

Improving Arterial Spin Labeling by Using Deep Learning¹

Ki Hwan Kim, MD
Seung Hong Choi, MD
Sung-Hong Park, PhD

¹From the Graduate School of Medical Science and Engineering (K.H.K., S.H.P.) and Department of Bio and Brain Engineering (S.H.P.), Korea Advanced Institute of Science and Technology, Room 1002, CMS (E16) Building, 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea; Department of Radiology, Seoul National University Hospital, Seoul, Republic of Korea (S.H.C.); Department of Radiology, Seoul National University College of Medicine, and Institute of Radiation Medicine, Seoul National University Medical Research Center, Seoul, Republic of Korea (S.H.C.); and Center for Nanoparticle Research, Institute for Basic Science, Seoul, Republic of Korea (S.H.C.). Received May 18, 2017; revision requested July 20; revision received September 8; accepted September 29; final version accepted October 12. **Address correspondence to S.H.P.** (e-mail: sunghongpark@kaist.ac.kr).

Study supported by Korea Health Industry Development Institute (HI16C1111) the Institute for Basic Science (IBS-R006-D1), the National Research Foundation of Korea (2015M3A9A7029740, 2016M3C7A1914002, 2017R1A2B2006526), and Seoul National University (Creative-Pioneering Researchers Program).

© RSNA, 2017

Purpose:

To develop a deep learning algorithm that generates arterial spin labeling (ASL) perfusion images with higher accuracy and robustness by using a smaller number of subtraction images.

Materials and Methods:

For ASL image generation from pair-wise subtraction, we used a convolutional neural network (CNN) as a deep learning algorithm. The ground truth perfusion images were generated by averaging six or seven pairwise subtraction images acquired with (a) conventional pseudocontinuous arterial spin labeling from seven healthy subjects or (b) Hadamard-encoded pseudocontinuous ASL from 114 patients with various diseases. CNNs were trained to generate perfusion images from a smaller number (two or three) of subtraction images and evaluated by means of cross-validation. CNNs from the patient data sets were also tested on 26 separate stroke data sets. CNNs were compared with the conventional averaging method in terms of mean square error and radiologic score by using a paired *t* test and/or Wilcoxon signed-rank test.

Results:

Mean square errors were approximately 40% lower than those of the conventional averaging method for the cross-validation with the healthy subjects and patients and the separate test with the patients who had experienced a stroke ($P < .001$). Region-of-interest analysis in stroke regions showed that cerebral blood flow maps from CNN (mean \pm standard deviation, 19.7 mL per 100 g/min \pm 9.7) had smaller mean square errors than those determined with the conventional averaging method (43.2 \pm 29.8) ($P < .001$). Radiologic scoring demonstrated that CNNs suppressed noise and motion and/or segmentation artifacts better than the conventional averaging method did ($P < .001$).

Conclusion:

CNNs provided superior perfusion image quality and more accurate perfusion measurement compared with those of the conventional averaging method for generation of ASL images from pair-wise subtraction images.

© RSNA, 2017

Arterial spin labeling (ASL) has been used increasingly in practice because of noninvasiveness, repeatability, advantages in quantification, and advancements in labeling and readout sequences (1–3), but it still has an inherently low signal-to-noise ratio and sensitivity to motion. For better signal-to-noise ratio and robustness to motion, several ASL techniques have been suggested, such as background suppression, automatic rejection of images affected by motion, and denoising (4–7). Despite these technical developments, sensitivity, robustness, and imaging time of ASL must be improved further for better clinical translation.

The convolutional neural network (CNN) is a class of deep artificial neural networks inspired by biologic processes of the visual cortex (8). CNNs respond to inputs through overlapping restricted regions called receptive fields of different neurons, similar to the cortical neurons responding to external stimuli in the entire visual

cortex through overlapping restricted visual fields of individual neurons. Distinct from traditional algorithms, CNNs can capture hierarchical contextual information automatically and can perform a specific task without defining image features (9). Applications of CNNs to medical imaging have rapidly increased during the last 2 years (10). CNNs are successfully implemented in radiographs to suppress noise (11). For reconstruction of low-dose computed tomographic (CT) images, CNN was superior to other methods including total variation and k-singular value decomposition (k-SVD), while maintaining the computational efficiency (12,13).

In this study, we propose a new deep learning approach to improve ASL image quality. ASL perfusion images usually are acquired by averaging a number of pair-wise subtractions, requiring a longer imaging time. To reduce imaging time, we aimed to develop CNNs that generate ASL perfusion images with higher accuracy and robustness by using a smaller number of pair-wise subtractions.

Materials and Methods

This was a retrospective study with approval from the institutional review board and waiver of the requirement for informed consent.

Convolutional Neural Networks

Each CNN was composed of a stack of multiple convolutional layers, where a batch normalization layer (14) and a rectified linear unit layer (15) were followed by every convolutional layer with a size of 3×3 (16). The feature map of a convolutional layer was treated as input to the subsequent convolutional

layer so that CNNs could extract a hierarchy of increasingly complex features. The proposed CNN framework consisted of two parallel pathways for extracting multiscale information efficiently (17,18): (a) a local pathway for extracting low-level features and (b) a global pathway for incorporating large-scale contextual information (Fig 1). The global pathway had four dilated convolutional layers with dilation factors of two, four, eight, and 16 to aggregate large scale information without losing resolution (19). Then, both local and global pathways were concatenated to integrate the multiscale contextual information. Finally, the output of the last convolutional layer (difference image) and average of the input were summed to generate the perfusion image. CNNs were trained to predict the difference between the average of input and the given ground truth, a procedure called residual learning (20,21). The residual learning approach was used to speed the learning process and improve the performance of CNNs by simplifying image generation (13,21). In the training stage, the parameters of CNNs were optimized to generate ground truth-like perfusion images, which could be regarded as a learning process to reduce noise and artifacts. In the test stage, the trained CNNs produced perfusion images by summing the output


Advances in Knowledge

- Arterial spin labeling (ASL) perfusion images can be generated with a convolutional neural network (CNN) from a small number of data acquisitions.
- Perfusion images from CNN were closer to the ground truth images than were those from the conventional averaging method in the whole brain, regardless of the type of MR imager and labeling scheme ($P < .01$).
- CNN could accurately quantify cerebral blood flow from multi-phase ASL images with fewer temporal encoding points than those of the ground truth images in patients who had experienced a stroke and was superior to those of the conventional averaging method ($P < .001$).
- CNN suppressed noise and motion and/or segmentation artifacts observable with the use of the conventional averaging method ($P < .001$).

Implication for Patient Care

- CNN may increase the utility of noninvasive ASL perfusion imaging in clinical practice by reducing imaging time needed to ensure sufficient image quality and by suppressing motion and/or segmentation artifacts.

<https://doi.org/10.1148/radiol.2017171154>

Content code: 

Radiology 2018; 000:1–9

Abbreviations:

ASL = arterial spin labeling
CNN = convolutional neural network

Author contributions:

Guarantors of integrity of entire study, K.H.K., S.H.P.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, K.H.K., S.H.C.; clinical studies, K.H.K.; statistical analysis, K.H.K.; and manuscript editing, all authors

Conflicts of interest are listed at the end of this article.

Figure 1

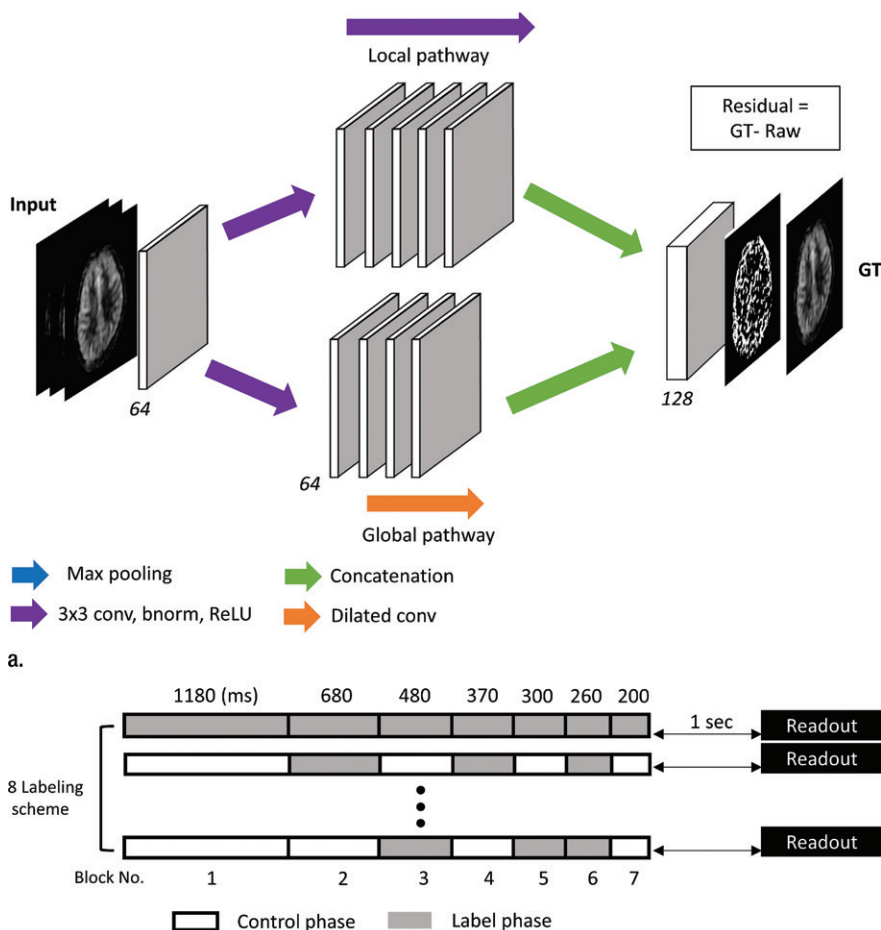


Figure 1: Schematic illustrations of a CNN framework and Hadamard-encoded pseudocontinuous ASL scheme. **(a)** CNN framework. The first convolutional layer (*conv*) received two or three subtraction images as input, followed by four convolutional layers in the global (orange arrow) and local (large purple arrow) pathways. Dilatation factors of four convolutional layers in the global pathway were two, four, eight, and 16. Parallel pathways were concatenated at a later step (green arrows). Residual image was acquired by subtracting the ground truth (*GT*) from average of input images. Number of filters is shown in italics. **(b)** Hadamard-encoded pseudocontinuous ASL scheme. Labeling duration was divided into seven blocks, and duration of blocks was designated for exponentially increasing postlabeling delay ranging from 1.0 sec to 3.3 sec. Hadamard encoding was used to make eight radiofrequency pulse combinations. *bnorm* = batch normalization, *ReLU* = rectified linear unit.

of the last convolutional layer and the average of the input with no information from the ground truth images.

Data Preparation

This study was performed separately for two different data sets as follows.

Data sets of healthy subjects.—CNN was applied to data sets acquired with a three-dimensional pseudocontinuous

ASL-balanced steady-state free-precession readout in seven healthy volunteers (five men and two women; mean age, 25 years; age range, 21–33 years) with no history of a brain-related disorder by using a 3-T magnetic resonance (MR) imager (MAGNETOM Trio; Siemens Healthineers, Erlangen, Germany) (22). All volunteers were recruited by means of an advertisement. Three ASL data

sets were collected from each subject by using three different labeling schemes: (a) a proximal control and proximal label, (b) proximal label and distal control with section selection gradients of 16 mT · m⁻¹, and (c) a control sequence with no radiofrequency preparation (22). Since these labeling schemes had slightly different labeling efficiency, they could be used to assess the effects of labeling efficiency on CNN. The imaging parameters were: repetition time msec/echo time msec, 3.96/1.74; field of view, 240 × 180 × 120 mm³; matrix size, 128 × 96 × 24; labeling duration, 1.5 sec; postlabeling delay, 1.0 sec; number of segments, six; flip angle, 30°; and centric phase-encoding order.

One ASL data set consisted of six labeling and control image pairs for pairwise subtraction. The top and bottom five sections were excluded because the brain tissue in the sections was small or absent. All six subtraction images were averaged for high-quality perfusion images, which were used as the ground truth. The proposed CNN and the conventional averaging method were used to generate perfusion images by taking only two randomly selected subtraction images. Since one ASL data set had six image pairs and two of these were randomly selected as inputs of the CNNs, 15 input combinations were generated from each ASL data set. We evaluated the performance of CNNs by means of *N*-fold cross-validation (*n* = 7): Data sets of six selected subjects and one remaining subject were used for training and testing, respectively. This process was repeated by selecting the test data set of the seven data sets seven times, and the results of CNN could be acquired from all seven subjects separately. The mean square error was calculated as a pixel-level error between the perfusion images generated with CNN and the ground truth. All mean square errors throughout the subjects were averaged to compare CNN to the conventional averaging method. The training time for each cross-validation study was approximately 7 hours, and the reconstruction time for one image was 4 msec.

Data sets of patients.—To test the performance of CNNs under various

conditions of labeling schemes and readouts, MR imagers, and brain diseases, the proposed CNN was evaluated by using data sets of Hadamard-encoded pseudocontinuous ASL, which was installed in another MR imager (Discovery MR750w; GE Healthcare, Milwaukee, Wis). All of the patients who underwent Hadamard-encoded pseudocontinuous ASL examinations with the GE MR imager in January and February 2016 for suspicion of perfusion abnormalities were included. The total number of included patients was 120. Among the 120 patients, six were excluded because of severe motion artifacts distorting or blurring anatomic structures, which were evaluated by a radiologist (K.H.K, with 7 years of experience). Finally, data sets of 114 patients were used in this study. Clinical characteristics of the 114 patients are described in Table 1. The readout sequence was a three-dimensional stack-of-spirals rapid acquisition with refocused echoes. Images rotated by 90° and 180° were added to the training set, allowing CNNs to be trained with additional views of the brain (23). The top and bottom three sections were excluded. The default parameters in experiments and simulations were: 5901/11; field of view, 240 × 240 × 140 mm³; matrix, 128 × 128 × 28; and section thickness, 4 mm. Hadamard-encoded radiofrequency pulse combinations were used with background suppression to generate eight baseline images, which were combined in different ways to reconstruct seven different perfusion images with postlabeling delays ranging from 1.0 to 3.3 seconds (Fig 1b) (24). By using the weighted delay measurement and the kinetic model, seven cerebral blood flow maps were quantified from the seven post-labeling delay perfusion images separately (25,26).

The ground truth was calculated by averaging all seven postlabeling delay data. Cerebral blood flow maps of the proposed CNN and the averaging method were produced by taking second, fourth, and sixth postlabeling delay data. The performance of CNNs was evaluated by averaging mean square

Table 1

Clinical Characteristics of the Patient Groups

Characteristic	Training and Validation Data Sets (n = 114)	Test Data Sets (n = 26)
Mean age (y)*	54.4 ± 16.1	68.0 ± 14.4
Sex		
Male	45	10
Female	69	16
Disease		
Acute and subacute infarction	19	26
Middle cerebral artery territory	13	19
Anterior cerebral artery territory	2	2
Posterior cerebral artery territory	2	2
Cerebellum	1	1
Multiple bilateral lesions	1	...
Basal ganglia	...	2
Tumor	10	...
Meningioma		
Left frontal lobe	1	...
Left parietal lobe	1	...
Bilateral temporal lobes	1	...
Pituitary adenoma	2	...
Glioma, frontal lobe	1	...
Lymphoma, left temporal lobe	1	...
Metastasis, bilateral multiple lesions	1	...
Schwannoma, left cerebellopontine angle	1	...
Central neurocytoma, right lateral ventricle	1	...
Moyamoya	5	...
Inflammatory disease	4	...
Aneurysm	3	...
Vascular malformation	3	...
Small vessel disease	2	...
Intracranial hemorrhage	1	...
Hydrocephalus	1	...
No relevant abnormal lesion	72	...

Note.—Unless otherwise indicated, data are number of patients.

* Data are means ± standard deviation.

errors from *N*-fold cross-validation (*n* = 6): Data sets of 95 selected patients and the remaining 19 patients were used for training and testing, respectively, which was repeated six times in order for data sets of all 114 patients to be tested at least once. The training time for each cross-validation study was approximately 19 hours, and the reconstruction time for one image was 7 msec.

After cross-validation, CNNs were trained by using the data sets of the 114 patients and were tested by using data sets of additional stroke patients. There were 98 patients who underwent ASL

examinations between March and July 2016 because of suspicion of cerebral ischemia. Among these 98 patients, 29 were confirmed as having experienced acute or subacute infarction. Among these 29 patients, three patients were excluded because brain anatomy could not be recognized because of severe artifacts. Eventually, data sets of 26 stroke patients were used to test the performance of CNNs.

Data Analysis

CNNs were developed by using software (MatConvNet function in Matlab; MathWorks, Natick, Mass) (27)

and tested with a graphic processor (GTX 1080; Nvidia, Santa Clara, Calif). They were trained with the stochastic gradient descent algorithm (28). All hyperparameters were optimized with cross-validation. Total epoch and learning rate were 300 and 5×10^{-7} , respectively.

Performance of CNNs was compared with that of the averaging method in terms of mean square errors by using the Wilcoxon signed-rank test, except for the whole volume analysis in the cross-validation of the patient data sets, where the paired *t* test was used. Mean square errors were calculated for all voxels of the test data sets, and then were averaged for each subject. The radiologic scores were also qualitatively assessed by using a four-point scale (1 = severe noise or motion artifact, 2 = moderate noise or motion artifact disturbing evaluation, 3 = mild noise and artifact not affecting evaluation, 4 = clear perfusion map) by a blinded radiologist (K.H.K) and were statistically compared by using the Wilcoxon signed-rank test. Regions that showed visually identifiable perfusion abnormalities were selected in the patient data sets. Two regions of interest representing normal and abnormal regions were manually drawn by a radiologist (K.H.K) on the images of the perfusion abnormalities with the largest lesion size. For the normal and pathologic regions of interest, concordance between cerebral blood flow maps from the ground truth and CNN or the conventional averaging method were evaluated with Bland-Altman analysis. *P* values of less than .05 were considered to indicate a statistically significant difference for all of the tests.

Results

Evaluation of Data Sets from Healthy Subjects

The proposed model provided perfusion images much closer to those of the ground truth (Fig 2). CNNs showed significantly lower mean square errors than did the averaging method, regardless of the labeling scheme

Figure 2

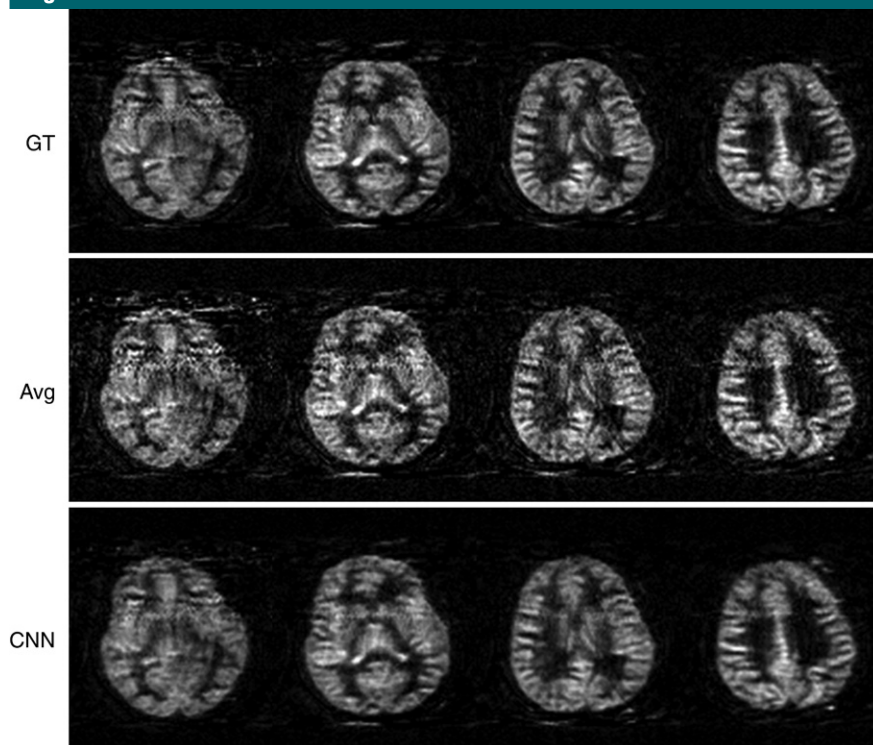


Figure 2: Representative results in a healthy 29-year-old man. Axial perfusion-weighted images are displayed for ground truth (GT), CNN, and the conventional averaging method (Avg).

(Fig 3a). CNNs showed significantly higher radiologic scores than those of the averaging method ($P < .001$) (Fig 3b), presumably because the CNN could learn to reduce noise and motion and/or segmentation artifacts in the training stage.

Evaluations in Patient Data Sets

In cross-validation, CNNs provided cerebral blood flow maps closer to those of the ground truth than the averaging method did, as shown for a representative subject in Figure 4. Mean square error values of CNNs were significantly lower than those of the averaging method in the entire data set ($n = 114$) and pathologic regions of interest ($n = 11$) (Table 2), indicating that CNNs provide more accurate cerebral blood flow maps than does the averaging method, regardless of the presence of abnormality. CNNs showed higher radiologic scores than those of the averaging method ($P < .001$) (Fig 5a).

Bland-Altman analysis showed narrower limits of agreement for CNN than for the averaging method, regardless of the presence of abnormality (Fig 5b, 5c).

The results of 26 patients who had experienced a stroke were consistent with those of the cross-validation (Fig 6). Mean square errors of CNNs was smaller than that of the averaging method in the whole image volumes and stroke regions (Table 2), indicating that CNNs maintain their performance in the stroke regions. CNNs showed significantly higher radiologic scores than the averaging method did ($P < .001$) (Fig 6a). Bland-Altman analysis in the regions of interest showed smaller mean bias and narrower limits of agreement for CNN (Fig 6b).

Discussion

In this study, we demonstrated that the proposed CNN approach allowed qualitative and quantitative improvement

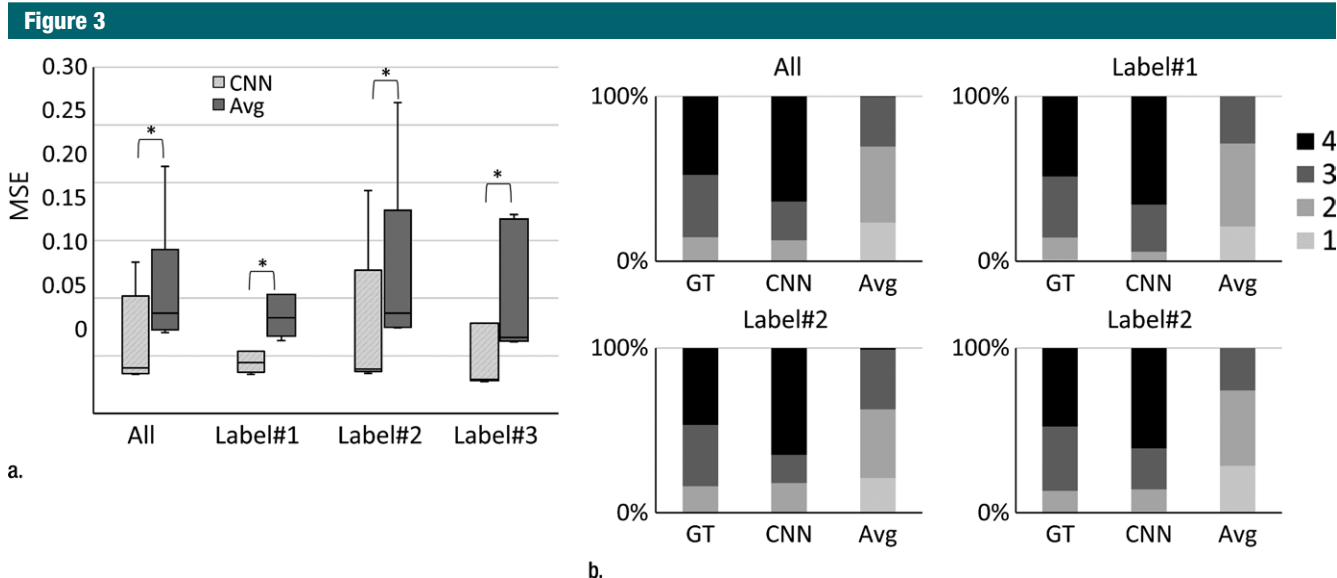


Figure 3: Comparison of mean square error (*MSE*) and radiologic score between CNN and averaging method (*Avg*) in healthy subjects. Comparisons of *MSE* (**a**) and radiologic score (**b**) were performed for all labeling schemes together (*ALL*) and separately for each labeling scheme (*labels 1, 2, and 3*). Radiologic scores are displayed with four gray colors. Vertical axis indicates percentage of each radiologic score. *GT* = ground truth, * = significant difference ($P < .05$).

on ASL perfusion images. CNNs received only two or three subtraction image pairs and then were trained to produce high-quality perfusion images. The trained CNNs outperformed the averaging method, regardless of MR imager, ASL technique, and presence of disease.

Since ASL results in a poor signal-to-noise ratio, multiple sequences are routinely performed for averaging to ensure sufficient image quality. Perfusion signal intensity values are generated by labeled arterial water, and neighboring areas are likely to receive blood from the same vessels. Therefore, considering perfusion signal intensity of neighbor voxels in the process of generating ASL images from multiple sequences would be helpful to improve ASL image quality. While the averaging method does not use spatial information, CNN can allow the multiscaled contextual information on ASL images to be inferred, improving the sensitivity and reliability of ASL.

As compared with the model-based methods, the learning-based methods rely on data used for training. We applied CNNs to ASL data sets acquired from various MR imagers, labeling schemes, image readouts, and healthy

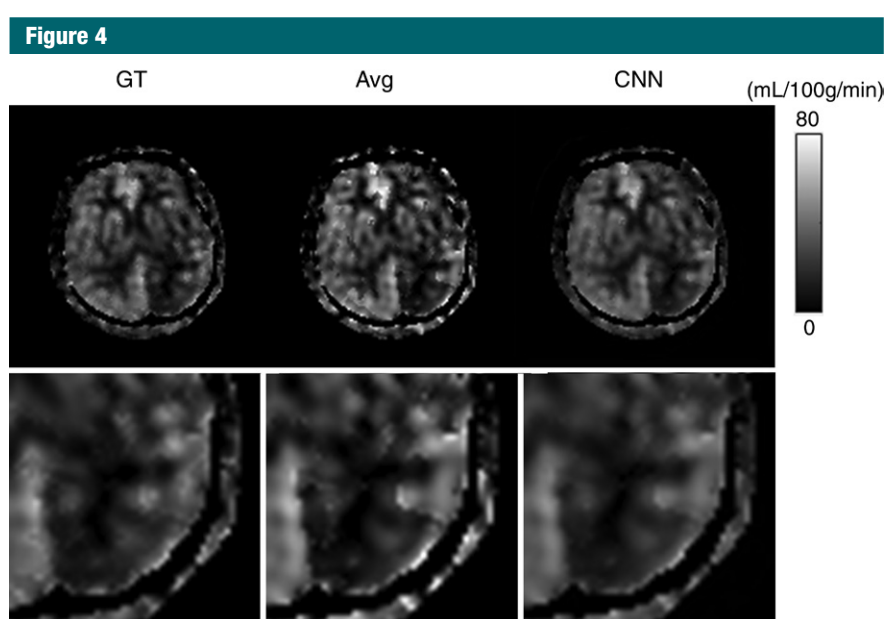


Figure 4: Representative results of CNN and the averaging method (*Avg*) in a 49-year-old man with acute cerebral infarction. Reconstructed cerebral blood flow images (top) and magnified images (bottom; magnification, $\times 2.5$) are displayed for ground truth (*GT*), CNN, and the conventional averaging method.

subjects and patients. When the labeling scheme, the readout sequence, or the MR imager is different between the training and test data sets, the performance of CNN can be reduced. However, the preferred ASL sequence

(labeling and readout) is typically fixed for one MR imager, and thus can be used for both training and real applications in each imager. Developing generalized CNNs applicable to any ASL sequence and any MR imager would

require training with bigger ASL data sets from various ASL sequences and MR imagers, while the ground truth images should be acquired with the same ASL sequence and with the same MR imager as those used to acquire each training data set.

CNNs reduced segmentation and/or motion artifacts. Since these artifacts caused by temporal inconstancy

can be suppressed with a higher number of averages, the ground truth was relatively robust to them. In comparison to model-based traditional methods, CNNs could be trained to suppress such artifacts (Fig 3) and outliers generated by means of a process of cerebral blood flow quantification (Fig 6), which is attributed to the high radiologic scores of CNNs.

In reality, the labeling efficiency of pseudocontinuous ASL can vary depending on labeling scheme, subject, labeling position, and off-resonance condition on the labeling plane. Sensitivity of CNNs to the labeling efficiency is a potential issue. We evaluated it by using three different labeling schemes in this study, which showed relatively consistent performance of CNNs for different labeling schemes.

We demonstrated that cerebral blood flow maps from three postlabeling delay data sets with CNNs were comparable to those from seven postlabeling delay data sets with the conventional averaging method. This implies reduction in imaging time for CNNs upon application to separate multiple postlabeling delay ASL studies. In the case of Hadamard encoding (as described in this study), the three postlabeling delay data sets have the potential to be generated from only four Hadamard-encoded pseudocontinuous ASL data sets rather

Table 2

Mean Square Error in Whole Volume and Pathologic Region for Patient Data Sets

Method	Cross-Validation		Stroke	
	Whole Volume (<i>n</i> = 114)	Pathologic Region (<i>n</i> = 11)	Whole Volume (<i>n</i> = 26)	Pathologic Region (<i>n</i> = 26)
Averaging method	54.8 ± 40.4	58.6 ± 46.7	57.1 ± 32.0	43.2 ± 29.8
CNN	20.4 ± 20.4 [†]	22.9 ± 20.0*	27.1 ± 18.5 [†]	19.7 ± 9.7 [†]

Note.—Data are means ± standard deviation.

* The paired *t* test showed a significant difference compared with averaging method (*P* < .01).

[†] The Wilcoxon signed-rank test showed a significant difference compared with the conventional averaging method (*P* < .001).

Figure 5

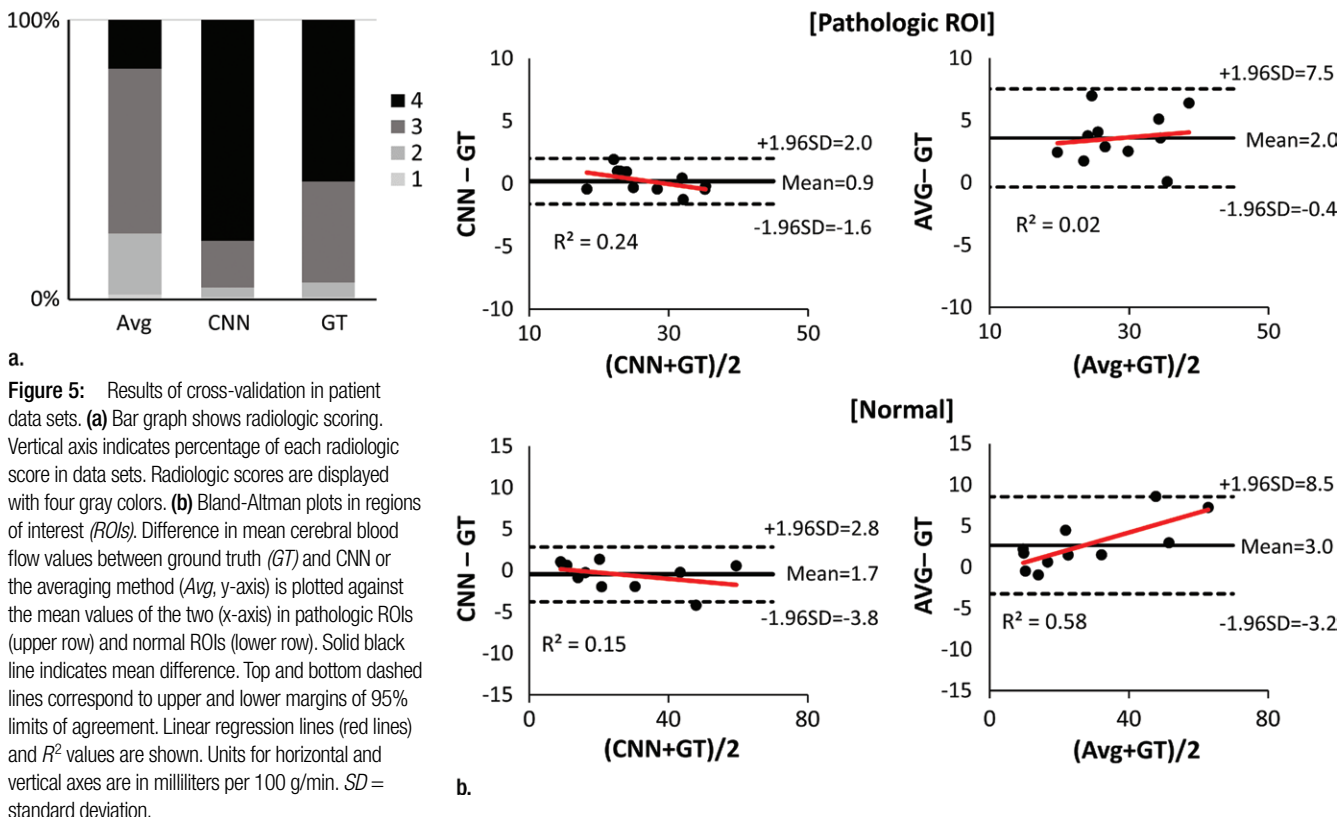
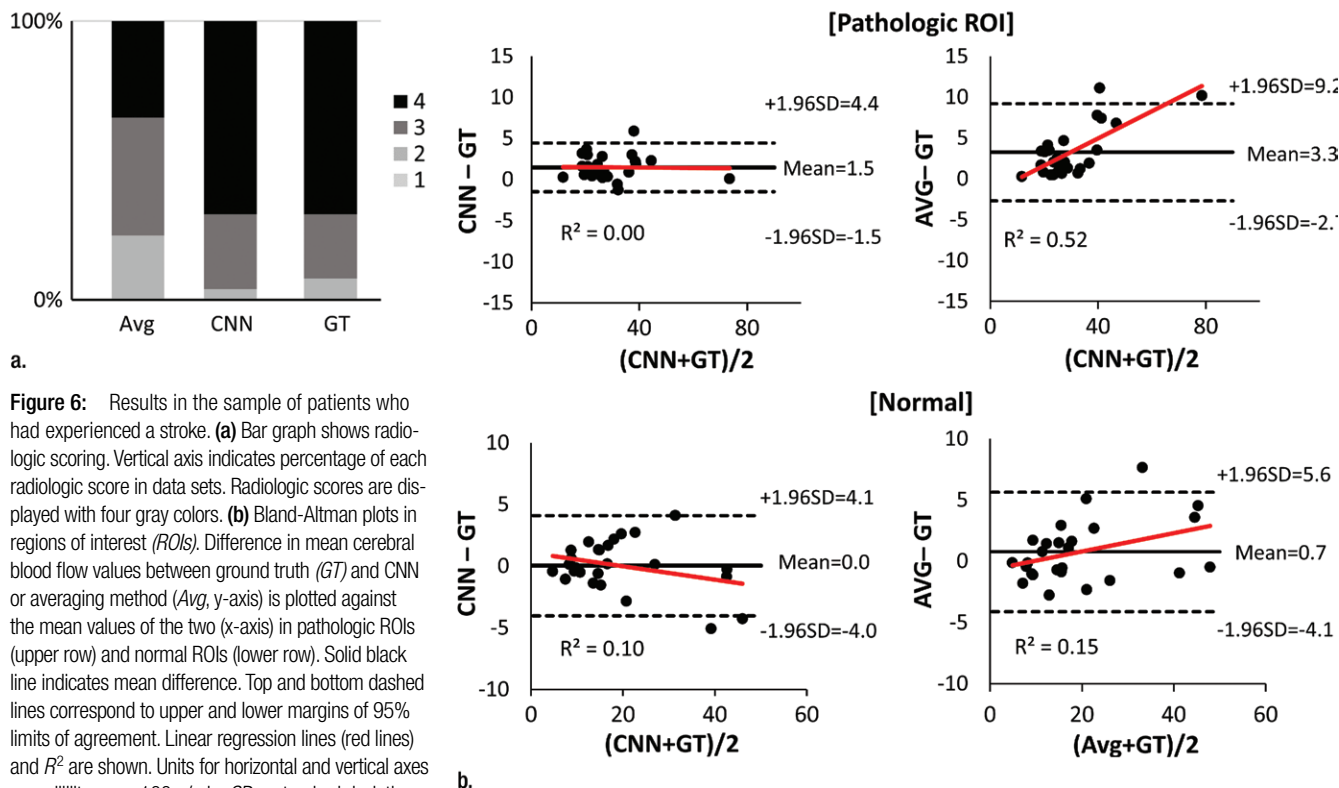


Figure 6



than eight. The reduction in number of Hadamard-encoded pseudocontinuous ASL data sets provides additional advantages of elongation of each labeling block for better labeling efficiency, higher temporal resolution, and less sensitivity to motion artifacts (29).

Our study had several limitations that need to be taken into account when interpreting the data. First, it included a relatively small number of patients in one institution. A multi-institutional study would be necessary to get data sets from a larger number of patients. Second, the evaluation of perceptual image quality with radiologic scoring is subjective. Although mean square error values for CNNs and the averaging method were provided, it may be difficult to fully assess the effects of the artifacts with mean square errors alone. Also, because images generated by deep learning algorithms often get blurred, observer scoring has been accepted as an auxiliary indicator

to evaluate the image quality in computer vision (30). Third, two separate CNNs were generated for the two different labeling schemes of conventional pseudocontinuous ASL and Hadamard-encoded pseudocontinuous ASL. Developing CNNs that can be applied to any ASL sequence regardless of labeling scheme, imaging parameter, readout sequence, and MR imaging vendor would require much bigger and more diverse data sets, which was beyond the scope of our study. Last, background suppression was used only for Hadamard encoded pseudocontinuous ASL (the patient data sets), but not for the conventional pseudocontinuous ASL (the healthy subjects). Although they worked well for both cases in this study, CNNs must be tested for the conventional pseudocontinuous ASL with background suppression.

In conclusion, CNN showed superior performance with ASL perfusion imaging compared with the conventional

averaging method, regardless of MR imager, labeling scheme, and readout scheme. The performance of CNN was demonstrated for data sets from a separate cohort of patients who had experienced a stroke.

Disclosures of Conflicts of Interest: K.H.K. disclosed no relevant relationships. S.H.C. disclosed no relevant relationships. S.H.P. disclosed no relevant relationships.

References

1. Detre JA, Leigh JS, Williams DS, Koretsky AP. Perfusion imaging. *Magn Reson Med* 1992;23(1):37-45.
2. Williams DS, Detre JA, Leigh JS, Koretsky AP. Magnetic resonance imaging of perfusion using spin inversion of arterial water. *Proc Natl Acad Sci U S A* 1992;89(1):212-216.
3. Alsop DC, Detre JA, Golay X, et al. Recommended implementation of arterial spin-labeled perfusion MRI for clinical applications: A consensus of the ISMRM perfusion study group and the European consortium

- for ASL in dementia. *Magn Reson Med* 2015;73(1):102–116.
4. Bibic A, Knutsson L, Ståhlberg F, Wirestam R. Denoising of arterial spin labeling data: wavelet-domain filtering compared with Gaussian smoothing. *MAGMA* 2010;23(3):125–137.
 5. Hu WT, Wang Z, Lee VM, Trojanowski JQ, Detre JA, Grossman M. Distinct cerebral perfusion patterns in FTL and AD. *Neurology* 2010;75(10):881–888
 6. Wells JA, Thomas DL, King MD, Connelly A, Lythgoe MF, Calamante F. Reduction of errors in ASL cerebral perfusion and arterial transit time maps using image de-noising. *Magn Reson Med* 2010;64(3):715–724
 7. Frouin F, Duteil S, Lesage D, Carlier PG, Herment A, Leroy-Willig A. An automated image-processing strategy to analyze dynamic arterial spin labeling perfusion studies. Application to human skeletal muscle under stress. *Magn Reson Imaging* 2006;24(7):941–951.
 8. Matsugu M, Mori K, Mitari Y, Kaneda Y. Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Netw* 2003;16(5-6):555–559.
 9. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521(7553):436–444.
 10. Greenspan H, van Ginneken B, Summers RM. Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique. *IEEE Trans Med Imaging*. 2016;35(5):1153–1159.
 11. Gondara L. Medical Image Denoising Using Convolutional Denoising Autoencoders. 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), 2016; 241–246.
 12. Chen H, Zhang Y, Zhang W, et al. Low-dose CT denoising with convolutional neural network. *ArXiv e-prints [serial online]*. 2016; vol 1610. <http://adsabs.harvard.edu/abs/2016arXiv161000321C>. Accessed October 1, 2016.
 13. Han YS, Yoo J, Ye JC. Deep Residual Learning for Compressed Sensing CT Reconstruction via Persistent Homology Analysis. *ArXiv e-prints [serial online]*. 2016; vol 1611. <http://adsabs.harvard.edu/abs/2016arXiv161106391H>. Accessed November 1, 2016.
 14. Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *arXiv preprint*. 2015;arXiv:1502.03167.
 15. Hinton GE, Srivastava N, Krizhevsky A, Sutskever I, Salakhutdinov RR. Improving neural networks by preventing co-adaptation of feature detectors. *ArXiv e-prints [serial online]*. 2012; vol 1207. <http://adsabs.harvard.edu/abs/2012arXiv1207.0580H>. Accessed July 1, 2012.
 16. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint*. 2014;arXiv:1409.1556.
 17. Havaei M, Davy A, Warde-Farley D, et al. Brain tumor segmentation with Deep Neural Networks. *Med Image Anal* 2017;35:18–31
 18. Kamnitsas K, Ledig C, Newcombe VF, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Med Image Anal* 2017;36:61–78
 19. Yu F, Koltun V. Multi-Scale Context Aggregation by Dilated Convolutions. *arXiv preprint*. 2015;arXiv:1511.07122.
 20. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. *ArXiv e-prints [serial online]*. 2015; vol 1512. <http://adsabs.harvard.edu/abs/2015arXiv151203385H>. Accessed December 1, 2015.
 21. Kim J, Lee JK, Lee KM. Accurate Image Super-Resolution Using Very Deep Convolutional Networks. *ArXiv e-prints [serial online]*. 2015; vol 1511. <http://adsabs.harvard.edu/abs/2015arXiv151104587K>. Accessed November 1, 2015.
 22. Han PK, Choi SH, Park SH. Investigation of control scans in pseudo-continuous arterial spin labeling (pCASL): Strategies for improving sensitivity and reliability of pCASL. *Magn Reson Med* 2017;78(3):917–929.
 23. van Dyk DA, Meng XL. The art of data augmentation. *J Comput Graph Stat* 2001;10(1):1–50.
 24. Wells JA, Lythgoe MF, Gadian DG, Ordidge RJ, Thomas DL. In vivo Hadamard encoded continuous arterial spin labeling (H-CASL). *Magn Reson Med* 2010;63(4):1111–1118.
 25. Wang DJ, Alger JR, Qiao JX, et al. Multi-delay multi-parametric arterial spin-labeled perfusion MRI in acute ischemic stroke - Comparison with dynamic susceptibility contrast enhanced perfusion imaging. *Neuroimage Clin* 2013;3:1–7.
 26. Dai W, Robson PM, Shankaranarayanan A, Alsop DC. Reduced resolution transit delay prescan for quantitative continuous arterial spin labeling perfusion imaging. *Magn Reson Med* 2012;67(5):1252–1265.
 27. Vedaldi A, Lenc K. MatConvNet - Convolutional Neural Networks for MATLAB. *ArXiv e-prints [serial online]*. 2014; vol 1412. <http://adsabs.harvard.edu/abs/2014arXiv1412.4564V>. Accessed December 1, 2014.
 28. Sinha NK, Griscik MP. Concerning Comments on a Stochastic Approximation Method. *IEEE Trans Syst Man Cybern B Cybern* 1973;SMC3(5):526.
 29. van Osch MJ, Teeuwisse WM, Chen Z, Suzuki Y, Helle M, Schmid S. Advances in arterial spin labelling MRI methods for measuring perfusion and collateral flow. *J Cereb Blood Flow Metab* 2017 Jan 1:271678X17713434 [Epub ahead of print].
 30. Ledig C, Theis L, Huszar F, et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. *arXiv preprint*. 2016;arXiv:1609.04802.