AN AGE ESTIMATION METHOD USING 3D-CNN FROM BRAIN MRI IMAGES

Masaru Ueda[⋆], *Koichi Ito*[⋆], *Kai Wu*[†], *Kazunori Sato*[‡], *Yasuyuki Taki*[‡], *Hiroshi Fukuda*[§] *and Takafumi Aoki*[⋆]

[⋆] Graduate School of Information Sciences, Tohoku University, Japan.
[†] South China University of Technology, China.
[‡] Institute of Divelopment, Aging and Cancer, Tohoku University, Japan.
[§] Tohoku Medical and Pharmaceutical University, Japan.

ABSTRACT

A specific pattern of morphological changes in the human brain is observed during the process of brain development and healthy aging. The age of subjects can be estimated from brain images by evaluating such patterns. This paper proposes an age estimation method using 3-Dimensional Convolutional Neural Network (3D-CNN) from brain T1-weighted images so as to fully utilize the potential of volume data. Through a set of experiments using over 1,000 T1-weighted images of healthy Japanese, we demonstrate that the proposed method exhibits better performance on age estimation than the conventional methods using handcrafted local features and 2D-CNN.

***Index Terms*—** MRI, T1-weighted image, age estimation, brain aging, deep learning, CNN

1. INTRODUCTION

The statistical analysis of brain MR images has presented that a specific pattern is observed in brain morphological changes during the process of brain development and healthy aging. Volumetric changes of brain tissues, i.e., gray matter (GM), white matter (WM) and Cerebrospinal fluid (CSF), are observed during the normal aging process [1, 2, 3, 4]. Taki et al. [4] have demonstrated that GM volume monotonically decreases, WM volume shows small changes, and CSF monotonically increases in contrast with GM volume during the aging process. This result allows us to estimate the age of subjects from T1-weighted images, which is one of MR images and has been used to analyze volumetric changes in brain sciences. In addition, the analysis between the estimated and actual age might become a biomarker for early identification and diagnostic support of age-related brain disorders, since neurodegenerative diseases such as Alzheimer's disease (AD) have caused the accelerated brain atrophy.

There is some literature on age estimation from T1-weighted images [5, 6, 7, 8]. These methods employ high-order features, all the voxels or combined information to estimate the age from T1-weighted images. In addition, only a few hundred MR images have been used to evaluate the accuracy of methods. We have proposed two approaches of age estimation [9, 10, 11, 12]. The first approach [9, 10, 12] employed brain local features defined by regional volume from 90 or 1,024 local regions of GM, WM and CSF parcellated by the automated anatomical labeling (AAL) atlas [13, 14]. Relevance vector regression (RVR) [15] was used to construct the age estimation model. The second approach [11] employed the 2-Dimensional Convolutional Neural Network (2D-CNN) to automatically extract features and estimate the age from T1-weighted images. The 2D-CNN-based method is higher prediction accuracy than other methods, while 3D features may not be utilized in age estimation, since 2D convolution was used to extract features from slice images of volume data.

In this paper, we propose an age estimation method using 3D-CNN [16] in order to fully utilize volume data in age estimation. We consider a new and simple 3D-CNN architecture for age estimation from T1-weighted images and some performance improvement techniques for 3D-CNN. We demonstrate that the proposed method exhibits efficient performance on age estimation compared with conventional methods through a set of experiments using T1-weighted images database collected by Aoba brain image research center project [17] and Sendai Tsurugaya project.

2. METHOD

The proposed method consists of preprocessing and age estimation. The following is the detail of each step and the training process with some techniques.

2.1. Preprocessing

Preprocessing has to be applied to T1-weighted images before age estimation. We empirically confirmed that preprocessing improves the accuracy of age estimation as well as 2D-CNN [11]. Preprocessing used in the proposed method consists of the following 3 steps. First, all the T1-weighted images are aligned with the standard template, i.e., the ICBM 152 template, using the procedure proposed by Good et al. [18]. Statistical Parametric Mapping 2 (SPM2)[1] and Voxel-Based Morphometry 2 (VBM2)[2] are used in this step, since the T1-weighted images used in the experiment are taken by a 0.5T MR scanner. Note that the newer version of SPM and VBM should be used when T1-weighted images are taken by a 1.5T or 3.0T MR scanner. Next, the aligned images are resized to $95 \times 79 \times 78$ voxels to reduce the computation time and the memory usage. Finally, the voxel values are normalized so as to have zero mean and unit variance.

2.2. Architecture

Fig. 1 shows a 3D-CNN architecture used in the proposed age estimation method. This architecture has 4 convolution blocks consisting of a 3D convolutional layer (kernel size: $3 \times 3 \times 3$, stride: 1),

[1] `http://www.fil.ion.ucl.ac.uk/spm/software/spm2/`
[2] `http://dbm.neuro.uni-jena.de/vbm/vbm2-for-spm2/`

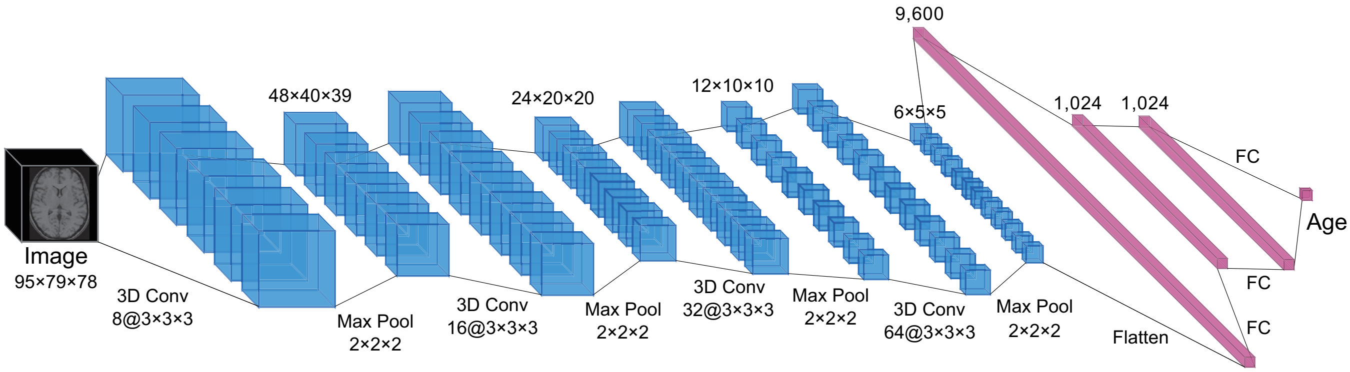

Fig. 1. 3D-CNN architecture used in the proposed age eatimation method.

a 3D batch-normalization layer, a rectified linear unit (ReLU) activation layer and a max-pooling layer (kernel size: $2 \times 2 \times 2$, stride: 2). The number of feature channels is 8, 16, 32, 64 for each block, respectively. The last 3 layers are fully-connected layers to combine the feature vectors. The output is a scalar value of the predicted age. We use TensorFlow and Keras to implement the 3D-CNN architecture in this paper.

2.3. Training

The hyper-parameters used in the experiment are shown in Table 1. This architecture is designed to be simple, taking into account the large computation time and the large memory usage in 3D-CNN. The network weights are trained so as to minimize the Mean Squared Error (MSE) using the Stochastic Gradient Descent (SGD) optimization with momentum.

The loss function L used in the proposed 3D-CNN is defined by

$$L = \frac{1}{N}\sum_{n}(t_n - y_n)^2 + \frac{1}{2}\lambda \mathbf{W^2}, \quad (1)$$

where N is the number of the data, t_n is the actual age of the n-th subject, y_n is the predicted age of the n-th subject, λ is the weight decay and $\mathbf{W}$ is the weights to be trained. The loss function includes the regularization term to prevent overfitting in the training. The regularization term is calculated by the weight decay and L2 norm of all the weights. The learning rate η is determined for every epoch as follows:

$$\eta = \frac{\eta_0}{1 + (epoch \times decay_\eta)}, \quad (2)$$

where η_0 is the initial value of learning rate, $decay_\eta$ is the learning rate decay and $epoch$ is the number of epochs.

All the training data have to be augmented to improve the accuracy of the proposed method. In the proposed method, we employ the three augmentation methods: (i) flipping, (ii) scaling and (iii) shifting. Note that the augmentation methods used in the proposed method add only realistic variation to T1-weighted images. We empirically confirmed that the estimation accuracy is decreased by other augmentation methods such as rotation, noise, deformation, etc. The method (i) flips input images horizontally with a probability of 0.5. The method (ii) randomly crops input images within $[0, 10]$-voxel around and resizes it to the original size. The method (iii) randomly shifts input images within $[-8, 8]$ voxels.

Table 1. Hyper-parameters for the proposed 3D-CNN.

Hyper-parameter	Value
Initial value of learning rate η_o	0.00005
Learning rate decay $decay_\eta$	0.0001
Weight decay λ	0.0005
Batch size	16
Momentum	0.9

3. EXPERIMENTS

This section describes experiments for evaluating the proposed method using a large-scale dataset.

3.1. Experimental Condition

T1-weighted MR images collected by the Aoba Brain Imaging Project (Aoba 1) in Sendai, Japan [17] and the Tsurugaya Project (Tsurugaya 1) in Sendai, Japan are used in the experiment. T1-weighted images were taken by the same 0.5T MR scanner (Signa contour, GE-Yokogawa Medical Systems, Tokyo) in both projects, where the image size is $256 \times 256 \times 124$ voxels. The image size is changed to $189 \times 157 \times 156$ voxels after alignment with the ICBM 152 template and then is resized to $95 \times 79 \times 78$ voxels.

The subjects of both projects were all healthy and had neither present illness nor a history of neurological disease, psychiatric disease, brain tumor or head injury. 1,101 subjects aged from 20 to 80 years from the dataset are used in the experiment. The age distribution of images used in the experiment is shown in Fig. 2. We randomly select 768 subjects for the training data and also select the remaining 333 subjects for test data. We evaluate the accuracy of age estimation using the Mean Absolute Error (MAE), the Root Mean Square Error (RMSE) and the correlation coefficient (Coor.). We compare the estimation accuracy of our proposed method with conventional methods: PCA [7], brain local features [9] and 2D-CNN [11].

3.2. Experimental Results

We evaluate the training techniques for the proposed method such as voxel normalization and data augmentation. Fig. 3 shows the training and testing errors with and without training techniques in the proposed method. In the case without data augmentation, the overfitting is observed as shown in Fig. 3 (a) and (c). In the case

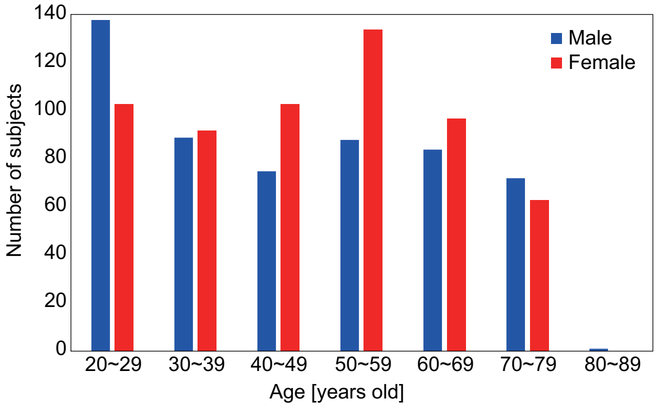

Fig. 2. Age distribution of Aoba and Tsurugaya dataset.

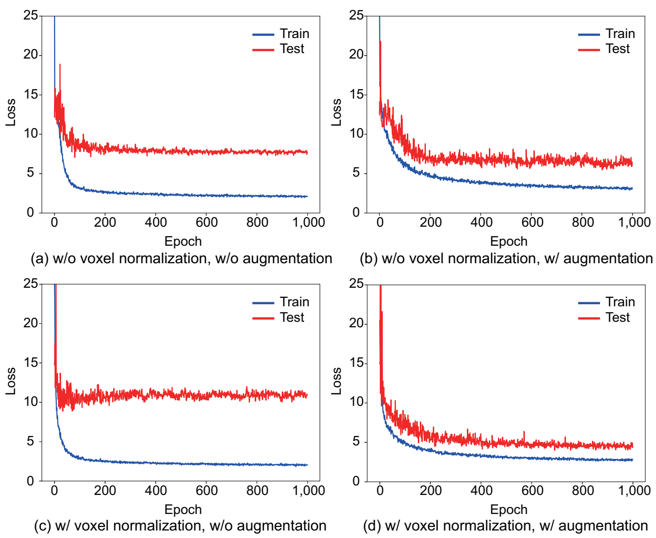

Fig. 3. Training and testing errors with and without training techniques in the proposed method.

without voxel normalization, the loss is not below 5 in the testing, while that is below 5 in the training as shown in Fig. 3 (a) and (b). In the case of the combined use of voxel normalization and data augmentation, there is no overfitting and the 3D-CNN model is well trained as shown in Fig. 3 (d).

Table 2 shows a summary of experimental results. The CNN-based methods exhibit higher estimation accuracy than the hand-crafted feature-based methods [7, 9]. The accuracy of the proposed method is much higher than that of 2D-CNN [11], since the 3D-CNN-based method can fully utilize volume data in age estimation.

We analyze effective local regions in the age estimation in order to discuss medical implication of the experimental result. We add the random noise to one local region of T1-weighted images and then estimate the age using the proposed method. If the MAE increases, the effectiveness of the masked local region is high. If the MAE decreases, its effectiveness is low. The masked region is determined by the 90 AAL atlas [13] or the 1,024 AAL atlas [14]. The use of the above procedure makes it possible to evaluate the effective local regions in the CNN-based approach although, in general, it is difficult for CNN-based approaches to explain which local features are effective.

Table 2. Sumarry of experimental results.

Method		MAE [y/o]	RMSE [y/o]	Corr.
Franke [7]	PCA	4.61	5.90	0.94
Kondo [9]	Local features	4.34	5.56	0.94
Huang [11]	2D-CNN	4.04	5.13	0.94
Proposed	3D-CNN	3.67	4.71	0.96

Fig. 4 shows effective local regions for each method. Effective local regions are concentrated in the frontal association area, the Wernicke's area, the angularis gyrus and the primary motor cortex. The frontal association area takes on the function of behavioral decision and working memory. The functions of the Wernicke's area include language comprehension, semantic processing, language recognition and language interpretation. The angularis gyrus takes on the functions related to language, spatial cognition, memory retrieval, attention and theory of mind. The above regions take on the high-order function compared with other regions, and hence are impaired with aging. The primary motor cortex exhibits high effectiveness in age estimation, although this region takes on the low-order function. The location of the primary motor cortex is close to that of the central sulcus. The central sulcus becomes dilated by atrophying the frontal area and the primary motor cortex also becomes dilated which looks like atrophy. On the other hand, ineffective local regions are concentrated in the parietal lobe and the occipital lobe. These regions take on the low-order function and hence are robust against aging. The above result corresponds to the statistical analysis of age-related morphological changes [19]. The proposed method shows a more clear trend than other methods, although all the methods indicate almost the same trend in the statistical analysis.

4. CONCLUSION

This paper proposed an age estimation method using 3D-CNN from T1-weighted images. The use of 3D-CNN makes it possible to fully utilize volume data in age estimation. We demonstrated that the proposed method exhibits efficient performance on age estimation compared with conventional methods through a set of experiments using the large-scale healthy Japanese MR images. We also investigated effective local regions in age estimation and demonstrated that our 3D-CNN model would be trained so as to correspond to the statistical analysis of age-related morphological change. In the future, we will apply the proposed method to early identification and diagnostic support of age-related brain disorders such as Alzheimer's disease.

5. REFERENCES

[1] T.L. Jernigan, S.L. Archibald, C. Fennema-Notestine, A.C. Gamst, J.C. Stout, J. Bonner, and J.R. Hesselink, "Effects of age on tissues and regions of the cerebrum and cerebellum," *Neurobiol. Aging*, vol. 22, pp. 581–594, 2001.

[2] J.S. Allen, J. Bruss, C.K. Brown, and H. Damasio, "Normal neuroanatomical variation due to age: The major lobes and a parcellation of the temporal region," *Neurobiol. Aging*, vol. 26, pp. 1245–1260, 2005.

[3] D. Terribilli, M.S. Ŝchaufelberger, F.L.S. Duran, M.V. Zanetti, P.K. Curiati, P.R. Menezes, M.S. Scazufca, E. Amaro Jr., C.C.

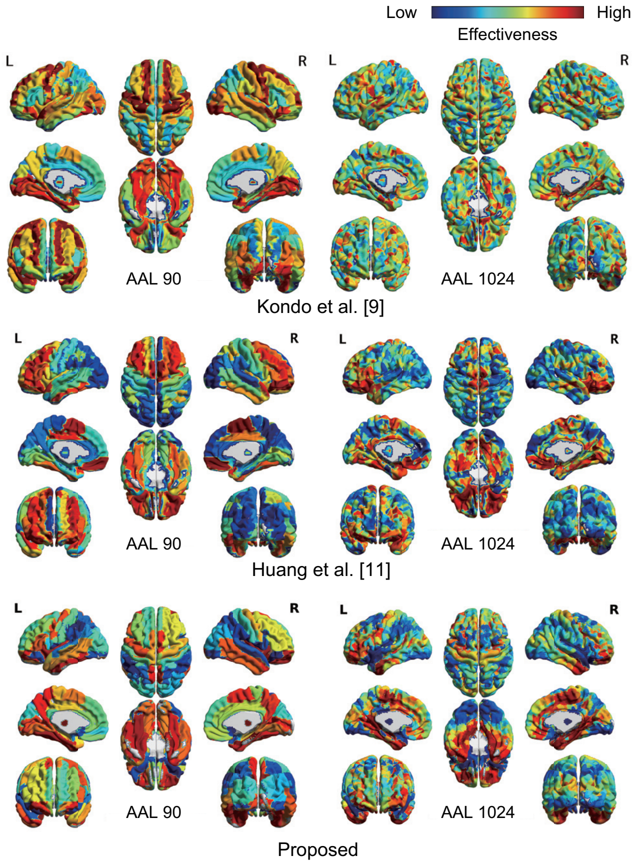

Fig. 4. Effective local regions in age estimation for each method.

Leite, and G.F. Busatto, "Age-related gray matter volume changes in the brain during non-elderly adulthood," *Neurobiol. Aging*, vol. 32, pp. 354–368, 2011.

[4] Y. Taki, B. Thyreau, S. Kinomura, K. Sato, R. Goto, R. Kawashima, and H. Fukuda, "Correlations among brain gray matter volumes, age, gender, and hemisphere in healthy individuals," *PLoS One*, vol. 6, no. 7, pp. e22734–1–e22734–13, Apr. 2011.

[5] Z. Lao, D. Shen, Z. Xue, B. Karacali, S. M. Resnick, and C. Davatzikos, "Morphological classification of brains via high-dimensional shape transformations and machine learning methods," *NeuroImage*, vol. 21, no. 1, pp. 46–57, Jan. 2004.

[6] H. Neeb, K. Zilles, and N. J. Shah, "Fully-automated detection of cerebral water content changes: Study of age- and gender-related H2O patterns with quantitative MRI," *NeuroImage*, vol. 29, no. 3, pp. 910–922, Feb. 2006.

[7] K. Franke, G. Ziegler, S. Kloppel, and C. Gaser, "Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: Exploring the influence of various parameters," *NeuroImage*, vol. 50, no. 3, pp. 883–892, Apr. 2010.

[8] J. Wang, W. Li, W. Miao, D. Dai, J. Hua, and H. He, "Age estimation using cortical surface pattern combining thickness with curvatures," *Med. Biol. Eng. Comput.*, vol. 52, no. 4, pp. 331–341, Apr. 2014.

[9] C. Kondo, K. Ito, K. Wu, K. Sato, Y. Taki, H. Fukuda, and T. Aoki, "An age estimation method using brain local features for T1-weighted images," *Proc. Int'l Conf. IEEE Eng. Med. Biol. Soc.*, pp. 666–669, Aug. 2015.

[10] R. Fujimoto, C. Kondo, K. Ito, K. Wu, K. Sato, Y. Taki, H. Fukuda, and T. Aoki, "Age estimation using effective brain local features from T1-weighted images," *Proc. Int'l Conf. IEEE Engineering in Medicine and Biology Society*, pp. 5941–5944, Aug. 2016.

[11] T. W. Huang, H. T. Chen, R. Fujimoto, K. Ito, K. Wu, K. Sato, Y. Taki, H. Fukuda, and T. Aoki, "Age estimation from brain MRI images using deep learning," *Proc. Int'l Symp. Biomed. Imaging*, pp. 849–852, Apr. 2017.

[12] R. Fujimoto, K. Ito, K. Wu, K. Sato, Y. Taki, H. Fukuda, and T. Aoki, "Brain age estimation from T1-weighted images using effective local features," *Proc. Int'l Conf. IEEE Engineering in Medicine and Biology Society*, pp. 3028–3031, July 2017.

[13] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI single-subject brain," *NeuroImage*, vol. 15, no. 1, pp. 273–289, Jan. 2002.

[14] A. Zalesky, A. Fornitoa, I.H. Hardinga, L. Cocchia, M. Yucela, C. Pantelisa, and E.T. Bullmorect, "Whole-brain anatomical networks: Does the choice of nodes matter?," *NeuroImage*, vol. 50, no. 3, pp. 970–983, Apr. 2010.

[15] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Machine Learning Research*, vol. 1, pp. 211–244, 2001.

[16] S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition," *IEEE Trans. Pattern Anal. Mech. Intell.*, vol. 35, no. 1, pp. 221–223, Jan. 2013.

[17] K. Sato, H. Fukuda, and R. Kawashima, "Neuroanatomical database of normal Japanese brains," *Neural Networks*, vol. 16, no. 9, pp. 1301–1310, Nov. 2003.

[18] C. D. Good, I. S. Johnsrude, J. Ashburner, R. N. A. Henson, and K. J. Friston, "A voxel-based morphometric study of ageing in 465 normal adult human brains," *NeuroImage*, vol. 14, no. 1, pp. 21–36, 2001.

[19] Y. Taki, R. Goto, A. Evans, A. Zijdenbos, P. Neelin, J. Lerch, K. Sato, S. Ono, S. Kinomura, M. Nakagawa, M. Sugiura, J. Watanabe, R. Kawashima, and H. Fukuda, "Voxel-based morphometry of human brain with age and cerebrovascular risk factors," *Neurobiol. Aging*, vol. 25, no. 4, pp. 455–463, Apr. 2004.