



Article Age Estimation from Brain Magnetic Resonance Images Using Deep Learning Techniques in Extensive Age Range

Kousuke Usui¹, Takaaki Yoshimura^{2,3,4,5}, Minghui Tang^{5,6} and Hiroyuki Sugimori^{4,5,7,*}

- ¹ Graduate School of Health Sciences, Hokkaido University, Sapporo 060-0812, Japan
- ² Department of Health Sciences and Technology, Faculty of Health Sciences, Hokkaido University, Sapporo 060-0812, Japan
- ³ Department of Medical Physics, Hokkaido University Hospital, Sapporo 060-8648, Japan
- ⁴ Global Center for Biomedical Science and Engineering, Faculty of Medicine, Hokkaido University, Sapporo 060-8648, Japan
- ⁵ Clinical AI Human Resources Development Program, Faculty of Medicine, Hokkaido University, Sapporo 060-8648, Japan
- ⁶ Department of Diagnostic Imaging, Faculty of Medicine and Graduate School of Medicine, Hokkaido University, Sapporo 060-8638, Japan
- ⁷ Department of Biomedical Science and Engineering, Faculty of Health Sciences, Hokkaido University, Sapporo 060-0812, Japan
- * Correspondence: sugimori@hs.hokudai.ac.jp; Tel.: +81-11-706-3410

Abstract: Estimation of human age is important in the fields of forensic medicine and the detection of neurodegenerative diseases of the brain. Particularly, the age estimation methods using brain magnetic resonance (MR) images are greatly significant because these methods not only are noninvasive but also do not lead to radiation exposure. Although several age estimation methods using brain MR images have already been investigated using deep learning, there are no reports involving younger subjects such as children. This study investigated the age estimation method using T1-weighted (sagittal plane) two-dimensional brain MR imaging (MRI) of 1000 subjects aged 5–79 (31.64 \pm 18.04) years. This method uses a regression model based on ResNet-50, which estimates the chronological age (CA) of unknown brain MR images by training brain MR images corresponding to the CA. The correlation coefficient, coefficient of determination, mean absolute error, and root mean squared error were used as the evaluation indices of this model, and the results were 0.9643, 0.9299, 5.251, and 6.422, respectively. The present study showed the same degree of correlation as those of related studies, demonstrating that age estimation can be performed for a wide range of ages with higher estimation accuracy.

Keywords: deep learning; age estimation; regression model; machine learning; ResNet-50

1. Introduction

The human brain develops and atrophies during growth and aging, resulting in structural changes in the brain [1,2]. These changes in the brain structure can be confirmed with the human eye using medical images. Particularly, noninvasive and radiation-free magnetic resonance (MR) imaging (MRI) of the brain is a useful method for confirming structural changes. Thus, human age can be estimated by using brain structures that change with age, and this type of age estimation technique can be applied to two major fields: forensic medicine and early detection of neurodegenerative diseases of the brain.

As for the application to the field of forensic science, both living and deceased subjects can be targeted [3,4]. In the case of the living subject, it is expected to be applied to situations where identification is required, such as asylum applications, criminal lawsuits, and youth sports [3], whereas in the case of the corpse, it is expected to be applied to the identification of the unidentified corpse [4]. In such applications, since previous studies have attempted to estimate the age with other parts of the body, such as the teeth, clavicles, hands, and knees,



Citation: Usui, K.; Yoshimura, T.; Tang, M.; Sugimori, H. Age Estimation from Brain Magnetic Resonance Images Using Deep Learning Techniques in Extensive Age Range. *Appl. Sci.* **2023**, *13*, 1753. https://doi.org/10.3390/app13031753

Academic Editors: László Szilágyi and Levente Adalbert Kovács

Received: 28 December 2022 Revised: 25 January 2023 Accepted: 28 January 2023 Published: 30 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). the reliability of the age estimation will be improved by using this method together with these techniques. As for the application to the early detection of neurodegenerative diseases of the brain, it is based on the fact that brain morphology changes with aging (atrophy due to linear decrease in gray matter volume) and these neurodegenerative diseases cause brain atrophy that is greater than that of normal aging [5]. Since patients with neurodegenerative diseases develop brain atrophy caused by systematic loss of nerve cells in the central nervous system, such changes can be confirmed through MRI and computed tomography (CT) images [6,7]. Therefore, patients with neurodegenerative diseases are assumed to show a larger discrepancy between their chronological age and estimated age based on their brain structures, and this discrepancy can be used for the early detection of neurodegenerative diseases. Furthermore, considering the aging of the population on a global scale and the fact that "aging" is considered a risk factor for neurodegenerative diseases, this type of technology is expected to be increasingly utilized in the future [8,9]. Early detection of these diseases is greatly significant since these diseases affect not only the patients themselves, but also the quality of life of their families [10,11].

In related works, regarding the methodology used to estimate age, in order to improve the identification of age-related developmental disorders as significant, an investigation was conducted based on a dental examination that included the determination of the X-ray image of the left hand and dental status and X-ray images of the dentition [12], however, if the skeletal formation of the hand has been completed, an additional X-ray examination of the clavicles was required. These results indicate that it is difficult to estimate age using the same physical indicators. In addition, another paper has been reported [13] that suggests a problem with the use of X-rays in age estimation because of the radiation exposure involved in age estimation, and as the international atomic energy agency (IAEA) regulates the use and possible abuse of X-rays, it is desirable to perform this non-invasive method. Moreover, a paper has been reported [14] that shows the usefulness of MRI-based methods as an alternative to age estimation methods that involve radiation exposure, and a report examining the degree to which imaging time can be reduced in age estimation methods using MR images of the hand and wrist joint. However, the method has the disadvantage that it is not applicable to a wide range of age groups because the number of subjects is small and the age range of the subjects is limited. In summarizing these reports, noninvasive and identical body part-based age estimation is required.

Recently, deep learning technology has continued to develop and has been applied to several medical image analyses [15–20] besides classification [21], including semantic segmentation [22,23], and object detection [24,25] using convolutional neural networks (CNNs) and regression tasks [26,27]. This study attempted to estimate human age from brain MR images using regression analysis. Although medical methods are more useful than nonmedical methods for human age estimation [28] and age estimation methods using brain MR images via deep learning [29–32] have already been reported, studies on young subjects such as minors have not been confirmed. However, no study on young subjects, such as minors, has been reported. Moreover, previous studies reported that sagittal images should be used for age estimation using two-dimensional (2D) brain MR images [32] and that central regions of the brain are better than peripheral regions [31] as the target regions. Therefore, this study aimed to examine the accuracy of age estimation by using 2D CNN, a type of deep learning technique, for a wider age range than previous studies, including from infants to elderly individuals, using brain MR images.

2. Materials and Methods

2.1. Subjects

Images from two public databases were used. One database is the Decoded Neurofeedback (DecNef) Project Brain Data Repository (https://bicr-resource.atr.jp/srpbsopen/ (accessed on 28 December 2022)), which is part of the Japanese Strategic Research Program of Brain Sciences (SRPBS) supported by the Japanese Advanced Research and Development Programs for Medical Innovation (AMED) [33]. Another one is the "Healthy Brain Network" published by the Child Mind Institute (http://fcon_1000.projects.nitrc.org/indi/ cmi_healthy_brain_network/ (accessed on 28 December 2022)) [34]. This study obtained T1-weighted images of 1000 subjects (men = 500, women = 500) aged 5–79 years from these databases (Table 1). The images were sourced from multiple centers. Therefore, the minimum to maximum width and height of the images used in this study ranged from 180–320 mm and 138–320 mm, respectively. When the same number of male and female data was obtained, care was taken to adjust the distribution of the number of persons by age to be as uniform as possible.

Table 1. Age information of subjects.

Carlainata			Age		
Subjects	Mean	SD	Median	Q1	Q3
Total: 1000 subjects	31.64	18.04	30.00	16.00	45.00
500 Male subjects	30.97	17.52	30.00	16.00	44.00
500 Female subjects	32.30	18.52	31.00	16.00	46.00

SD: standard deviation, Q1: first quartile, Q3: third quartile.

2.2. Preprocessing

Since all data acquired in this study were in the Neuroimaging Informatics Technology Initiative (NIfTI) format and this study targets 2D images (sagittal plane), the data were converted to Joint Photographic Experts Group (JPEG) format. The accuracy of age estimation using 2D brain MR images is higher when the central region of the brain is targeted than the peripheral regions. Therefore, in this study, we used a total of 11 images from each subject, consisting of the central value of the total number of images after conversion from NIFTI format to JPEG format and five images before and after the central value.

Since the lower limit of the subject age was set lower than that of previous studies, it cannot be denied that the age estimation in the present study is based on the size of the cranium and brain, not on the structural differences of the brain. Since the proportion of the brain in the MRI image can be changed by setting the field of view (FOV) and other factors, the size of the head in the MRI image tends to be the same to some extent, if we simply consider the size of the head in the image. However, considering the fact that there is a difference in the standardization of the head size, it cannot be eliminated that the age estimation is based only on the size of the cranium and the brain. To exclude the influence of the head size, the data augmentation of brain size rescaling was performed. Since the results of the actual measurement of head standards (Figures 1 and 2) and images that included a part of the brain (Figure 3) are missing due to magnification, scaling factors of data augmentations as 0.8, 0.9, and 1.0 were conducted in this study. The measured angles of the subcallosal line (the line connects the inferior surface of the genu of the corpus callosum to the inferior surface of the splenium of the corpus callosum) showed normality in the variation of the angle of the head (mean = -1.0°). Therefore, in this study, data augmentation on rotation was not conducted. Data augmentation by scaling down or equalizing was conducted.

2.3. Architecture and Training

To construct a regression model for estimating chronological age (CA) from brain MR images, we developed a regression CNN based on ResNet-50 [35]. Originally, ResNet-50 has a structure in which the softmax layer outputs the input data as values ranging from 0 to 1, and the classification layer shows the final classification result. However, in this study, since the estimated age is the output, the regression model is created by introducing a regression layer instead of these layers, which can output the estimated value before it is converted to a value between 0 and 1 by the softmax layer. Using this model, we created three age estimation models (Table 2). The CNN conditions used during training were the stochastic gradient descent with momentum optimizer, and the initial learning rate, max Epochs, and mini-batch size were set to 0.0001, 10, and 512, respectively.



Figure 1. Measured head standard in a brain MR image. (MH: Measured Height, IH: Image Height, MW: Measured Width, IW: Image Width).



Figure 2. Measured results of head standard. (MH: Measured Height, IH: Image Height, MW: Measured Width, IW: Image Width).



Figure 3. Images missing due to data augmentation. *: magnification factor.

Table 2. Age estimation model.

Model	Content
ALL_Model	Training/evaluation using images of a total of 1000 subjects
Male_Model	Training/evaluation using images of 500 male subjects
Female_Model	Training/evaluation using images of 500 female subjects

A fivefold cross-validation (train:test = 4:1) was used in training the data. In this study, we defined a subset as one of the five divisions of the dataset, and we adjusted the distribution of the number of persons in each age group to be as uniform as possible before creating the subset, so as not to bias the age distribution within each dataset. These pre-adjustments and the five-fold cross validation are performed to avoid overfitting of specific data to the created predictor. Table 3 shows the software and computer specification used for the training of these CNNs.

 Table 3. Software/equipment used in the study.

Environment	Content
Software	MATLAB 2021a (developed by MathWorks)
OS	Windows 10
CPU	Intel core i9-10980XE 3.0 GHz
GPU	NDIVIA RTX A6000 48 GB \times 2
Memory	DDR4 2933 Quad-Channel 64 GB

2.4. Evaluating the Created Regression Models

The regression models were used to estimate CA from brain MR images in the test data. Since training and evaluation were conducted using fivefold cross-validation in this study, the average performance of the five regression models was evaluated as the performance of this method.

Mean absolute error (MAE), root mean squared error (RMSE), correlation coefficient (R), and coefficient of determination (R^2) were used as evaluation indices. Since linear regression is used in this study, we calculated the R^2 value by squaring R which was

obtained from linear regression and used it as an evaluation index. R, MAE, and RMSE are calculated using the following equations:

$$R = \frac{\frac{1}{n}\sum_{i=1}^{n} \left(y_{obs, i} - \overline{y_{obs}}\right) \left(y_{pred, i} - \overline{y_{pred}}\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} \left(y_{obs, i} - \overline{y_{obs}}\right)^{2} \times \sqrt{\frac{1}{n}\sum_{i=1}^{n} \left(y_{pred, i} - \overline{y_{pred}}\right)^{2}}}$$

$$MAE = \frac{\sum_{i} \left|y_{obs, i} - y_{pred, i}\right|}{n}$$

$$RMSE = \sqrt{\frac{\sum_{i} \left(y_{obs, i} - y_{pred, i}\right)^{2}}{n}}$$
(1)

where $y_{obs,i}$ is the real age, $y_{pred,i}$ is the estimated age, $\overline{y_{obs}}$ is the average of $y_{obs,i}$, $\overline{y_{pred}}$ is the average of $y_{pred,i}$, and *n* is the number of samples in the case.

2.5. Evaluation of the Significant Region in Age Estimation

The results of previous studies showed that central regions of the brain have a better effect on age estimation than peripheral regions of the brain [31]. A detailed examination of the visualization map of the basis for this report shows that the MAE is smaller in the corpus callosum region. This indicates that the corpus callosum region is a useful region for age estimation.

In this study, based on the results of the previous study, we created an image in which the corpus callosum region is removed (Figure 4). By determining how the accuracy of the estimation changes before and after the removal of the corpus callosum region, it is possible to examine whether the corpus callosum region is a useful region for age estimation in this method. As a specific analysis procedure, we used the method for examining significant differences as described in Section 2.6.

2.6. Statistical Analysis

In this method, three adjustments were considered when collecting data and creating a dataset/subset: (1) adjusting the distribution of the number of persons by age to be as uniform as possible between men and women, (2) adjusting the number of data to be as uniform as possible between men and women, and (3) adjusting the distribution of the number of persons by age to be as uniform as possible between each subset. Therefore, the five age estimation models A, B, C, D, and E, which were created using the fivefold cross-validation, are models that each learn the same number of data with similar age distributions, which contributes to showing similar age distribution trends for both Male_Model A–E and Female_Model A–E.

Based on this trend, we conducted a statistical significance study using the absolute error (AE) of each test data. The AE of 200 subjects for ALL_Model and the AE of 100 subjects for Male_Model and Female_Model were used to confirm the normality by the Shapiro–Wilk test, and when the normality was found, the F test and *t*-test were used to examine the statistical significance. In the case of nonnormality, the Wilcoxon rank sum test was used to examine the significance of the difference. A *p* value less than 0.05 was considered statistically significant. Specifically, we used this method to examine whether there is a statistically significant difference in the accuracy of age estimation between Male_Model and Female_Model and whether there is a statistically significant difference in the accuracy of age estimation before and after the deletion of the corpus callosum region. In the former case, a total of 25 patterns (cross comparison for five patterns of Male_Model (Male_Model A–E) and five patterns of Female_Model (Female_Model A–E)) were examined for statistical significance using AE. Conversely, in the latter case, since the images before and after the deletion), we examined the statistical significance of AE

using five patterns each of ALL_Model, Male_Model, and Female_Model (15 patterns in total). The test data used for the input are images without data augmentation processing and images from which the corpus callosum region is removed.



Figure 4. The process of the statistical analysis on the accuracy of age estimation before and after the removal of the corpus callosum (CC) region.

3. Results

3.1. Evaluation of the Accuracy of the Age Estimation Model

Accuracy evaluations for the ALL_Model, Male_Model, and Female_Model age estimation models are shown in Figures 5–8. As for the R and R² value, which indicate the degree of correlation, the accuracy of both age estimation models improved before and after data augmentation. Specifically, the R and R² value were 0.9643 and 0.9299, respectively, for ALL_Model, and 0.94582 and 0.93945, respectively, for Male_Model and Female_Model, and 0.89513 and 0.88257, respectively, for R². As for MAE and RMSE, which indicate the error between the actual age and estimated age, the accuracy improved before and after data augmentation for ALL_Model and Female_Model, but the results for Male_Model were almost the same before and after data augmentation. Specifically, the MAE and RMSE of the ALL_Model were 5.251 and 6.422 years, respectively, whereas the MAE and RMSE of the Male_Model and Female_Model were 6.323 and 5.766 years, respectively, and 7.734 and 7.006 years, respectively. Figure 9 shows examples of age estimation from brain MR images using this method. The upper two rows show examples of highly accurate age estimation, and the lower two rows show examples with estimation errors of about MAE or larger.



Figure 5. Results for correlation coefficient (R). (DA: Data Augmentation).



Figure 6. Results for the coefficient of determination (R²). (DA: Data Augmentation).



Figure 7. Results for mean absolute error (MAE). (DA: Data Augmentation).



Figure 8. Results for root mean squared error (RMSE). (DA: Data Augmentation).



Figure 9. Typical example of this age estimation method. (CA, chronological age; EA, estimated age).

3.2. Evaluation of the Significant Region in Age Estimation

Table 4 shows the results of the Shapiro–Wilk test for the data on AEs obtained in this study. In Table 4, since there is no normality in the data on AEs obtained in this study, it was decided that the statistical significance be examined by using the Wilcoxon rank sum test. Table 5 shows the results of the Wilcoxon rank sum test. This table shows that although there is one case each for Male_Model and Female_Model in which there is no statistically significant difference. In all other cases, there is a statistically significant difference in the accuracy of age estimation before and after the deletion of the corpus callosum region. In sum, the overall results show that the accuracy of age estimation significantly decreases before and after the removal of the corpus callosum region. Table 6 shows the specific values of MAE and other parameters.

Table 4. Results of the Shapiro–Wilk test.

	ALL_Model		Male_Model		Female_Model	
	With CC	Without CC	With CC	Without CC	With CC	Without CC
Model_A Model_B Model_C Model_D Model_E	$p < 0.0001 \\ p < 0.0001$	p < 0.0001 p = 0.0003 p = 0.0003 p = 0.0001 p = 0.0005	$p < 0.0001 \\ p < 0.0001$	p = 0.0024 p = 0.0013 p = 0.0289 p = 0.0090 p = 0.0004	$p < 0.0001 \\ p = 0.0145 \\ p < 0.0001 \\ p = 0.0003 \\ p < 0.0001$	p = 0.0010 $p = 0.0644$ $p < 0.0001$ $p = 0.0002$ $p < 0.0001$

CC: corpus callosum.

	ALL_Model	Male_Model	Female_Model
	<i>p</i> Value	p Value	p Value
Model_A	< 0.0001	0.0001	0.2523
Model_B	< 0.0001	0.0825	< 0.0001
Model_C	< 0.0001	< 0.0001	< 0.0001
Model_D	< 0.0001	0.0025	< 0.0001
Model_E	< 0.0001	0.0048	0.0046

Table 5. Results of the Wilcoxon rank sum test.

Table 6. Change in accuracy before and after deletion of the corpus callosum region.

	ALL_Model		Male_Model		Female_Model	
	With CC	Without CC	With CC	Without CC	With CC	Without CC
MAE	5.2508	7.7777	6.3234	8.0344	5.7555	8.4124
RMSE	6.4215	9.2961	7.7338	9.6469	7.0057	10.201
Radius	0.96425	0.93054	0.94582	0.91073	0.93945	0.91764
Radius ²	0.92986	0.86602	0.89513	0.82725	0.88257	0.84228
Slope	0.78307	0.62174	0.75018	0.63040	0.75067	0.63020
Intercept	4.7428	11.520	4.9793	11.436	6.0545	9.3505

CC: corpus callosum.

3.3. Results of Statistical Analysis on the Accuracy of Age Estimation between Male and Female Age Estimation Models

As shown in Table 4, the Wilcoxon rank sum test should be used in statistical analysis using AE. Therefore, Table 7 shows the results of the Wilcoxon rank sum test to examine whether there is a statistically significant difference in the accuracy of age estimation between men and women. Table 7 shows that in 17 of 25 cases, the p value is higher than the significance level of 0.05, indicating that there is no statistically significant difference between the two cases. In the remaining eight cases, the p values are below the significance level of 0.05, indicating that there is a statistically significant difference. These results indicate that there is no statistically significant difference in the age-estimated sex between men and women when evaluated as a whole.

Table 7. Wilcoxon rank sum test.

		Female_Model				
		Model_A	Model_B	Model_C	Model_D	Model_E
Male_Model	Model_A Model_B Model_C Model_D Model E	p = 0.0007 p = 0.0817 p < 0.0001 p = 0.0537 p = 0.1766	p = 0.6734 p = 0.0328 p = 0.5071 p = 0.0513 p = 0.0017	p = 0.7517 p = 0.1345 p = 0.1957 p = 0.1949 p = 0.0201	p = 0.9815 p = 0.0808 p = 0.2895 p = 0.1147 p = 0.0068	p = 0.0037 p = 0.2647 p = 0.0001 p = 0.1889 p = 0.3877

4. Discussion

To the best of our knowledge, this is the first study to examine age estimation methods based on brain MR images using deep learning over a wide age range from young people, including infants, to elderly individuals. Therefore, we were concerned that the age estimation would be affected by the difference in the head standard due to the age range of the target population, but we did not observe any decrease in accuracy even after data augmentation by scaling down/equalizing the images. In fact, the accuracy of age estimation improved as the number of data increased with the implementation of data augmentation, and we expect further improvement in accuracy by adding more data, including data in age ranges where data are lacking, and increasing the number of data through further implementation of data augmentation methods. A further increase in the number of data is expected to improve the accuracy.

Compared with previous studies, the results of MAE are inferior to those of the previous studies, whereas the results of correlation coefficients are comparable to those of previous studies. Jiang et al. [32] reported the results of age estimation using 2D images segmented into gray matter and white matter in the axial, sagittal, and coronal sections. The results show that the best estimation accuracy is obtained for the combination of sagittal and GM, recording 3.57 years, which exceeds the MAE of 5.25 years in the present study. Hepp et al. [31] also reported an age estimation method for 10,691 subjects, which is 10 times larger than the present study, although using 3D brain MR images. In this report, the MAE was 3.21 years, which is approximately two years less than that of the present study. This may be due to the fact that more features can be obtained from each subject using 3D images rather than using 2D images and that more features can be obtained from each subject per age than in the present study because more subjects were included in the study. The reason for this may be that the number of features that can be obtained from each age group is larger than that of the present study. Conversely, Masaru et al. [30] reported an age estimation method using 3D brain MR images for 1101 subjects aged 20–80 years. In this report, the correlation coefficient was 0.96, which was equivalent to that of the present study, although the MAE was 3.67 years, which was approximately 1.6 years smaller than that of the present study. Although the differences in MAE may be attributed to the differences in the number of subjects and the image format, an overview of the estimated age against CA suggests that this method already has the performance to show the same trend as the previous studies.

The abovementioned comparison with previous studies suggests that the estimation accuracy of the present method will be further improved with an increase in the number of data and similar analysis using 3D-CNN.

Next, as for the analysis of the significant region, the accuracy of age estimation significantly decreased before and after the deletion of the corpus callosum region in the overall evaluation of this method. This result suggests that the corpus callosum is a useful region for age estimation, and the results show a similar trend to those obtained in previous studies. The corpus callosum connects the left and right cerebral hemispheres, which are separated by the median longitudinal fissure, and therefore plays a role in the transmission of information between the two hemispheres, as it is composed of more than 190 million axons. Cognitive functions are affected by cerebral corpus callosum dysplasia and cerebral corpus callosum disconnection [36,37], and considering this, it can be said that the corpus callosum plays a role in brain functions corresponding to higher brain dysfunction and cognitive dysfunction. Thus, taking into account that brain atrophy and cognitive decline progress with aging, it is considered that some morphological changes have occurred in the corpus callosum region due to aging and that the AI focused on these changes to estimate the age of the subjects. However, although we focused only on the corpus callosum region in this study, we cannot deny that other regions may have had a significant impact on age estimation. Therefore, the influence of other regions on age estimation should be examined by using heat mapping techniques, such as Grad-CAM [38] and occlusion sensitivity [39].

Finally, regarding the accuracy of age estimation between Male_Model and Female_Model, no statistically significant difference in estimation accuracy was found in 17 of 25 cases in this study, and significant differences were found in eight cases. Therefore, since no statistically significant differences were found overall, the results suggest the possibility of data collection methods that focus on age distribution regardless of sex in the field of research on age estimation by combining deep learning techniques and brain MR images. For example, the results indicate that the brain MR images of a "42-year-old man" and a "42-year-old woman" can be treated as data of a 42-year-old person. Therefore, this result supports the fact that it is possible to collect data more easily without falling into the data shortage caused by the consideration of the sex ratio at the time of data collection, as in the present study.

However, in eight of 25 cases (32%), there was a significant difference in the accuracy of age estimation. There are various theories regarding anatomical differences in brain structure. In particular, morphological differences in the corpus callosum between men and

women were raised by Lacoste-Utamsing and Holloway [40], but their truth or falsehood has been debated due to the recent spread of MRI and CT equipment [41–43]. Therefore, it is difficult to explain the eight cases in which significant differences were found between men and women because of the difference in brain structures between sexes, but some structural differences indeed caused the significant differences. Therefore, it is important to confirm how the current results will change when the volume of data is increased and to examine how the results are represented by heat mapping techniques, such as Grad-CAM and occlusion sensitivity, as described above.

There was a limitation in this study. Even though ResNet-50 was used in this study, original network models, rather than pre-trained models, are considered to be important in order to create an age estimation model with higher performance. Considering the fact that previous study [17] has already confirmed the use of original networks, we believe that it will be necessary to consider the network model to be used in order to create an age estimation model that shows higher performance in the future.

5. Conclusions

Although the MAE of this method is lower than that of the previous study, the degree of correlation is comparable to that of the previous study. Since the degree of correlation is already high, it is expected that the estimation accuracy will be higher and the age estimation model with higher generalization performance will become closer by increasing the volume of data through the implementation of the data collection method suggested in this study and further data augmentation methods. In this study, chronological age (CA) estimation was performed using deep learning techniques on brain MR images from individuals across a wide range of ages from young to old, including infants, who had not been included in previous studies. The results showed that age estimation, including young people such as infants, showed correlations equivalent to those in previous studies. In addition, it would be essential to investigate a method to perform age estimation on 3D brain MR images using 3D-CNN in order to improve the accuracy of age estimation in the future.

Author Contributions: K.U. contributed to the data analysis, algorithm construction, and writing and editing of the manuscript. T.Y. and M.T. reviewed and edited the manuscript. H.S. proposed the idea and contributed to the data acquisition, performed supervision and project administration, and reviewed and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The created models in this study are available on request from the corresponding author. The source code of this study is available at https://github.com/MIAlaboratory/AgeEstimation (accessed on 16 December 2022).

Acknowledgments: The authors would like to thank the laboratory members of the Medical Image Analysis Laboratory and Yoshimura's laboratory for their help.

Conflicts of Interest: The authors declare that no conflict of interest exists.

References

- Whitaker, K.J.; Vértes, P.E.; Romero-garcia, R.; Moutoussis, M.; Prabhu, G. Adolescence Is Associated with Genomically Patterned Consolidation of the Hubs of the Human Brain Connectome. *Proc. Natl. Acad. Sci. USA* 2016, *113*, 9105–9110. [CrossRef] [PubMed]
- Fjell, A.M.; Walhovd, K.B.; Fennema-notestine, C.; Mcevoy, L.K.; Hagler, D.J.; Holland, D.; Brewer, J.B.; Dale, A.M. One-Year Brain Atrophy Evident in Healthy Aging. J. Neurosci. 2009, 29, 15223–15231. [CrossRef]
- der Mauer, M.A.; van Well, E.J.; Herrmann, J.; Groth, M.; Morlock, M.M.; Maas, R.; Säring, D. Automated Age Estimation of Young Individuals Based on 3D Knee MRI Using Deep Learning. *Int. J. Legal Med.* 2021, 135, 649–663. [CrossRef] [PubMed]

- 4. Willems, G.; Moulin-Romsee, C.; Solheim, T. Non-Destructive Dental-Age Calculation Methods in Adults: Intra- and Inter-Observer Effects. *Forensic Sci. Int.* 2002, 126, 221–226. [CrossRef] [PubMed]
- Good, C.D.; Johnsrude, I.S.; Ashburner, J.; Henson, R.N.A.; Friston, K.J.; Frackowiak, R.S.J. A Voxel-Based Morphometric Study of Ageing in 465 Normal Adult Human Brains. *Neuroimage* 2001, 14, 21–36. [CrossRef]
- Seeley, W.W.; Crawford, R.K.; Zhou, J.; Miller, B.L.; Greicius, M.D. Neurodegenerative Diseases Target Large-Scale Human Brain Networks. *Neuron* 2009, 62, 42–52. [CrossRef]
- Bin Zahid, A.; Mikheev, A.; Srivatsa, N.; Babb, J.; Samadani, U.; Rusinek, H. Accelerated Brain Atrophy on Serial Computed Tomography: Potential Marker of the Progression of Alzheimer Disease. J. Comput. Assist. Tomogr. 2016, 40, 827–832. [CrossRef]
- 8. Chang, A.Y.; Skirbekk, V.F.; Tyrovolas, S.; Kassebaum, N.J.; Dieleman, J.L. Measuring Population Ageing: An Analysis of the Global Burden of Disease Study 2017. *Lancet Public Health* **2019**, *4*, e159–e167. [CrossRef]
- Tyas, S.L.; Manfreda, J.; Strain, L.A.; Montgomery, P.R. Risk Factors for Alzheimer's Disease: A Population-Based, Longitudinal Study in Manitoba, Canada. Int. J. Epidemiol. 2001, 30, 590–597. [CrossRef]
- Revell, A.J.; Caskie, G.I.L.; Willis, S.L.; Schaie, K.W. Factor Structure and Invariance of the Quality of Life in Alzheimer's Disease (Qol-Ad) Scale. *Exp. Aging Res.* 2009, 35, 250–267. [CrossRef]
- Wimo, A.; Jönsson, L.; Bond, J.; Prince, M.; Winblad, B. The Worldwide Economic Impact of Dementia 2010. *Alzheimer's Dement*. 2013, 9, 1–11.e3. [CrossRef] [PubMed]
- 12. Schmeling, A.; Grundmann, C.; Fuhrmann, A.; Kaatsch, H.-J.; Knell, B.; Ramsthaler, F.; Reisinger, W.; Riepert, T.; Ritz-Timme, S.; Rösing, F.W.; et al. Criteria for Age Estimation in Living Individuals. *Int. J. Legal Med.* **2008**, 122, 457–460. [CrossRef] [PubMed]
- De Tobel, J.; Bauwens, J.; Parmentier, G.I.L.; Franco, A.; Pauwels, N.S.; Verstraete, K.L.; Thevissen, P.W. Magnetic Resonance Imaging for Forensic Age Estimation in Living Children and Young Adults: A Systematic Review. *Pediatr. Radiol.* 2020, 50, 1691–1708. [CrossRef] [PubMed]
- 14. Widek, T.; Genet, P.; Ehammer, T.; Schwark, T.; Urschler, M.; Scheurer, E. Bone Age Estimation with the Greulich-Pyle Atlas Using 3T MR Images of Hand and Wrist. *Forensic Sci. Int.* **2021**, *319*, 110654. [CrossRef] [PubMed]
- 15. Sugimori, H.; Sugiyama, T.; Nakayama, N.; Yamashita, A. Development of a Deep Learning-Based Algorithm to Detect the Distal End of a Surgical Instrument. *Appl. Sci.* **2020**, *10*, 4245. [CrossRef]
- 16. Sugimori, H. Evaluating the Overall Accuracy of Additional Learning and Automatic Classification System for CT Images. *Appl. Sci.* **2019**, *9*, 682. [CrossRef]
- 17. Manabe, K.; Asami, Y.; Yamada, T.; Sugimori, H. Improvement in the Convolutional Neural Network for Computed Tomography Images. *Appl. Sci.* 2021, *11*, 1505. [CrossRef]
- 18. Sugimori, H.; Hamaguchi, H.; Fujiwara, T.; Ishizaka, K. Classification of Type of Brain Magnetic Resonance Images with Deep Learning Technique. *Magn. Reson. Imaging* **2021**, *77*, 180–185. [CrossRef] [PubMed]
- 19. Yoshimura, T.; Nishioka, K.; Hashimoto, T.; Mori, T.; Kogame, S.; Seki, K.; Sugimori, H.; Yamashina, H.; Nomura, Y.; Kato, F.; et al. Prostatic Urinary Tract Visualization with Super-Resolution Deep Learning Models. *PLoS ONE* **2022**, *18*, e0280076. [CrossRef]
- 20. Asami, Y.; Yoshimura, T.; Manabe, K.; Yamada, T.; Sugimori, H. Development of Detection and Volumetric Methods for the Triceps of the Lower Leg Using Magnetic Resonance Images with Deep Learning. *Appl. Sci.* **2021**, *11*, 12006. [CrossRef]
- Mohsen, H.; El-Dahshan, E.-S.A.; El-Horbaty, E.-S.M.; Salem, A.-B.M. Classification Using Deep Learning Neural Networks for Brain Tumors. *Future Comput. Inform. J.* 2018, 3, 68–71. [CrossRef]
- Chen, H.; Dou, Q.; Yu, L.; Qin, J.; Heng, P.A. VoxResNet: Deep Voxelwise Residual Networks for Brain Segmentation from 3D MR Images. *Neuroimage* 2018, 170, 446–455. [CrossRef] [PubMed]
- Myronenko, A.; Hatamizadeh, A. Robust Semantic Segmentation of Brain Tumor Regions from 3D MRIs. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2020; pp. 82–89. [CrossRef]
- Zhou, Z.; Sanders, J.W.; Johnson, J.M.; Gule-Monroe, M.K.; Chen, M.M.; Briere, T.M.; Wang, Y.; Bum Son, J.; Pagel, M.D.; Li, J.; et al. Computer-Aided Detection of Brain Metastates in T1-Weighted MRI for Stereotactic Radiosurgery Using Deep Learning Single-Shot Detectors. *Radiology* 2020, 295, 407–415. [CrossRef] [PubMed]
- 25. Chegraoui, H.; Philippe, C.; Dangouloff-ros, V.; Grigis, A.; Calmon, R.; Boddaert, N.; Frouin, F.; Grill, J.; Frouin, V. Object Detection Improves Tumour Segmentation in MR Images of Rare Brain Tumours. *Cancers* **2021**, *13*, 6113. [CrossRef] [PubMed]
- Jonsson, B.A.; Bjornsdottir, G.; Thorgeirsson, T.E.; Gudbjartsson, D.F.; Stefansson, H.; Stefansson, K.; Ulfarsson, M.O. Brain Age Prediction Using Deep Learning Uncovers Associated Sequence Variants. *Nat. Commun.* 2019, 10, 5409. [CrossRef] [PubMed]
- 27. Ichikawa, S.; Hamada, M.; Sugimori, H. A Deep-Learning Method Using Computed Tomography Scout Images for Estimating Patient Body Weight. *Sci. Rep.* 2021, *11*, 15627. [CrossRef]
- Hagen, M.; Schmidt, S.; Rudolf, E.; Schmeling, A. The Significance of Sociopedagogic Age Estimation in Comparison to Forensic Age Assessment. *Rechtsmedizin* 2020, 30, 233–240. [CrossRef]
- Rossi, A.; Vannuccini, G.; Andreini, P.; Bonechi, S.; Giacomini, G.; Scarselli, F.; Bianchini, M. Analysis of Brain NMR Images for Age Estimation with Deep Learning. *Procedia Comput. Sci.* 2019, 159, 981–989. [CrossRef]
- Ueda, M.; Ito, K.; Wu, K.; Sato, K.; Taki, Y.; Fukuda, H.; Aoki, T. An age estimation method using 3D-CNN from brain MRI images. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019.

- Hepp, T.; Blum, D.; Armanious, K.; Schölkopf, B.; Stern, D.; Yang, B.; Gatidis, S. Uncertainty Estimation and Explainability in Deep Learning-Based Age Estimation of the Human Brain: Results from the German National Cohort MRI Study. *Comput. Med. Imaging Graph.* 2021, 92, 101967. [CrossRef]
- Jiang, H.; Guo, J.; Du, H.; Xu, J.; Qiu, B. Transfer Learning on T1-Weighted Images for Brain Age Estimation. *Math. Biosci. Eng.* 2019, 16, 4382–4398. [CrossRef]
- Tanaka, S.C.; Yamashita, A.; Yahata, N.; Itahashi, T.; Lisi, G.; Yamada, T.; Ichikawa, N.; Takamura, M.; Yoshihara, Y.; Kunimatsu, A.; et al. A Multi-Site, Multi-Disorder Resting-State Magnetic Resonance Image Database. *Sci. Data* 2021, *8*, 227. [CrossRef] [PubMed]
- Alexander, L.M.; Escalera, J.; Ai, L.; Andreotti, C.; Febre, K.; Mangone, A.; Vega-Potler, N.; Langer, N.; Alexander, A.; Kovacs, M.; et al. An Open Resource for Transdiagnostic Research in Pediatric Mental Health and Learning Disorders. *Sci. Data* 2017, *4*, 1–26. [CrossRef] [PubMed]
- 35. Kaiming, H.; Xiangyu, Z.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [CrossRef]
- Marco, E.; Harrell, K.; Brown, W.; Hill, S.; Jeremy, R.; Kramer, J.; Sherr, E.; Paul, L. Processing Speed Delays Contribute to Executive Function Deficits in Individuals with Agenesis of the Corpus Callosum. J. Int. Neuropsychol. Soc. 2012, 18, 521–529. [CrossRef]
- 37. Huang, X.; Du, X.; Song, H.; Zhang, Q.; Jia, J.; Xiao, T.; Wu, J. Cognitive Impairments Associated with Corpus Callosum Infarction: A Ten Cases Study. *Int. J. Clin. Exp. Med.* **2015**, *8*, 21991–21998. [PubMed]
- Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations From Deep Networks via Gradient-Based Localization. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 618–626.
- Zeiler, M.D.; Fergus, R. Visualizing and Understanding Convolutional Networks. In Computer Vision–ECCV 2014; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2014; ISBN 9783319105895.
- 40. De Lacoste-Utamsing, C.; Holloway, R.L. Sexual Dimorphism in the Human Corpus Callosum. *Am. Assoc. Adv. Sci.* **1982**, 216, 1431–1432. [CrossRef]
- Joel, D.; Berman, Z.; Tavor, I.; Wexler, N.; Gaber, O.; Stein, Y.; Shefi, N.; Pool, J. Sex beyond the Genitalia: The Human Brain Mosaic. Proc. Natl. Acad. Sci. USA 2015, 112, 15468–15473. [CrossRef]
- Ritchie, S.J.; Cox, S.R.; Shen, X.; Lombardo, M.V.; Reus, L.M.; Alloza, C.; Harris, M.A.; Alderson, H.L.; Hunter, S.; Neilson, E.; et al. Sex Differences in the Adult Human Brain: Evidence from 5216 UK Biobank Participants. *Cereb. Cortex* 2018, 28, 2959–2975. [CrossRef]
- 43. Ingalhalikar, M.; Smith, A.; Parker, D.; Satterthwaite, T.D.; Elliott, M.A.; Ruparel, K.; Hakonarson, H.; Gur, R.E.; Gur, R.C.; Verma, R. Sex Differences in the Structural Connectome of the Human Brain. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 823–828. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.