



OPEN

A multinational study of deep learning-based image enhancement for multiparametric glioma MRI

Yae Won Park^{1,9}, Roh-Eul Yoo^{2,9}, Ilah Shin³, Young Hun Jeon², Kanwar Partap Singh⁴, Matthew Dongwoo Lee⁴, Sohyun Kim⁵, Kevin Yang⁵, Geunu Jeong⁵, Leeha Ryu⁶, Kyunghwa Han¹, Sung Soo Ahn¹✉, Seung-Koo Lee¹, Rajan Jain^{4,7} & Seung Hong Choi^{2,8}✉

This study aimed to validate the utility of commercially available vendor-neutral deep learning (DL) image enhancement software for improving the image quality of multiparametric MRI for gliomas in a multinational setting. A total of 294 patients from three institutions (NYU, Severance, and SNUH) who underwent glioma MRI protocols were included in this retrospective study. DL image enhancement was performed on T2-weighted (T2W), T2 FLAIR, and postcontrast T1-weighted (T1W) imaging using commercially available DL image enhancement software. Signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) were calculated for both conventional and DL-enhanced images. Three neuroradiologists, one from each institution, independently evaluated the following image quality parameters in both images using a 5-point scale: overall image quality, noise, gray–white matter differentiation, truncation artifact, motion artifact, pulsation artifact, and main lesion conspicuity. The quantitative and qualitative image parameters were compared between conventional and DL-enhanced images. Compared with conventional images, DL-enhanced images showed significantly higher SNRs and CNRs in T2W, T2 FLAIR, and postcontrast T1W imaging (all $P < 0.001$). The average scores of radiologist assessments in overall image quality, noise, gray–white matter differentiation, and main lesion conspicuity were significantly higher for DL-enhanced images than conventional images in T2W, T2 FLAIR, and postcontrast T1W imaging (all $P < 0.001$). Regarding artifacts, truncation artifacts decreased (all $P < 0.001$), while pre-existing motion and pulsation artifacts were not further exaggerated in most structural MRI sequences. In conclusion, DL image enhancement using commercially available vendor-neutral software improved image quality and reduced truncation artifacts in multiparametric glioma MRI.

Abbreviations

CI	Confidence interval
CNR	Contrast-to-noise ratio
DL	Deep learning
SNR	Signal-to-noise ratio
T1W	T1-weighted
T2W	T2-weighted
WHO	World Health Organization

¹Department of Radiology and Research Institute of Radiological Science, Center for Clinical Imaging Data Science, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 120-752, Korea. ²Department of Radiology, Seoul National University Hospital, Seoul National University College of Medicine, 101, Daehangno, Jongno-gu, Seoul 03080, Republic of Korea. ³Department of Radiology, Seoul St. Mary's Hospital, College of Medicine, Seoul, Korea. ⁴Department of Radiology, New York University Grossman School of Medicine, 550 1st Ave, New York, NY, USA. ⁵Airs Medical, Seoul, Korea. ⁶Department of Biostatistics and Computing, Yonsei University Graduate School, Seoul, Korea. ⁷Department of Neurosurgery, New York University Grossman School of Medicine, 550 1st Ave, New York, NY, USA. ⁸Center for Nanoparticle Research, Institute for Basic Science, and School of Chemical and Biological Engineering, Seoul National University, 1, Gwanak-ro, Gwanak-gu, Seoul 302-909, Republic of Korea. ⁹Yae Won Park and Roh-Eul Yoo contributed equally to this work. ✉email: sungssoo@yuhs.ac; verocay1@snu.ac.kr

Diffuse gliomas account for approximately 26% of brain tumors and are the most common primary intra-axial brain tumors in adults¹. Brain MRI is the fundamental diagnostic imaging modality for preoperative, immediate postoperative, and follow-up assessments in patients with diffuse gliomas². The recommended consensus MRI protocol includes T2-weighted (T2W), T2 FLAIR, 3D precontrast and postcontrast T1-weighted (T1W) imaging as well as diffusion-weighted imaging³. Accurate image interpretation is crucial in establishing an initial preliminary diagnosis, predicting prognosis, optimizing treatment planning, and changing treatment regimens in case of tumor recurrence or progression. Thus, a method to enhance image quality while maintaining clinical and quantitative integrity in a multinational, multi-vendor setting would be useful for routine clinical practice. Furthermore, MRIs that fail to meet the standards not only impede clinical assessment but may also hinder research in broader contexts, including clinical trials and multicenter studies.

Various imaging acquisition and reconstruction techniques, such as parallel imaging or compressed sensing, have been developed to reduce scan time while maintaining image quality; however, these conventional methods require a long computational time for iterative reconstruction and may result in suboptimal image quality with undersampling artifacts^{7–9}. Recently, deep learning (DL) has emerged as a paradigm-shifting tool that can be used as an alternative or in conjunction with preexisting methods for MRI scan acceleration to improve image quality while decreasing computational power and reconstruction time^{10–13}.

Unlike conventional acceleration and reconstruction methods, which modify imaging parameters at the cost of reduced image clarity, DL methods enable improvements in image quality from undersampled data by learning complex relationships between the undersampled and fully-sampled data. These methods may operate either directly on the k-space data or on the image domain with the primary functions of reducing noise and enhancing resolution. Multiple published works have presented the clinical efficacy of DL image enhancement for preserving image quality in accelerated brain imaging or improving image quality for standard-of-care imaging including thin-section imaging^{12,14–17}. However, their application to glioma imaging and their effect on radiologists' image interpretations have not yet been explored. Furthermore, most studies have focused on data obtained from a single institution, with a limited set of images acquired under uniform acquisition parameters. Accordingly, it is necessary to rigorously validate whether the DL model consistently shows robust results under various imaging scenarios from multiple institutions with various MR scanners and imaging protocols in order to ensure its translation to routine clinical practice.

Therefore, our study aimed to validate the utility of commercially available vendor-neutral DL image enhancement software for improving the image quality of multiparametric MRI for gliomas in a multi-center, multi-vendor, and multi-reader manner.

Methods

Standard protocol approvals, registrations, and patient consents

This multinational retrospective study was approved by the Institutional Review Boards of the New York University (NYU) Langone Health, Severance Hospital, and Seoul National University Hospital (SNUH) (IRB no.: i24-01671, 2308-069-1457, and 2023-1948-004); the requirement for patient consent was waived owing to the retrospective study design. The study was conducted in accordance with the Declaration of Helsinki.

Patient enrollment

A total of 300 consecutive patients (100 glioma patients from each of the three institutions (NYU, Severance, and SNUH) who underwent a dedicated MRI protocol for the evaluation of gliomas between April 2007 and June 2023 were initially enrolled. The inclusion criteria were as follows: (1) Patients who underwent preoperative glioma MRI protocols between April 2007 and June 2023 and were pathologically confirmed with gliomas and (2) aged ≥ 18 years. Cases where the required imaging studies or medical records were unavailable or inappropriate for this study were excluded.

MRI protocol

MRI examinations were performed on 21 different 1.5T and 3.0T systems from major MR scanner vendors (GE, Philips, and Siemens). The glioma MRI protocols from each institution were in accordance with the standardized imaging protocol consensus recommendations¹⁸ including precontrast T1W, T2W, T2 FLAIR, and 3D postcontrast T1W imaging. The imaging parameters routinely used in participating institutions as part of their clinical practice varied across institutions. The specific imaging parameters for each of the scanners are provided in Supplementary Tables 1, 2, and 3.

Molecular classification

All tissues were classified and graded according to the 2021 WHO classification¹⁹. Isocitrate dehydrogenase (IDH) mutation and 1p/19q codeletion status were routinely assessed according to institutional protocols.

Deep learning-based image enhancement

Commercially available MRI enhancement software (SwiftMR, v3.0.3.0, AIRS Medical) was utilized for image denoising and enhancing spatial resolution in this study. The software was applied to the obtained images in a post-processing manner in the Digital Imaging and Communications in Medicine (DICOM) domain. The software's DL algorithm is based on the U-Net architecture²⁰. The foundational U-net consists of an initial convolutional block with 64 output channels including four stages of down-sampling and up-sampling. The model also cascades 18 convolutional blocks with layers for down-sampling, up-sampling and feature concatenation, summing up to three convolutional layers.

The software performs MR image enhancement in a multi-dimensional approach due to the training inputs generated through multi-dimensional degradation from raw k-space data. The training process involved combining noise addition and multiple patterns of undersampling, such as uniform, random, $k_{\max}$, partial Fourier and elliptical undersampling. To account for the network's complexity of learning processes from various inputs, contextual data for auxiliary input, such as the acquisition parameters defining the k-space sampling and expected noise reduction factors for each training pair, were also part of the model's architecture. The modified U-Net architecture, the Context-Enhanced U-Net (CE U-Net), integrates a dynamic modulation pathway in order to utilize contextual data as auxiliary input. The images used for model training and those acquired for this study were mutually exclusive. Detailed information regarding the software can be found in a previous article²¹.

The computational time for the image processing was 3 s for T2W, 3 s for T2 FLAIR, and 35 s for postcontrast T1W images.

Quantitative evaluation

Signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) were calculated for both conventional and DL-enhanced images by a neuroradiologist (Y.W.P. with 13 years of experience). For SNR calculation, regions of interest (ROIs) were drawn at the putamen (S_1) and internal capsule (S_2) on T2W and T2 FLAIR images and at the caudate nucleus (S_1) and corona radiata (S_2) on postcontrast T1W images. CNRs were calculated based on the signal differences between the putamen (S_1) and internal capsule (S_2) in T2W and T2 FLAIR images and between the caudate nucleus (S_1) and corona radiata (S_2) in postcontrast T1W images. Subsequently, the following formulas were used to calculate the SNRs (SNR_1 and SNR_2) and CNRs, respectively:

$$SNR_1 \text{ or } SNR_2 = \frac{\text{Mean signal intensity at } S_1 \text{ or } S_2}{\text{SD of the background noise}}$$

$$CNR = \frac{|S_1 - S_2|}{\text{SD of the background noise}}$$

Qualitative evaluation

Three board-certified neuroradiologists from three hospitals (I.S., K.P.S., and Y.H.J., with 10, 2 and 7 years of experience, respectively) performed the qualitative evaluations. To ensure consistency in qualitative evaluation, consensus thresholds were established among the three readers during a calibration session using example cases prior to formal image review. Each reviewer was asked to review both conventional and DL-enhanced images obtained from the other two institutions to avoid bias. All images were anonymized and blinded in terms of patient identification, institutional information and origin, and were presented to the readers in a randomized order. An online image storing and viewing platform (Pacsbin, Orion Medical Technologies, LLC), a fully-featured picture archiving and communication system (PACS) environment to the web complying with the Health Insurance Portability and Accountability Act (HIPAA), was utilized to facilitate multi-center data accessibility.

Each reviewer independently evaluated T2W, T2 FLAIR, postcontrast T1W with the following image quality criteria in conventional and DL-enhanced images based on a 5-point Likert scale: overall image quality, noise, gray–white matter differentiation, truncation artifact, motion artifact, pulsation artifact, and main lesion conspicuity. The overall image quality, noise, and gray–white matter differentiation was assessed as follows: 1, unacceptable; 2, poor; 3, acceptable; 4, good; and 5, excellent or ideal. Three parameters related to image artifacts (truncation, motion, and pulsation artifacts) were evaluated as follows: 1, unreadable motion artifact, images of non-diagnostic quality; 2, severe artifact, images degraded but interpretable; 3, moderate artifact with some, but not severe, effect on diagnostic quality; 4, minimal artifact, no effect on diagnostic quality; and 5, no artifact. Truncation artifacts appear as a series of parallel lines at regions with abrupt and intense signal changes, while motion artifacts are caused by random movement during the image acquisition process, and pulsation artifacts are caused by periodic motion such as respiratory or vascular pulsation. The lesion conspicuity was scored as follows: 1, unable to see; 2, blurry but visualized; 3, acceptable; 4, good; and 5, excellent.

Statistical analysis

Statistical analysis was performed by biostatisticians (L.R. and K.H. with 2 and 16 years of experience, respectively) using R studio (version 4.3.3). The normality of the data was assessed using the Shapiro–Wilk test. The clinical and imaging characteristics of patients from three institutions were compared using the Chi-square for categorical variables and Kruskal–Wallis test for continuous variables. The interobserver agreement was evaluated using the weighted Cohen kappa test (κ) and percent agreement. A κ value ≤ 0.20 indicated slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.81–1.00, almost perfect agreement. The percent agreement was additionally evaluated to comprehensively evaluate the interobserver agreement²². A generalized linear mixed model using a cumulative link function was employed; the reviewers were treated as a fixed effect while patients and the hospitals to which they belonged were treated as random effects. Owing to the exploratory nature of this study, adjustment for multiple testing was not performed.

Results

Patient characteristics

Of 300 consecutive patients with gliomas, five patients from NYU dataset were excluded due to incomplete scan range and one patient from Severance dataset was excluded due to image upload failure to an online image-sharing platform. As a result, this study included 294 glioma patients (mean age, 55.7 ± 15.7 years, range 16–86

years) comprising 168 males and 126 females. There were 30 (10.2%), 61 (20.7%), and 182 (61.9%) patients with oligodendroglioma, IDH-mutant astrocytoma, and IDH-wildtype glioblastoma, respectively, while the remaining 21 (7.1%) patients had other miscellaneous types of gliomas such as pediatric-type diffuse high-/low-grade gliomas or circumscribed astrocytic gliomas. Patients were classified as WHO grade 4 ($n=216$, 73.5%), WHO grade 2 ($n=39$, 13.3%), WHO grade 3 ($n=35$, 11.9%), and WHO grade 1 ($n=3$, 1.0%). A definite WHO grade was not pathologically defined in one patient (0.3%).

Although the distribution of sex was not statistically different between the three institutions, age ($P=0.003$), tumor type ($P<0.001$), and WHO grade ($P=0.004$) showed significant differences between three institutions, reflecting the heterogeneity of real-world datasets. The patient characteristics in the multinational dataset are summarized in Table 1.

Quantitative evaluation

Both SNR_1 and SNR_2 values were significantly higher in DL-enhanced images than in the conventional images for T2W, T2 FLAIR, and postcontrast T1W imaging (all $P<0.001$) (Fig. 1a and b). Subsequently, the CNR values of DL-enhanced images were also significantly higher than those of the conventional images across all imaging sequences (all $P<0.001$) (Fig. 1c). The measured values for SNR and CNR are shown in Table 2.

Qualitative evaluation

The results of qualitative evaluation for both conventional and DL-enhanced images are summarized in Fig. 1d; Table 3. The average radiologist assessments of overall image quality, noise, and gray–white matter differentiation were significantly higher in DL-enhanced images than in the conventional images for T2W, T2 FLAIR, and postcontrast T1W imaging (all $P<0.001$). In terms of artifacts, DL-enhanced images showed significantly fewer truncation artifacts than conventional images in T2W, T2 FLAIR, and postcontrast T1W imaging (all $P<0.001$). DL-enhanced images also showed significantly fewer motion artifacts than conventional images in T2W and postcontrast T1W imaging (all $P<0.001$). However, no significant difference was observed for motion artifact on T2 FLAIR ($P=0.721$) between DL-enhanced and conventional images. DL-enhanced images showed mixed results in terms of pulsation artifacts. There were significantly fewer pulsation artifacts in postcontrast T1W imaging compared to conventional images ($P<0.001$), while the opposite result was observed in T2 FLAIR ($P<0.001$). No significant difference was observed for pulsation artifact in T2W imaging ($P=0.052$). The main lesion conspicuity was significantly higher in DL-enhanced images than in conventional images for T2W, T2 FLAIR, and postcontrast T1W imaging (all $P<0.001$). The weighted Cohen κ ranged from slight to substantial agreement (Supplementary Table 4). Specifically, one reader consistently scored more generously, leading to systematic discrepancies. Conventional and DL-enhanced images of representative cases with gliomas are shown in Figs. 2,3,4,5.

Discussion

In this study, we investigated the utility of using a commercially available DL image enhancement software for glioma patients in a large multinational dataset. MRIs often suffer from low SNR and CNR along with image artifacts when performed within a clinically feasible scan duration with conventional acceleration techniques. Our study validated that DL image enhancement can improve image quality and reduce truncation artifacts in routine multiparametric glioma MRI protocols without exaggerating motion or pulsation artifacts among multiple vendors and scanner models.

As the training dataset for commercially available vendor-neutral DL image enhancement software included MRIs collected from multiple vendors, scanner models, and various imaging sequences with different acquisition

	Total ($n=294$)	NYU ($n=95$)	Severance ($n=99$)	SNUH ($n=100$)	P -value*
Age (years)	55.7 $\pm$ 15.7	54.3 $\pm$ 16.4	59.7 $\pm$ 14.4	53.0 $\pm$ 15.5	0.003
Male Sex	168 (57.1)	55 (57.9)	62 (62.6)	51 (51.0)	0.249
Tumor type					<0.001
Oligodendroglioma	30 (10.2)	11 (11.6)	2 (2.0)	17 (17.0)	
IDH-mutant astrocytomas	61 (20.7)	14 (14.7)	32 (32.3)	15 (15.0)	
IDH-wildtype glioblastoma	182 (61.9)	70 (73.7)	60 (60.6)	52 (52.0)	
Other glioma	21 (7.1)	0 (0)	5 (5.1)	16 (16.0)	
WHO grade					0.004
Grade 1	3 (1.0)	0 (0)	0 (0)	3 (3.0)	
Grade 2	39 (13.3)	17 (17.9)	7 (7.1)	15 (15.0)	
Grade 3	35 (11.9)	16 (16.8)	5 (5.1)	14 (14.0)	
Grade 4	216 (73.5)	62 (65.3)	87 (87.9)	67 (67.0)	
NA	1 (0.3)	0 (0)	0 (0)	1 (1.0)	

Table 1. Patient characteristics in the multinational dataset. Data are expressed as the mean $\pm$ standard deviation or numbers with percentages in parentheses. * P -value indicates statistical significance among the three institutions. IDH = isocitrate dehydrogenase; NA = not available; NYU = New York University; SNUH = Seoul National University Hospital; WHO = World Health Organization.

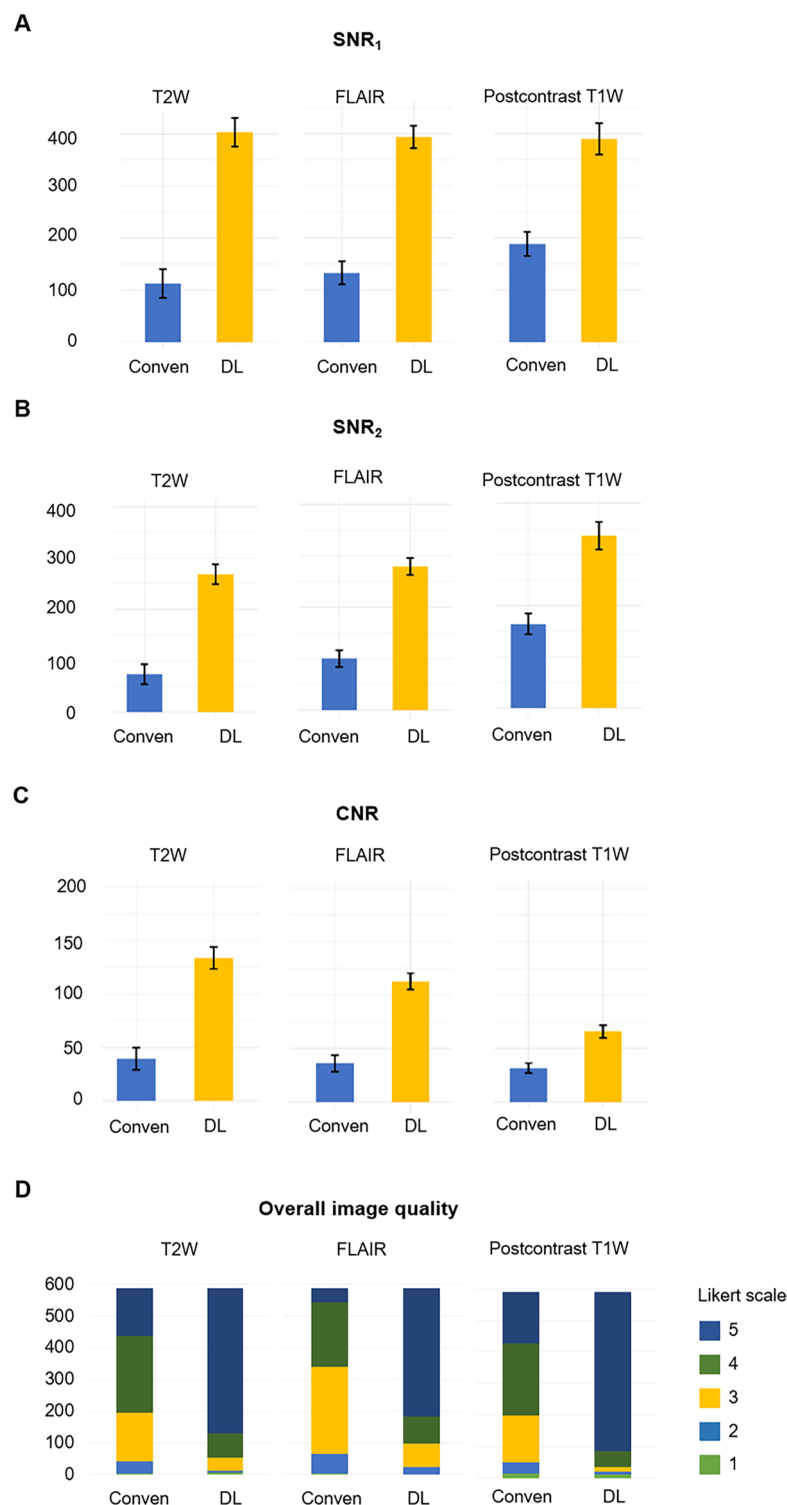


Fig. 1. Comparison of quantitative and qualitative imaging quality scores between conventional and DL-enhanced images. Mean values of (A) SNR₁, (B) SNR₂, and (C) CNR in conventional and DL-enhanced images. Error bars indicate 95% confidence intervals. (D) Comparison of the overall image quality between conventional and DL-enhanced axial T2-weighted, FLAIR, and postcontrast 3D T1-weighted images. Each image in 294 patients was rated on a 5-point Likert scale by 2 readers. CNR = contrast-to-noise ratio; Conven = conventional; DL = deep learning; SNR = signal-to-noise ratio; T1W = T1-weighted; T2W = T2-weighted.

	T2W			T2 FLAIR			Postcontrast T1W		
	Conventional	DL-enhanced	P value	Conventional	DL-enhanced	P value	Conventional	DL-enhanced	P value
SNR ₁ [*]	112.1 (84.5, 139.8)	403.6 (376.0, 431.2)	<0.001	132.8 (110.9, 154.7)	394.2 (372.2, 416.0)	<0.001	188.2 (165.4, 211.0)	390.2 (359.8, 420.6)	<0.001
SNR ₂ [*]	73.9 (54.4, 93.3)	268.8 (249.3, 288.2)	<0.001	100.1 (83.7, 116.6)	279.3 (262.9, 295.8)	<0.001	164.3 (144.0, 184.6)	336.6 (309.7, 363.5)	<0.001
CNR [†]	39.4 (29.1, 49.7)	133.8 (123.5, 144.2)	<0.001	36.1 (28.4, 43.9)	112.8 (105.1, 120.5)	<0.001	31.8 (27.2, 36.3)	65.9 (59.8, 71.9)	<0.001

Table 2. Comparison of SNR and CNR between conventional and DL-enhanced images. Data are mean values with 95% confidence intervals in parentheses. ^{*} SNR₁ or SNR₂ = (Mean signal intensity at S₁ or S₂)/(SD of the background noise) where S₁ and S₂ represent the putamen and internal capsule for T2W and T2 FLAIR images and the caudate nucleus and corona radiata for postcontrast T1W images. [†] CNR = |S₁-S₂|/(SD of the background noise). CNR = contrast-to-noise ratio; DL = deep learning; SNR = signal-to-noise ratio; T1W = T1-weighted; T2W = T2-weighted.

	T2W			T2 FLAIR			Postcontrast T1W		
	Conventional	DL-enhanced	P value	Conventional	DL-enhanced	P value	Conventional	DL-enhanced	P value
Overall image quality	3.84 (3.63, 4.05)	4.28 (4.08, 4.48)	<0.001	3.36 (3.24, 3.49)	3.69 (3.56, 3.83)	<0.001	3.84 (3.68, 3.99)	4.56 (4.46, 4.67)	<0.001
Noise	3.83 (3.73, 3.92)	4.42 (4.34, 4.50)	<0.001	3.40 (3.34, 3.46)	4.19 (4.14, 4.25)	<0.001	3.62 (3.46, 3.78)	4.64 (4.54, 4.75)	<0.001
Gray-white matter differentiation	3.99 (3.87, 4.10)	4.28 (4.17, 4.38)	<0.001	3.63 (3.50, 3.77)	3.95 (3.82, 4.08)	<0.001	4.20 (4.09, 4.32)	4.65 (4.54, 4.76)	<0.001
Truncation artifact	3.97 (3.59, 4.35)	4.25 (3.89, 4.60)	<0.001	3.59 (3.46, 3.72)	3.89 (3.77, 4.01)	<0.001	3.84 (3.57, 4.12)	4.38 (4.17, 4.59)	<0.001
Motion artifact	4.44 (4.12, 4.76)	4.58 (4.30, 4.87)	<0.001	4.05 (3.92, 4.17)	4.04 (3.91, 4.16)	0.721	4.26 (3.99, 4.53)	4.67 (4.47, 4.88)	<0.001
Pulsation artifact	4.48 (4.44, 4.52)	4.53 (4.48, 4.57)	0.052	4.07 (3.89, 4.26)	4.06 (3.87, 4.24)	<0.001	4.33 (4.20, 4.46)	4.52 (4.41, 4.63)	<0.001
Main lesion conspicuity	4.24 (4.04, 4.44)	4.44 (4.24, 4.63)	<0.001	3.87 (3.82, 3.92)	4.08 (4.02, 4.13)	<0.001	4.18 (4.02, 4.34)	4.57 (4.42, 4.71)	<0.001

Table 3. Comparison of qualitative scores between conventional and DL-enhanced images. Data are mean values with 95% confidence intervals in parentheses. Higher scores indicate better overall image quality, including lower noise and artifacts, as well as improved lesion conspicuity. DL = deep learning; T1W = T1-weighted; T2W = T2-weighted.

parameters²¹ we expected the DL image enhancement algorithm to be vendor-neutral and robust in our glioma dataset. Analysis comparing the SNR and CNR improvements after DL application to the conventional MRI showed a similar trend of SNR and CNR improvement across all institutions and imaging sequences, even though the acquisition scenarios were heterogeneous. Such results suggest that the DL image enhancement algorithm has learned generalizable representations of diverse noise patterns to produce high-quality images across over 20 different MRI scanners and acquisition schemes. To our knowledge, no previous publications explored the robustness of DL image reconstruction or enhancement algorithms to the extent of this work. Single-center or single-vendor studies may demonstrate the initial feasibility of a specific DL reconstruction or enhancement algorithm but lack the rigorous validation process necessary to facilitate clinical translation. DL algorithms that aim to influence the initial steps in routine clinical practice, such as image acquisition, require stringent evaluation because the resulting image quality can directly impact subsequent clinical assessments and decisions.

The reduction in truncation artifacts further highlights the potential of DL image enhancement in clinical settings, which may be attributed to slice-direction resolution enhancement by the algorithm²¹. As the DL model was trained on pairs of low and high spatial resolution images, it not only boosts apparent resolutions but also reduces truncation artifacts that become more prominent at lower resolutions. A previous study has also shown significant reduction in truncation artifacts with another DL algorithm²³. In conventional image reconstruction, filters are used to reduce noise and truncation artifacts, but this often leads to lower effective spatial resolution and blurred images. In contrast, our results show that DL image enhancement effectively reduces truncation artifacts and minimizes image noise while improving lesion conspicuity. Regarding motion and pulsation, common criticisms of DL image enhancement include artifact generation (hallucination) and exaggeration. However, our results suggest that the pre-existing motion and pulsation artifacts were not exaggerated in most structural MRI sequences.

From a practical point of view, one of the promising applications of DL image enhancement in a post-processing manner would be to reduce the scan time while maintaining or even improving the MRI quality in glioma patients. In patients with glioma, the capacity of anatomical MRI to accurately reveal the underlying tumor

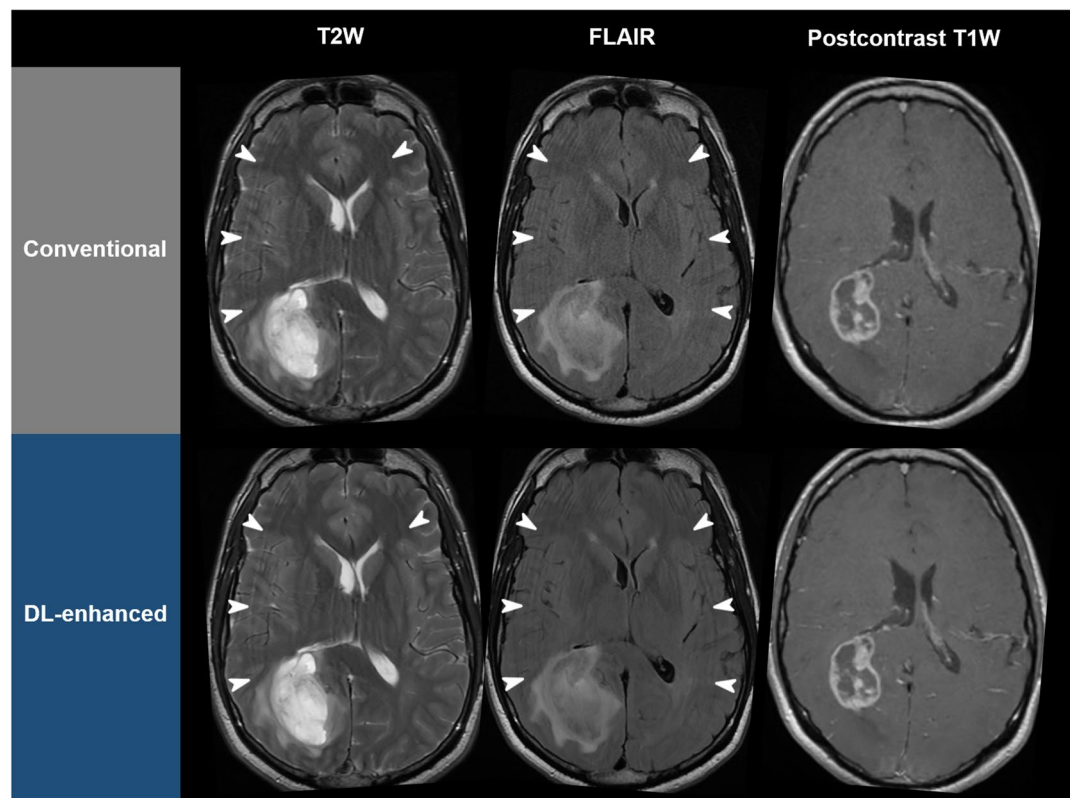


Fig. 2. A representative case of a 61-year-old female with IDH-wildtype glioblastoma. DL-enhanced T2-weighted, FLAIR, and postcontrast 3D T1-weighted images (lower row) show better overall image quality with less noise and better gray-white matter differentiation as compared with conventional images (upper row). Main mass with complex inner architecture appears more conspicuous in DL-enhanced images (lower row) than in conventional images (upper row). Note that motion artifacts (arrowheads) are not exaggerated in DL-enhanced images. DL = deep learning; T1W = T1-weighted; T2W = T2-weighted.

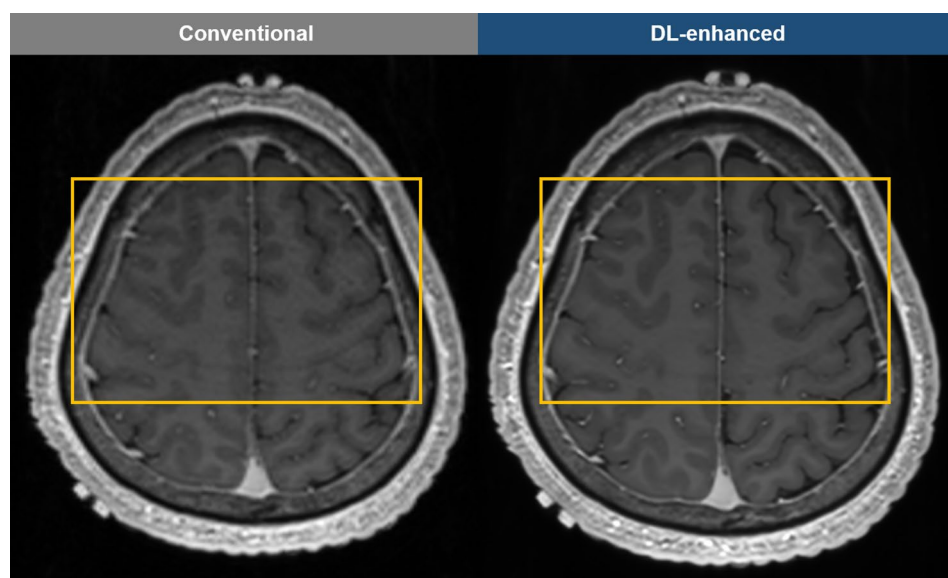


Fig. 3. A representative case of a 52-year-old male with IDH-wildtype glioblastoma. There is reduction of truncation artifacts (yellow box) on DL-enhanced postcontrast T1-weighted images compared to conventional images.

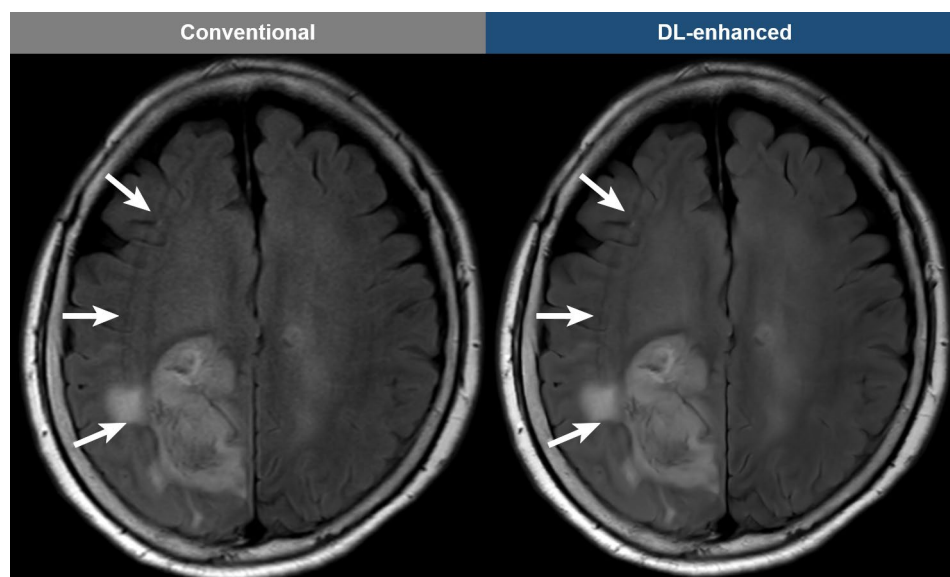


Fig. 4. A representative case of a 76-year-old female with IDH-wildtype glioblastoma. There is no exaggeration of pre-existing motion artifacts (arrows) on DL-enhanced FLAIR images compared to conventional images.

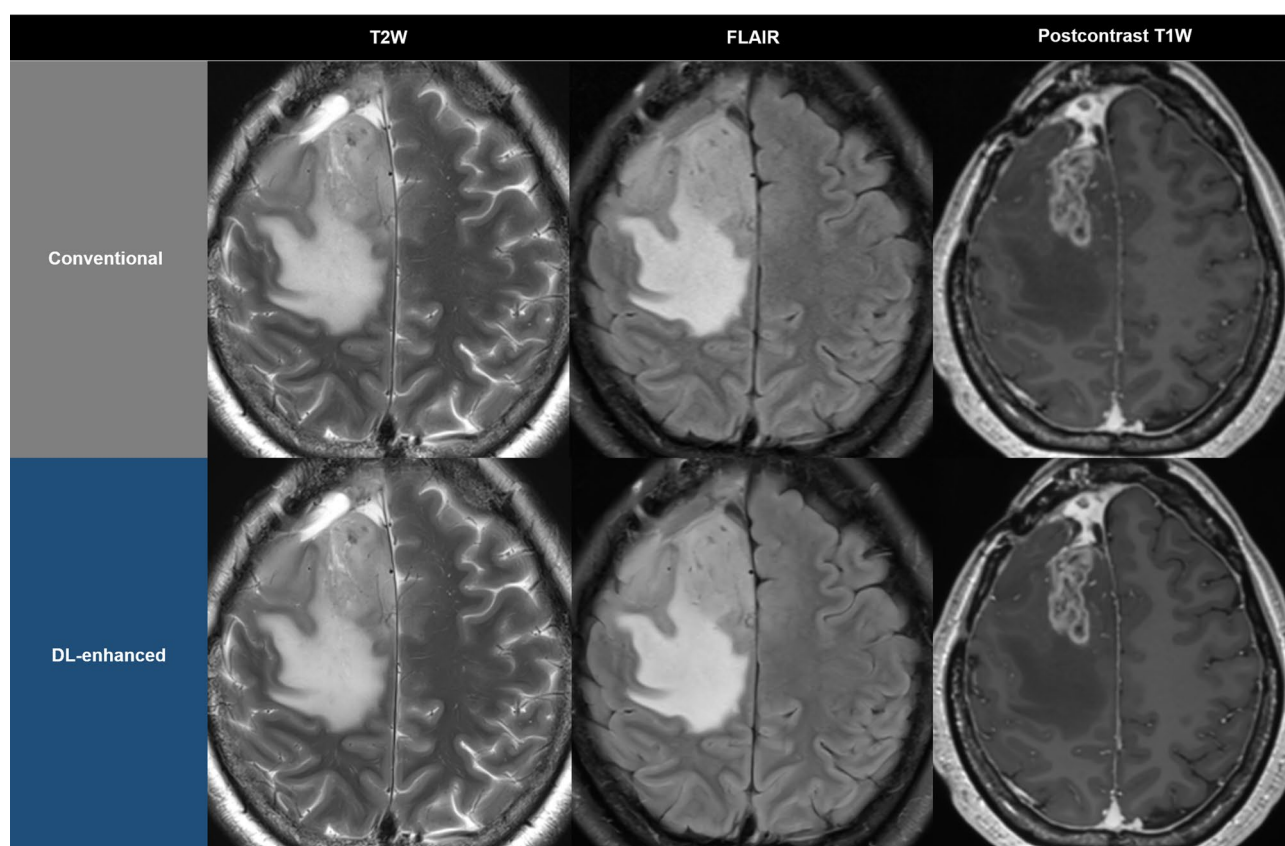


Fig. 5. A representative case of a 26-year-old female with WHO grade 4 IDH-mutant astrocytoma. A main T2 hyperintense mass with heterogeneous contrast enhancement in the right frontal lobe appears more conspicuous in DL-enhanced images (lower row) than in conventional images (upper row) (especially on T2 FLAIR and postcontrast T1W images). No new artifacts (hallucinations) were detected in DL-enhanced images, even in the presence of anatomical distortions due to extensive perilesional edema and postoperative changes at the right frontal lobe.

biology is limited, and advanced techniques such as perfusion-weighted imaging and diffusion-weighted imaging are routinely acquired at most institutions³ while MR spectroscopy and amide proton transfer imaging may also have additional clinical value²⁴. As routinely including all of these various imaging sequences would increase scan time, multiple acceleration techniques have been proposed to decrease the scan time while maintaining image quality. However, conventional acceleration techniques such as parallel imaging, simultaneous multi-slice acquisition, compressed sensing, and adjusting imaging acquisition parameters such as the receiver bandwidth, number of excitations, and in-plane/through-plane resolution generally reduce the SNR and/or spatial resolution, degrading image quality^{25–27}. In contrast, emerging DL image enhancement techniques have potential to reduce artifacts while improving image quality. Based on these results, we anticipate that this approach could lead to a reduction in scan time in actual clinical environments.

Our study had some limitations. First, the retrospective nature of this study may introduce selection bias, and while we included data from multiple international institutions, the findings may need further validation in prospective multicenter studies to ensure broader applicability. Second, our study did not investigate the utility of DL image enhancement in sequences such as diffusion-weighted imaging, perfusion-weighted imaging, or susceptibility-weighted imaging. As these are also commonly acquired sequences in glioma imaging that reflect the underlying tumor physiology³ the utility of DL image enhancement should also be assessed in these sequences in future studies. Third, SNR and CNR were calculated on a single slice per sequence, which may have led to errors for images with parallel imaging. Fourth, while no adjustment for multiple testing was performed due to the exploratory nature of the study, this may have resulted in an overestimation of statistical significance.

Conclusions

In conclusion, DL image enhancement using commercially available vendor-neutral software improved image quality and reduced truncation artifacts in multiparametric glioma MRI. This approach may be feasible and useful for clinical evaluation in glioma MRI.

Data availability

The datasets generated or analyzed during the study are available from the corresponding authors upon a reasonable request.

Received: 20 May 2025; Accepted: 28 August 2025

Published online: 25 September 2025

References

- Ostrom, Q. T. et al. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the united States in 2016–2020. *Neuro Oncol.* **25**, iv1–iv99. <https://doi.org/10.1093/neuonc/noad149> (2023).
- Weller, M. et al. EANO guidelines on the diagnosis and treatment of diffuse gliomas of adulthood. *Nat. Rev. Clin. Oncol.* **18**, 170–186. <https://doi.org/10.1038/s41571-020-00447-z> (2021).
- Ellingson, B. M. et al. Consensus recommendations for a standardized brain tumor imaging protocol in clinical trials. *Neuro Oncol.* **17**, 1188–1198. <https://doi.org/10.1093/neuonc/nov095> (2015).
- Ahn, S. S. & Cha, S. Pre- and Post-Treatment imaging of primary central nervous system tumors in the molecular and genetic era. *Korean J. Radiol.* **22**, 1858–1874. <https://doi.org/10.3348/kjr.2020.1450> (2021).
- Vollmuth, P. et al. A radiologist's guide to IDH-Wildtype glioblastoma for efficient communication with clinicians: part II—Essential information on Post-Treatment imaging. *Korean J. Radiol.* **26**, 368–389. <https://doi.org/10.3348/kjr.2024.0983> (2024).
- Vollmuth, P. et al. A radiologist's guide to IDH-Wildtype glioblastoma for efficient communication with clinicians: part I—Essential information on preoperative and immediate postoperative imaging. *Korean J. Radiol.* **26**, 246–268. <https://doi.org/10.3348/kjr.2024.0982> (2025).
- Kiryu, S. et al. Clinical impact of deep learning reconstruction in MRI. *Radiographics* **43**, e220133. <https://doi.org/10.1148/rq.220133> (2023).
- Jaspan, O. N., Fleysher, R. & Lipton, M. L. Compressed sensing MRI: a review of the clinical literature. *Br. J. Radiol.* **88**, 20150487. <https://doi.org/10.1259/bjr.20150487> (2015).
- Deshmane, A., Gulani, V., Griswold, M. A. & Seiberlich, N. Parallel MR imaging. *J. Magn. Reson. Imaging* **36**, 55–72. <https://doi.org/10.1002/jmri.23639> (2012).
- Sriram, A. et al. in *Medical Image Computing and Computer Assisted Intervention—MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, Proceedings, Part II* 23. 64–73 (Springer). (2020).
- Seo, M. et al. Deep learning improves quality of intracranial vessel wall MRI for better characterization of potentially culprit plaques. *Sci. Rep.* **14**, 18983. <https://doi.org/10.1038/s41598-024-69750-4> (2024).
- Kim, M. et al. Thin-Slice pituitary MRI with deep Learning-based reconstruction: diagnostic performance in a postoperative setting. *Radiology* **298**, 114–122. <https://doi.org/10.1148/radiol.2020200723> (2021).
- Chen, Z. et al. Deep learning for image enhancement and correction in magnetic resonance imaging—state-of-the-art and challenges. *Journal of Digital Imaging* **36**, 1, 204–230 (2023).
- Bash, S. et al. Deep learning enables 60% accelerated volumetric brain MRI while preserving quantitative performance: A prospective, multicenter, multi-reader trial. *AJNR Am. J. Neuroradiol.* **42**, 2130–2137. <https://doi.org/10.3174/ajnr.A7358> (2021).
- Oshima, S. et al. Denoising approach with deep learning-based reconstruction for neuromelanin-sensitive MRI: image quality and diagnostic performance. *Jpn J. Radiol.* **41**, 1216–1225. <https://doi.org/10.1007/s11604-023-01452-9> (2023).
- Suh, P. S. et al. Improving diagnostic performance of MRI for Temporal lobe epilepsy with deep Learning-Based image reconstruction in patients with suspected focal epilepsy. *Korean J. Radiol.* **25**, 374–383. <https://doi.org/10.3348/kjr.2023.0842> (2024).
- Seo, M. et al. Deep learning improves quality of intracranial vessel wall MRI for better characterization of potentially culprit plaques. *Scientific Reports* **14**, 1, 18983 (2024).
- Kaufmann, T. J. et al. Consensus recommendations for a standardized brain tumor imaging protocol for clinical trials in brain metastases. *Neuro Oncol.* **22**, 757–772. <https://doi.org/10.1093/neuonc/noaa030> (2020).
- Louis, D. N. et al. The 2021 WHO classification of tumors of the central nervous system: a summary. *Neuro Oncol.* **23**, 1231–1251. <https://doi.org/10.1093/neuonc/noab106> (2021).
- Ronneberger, O., Fischer, P. & Brox, T. in *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, proceedings, part III* 18. 234–241 (Springer). (2015).

21. Jeong, G., Kim, H., Yang, J., Jang, K. & Kim, J. All-in-One Deep Learning Framework for MR Image Reconstruction. *arXiv preprint arXiv:2405.03684* (2024).
22. McHugh, M. L. Interrater reliability: the kappa statistic. *Biochem. Med. (Zagreb)*. **22**, 276–282 (2012).
23. Kim, S. H. et al. Deep learning reconstruction in pediatric brain MRI: comparison of image quality with conventional T2-weighted MRI. *Neuroradiology* **65**, 207–214, (2023). <https://doi.org/10.1007/s00234-022-03053-1>
24. Galldiks, N. et al. Challenges, limitations, and pitfalls of PET and advanced MRI in patients with brain tumors: A report of the PET/RANO group. *Neuro Oncol.* **26**, 1181–1194. <https://doi.org/10.1093/neuonc/noae049> (2024).
25. Fujita, S. et al. Accelerated isotropic multiparametric imaging by high Spatial resolution 3D-QALAS with compressed sensing: A phantom, volunteer, and patient study. *Invest. Radiol.* **56**, 292–300. <https://doi.org/10.1097/rli.0000000000000744> (2021).
26. Saccenti, L. et al. Brain tissue and Myelin volumetric analysis in multiple sclerosis at 3T MRI with various in-plane resolutions using synthetic MRI. *Neuroradiology* **61**, 1219–1227. <https://doi.org/10.1007/s00234-019-02241-w> (2019).
27. Bauer, S., Markl, M., Honal, M. & Jung, B. A. The effect of reconstruction and acquisition parameters for GRAPPA-based parallel imaging on the image quality. *Magn. Reson. Med.* **66**, 402–409. <https://doi.org/10.1002/mrm.22803> (2011).

Acknowledgements

This study received technical support from AIRS Medical Inc.

Author contributions

S.S.A., S.-K.L., and S.H.C. conceived and designed the study. Y.W.P., R.-E.Y., I.S., Y.H.J., K.P.S., M.D.L., S.K., K.Y., S.S.A., S.-K.L., R.J., and S.H.C. collected and assembled the data. Y.W.P., R.-E.Y., and S.K. analyzed and interpreted the data. L.R. and K.H. performed statistical analysis. Y.W.P., R.-E.Y., M.D.L., S.K., K.Y., S.S.A., and S.H.C. wrote the main manuscript text. G.J. participated in the interpretation of data and critical revision of the manuscript. All authors reviewed and approved the final manuscript.

Funding

This work was supported by the Korea Medical Device Development Fund grant funded by the Korean government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health & Welfare, the Ministry of Food and Drug Safety) (Project Number: RS-2023-00224382). This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2023R1A2C3003250). This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (RS-2025-00515423). This study was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (RS-2023-00242754 and RS-2023-00207783), by Samsung Research Funding & Incubation Center of Samsung Electronics under Project Number SRFC-IT2201-04, by SNUH GE center grant (1820230040), by Korea Basic Science Institute (National Research Facilities and Equipment Center) grant funded by the Ministry of Education (RS-2024-00435727), by SNU Research Grant (1000-20240004), and by the Institute for Basic Science (IBS-R006-D1).

Declarations

Competing interests

Dr. Seung Hong Choi is an external director of AIRS Medical Inc. All the other authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-17993-0>.

Correspondence and requests for materials should be addressed to S.S.A. or S.H.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025