would help obtain basic features of our dataset. This approach, which involves the usage of the weights of the pre-trained model, would save time. With this aim, using the weights of the pre-trained model, the last fully connected layers are retrained on a new dataset with a different aim. Thus, it is aimed to obtain high performance quickly. The primary approach of transfer learning may be seen in Figure 5.

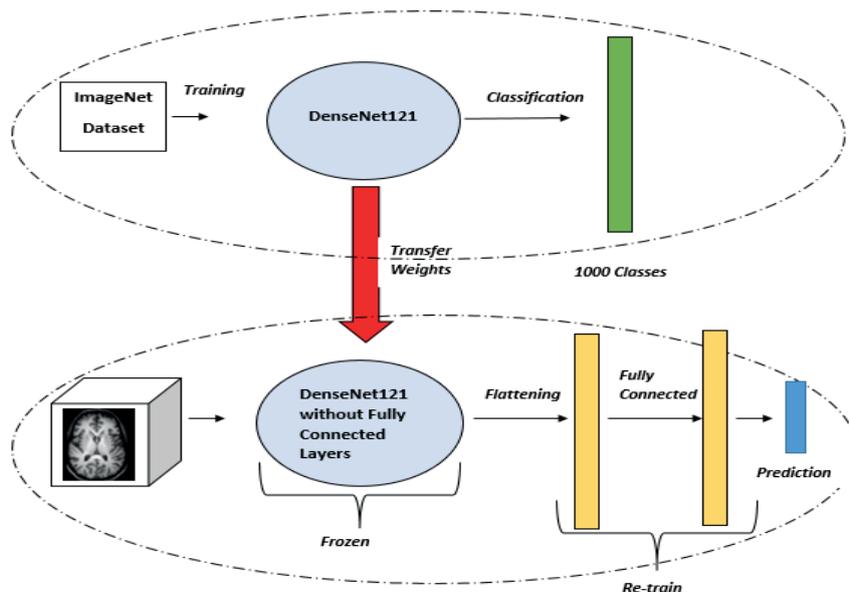


Figure 5. Transfer Learning approach.

In this study, as seen in Figure 5, the pre-trained DenseNet121 model weights with the ImageNet dataset were used. The last fully connected layers of the DenseNet121 model were extracted, and the remaining layers were frozen. This new model was fed with the brain scan dataset. While the weights of the pre-trained model were used for the frozen layers, the weights of the last layers were changed in each epoch for the new dataset, and the aim was to obtain a result suitable for the problem.

3.2 Proposed Method

In this section, firstly, the primary approach to the Brain Age Estimation task is given with its reasons (Figure 6). Then each problem and its solutions are explained in detail.

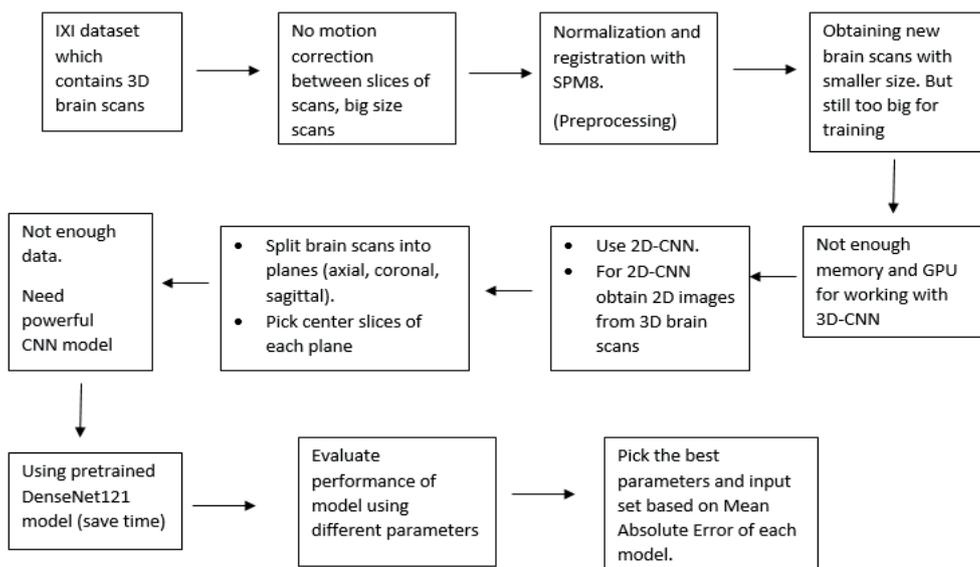


Figure 6. The primary approach to the BAE task in the proposed method.

As said earlier, the IXI dataset that was used in this study contains 3D brain scan images. To train a model with these three-dimensional (3D) images, we need 3D convolutions instead of a 2D one to be able to extract the feature maps of images. The difference between 3D and 2D convolution is the size of the filters. In 2D convolution, the filters used in the convolution

process are also in two dimensions. However, unlike 2D, in 3D convolution, three-dimensional filters are used. Because of this difference, models with 3D convolutions require more memory and time. For this study, since we could not meet this requirement, we decided to split these brain scans into planes and not use 3D convolutions. After splitting the brain scans into their planes separately, we still had images in 3D. As said earlier in Section 2.2, for the IXI dataset, we had 79 95x68 2D sagittal images, 95 79x68 2D coronal images and 68 79x95 2D axial images. For example, for the axial plane, if we put in order 68 images with the dimensions of 79x95, we would obtain a 3D image whose width, height and number of channels are 79, 95 and 68, respectively. As a result, for each patient, a 3D image would be obtained with 2D single-channel images lined up consecutively, as seen in Figure 7. Each slice in the axial plane represents parts of the brain from top to bottom.

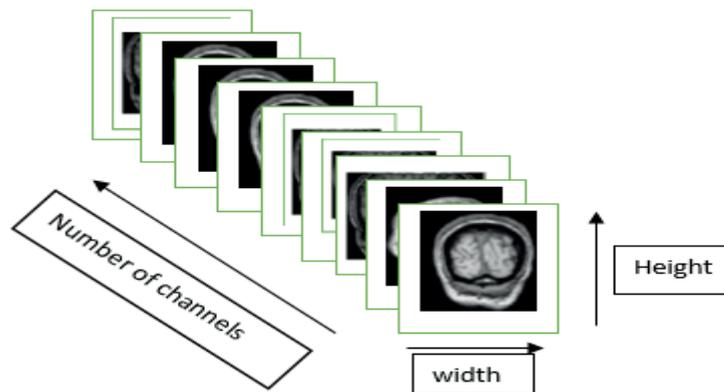


Figure 7. Slices of the axial plane of a brain scan.

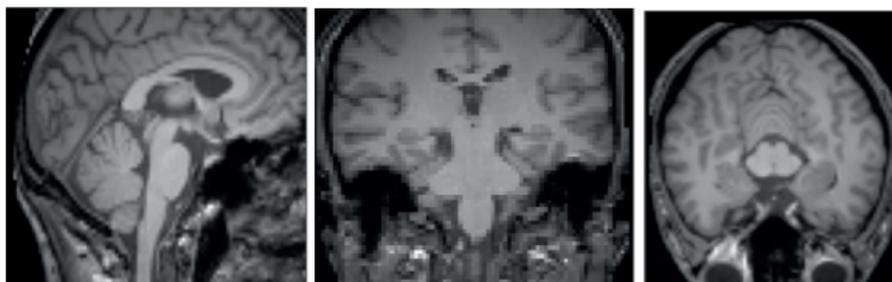
Since we still had 3D images after splitting the 3D brain scan images into their planes, we decided to use only the center slice of each plane. So, since 2D images would be obtained, 3D convolutions were not needed anymore. So, this study consisted of three parts since the brain scan images consisted of three planes. For the first part, the inputs were the center images of the sagittal plane. In the second part, the inputs were the center images of the coronal plane. Lastly, in the third part, the inputs were the center images of the axial plane. The slices used for each plane may be seen in Table 1.

Table 1

Details of the inputs of the model.

Plane	Total Number of Slices	Center Slice	Dimensions
Sagittal	79	40	95x68
Coronal	95	48	79x68
Axial	68	34	79x95

The center slices of each plane selected as the inputs of the model were in grayscale. So, they had just one channel each. However, due to its structure, the DenseNet architecture needs three-channel images as input. To solve this problem, the center slices were replicated three times for each patient to create three-channel images. For each input set (sagittal, coronal, axial), example images are shown in Figure 8.



(a) sagittal (b) coronal (c) axial

Figure 8. An example of the center slice of each brain plane.

After obtaining input sets using different planes of the brain, these input sets were trained on the DenseNet121 model using the transfer learning approach. As seen in Figure 5 in Section 3.1, the weights and biases of the pre-trained DenseNet121 model with the ImageNet dataset were loaded, and instead of training the entire model from the beginning, only the last layers were trained using the brain scan image dataset. Except for these layers, all layers were frozen. The aim that was thought to be achieved with this approach was to save a significant amount of time during the training process.

For DenseNet121 using the transfer learning approach, the last fully connected layers of the DenseNet121 model were extracted. The global Average Pooling process was applied to obtain the last feature map of the DenseNet model. After this layer, this model contained two fully connected layers. Each layer had a ReLU activation function and a dropout process with a 0.3 rate to prevent overfitting. In the end, a value that represented brain age was obtained. The structure of the DenseNet121 model with transfer learning may be seen in Figure 9. For each input set, 20% of the entire dataset was used in the testing process, and the rest was used in the training process. Among the training images, 10% were reserved for validation. The number of images for each input set and each process is given in Table 2.

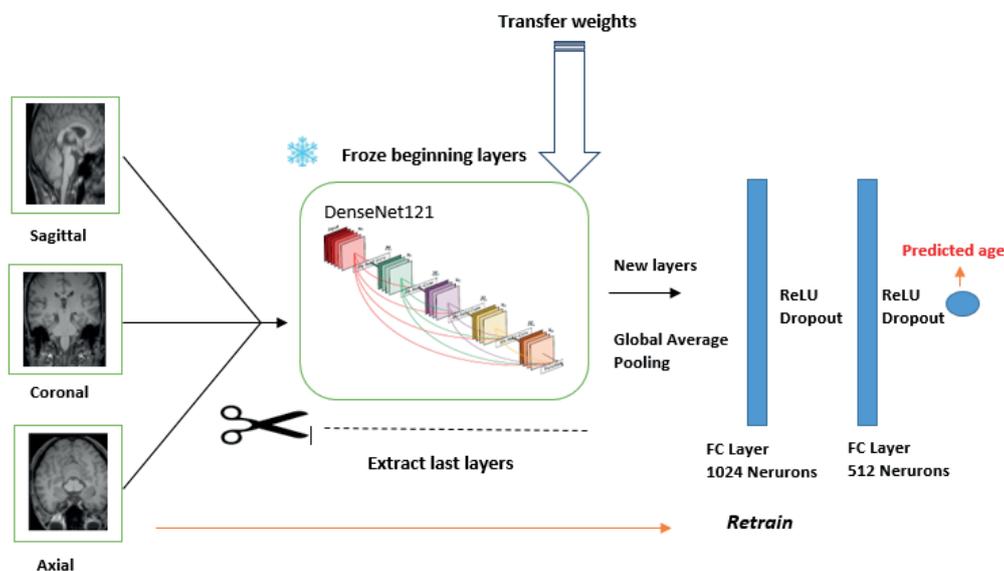


Figure 9. Flowchart of the proposed model.

Table 2

The number of images used in Training/Testing/Validation processes.

Input Set	Total	Training/Validation	Testing
Sagittal	563	360/40	163
Coronal	563	360/40	163
Axial	563	360/40	163

After creating a model to train and test the dataset, it is needed to set some hyperparameters to use in the model during training. These parameters are the optimizer, learning rate value, batch size and number of the epoch. The detailed information about these parameters used in this study is given in Table 3.

Table 3

Hyperparameters that were used in this study.

Hyperparameter	Value / Type
Optimizer	Adam, Adamax, Adagrad
Learning Rate	ReduceLROnPlateau
Batch Size	8
Epoch	100
Shuffle	True

As seen in Table 3, some different optimizers such as “Adam,” “Adamax” and “Adagrad” were used to see which provided the best results for the Brain Age Estimation (BAE) task. Instead of setting a certain value for the learning rate, it was

preferred to use a variable learning rate. The working principle of the “ReduceLROnPlateau” parameter is to reduce the learning rate value based on the set rate. This means, if a metric stops improvement, the learning rate value is reduced. In this study, the model reduced the learning rate for every five epochs by 0.3 if no improvements were seen in the validation loss. The optimization function and the learning rate are two important parameters in the training process. Since these parameters are used in the process of updating weights during training, it is very important to use appropriate values and proper functions to achieve successful results. Moreover, instead of training the entire training set at once, the training set is divided into batches. Each batch has 8 units of data. All training processes are repeated 100 times for the entire dataset, which means the number of the epochs is 100. Moreover, the dataset is shuffled before each epoch to be sure that each batch has different data units for each epoch.

As a summary, first of all, different input sets were obtained from the 3D brain scan images by splitting them into their planes individually. These input sets were then fed to the DenseNet121 model. Since training this model would take a long time, instead of training the entire model from the beginning, only the layers were trained. For the top layers, the weights of the pre-trained DenseNet121 model obtained from the ImageNet dataset were used. This training process was carried out with different optimization functions for each dataset. In the end, a value that represents brain age was obtained for each brain slice (Figure 10).

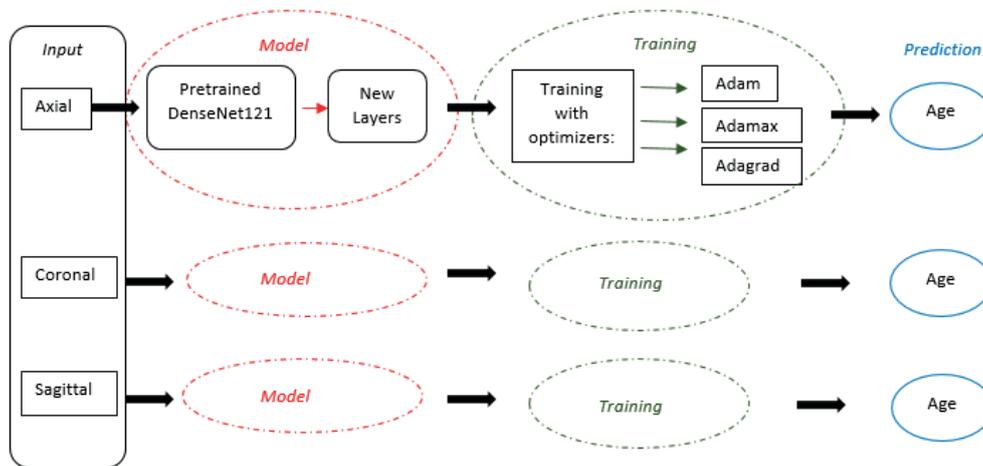


Figure 10. Summary of the proposed method.

4. RESULTS

This section reports the results of the proposed method. The proposed method was on three different input sets with three different optimization functions. For each experiment, the Mean Absolute Error (MAE) metric was used to perform test evaluation (see Eq. 1).

$$MAE = \frac{\sum_{i=1}^N |y - \hat{y}|}{N} \quad (1)$$

In Eq. 1 above, \hat{y} is the predicted age obtained from the proposed model; y is the real age obtained from the IXI dataset; N is the number of samples. Each table and the figure below each table show the results of the proposed model, training time and learning curves when the inputs were from the axial, coronal and sagittal planes, respectively.

Table 4

Results of the proposed method for the Axial Plane

Optimizers	MAE	Time per epoch (s)	Training time (m)
Adam	8.82	3.28	5.5
Adamax	7.66	3.19	5.3
Adagrad	7.22	2.9	4.8

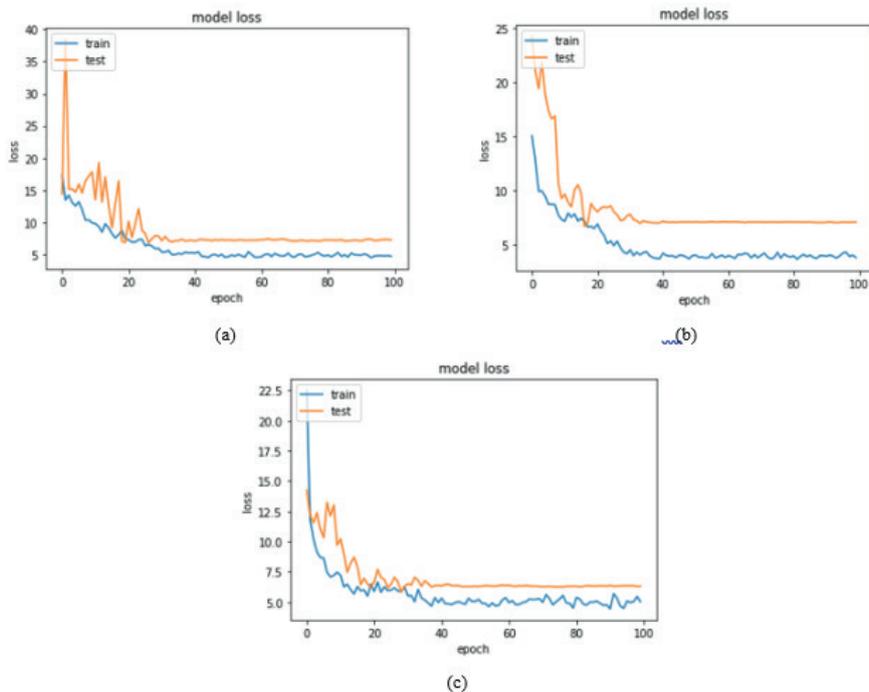


Figure 11. Axial plane learning curves of the model for the optimizers: (a) Adam, (b) Adamax, (c) Adagrad.

Table 5

Results of the proposed method for the Coronal Plane

Plane	MAE	Time per epoch (s)	Training time (m)
Adam	7.2	2.9	4.8
Adamax	7.62	3	5.1
Adagrad	7.45	2.7	4.05

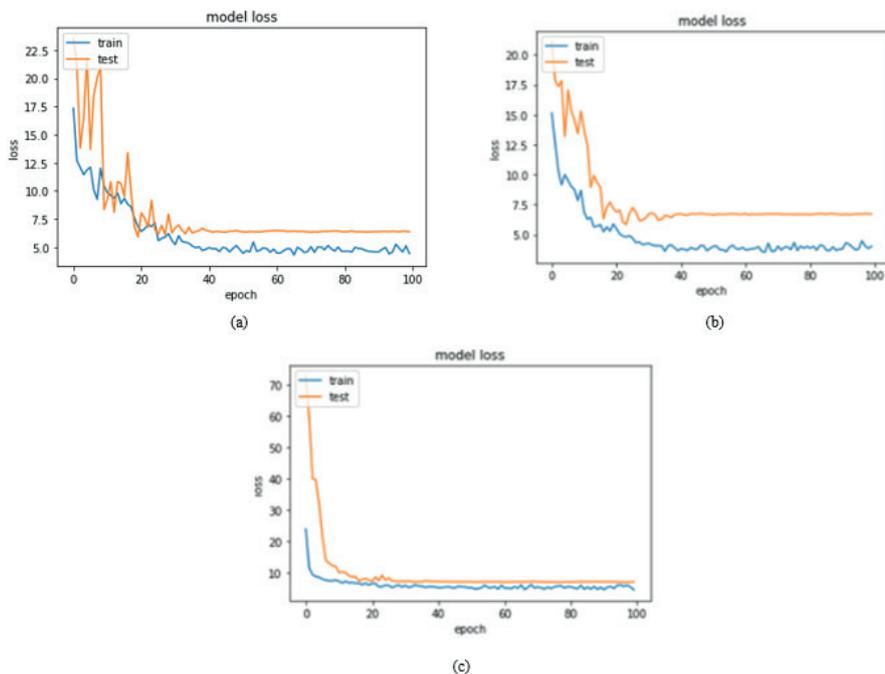


Figure 12. Coronal plane learning curves of the model for the optimizers: (a) Adam, (b) Adamax, (c) Adagrad.

Table 6
Results of the proposed method for the Sagittal Plane

Plane	MAE	Time per epoch (s)	Training time (m)
Adam	6.9	3.08	5.13
Adamax	6.3	3.21	5.35
Adagrad	7.93	3.01	5.02

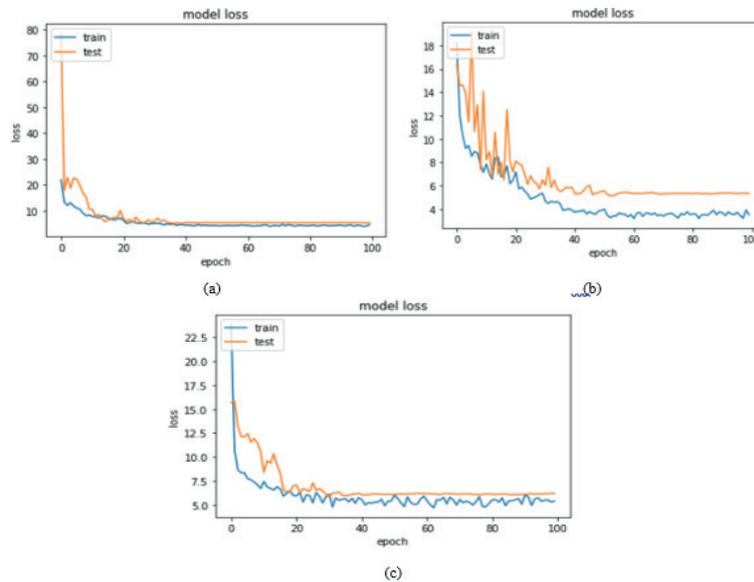


Figure 13. Sagittal plane learning curves of the model for the optimizers: (a) Adam, (b) Adamax, (c) Adagrad.

As seen from all result tables and figures, the best result was the lower value of MAE obtained from the axial plane when the optimizer was Adamax. The training time of all models for each input set and optimizer was shorter than 5.5 minutes, which means the proposed algorithm was quite fast.

Finally, Table 7 compares the results of the proposed method to the ones of other methods used on different datasets. It may be seen that the best MAE value, which means deviation from the actual age, was about 1.43 years on the PNC dataset.

Table 7
Comparison of the results and the training times of the proposed method to other methods.

Study	Method	Dataset	MAE	Time per epoch (s)	Training time (m)
(Cole et al., 2017b)	3D-CNN	IXI	6.76	1200	4200
(Cole et al., 2017b)	Applied by (Rossi et al., 2019) 3D-CNN	PNC	1.43	N/A	N/A
(Rossi et al., 2019)	2D-CNN (using all slices)	IXI	6.99	240	270
(Rossi et al., 2019)	2D-CNN inspired by ResNet (using all slices)	IXI	6.61	300	420
(Rossi et al., 2019)	Combining VGG-16 and BLSTM	IXI	5.94	60	99.6
(Rossi et al., 2019)	Combining ResNet50 and BLSTM	IXI	6.86	60	94.8
(Rossi et al., 2019)	Combining DenseNet121 and BLSTM	IXI	6.17	60	90
(Jonsson et al., 2019)	Combining some 3D-CNN	IXI and UK Biobank	4.149	N/A	N/A
(Peng et al., 2021)	3D-CNN	UK Biobank	2.14	N/A	N/A
This Study	2D-CNN using just center slices	IXI	6.3	3.21	5.35

As seen in Table 7, some comparable results can be achieved in a very short training time thanks to the simple method proposed in this study. Even if the approach of using 3D-CNN shows better prediction results, the method used in this study can outperform some complex models and 2D-CNNs which use all slices of brain scans as the input.

5. CONCLUSION

The main purpose of this study was to estimate the age of the person with high performance in the shortest time possible using brain MR images. Another important goal of this study was to not require too much memory and computing power while achieving a high performance level. With this aim, each brain scan in the IXI dataset was divided into planes. Each scan has axial, coronal and sagittal planes, which represent some specific areas of the brain. After obtaining three different planes, the center slices of each plane were reserved as the input for the proposed model. Thus, the need for 3D models was eliminated. Since the dataset did not contain too many brain scan images, the weights of the pre-trained DenseNet121 model were used for the model. The model whose last layers were changed was retrained using these weights with a new input set. This approach eliminated the disadvantage of the small number of data units and the high computational power required for a powerful model. BAE is an important task to estimate the ages of individuals from their brain MR images and to detect some neurodegenerative diseases such as Alzheimer's and Parkinson's. Therefore, estimating brain age in the shortest time possible with the lowest loss is the primary purpose of BAE tasks. In this study, with the proposed method, the best test MAE result that was obtained was 6.3 using the sagittal planes as the inputs and the Adamax optimizer in the DenseNet121 model. Moreover, the training time for the parameters that provided the lowest loss was just 5.35 minutes, 3.21 seconds for each epoch. The obtained results showed that the proposed method in this study is comparable to similar studies on the task of BAE. This method provides good results in a short time with low computational power and requirements. These advantages make this method usable and applicable for everybody else.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors have no conflict of interest to declare.

Author Contributions: Conception/Design of Study- M.B.D., Ş.Y.; Data Acquisition- M.B.D.; Data Analysis/Interpretation- M.B.D., M.G.; Drafting Manuscript- M.B.D., Ş.Y.; Critical Revision of Manuscript-M.G.; Final Approval and Accountability- M.B.D., Ş.Y., M.G.

Grant Support: The authors declared that this study has received no financial support.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemiştir.

Yazar Katkıları: Çalışma Konsepti/Tasarım- M.B.D., Ş.Y.; Veri Toplama-M.B.D.; Veri Analizi/Yorumlama- M.B.D., M.G.; Yazı Taslağı- M.B.D., Ş.Y.; İçeriğin Eleştirel İncelemesi-M.G.; Son Onay ve Sorumluluk- M.B.D., Ş.Y., M.G.

Finansal Destek: Yazarlar bu çalışma için finansal destek almadığını beyan etmiştir.

References/Kaynaklar

- Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua*, 1800–1807. <https://doi.org/10.1109/CVPR.2017.195>
- Cole, J. H., Poudel, R. P. K., Tsagkrasoulis, D., Caan, M. W. A., Steves, C., Spector, T. D., & Montana, G. (2017a). Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage*, *163*, 115–124. <https://doi.org/10.1016/j.neuroimage.2017.07.059>
- Cole, J. H., Poudel, R. P. K., Tsagkrasoulis, D., Caan, M. W. A., Steves, C., Spector, T. D., & Montana, G. (2017b). Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage*, *163*, 115–124. <https://doi.org/10.1016/j.neuroimage.2017.07.059>
- Darici, M. B., Dokur, Z., & Olmez, T. (2020). Pneumonia Detection and Classification Using Deep Learning on Chest X-Ray Images. *International Journal of Intelligent Systems and Applications in Engineering*, *8(4)*, 177–183.
- Darici, M. B. (2020). *Göğüs kafesi röntgen görüntülerinde derin öğrenme metoduyla zatürre hastalığının tanısı* (Vol. 21, Issue 1). <https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp>
- Doi, K. (2007). Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, *31(4–5)*, 198–211. <https://doi.org/10.1016/j.compmedimag.2007.02.002>
- Franke, K., & Gaser, C. (2012). Longitudinal changes in individual BrainAGE in healthy aging, mild cognitive impairment, and Alzheimer's Disease. *GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry*, *25(4)*, 235–245. <https://doi.org/10.1024/1662-9647/a000074>
- Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., & Zijdenbos, A. (1999). <Giedd1999B.Pdf>. *Neuropsychol, Dev.*
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-Decem*, 770–778. <https://doi.org/10.1109/CVPR.2016.90>

- Hong, J., Feng, Z., Wang, S. H., Peet, A., Zhang, Y. D., Sun, Y., & Yang, M. (2020). Brain Age Prediction of Children Using Routine Brain MR Images via Deep Learning. *Frontiers in Neurology*, *11*(October), 1–13. <https://doi.org/10.3389/fneur.2020.584682>
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *ArXiv*.
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua*, 2261–2269. <https://doi.org/10.1109/CVPR.2017.243>
- Huang, T., Chen, H., Fujimoto, R., Ito, K., Wu, K., Sato, K., Taki, Y., Fukuda, H., & Aoki, T. (2017). AGE ESTIMATION FROM BRAIN MRI IMAGES USING DEEP LEARNING Department of Computer Science , National Tsing-Hua University , Taiwan Graduate School of Information Science , Tohoku University , Japan South China University of Technology , China Institute of D. *Conference: 2017 IEEE 14th International Symposium on Biomedical*, *2*(1), 849–852.
- Ito, K., Fujimoto, R., Huang, T. W., Chen, H. T., Wu, K., Sato, K., Taki, Y., Fukuda, H., & Aoki, T. (2018). Performance Evaluation of Age Estimation from T1-Weighted Images Using Brain Local Features and CNN. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2018-July*, 694–697. <https://doi.org/10.1109/EMBC.2018.8512443>
- Jia Deng, Wei Dong, Socher, R., Li-Jia Li, Kai Li, & Li Fei-Fei. (2009). *ImageNet: A large-scale hierarchical image database*. 248–255. <https://doi.org/10.1109/cvprw.2009.5206848>
- Jiang, H., Lu, N., Chen, K., Yao, L., Li, K., Zhang, J., & Guo, X. (2020). Predicting Brain Age of Healthy Adults Based on Structural MRI Parcellation Using Convolutional Neural Networks. *Frontiers in Neurology*, *10*(January). <https://doi.org/10.3389/fneur.2019.01346>
- Jonsson, B. A., Bjornsdottir, G., Thorgeirsson, T. E., Ellingsen, L. M., Walters, G. B., Gudbjartsson, D. F., Stefansson, H., Stefansson, K., & Ulfarsson, M. O. (2019). Brain age prediction using deep learning uncovers associated sequence variants. *Nature Communications*, *10*(1), 1–10. <https://doi.org/10.1038/s41467-019-13163-9>
- Krizhevsky, A. (2009). Learning Multiple Layers of Features from Tiny Images. *Asha*, *34*(4).
- Lenroot, R. K., & Giedd, J. N. (2006). Brain development in children and adolescents: Insights from anatomical magnetic resonance imaging. *Neuroscience and Biobehavioral Reviews*, *30*(6), 718–729. <https://doi.org/10.1016/j.neubiorev.2006.06.001>
- Levakov, G., Rosenthal, G., Shelef, I., Raviv, T. R., & Avidan, G. (2020). From a deep learning model back to the brain—Identifying regional predictors and their relation to aging. *Human Brain Mapping*, *41*(12), 3235–3252. <https://doi.org/https://doi.org/10.18201/ijisae.2020466310>
- Li, H., Satterthwaite, T. D., & Fan, Y. (2018). Brain age prediction based on resting-state functional connectivity patterns using convolutional neural networks. *ArXiv, Isbi*, 101–104.
- Liem, F., Varoquaux, G., Kynast, J., Beyer, F., Kharabian Masouleh, S., Huntenburg, J. M., Lampe, L., Rahim, M., Abraham, A., Craddock, R. C., Riedel-Heller, S., Luck, T., Loeffler, M., Schroeter, M. L., Witte, A. V., Villringer, A., & Margulies, D. S. (2017). Predicting brain-age from multimodal imaging data captures cognitive impairment. *NeuroImage*, *148*(November 2016), 179–188. <https://doi.org/10.1016/j.neuroimage.2016.11.005>
- Luders, E., Gaser, C., Narr, K. L., & Toga, A. W. (2009). Why sex matters: Brain size independent differences in gray matter distributions between men and women. *Journal of Neuroscience*, *29*(45), 14265–14270. <https://doi.org/10.1523/JNEUROSCI.2261-09.2009>
- Netzer, Y., Wang, T., Coates, A., Bissacco, A., Wu, B., & Ng, A. Y. (1952). PROFESSOR V.N. SHamov. *Voprosy Neirokhirurgii*, *16*(5), 9–13.
- Pardakhti, N., & Sajedi, H. (2020). Brain age estimation based on 3D MRI images using 3D convolutional neural network. *Multimedia Tools and Applications*, *79*(33–34), 25051–25065. <https://doi.org/10.1007/s11042-020-09121-z>
- Peng, H., Gong, W., Beckmann, C. F., Vedaldi, A., & Smith, S. M. (2021). Accurate brain age prediction with lightweight deep neural networks. *Medical Image Analysis*, *68*, 101871. <https://doi.org/10.1016/j.media.2020.101871>
- Rokicki, J., Wolfers, T., Nordhøy, W., Tesli, N., Quintana, D. S., Alnæs, D., Richard, G., de Lange, A. M. G., Lund, M. J., Norbom, L., Agartz, I., Melle, I., Nærland, T., Selbæk, G., Persson, K., Nordvik, J. E., Schwarz, E., Andreassen, O. A., Kaufmann, T., & Westlye, L. T. (2020). Multimodal imaging improves brain age prediction and reveals distinct abnormalities in patients with psychiatric and neurological disorders. *Human Brain Mapping*, *August 2020*, 1714–1726. <https://doi.org/10.1002/hbm.25323>
- Rossi, A., Vannuccini, G., Andreini, P., Bonechi, S., Giacomini, G., Scarselli, F., & Bianchini, M. (2019). Analysis of brain NMR images for age estimation with deep learning. *Procedia Computer Science*, *159*, 981–989. <https://doi.org/10.1016/j.procs.2019.09.265>
- Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–14.
- Symposium, I., & Imaging, B. (2019). AN AGE ESTIMATION METHOD USING 3D-CNN FROM BRAIN MRI IMAGES Graduate School of Information Sciences , Tohoku University , Japan . South China University of Technology , China . Institute of Dvelopment , Aging and Cancer , Tohoku University , Japan . *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Isbi*, 380–383.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 07-12-June*, 1–9. <https://doi.org/10.1109/CVPR.2015.7298594>
- Tang, J., Rangayyan, R. M., Xu, J., El Naqa, I. E., & Yang, Y. (2009). Computer-aided detection and diagnosis of breast cancer with mammography: Recent advances. *IEEE Transactions on Information Technology in Biomedicine*, *13*(2), 236–251. <https://doi.org/10.1109/TITB.2008.2009441>
- Wang, J., Li, W., Miao, W., Dai, D., Hua, J., & He, H. (2014). Age estimation using cortical surface pattern combining thickness with curvatures. *Medical and Biological Engineering and Computing*, *52*(4), 331–341. <https://doi.org/10.1007/s11517-013-1131-9>
- Yin, T. K., & Chiu, N. T. (2004). A computer-aided diagnosis for locating abnormalities in bone scintigraphy by a fuzzy system with a three-step minimization approach. *IEEE Transactions on Medical Imaging*, *23*(5), 639–654. <https://doi.org/10.1109/TMI.2004.826355>

