Supplementary Methods

Training data

A total of 8,661 diffusion-weighted images of patients with ischemic stroke from 10 university hospitals in Republic of Korea were collected. Using the validated Image_QNA program, ischemic lesions were outlined by experienced researchers and supervised by an experience vascular neurologist.

Artificial intelligence algorithm

Three-dimensional (3D) U-Net[1] was used to train the segmentation model. To compare model performance based on the training dataset size, the training and validation datasets were subsampled by factor of 10%/20%/50%/100%. To train the model, each subsampled dataset was divided into training and validation sets at an 8:2 ratio. Data augmentation was used during the training process to prevent overfitting and alleviate the domain shift problem. An augmentation algorithm was developed using TorchIO (https://torchio.readthedocs.io/), which is a medical imaging library written in Python.

Data augmentation

Random augmentation was used in real-time during the training process to prevent overfitting. Four types of augmentation were used: slice-wise affine transformation, magnetic resonance imaging (MRI) bias field artifact simulation, axis flip, and gamma/contrast change.

3D U-Net

The model consisted of an encoder/decoder layer that performed maximum pooling four times, with feature sizes set to 12, 24, 48, 96, and 192 in each step. The kernel size of the convolutional layer was set as 3×3×3. The Conv3d model weights were initialized with the He normal and the ConvTranspose3d model weights with Xavier uniform. For model training, the Adam optimizer was used with a batch size of two and an exponential cyclic learning rate oscillating between 1e-5 and 1e-4. In addition, to address the class imbalance caused by a small lesion size compared to the total brain volume, focal Tversky loss with parameters $\alpha=0.6$, $\beta=0.4$, and $\gamma=4/3$ was used.

Environment

Python 3.7.9/3.8.13 (https://www.python.org/), pytorch 1.12.0, torchvision 0.13.0, pandas 1.2.4, numpy 1.19.5/1.22.3, scipy 1.4.1/1.6.3, scikit-image 0.15.0/ 0.18.1, SimpleITK 2.1.1, and pydicom 2.1.2 were used for all experimental procedures, including preprocessing and model development. Intel Xeon Silver 4314 @2.40 GHz, 640 GB RAM, and NVIDIA Quadro RTX A6000 48 GB GDDR6 were used to train the models.

Bland-Altman plot

Bland-Altman plot was used to compare automatically and manually segmented volumes. The percentage of the difference was calculated as follows:

$$\%\ \text{Difference} = \frac{|A-B|}{(A+B)/2} \times 100,$$

where A and B indicate manually or automatically segmented infarct volumes, respectively.

Supplementary Discussion

Although deep learning-based infarct segmentation performed well, JBS-01K overlooked some ischemic lesions. As illustrated in Supplemental Figure 2, JBS-01K did not detect small, faint lesions on DWIs that could have been acquired early after symptom onset. In addition, the missed infarcts in the cingulate gyrus (Supplementary Figure 2C) indicated that infarcts in uncommon locations may be difficult to learn using deep learning because of the small number of cases in the training dataset. Future studies should incorporate clinical symptoms and DWIs into deep-learning models, similar to the clinicians. Adding missed deep learning cases to the training data may also improve model performance. However, the addition of small indistinct ischemic lesions may increase the frequency of false-positive predictions made by deep learning.

Supplementary References

1. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: Ourselin S, Joskowicz L, Sabuncu M, Unal G, Wells W, editors. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: Proceedings of the 19th International Conference; 2016 October 17-21; Athens, Greece. Cham: Springer International Publishing; 2016. p. 424-432.
2. Albers GW, Marks MP, Kemp S, Christensen S, Tsai JP, Ortega-Gutierrez S, et al. Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging. *N Engl J Med* 2018;378: 708-718.
3. Nogueira RG, Jadhav AP, Haussen DC, Bonafe A, Budzik RF, Bhuva P, et al. Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. *N Engl J Med* 2018;

378:11-21.
4. Campbell BC, Mitchell PJ, Kleinig TJ, Dewey HM, Churilov L, Yassi N, et al. Endovascular therapy for ischemic stroke with perfusion-imaging selection. *N Engl J Med* 2015;372:1009-1018.

Supplementary Table 1. Baseline characteristics comparison between included and excluded patients

	Included (n=414)	Excluded (n=259)	*P*
Age (yr)	70.1±12.4	68.1±12.4	0.043
Male sex	252 (60.9)	158 (61.0)	0.972
Admission NIHSS scores	3 (1–7)	1 (0–4)	<0.001*
Previous stroke	70 (16.9)	53 (20.5)	0.246
Hypertension	236 (57.0)	153 (59.1)	0.597
Diabetes	136 (32.9)	75 (29.0)	0.290
Hyperlipidemia	61 (14.7)	41 (15.8)	0.700
Current smoking	101 (24.4)	69 (26.6)	0.514
Coronary artery disease	52 (12.6)	30 (11.6)	0.706
Atrial fibrillation	109 (26.3)	45 (17.4)	0.007
Subtype			0.003
Large artery atherosclerosis	125 (30.2)	107 (41.3)	
Small vessel occlusion	57 (13.8)	27 (10.4)	
Cardioembolism	89 (21.5)	32 (12.4)	
Undetermined	138 (33.3)	87 (33.6)	
Other-determined	5 (1.2)	6 (2.3)	
Time from LKW to admission (h)	8.8 (3.0–20.5)	12.6 (3.4–44.8)	0.001*
Time from LKW to image (h)	12.4 (5.2–28.3)	17.7 (7.9–56.4)†	<0.001*
Intravenous thrombolysis	52 (12.6)	22 (8.5)	0.101
Endovascular treatment	40 (9.7)	20 (7.7)	0.390

Data were presented as mean±SD, median (interquartile range), or number (percentage).
NIHSS, National Institutes of Health Stroke Scale; LKW, last known well.
*Rank-sum test was used; †Data were missing in 15 patients.

Supplementary Table 2. Accuracy of patient classification according to cutoff using endovascular trials

Trial	LKW to image (h)	No. of patients	Category (mL)	JBS-01K			RAPID MRI		
				Correct	Incorrect	Accuracy (%)	Correct	Incorrect	Accuracy (%)
DEFUSE-3[2]	6–16	127	<70	116	0	100.0	115	1	99.2
			≥70	11	0		11	0	
DAWN[3]	6–24	171	<21	132	1	96.5	125	8	88.9
			21–31	10	2		6	6	
			31–51	5	1		2	4	
			≥51	18	2		19	1	
EXTEND-IA[4]	<4.5	86	<70	77	0	98.8	74	3	96.5
			≥70	8	1		9	0	

DEFUSE-3, Endovascular Therapy Following Imaging Evaluation for Ischemic Stroke; DAWN, DWI or CTP Assessment With Clinical Mismatch in the Triage of Wake-Up and Late Presenting Strokes Undergoing Neurointervention With Trevo; EXTEND-IA, Extending the Time for Thrombolysis in Emergency Neurological Deficits With Intra-Arterial Therapy; LKW, last known well; DWI, diffusion-weighted imaging; CTP, computed tomography perfusion.

https://doi.org/10.5853/jos.2023.02145