Significant Dimension Reduction of 3D Brain MRI using 3D Convolutional Autoencoders

Hayato Arai*, Yusuke Chayama*, Hitoshi Iyatomi, and Kenichi Oishi
for the Alzheimer's Disease Neuroimaging Initiative

***Abstract*— Content-based image retrieval (CBIR) is a technology designed to retrieve images from a database based on visual features. While the CBIR is highly desired, it has not been applied to clinical neuroradiology, because clinically relevant neuroradiological features are swamped by a huge number of noisy and unrelated voxel information. Thus, effective dimension reduction is the key to successful CBIR. We propose a novel dimensional compression method based on 3D convolutional autoencoders (3D-CAE), which was applied to the ADNI2 3D brain MRI dataset. Our method succeeded in compressing 5 million voxel information to only 150 dimensions, while preserving clinically relevant neuroradiological features. The RMSE per voxel was as low as 8.4%, suggesting a promise of our method toward the application to the CBIR.**

I. INTRODUCTION

Neurological diseases are the most common cause of disability in the world [1]. Neuroradiology plays an essential role in the diagnosis and management of neurological disorder. Among various neuroimaging modalities, magnetic resonance image (MRI) of the brain has been widely used because of the capability to delineate various types of tissues, anatomical structures, and pathologies within the brain non-invasively. While huge amount of brain MRIs are routinely scanned for diagnosis or evaluation of the neurological diseases and stored in the medical Picture Archive and Communication System (PACS), usually only linguistic encoding from the radiologists is stored in the searchable patient record. Once reviewed by physicians and utilized for the medical decision making, the stored MRIs are rarely used for the secondary purpose, although this Big Data that consists with huge amount of brain MRIs and corresponding medical records (MRI Big Data) has potential to provide new evidences about diagnosis and treatment of various diseases.

Modern medicine is evidence based. To improve current medical practice, efforts have been made to extract clinically relevant information from existing clinical records. The clinical MRI Big Data, from which diagnostic or prognostic information could be obtained from MRI-derived features, is expected to play an essential role in the effort to learn from past clinical cases to improve cure and care of the future patients. To fully exploit the value of the MRI Big Data, data searching technology is indispensable. Currently, text-based searching is widely used for the retrieval of medical images. To retrieve images of interest from a database, users are expected to have a fair amount of knowledge to correctly describe image features. However, the verbalization itself is often demanding, simply because there is too much amount of information included in one image. Therefore, a technology designed to retrieve images from a database based on image-derived features, which is generally called content-based image retrieval (CBIR), is strongly desired.

Several attempts have been made to apply the CBIR techniques on 3D images [2], [3], although the applicability to a general CBIR system is not fully evaluated and therefore still an open issue. Particularly for the three dimensional and high-resolution brain MRI, "curse of dimensionality" problem plagues many CBIR studies, i.e., the number of dimension, is very large, while that of available cases is extremely small with respect to that dimension. From the viewpoint of machine learning, high dimensional data is difficult to perform discrimination, regression, or clustering which is deeply involved in CBIR techniques, and an appropriate dimension reduction is essential. However, there are not many explicit dimensional reduction techniques for these diagnostic modalities. Yaacoub et al. [4] proposed dimensional reduction on MRI image with singular value decomposition (SVD) and least square estimation. Their method showed superior performance to other methods they gave, however almost 40% dimensional reduction at the expense of 20% reconstruction error in average is unsatisfactory for our purpose. Lyra-Leite et al. [5] obtained low-rank representation of MRI image with SVD and attained reasonable compression performance while preserving visual features. However, since they focused only per-image compression on 2D image slice i.e. SVD was performed on per slice, obtained basis vector is different among images, this can not be used for our general dimensional compression. Compressed sensing [6], [7] estimates many unknown parameters using much less observations under the assumption that they have sparse representation in some orthogonal basis. CS finds similarities with dimensional reduction from the viewpoint of utilizing intrinsic low dimensionality of data. CS has mainly used for reconstruction of images; shortening the imaging time of CT/MRI and improving image quality. To the best of our knowledge, on the other hand, CS is not applied for dimensional reduction for already reconstructed images.

Meanwhile image recognition performance of recent machine learning techniques, especially convolutional neural

*These two authors contributed equally.

Hayato Arai, Yusuke Chayama, and Histoshi Iyatomi are with the Department of Applied Informatics, Faculty of Science and Engineering, Hosei University, Tokyo, Japan (e-mail: {yusuke.chayama.2t@stu., hayato.arai.5t@stu., iyatomi@}hosei.ac.jp).

Kenichi Oishi is with the Department of Radiology, Johns Hopkins University School of Medicine, MD, USA (email: koishi2@jhmi.edu).

978-1-5386-3646-6/18/$31.00 ©2018 IEEE

networks (CNNs) have improved drastically in the field of computer vision [8], [9]. Improvement in computational resources have enabled manipulation of high resolution images, and automated recognition of visual features that doctors note during diagnosis of medical images have become pragmatic. Furthermore, by extending the function of CNNs that have shown promising results on 2D images to 3D objects, we are able to capture 3D objects not as continuous slices of 2D images, but objects with information of all three spatial dimensions. Recently, 3D-CNN began to be used for analysis of 3D modalities; mainly for segmentation of organs or involved areas in CT [10] and MRI [11], [12]. Hosseini-Asl et al. [13] introduces 3D-CNN approach on a classification task of predicting Alzheimer's disease (AD) using 3D brain MRI. This method has achieved promising results, exceeding conventional methods that uses image slice as inputs. However, classifiers specifying to limited disease types are not applicable to search systems that need adaptation to biologically and pathologically heterogeneous clinical cases. By constructing a feature extractor that captures features of the entire brain, we are able to acquire a practical feature representation of the input. Furthermore, in order to quantitatively decide the similarity of the patient's brain to past cases, extracted features must be grouped into sets based on distance metrics by a clustering process.

In this paper, we propose a new and significant dimensional compression method for 3D MRI images using 3D convolutional autoencoders (3D-CAE) as an important element for realization of their CBIR system. Our 3D-CAE is basically composed of two 3D-CNNs in a mirror image state. The results were evaluated visually by a board-certified neurologist who has extensive experience in neuroradiology, and also quantitatively evaluated by multiple criteria.

Key contribution of this paper are: (1) To our best knowledge, there have been no practical dimensional reduction for 3D brain MRI and we accomplished it with 3D-CAE and (2) With our method, 3D brain MRI having around 5 million dimension data is compressed to only 150 dimensions, while important clues for image diagnosis for typical diseases can be preserved.

II. METHOD

A. Datasets and Preprocessing

The 3D brain MRI data used for training and evaluation of the 3D-CAE are a part of the ADNI2 dataset [14]. For training we used 112 amyloid-positive Alzhiemer's disease (AD) cases and 146 amyloid-negative cognitively normal (CN), whereas for evaluation we used 152 cases of patients diagnosed as Early Mildly Cognitive Impairment (EMCI). The MRICloud (www.MRICloud.org) was used to extract the brain area, to correct intensity inhomogeneity, and to linearly transform the brain area to the Montreal Neurological Institute (MNI) space. Technical details are available in [15]. The volumetric data was then converted to a 3D tensor with float values per filled voxel and 0 otherwise. Additional zero-padding was done to effectively feed the data to our 3D-CAE, converting the tensor size from $150 \times 184 \times 140$ to $160 \times 192 \times 160$.

B. 3D convolutional autoencoder (3D-CAE)

We took notice of the 3D-CAEs capability to extract spatial features of the input that are practical for diagnosis. Given a volumetric representation of the brain MRI as an input, our 3D-CAE extracts visual features and compresses to arbitrary dimensional representation. Autoencoder is an unsupervised learning algorithm that learns an identity mapping of the input by minimizing the loss function between the input and its reconstructed output. It consists of encoding and decoding, where the encoder will map the input data $\mathbf{x} \in \mathcal{R}^D$ to a feature vector $\mathbf{y} \in \mathcal{R}^d$ and the decoder will reconstruct back to the original data space $\mathbf{z} = \hat{\mathbf{x}} \sim \mathbf{x} \in \mathcal{R}^D$ ($\mathcal{D} \gg d$).

We will start the simplest example of autoencoder model having a single encoder and decoder. The embedding $\mathbf{y}$ is obtained from $\mathbf{x}$ by the following equation.

$$\mathbf{y} = f(W\mathbf{x} + \mathbf{b}) \tag{1}$$

Where, W and $\mathbf{b}$ are learnable weights and the function noted f, is often a non-linear activation function such as the ReLU function expressed as $f(x) = \max(0, x)$. Reverse mapping to obtain the reconstructed output $\hat{\mathbf{x}}$ is calculated in a similar manner, taking $\mathbf{y}$ as the input. Here $\tilde{W}$ and $\tilde{\mathbf{b}}$ are weights used in the decoder.

$$\hat{\mathbf{x}} = \tilde{f}(\tilde{W}\mathbf{y} + \tilde{\mathbf{b}}) \tag{2}$$

Generally, mean squared error (MSE) is used as a loss function for autoencoders.

$$E(W) = \frac{1}{D}\sum_{d=1}^{D} ||\mathbf{x}_d - \hat{\mathbf{x}}_d||^2 \tag{3}$$

It is important to note that the decoder reconstructs the original input by using the embedding compressed by the encoder. If the reconstruction $\hat{\mathbf{x}}$ that is obtained by a trained AE has a certain degree of similarity with the input $\mathbf{x}$, this will conclude that the decoder reconstructed by the data solely from the low dimensional representation $\mathbf{y}$. Therefore, embedding $\mathbf{y}$ captures the visual feature of its input while representing the input in a smaller dimension.

CAE inherits the concept of the AE, with an exception that the layers consist of convolution and pooling layers, components of the CNN widely known for achieving state-of-the-art results in the field of computer vision.

Local connectivity of convolution layers enable the CAE to extract local and hierarchical features, ultimately capturing the global feature of the input by combining the local features. Thus local connections require much less computational cost than full connections. Pooling layers are used to reduce the input size and to add robustness to shift and position variance. 3D-CAE is an extended CAE composed of 3D convolution and pooling layers, applicable to volumetric data. Recently, 3D-CAE was applied for noise elimination

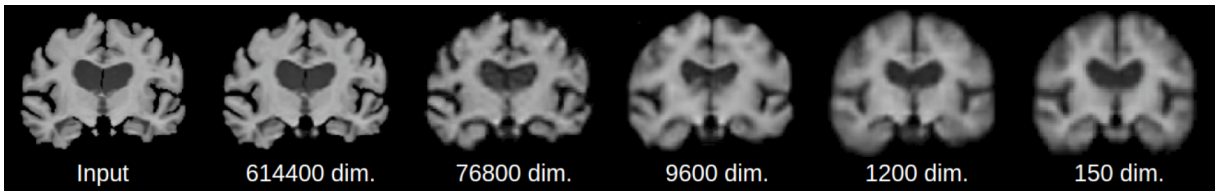

Fig. 1. Input and reconstructed images of varying embedding dimensions

TABLE I
RMSE AND SNR VALUES OF DATA ON 3D-CAE ARCHITECTURES

Depth	Dimension	RMSE [%]		SNR [dB]	
		AD + CN	EMCI	AD + CN	EMCI
1	614,400	1.91	1.95	25.23	25.01
2	76,800	4.60	4.57	15.90	15.91
3	9,600	6.18	6.14	13.19	13.22
4	1,200	7.49	7.66	11.03	10.63
5	150	7.86	8.45	10.64	9.72

of Low Dose CT and revealed their remarkable performance comparable to state-of-the-art methodologies [16].

The encoder of our 3D-CAE consists of five pooling layers each having convolution layers. The decoder reconstructing the volumetric data from the compressed feature representation adopts mirrored architecture, using unpooling and deconvolutional layers for corresponding layers. Each convolutional layer consists of varying kernel numbers, and the innermost kernel number determines the dimension of the embedding.

Since position correction has been performed on the data during preprocessing, inner convolution layers were specifically stacked in order to extract more global features of the input brain MRI. Kernel numbers are fixed to 27 until the innermost layer of the encoder and decoder, where it is set to one to squeeze the embedding dimension.

C. Regularization

Generally, the designer of the autoencoder would apply constraints to the model. This constraint is applied to reduce the number of parameters within the autoencoder, leading to a lower degree of freedom thus higher generalization to unknown data. We implemented a pseudo tied weights for the 3D-CAE model, in which each weight on the filter of the decoder are considered as a set, and the matrix having these sets is transposed to share the weights with the corresponding layer of its encoder counterpart.

III. EXPERIMENTS

We prepared five 3D-CAE architectures with varying depths and feature dimensions, and evaluated their reconstruction performance by calculating the root mean squared error (RMSE) and signal-to-noise-ratio (SNR) of the input and reconstructed output. The 3D-CAE were modified such that the final embedding will have the dimensions d = 614,400, 76,800, 9,600, 1,200 and 150. The RMSE and SNR of the reconstructed outputs are shown in Table I.

From Table I, increase in RMSE and decline in SNR can be observed with decreasing dimensionality d.

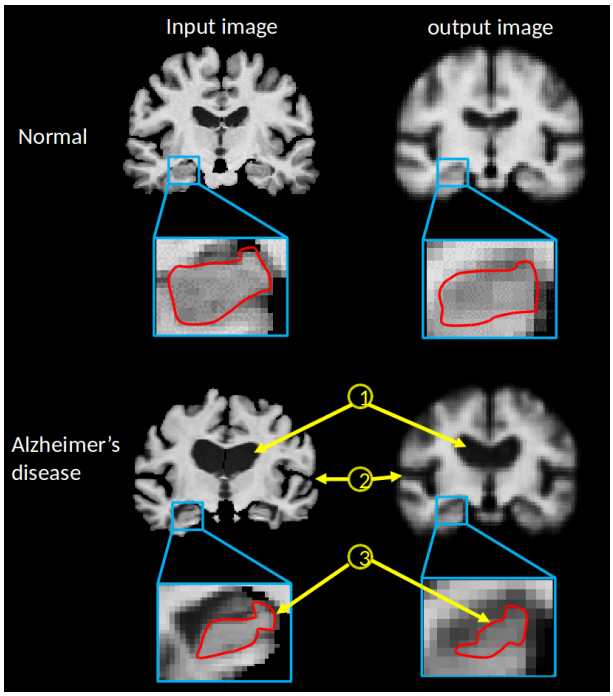

Fig. 2. Visual assessment on image slices

Fig. 1 shows the visualization of the original input and corresponding reconstructed image for each embedding dimension. Reconstructed volumetric data were converted to 256-greyscale and sliced to create the coronal plane images.

As illustrated by Fig. 1, the lower the dimension was, the more blurred the reconstructed images were. The brain surface lost the sharp contour from 9,600 dimensions and lower.

Since we aim to utilize the embedding for clustering and search tasks, it is ideal for the embedding to take a small dimension. Therefore to further evaluate the validity of the embedding with the smallest dimension, additional visual assessment was performed by a neurologist on the image slices.

Representative results from two indivisuals, one with normal cognition (upper row) and another with Alzheimer's disease (lower row), are demonstrated in Fig. 2. The images reconstructed from 150 dimensions (output images, right column) demonstrated that the known anatomical features related to Alzheimer's disease, such as 1 ventricular enlargement, 2 widening of the Sylvian fissure, and 3 hippocampal atrophy, are well preserved in the Alzheimer's disease. On the other hand, cortical area, gyrification of which is unique to each individual, is blurred in the output images.

Detailed line profiles of the region of interest are shown in Fig. 3. Line profiles were taken from lateral ventricle,

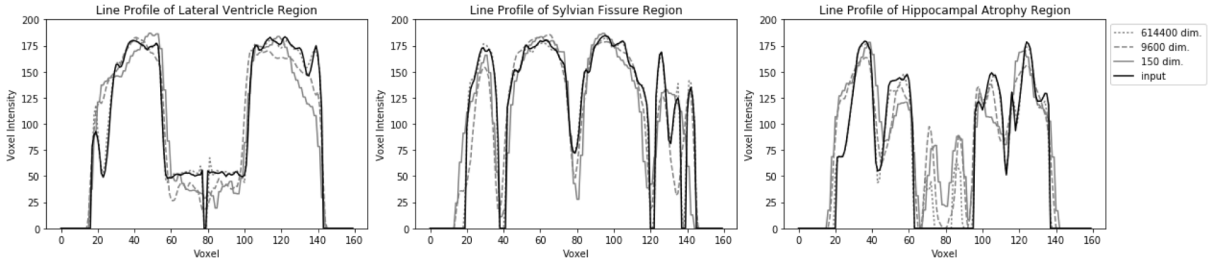

Fig. 3. Line profiles of region of interest

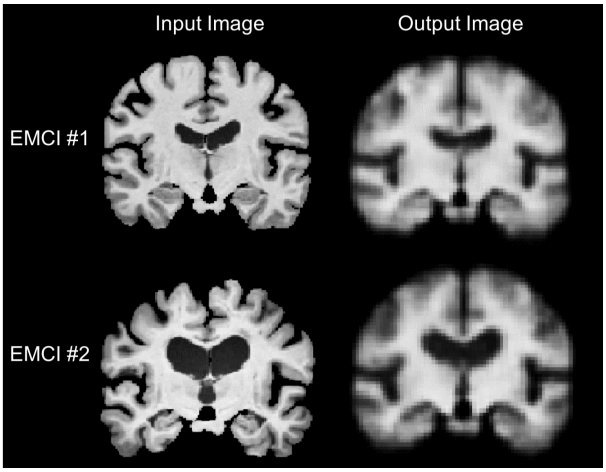

Fig. 4. Reconstructed images of EMCI data (not trained)

Sylvian fissure, and hippocampal atrophy regions labeled as 1, 2, and 3 in Fig 2, respectively. As illustrated in Fig. 3, the reconstructed input from 150 dimension representation roughly traces its original input. Line profile of the hippocampal atrophy region shows the reconstruction of the brainstem is relatively poor compared to other regions. This is due to difficulty of image alignment for this region. We need to address this issue in further study.

Furthermore, our 3D-CAE has the ability to generalize well with unknown data excluded during training of the network. The input and reconstructed image slices of two EMCI patients are shown on Fig. 4, where the top image shows EMCI case visually similar to a cognitively normal patient's brain, and bottom showing a case with brain atrophy. As illustrated on Fig. 4, we can tell that the model generalizes well with unknown data, extracting unique features of the input. From the abovementioned, we built the basis of CBIR systems as the outputs reconstructed by a significantly compressed representation preserves critical information for diagnosis. Our dimension reduction method could be used for various machine learning applications.

IV. CONCLUSIONS

In this paper we proposed a dimension reduction method using 3D-CAE and reduced the input dimension of the volumetric input from 4,915,200 to 150 while preserving features that radiologists would note during diagnosis. We evaluated our model by visual and quantitative analysis. Our results show promising establishment of a medical CBIR system and various machine learning applications using deep representations.

V. ACKNOWLEDGEMENT

This research was partially supported by the Ministry of Education, Culture, Science and Technology of Japan (Grant inAid for Fundamental research program (C), 17K8033, 2017-2020) and the Fakhri Rad Brite Star award from the Department of Radiology Johns Hopkins University School of Medicine.

REFERENCES

[1] J.J.Sejvar, "Global burden of neurological disease: whats in a name," *Lancet Neurology*, vol.16, no.11, pp.858-859, 2017.

[2] N.Huang, W.Yang, M.Yu et al., "Retrieval of Brain Tumors with Region-Specific Bag-of-Visual-Words Representations in Contrast-Enhanced MRI Images," *Computational and Mathematical Methods in Medicine*, no. 280538, 2012.

[3] S.Murala, Q.M.J.Wu, "Local Mesh Patterns Versus Local Binary Patterns: Biomedical Image Indexing and Retrieval," *IEEE Journal of Biomedical and Health informatics*, vol.18, no.3, pp.929-938, 2014.

[4] F. Yaacoub, A.Abche, and E.Karam, "MRI Reconstruction Using SVD in the Least Square Sense," *IEEE Proc. Computer-based Medical Systems*, pp.47-49, 2008.

[5] D.M.Lyra-Leite, J.P.C.K.Costa, and L.K.A.Carvalho, "Improved MRI reconstruction and denoising using SVD-based low-rank approximation," *IEEE Proc. Engineering Applications (WEA)*, 2012.

[6] D.L. Donoho, "Compressed Sensing," *IEEE Trans. Information Theory*, vol. 52, no. 4, 2006.

[7] U. Gamper, P. Boesiger, S. Kozerke, "Compressed sensing in dynamic MRI," *Magnetic Resonance in Medicine*, vol.59, no.2, pp.365-373, 2008.

[8] K. Simonyan and A. Zisserman, "Very deep convolutional neural network for large-scale image recognition," *CoRR*, abs/1409.1556, 2014.

[9] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Imagenet Recoginition," *Proc. International Conference on Computer Vision and Pattern Recognition (CVPR2016)*, pp. 770-778, 2016.

[10] Q.Dou, C. Lequan, H.Chen et. al., "3D deeply supervised network for automated segmentation of volumetric medical images," *Medical Image Analysis*, vol.41, pp.40-51, 2017.

[11] Q. Dou, C. Qi, Y. Hao et al. "Automatic Detection of Cerebral Microbleeds From MR Images via 3D Convolutional Neural Networks," *IEEE Trans. Medical Imaging*, vol.35, no.5, pp1182-1195, 2016.

[12] K. Kamnitsas, C. Ledig, V.F.J. Newcombe et al., "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Medical Image Analysis*, vol.36, pp.61-78, 2017.

[13] E. Hosseini-Asl, R. Keynto and A. El-Baz, "Alzheimer's disease diagnostics by adaptation of 3d convolutional network," *International Conference of Image Processing (ICIP)*, pp. 126130, 2016.

[14] MR Image Data, ADNI - USC Mark and Mary Stevens Neuroimaging and Informatics Institute, 2018, http://adni.loni.usc.edu/data-samples/mri/

[15] X. Tang, K. Oishi, AV. Faria et al., "Bayesian Parameter Estimation and Segmentation in the Multi-Atlas Random Orbit Model," *PLoS One* 8(6), 2013.

[16] H.Chen, Z.Hu, K. Yi et al., "Low-Dose CT With a Residual Encoder-Decoder Convolutional Neural Network," *IEEE Trans. Medical Imaging*, vol.36, no.12, pp.2524-2535, 2017.