JOURNAL *of*
THE KOREAN SOCIETY *of*
RADIOLOGY

Adaptation of Deep Learning Image Reconstruction for Pediatric Head CT: A Focus on the Image Quality

소아용 두부 컴퓨터단층촬영에서 딥러닝 영상 재구성 적용: 영상 품질에 대한 고찰

Nim Lee, MD[1,2] (iD), Hyun-Hae Cho, MD[1*] (iD),
So Mi Lee, MD[3] (iD), Sun Kyoung You, MD[4] (iD)

[1]Department of Radiology, Medical Research Institute, College of Medicine, Ewha Womans University Mokdong Hospital, Seoul, Korea
[2]Department of Radiology, Research Institute of Radiological Science, Severance Hospital, Yonsei University, College of Medicine, Seoul, Korea
[3]Department of Radiology, School of Medicine, Kyungpook National University, Kyungpook National University Chilgok Hospital, Daegu, Korea
[4]Department of Radiology, Chungnam National University Hospital, Chungnam National University College of Medicine, Daejeon, Korea

ORCID iDs
Nim Lee (iD) https://orcid.org/0000-0002-5781-1424
Hyun-Hae Cho (iD) https://orcid.org/0000-0002-4865-2601
So Mi Lee (iD) https://orcid.org/0000-0002-2073-8198
Sun Kyoung You (iD) https://orcid.org/0000-0002-1026-5809

Purpose To assess the effect of deep learning image reconstruction (DLIR) for head CT in pediatric patients.
Materials and Methods We collected 126 pediatric head CT images, which were reconstructed using filtered back projection, iterative reconstruction using adaptive statistical iterative reconstruction (ASiR)-V, and all three levels of DLIR (TrueFidelity; GE Healthcare). Each image set group was divided into four subgroups according to the patients' ages. Clinical and dose-related data were reviewed. Quantitative parameters, including the signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR), and qualitative parameters, including noise, gray matter-white matter (GM-WM) differentiation, sharpness, artifact, acceptability, and unfamiliar texture change were evaluated and compared.
Results The SNR and CNR of each level in each age group increased among strength levels of DLIR. High-level DLIR showed a significantly improved SNR and CNR ($p < 0.05$). Sequential reduction of noise, improvement of GM-WM differentiation, and improvement of sharpness was noted among strength levels of DLIR. Those of high-level DLIR showed a similar value as that with ASiR-V. Artifact and acceptability did not show a significant difference among the adapted levels of DLIR.
Conclusion Adaptation of high-level DLIR for the pediatric head CT can significantly reduce image

noise. Modification is needed while processing artifacts.

INTRODUCTION

There are increasing concerns about radiation exposure during medical imaging examinations, particularly in pediatric patients undergoing CT, who are relatively sensitive to radiation (1-5). Although efforts are being made to reduce the examination itself, it does not decrease much due to unclear symptoms and past history of the pediatric patient, and for this reason, head CT is often performed in emergency situations in children (3-5).

Therefore, in response to these concerns for radiation dose reduction, CT image reconstruction methods were developed from conventional filtered back projection (FBP) to iterative reconstruction (IR) (1-3, 5-8). Moreover, widely used variable IR techniques have been successful and useful in reducing radiation dose recently. However, image texture, spatial resolution and lesion detectability remain challenging (1, 2, 5-9).

Recently, deep learning techniques have been spotlighted in the medical imaging field to detect lesions, segmentation, characterization and image reconstruction (1, 2, 6-8). Adaptation of the deep learning algorithm for image reconstruction is expected to lower the image noise and increase the spatial resolution simultaneously (1, 2, 6-8). Therefore, there have been many recent studies using variable deep learning image reconstruction (DLIR), including phantom study (1, 2, 6-8). However, a majority of these studies have been performed in adults, with a limited information about children (1, 2, 7-11). Therefore, the purpose of this study was to evaluate image quality using DLIR of pediatric head CT data, compared with conventional FBP and IR methods.

MATERIALS AND METHODS

Approval of the Institutional Review Board was obtained and the requirement for subject/parent informed consent was waived (IRB No. 2020-08-032-001).

STUDY DESIGN

A 512-slice CT scan (Revolution; GE Healthcare, Milwaukee, WI, USA) was installed in our hospital in February 2019, and all CT images were displayed after FBP and IR using adaptive statistical iterative reconstruction (ASiR)-V (GE Healthcare, Waukesha, WI, USA). After DLIR of this vendor received approval from the Food and Drug Administration (FDA) as a commercialized deep learning reconstruction method, adaptation of DLIR (TrueFidelity; GE Healthcare) was followed at our center in December 2019.

Before adaptation for real clinical examination, reconstructed images using DLIR with conventional dose setting for head CT of our center was analyzed using an American Association of Physicists in Medicine CT performance phantom (Model 76-410-4,130; Nuclear Associates,

Carle Place, NY, USA). All parameters were noted to be acceptable for clinical adaptation.

For exact comparison, we included pediatric patients who encountered to our emergency center and underwent non-contrast head CT studies from December 2019 to August 2020. Among those studies for follow-up in the same patients were excluded, and there were no studies where obtaining an appropriate region of interest (ROI) was impossible owing to pathology or severe artifacts. Therefore, a total of 126 studies were finally included.

CLINICAL DATA

Owing to different protocols, each image set group was divided into four subgroups according to the patient's age: less than 1 year, 1–6 years, 6–12 years and more than 12 years. Clinical data including age on examination date, sex, body weight, height and body mass index were recorded for each patient. The reasons for examination were also recorded.

CT SCAN IMAGE ACQUISITION

We used a non-enhanced head CT protocol with 100 kVp and a reference mA of 250 (effective mAs of 125) in younger than 1 year, 290 (effective mAs of 145) in ages 1–6 years. And 120 kVp and a reference mA of 200 (effective mAs of 100) in ages 6–12 years, or 270 (effective mAs of 135) in older than 12 years. Slice thickness was about 2 mm and coronal and sagittal reformation was done with 2 mm reconstruction interval with 0.625 mm thickness source data.

DATA RECONSTRUCTION

In those patients, image sets for study was concluded with studies reconstructed by FBP (FBP group), ASiR-V (ASiR-V group) and DLIR (DLIR group). Among the ASiR-V level, 30% ASiR-V was adapted for the patient with under age 12 and 40% ASiR-V was adapted for the older patients based on the previous studies (5, 12). All three level (low, medium and high) of DLIR were used and compared, which were blinded for all clinical data, medical information and radiation dose data.

All CT image data sets were displayed at the picture archiving and communication systems (PACS) workstation (G3 PACS; G3 Infinitt PACS; Infinitt Healthcare, Seoul, Korea) with brain setting of 80 Hounsfield units (HU) window width and 40 HU window level.

RADIATION DOSE

We collected estimated CT dose index volume (CTDIvol), size specific dose estimates (SSDE) and dose length product (DLP) data for all scans. Those data were automatically calculated by the CT machine during examination and displayed on the PACS. Effective dose was calculated using the conversion factor in International Commission on Radiological Protection report 2007 (13, 14), and adapted according to the mean age of age group. The conversion factor for 100 V was used for age group under 6 years old; 0.0054 for patients younger than 1 year and 0.0035 for ages 1 to 6 years. And conversion factor for 120 V was adapted for patients older than 6 years old ; 0.0027 for ages 6 to 12 years and 0.0019 for those older than 12 years.

QUANTITATIVE IMAGE QUALITY

Objective analysis for quantitative image quality was done by calculating signal-to-noise ra-

tio (SNR) and contrast-to-noise ratio (CNR). The average signal attenuation is measured by using mean HU within the ROI. The standard deviation within the ROI was calculated and considered as parameters for noise.

To reduce bias, two pediatric radiologists (H.-H.C. and S.K.Y., each with 10 years of experience with CT), who were blinded for all clinical data, medical information and radiation dose data.

Each observer drew regions of interest in each area, level for each image set, and the mean value was used for final analysis. We measured noise in air by averaging the standard deviation of three 30-mm^2 circular regions of interest that were drawn at three different places in the air on the lateral ventricle level. We drew a 10-mm^2 ROI for noise of cerebrospinal fluid on the level of lateral ventricle and drew three different regions of interest within the white matter (WM) of centrum semiovale for noise in the WM on the level of high convexity. These measuring methods were adapted from our previous study (4).

With the measured data, SNR of air, cerebrospinal fluid and WM were calculated by the following equation:

$$SNR = \frac{\text{Mean HU of tissue in ROI}}{\text{SD of HU in ROI}}$$

where Mean and SD refer to the mean and standard deviation, respectively, of Hounsfield numbers and ROI refer to region of interest of measured level.

CNR was calculated at three different levels: high convexity, lateral ventricle and infratentorial. We drew six 4-mm^2 circular regions of interest at each level of the centrum semiovale of high convexity level and infratentorial level, three of which were placed in the WM, and three of which were placed in the cortical gray matter (GM). We averaged the mean density and standard deviation of measured data, as described in the previous studies for pediatric head CT (4).

To calculate CNR of the lateral ventricle level, three regions of interest for GM were placed within the deep GM, such as the thalamus or basal ganglia. To reduce the bias made by measuring at different levels of the different patients, we placed regions of interest at the high convexity level where the central gyrus appears longest, the lateral ventricle level as the 3rd ventricle appearance level, and the infratentorial level as the level where the 4th ventricle shows an inverted U shaped for standardization of measurement level.

The CNR was calculated at each of the three levels—high convexity, lateral ventricle and infratentorial levels—by using the following equation:

$$CNR = \frac{\text{Mean}_{\text{gray matter}} - \text{Mean}_{\text{white matter}}}{\sqrt{\text{SD}^2_{\text{gray matter}} + \text{SD}^2_{\text{white matter}}}}$$

where Mean and SD refer to the mean and standard deviation, respectively, of Hounsfield numbers.

QUALITATIVE IMAGE QUALITY

Qualitative image quality was rated with reference to subjective noise, GM-WM differentiation of the supratentorial and infratentorial levels, sharpness, artifact, acceptability and im-

age texture. Among the parameters, sharpness was defined according to how clearly the margin of ventricles was defined. Artifact was defined based on whether there exist motion or beam hardening artifacts and how severely displayed. Image texture was defined by how severely "blotchy" or "plastic-like texture" is noted when it was compared to the FBP. If there exist severe texture change with unfamiliar smoothness, the lower score was recorded. Analysis for qualitative image quality was conducted by two attending pediatric radiologists (H.-H. C and S.M.L., each with 10 years of clinical experience). To reduce bias, all identifying data were removed for each image set and then image sets were reordered randomly before analysis. Interobserver agreement between the two radiologists was also calculated.

STATISTICAL ANALYSIS

To compare the quantitative parameters of DLIR high, medium and low level were compared with that of ASiR-V group, the Wilcoxon signed rank test was used for the results of the 1 year and younger group and 1 to 6 years age group. The paired Student's *t* test was used for the analysis of those results of ages 6 to 12 years, as well as for the group 12 years and older.

To compare the results of the qualitative analysis of image quality, those results of DLIR high, medium and low level were compared with that of ASiR-V group by using McNemar's test in age group of 1 year and younger group and 1 to 6 years age group. To compare the results of the those qualitative analysis of age group 6 to 12 years and 12 years and older, the Kruskal Wallis and analysis of variation (ANOVA) was used. Interobserver agreement was assessed with the weighted Cohen's kappa statistic. All statistical analyses were conducted using commercially available software (SPSS Statistics, Version 19.0; IBM Corp., Armonk, NY, USA).

RESULTS

CLINICAL DATA AND RADIATION DOSE

The study group of 126 patients (male:female = 78:48, mean age: 109.5 months; range: 4 month-204 month from December 2019 to August 2020) were included. Those were divided into 4 subgroups according to the patient's age due to different scan protocols. Table 1 summarizes the clinical data and radiation dose results of each subgroups.

Among those patients, 83 patients were referred due to head trauma, 22 patients due to headache, 14 patients due to seizure, 4 patients due to syncope and 3 patients for other nonspecific reasons.

Table 1. Comparison of Clinical Factors of Patients Based on Age Groups

Age Group	N	Mean Age (months)	Body Weight (kg)	Height (cm)	CT Dose Index Volume (mGy)	DLP (mGy*cm)	Effective Dose (mSv)	SSDE*
< 1 year	8	6.3	7.93	59.3	10.5	135.1	0.73	12.29
1–6 years	28	37.7	14.7	99.6	12.5	197.9	0.69	26.00
6–12 years	48	117.9	42.2	147.4	14.0	242.1	0.85	28.14
> 12 years	42	182.2	68.0	167.8	17.7	297.5	0.42	36.82

*SSDE was calculated based on the American Association of Physicists in Medicine report No. 204.
DLP = dose length product, SSDE = size specific dose estimates

Table 2. Comparison of the SNR of Image Sets

	FBP	ASiR-V	DLIR-Low	DLIR-Medium	DLIR-High	p-Value*		
						DLIR-Low	DLIR-Medium	DLIR-High
High convexity level WM								
< 1 year	4.50 ± 1.41	5.90 ± 2.24	4.90 ± 1.92	5.79 ± 2.43	9.88 ± 2.62	0.017	0.779	0.012
1–6 years	4.21 ± 2.11	6.27 ± 2.32	4.90 ± 2.63	5.88 ± 2.31	8.89 ± 1.41	< 0.001	0.076	< 0.001
6–12 years	3.25 ± 1.97	4.80 ± 2.56	4.05 ± 2.44	4.97 ± 1.83	6.50 ± 1.72	< 0.001	0.572	< 0.001
> 12 years	3.47 ± 2.39	5.56 ± 1.93	4.24 ± 1.14	5.33 ± 2.56	7.84 ± 2.18	< 0.001	0.244	< 0.001
Lateral ventricle level WM								
< 1 year	5.27 ± 1.63	7.37 ± 1.56	5.98 ± 1.24	7.13 ± 3.12	12.58 ± 2.44	0.017	0.092	0.012
1–6 years	3.55 ± 1.28	4.77 ± 1.24	4.11 ± 1.93	4.67 ± 1.67	6.37 ± 3.42	0.308	0.554	< 0.001
6–12 years	3.95 ± 1.96	5.92 ± 1.12	4.93 ± 1.29	6.06 ± 2.31	8.40 ± 2.92	< 0.001	0.503	< 0.001
> 12 years	4.35 ± 2.44	6.63 ± 1.42	5.23 ± 1.61	6.85 ± 1.61	10.17 ± 1.38	< 0.001	0.153	< 0.001
Infratentorial level WM								
< 1 year	4.95 ± 1.28	5.74 ± 0.84	5.13 ± 0.92	6.09 ± 2.13	10.15 ± 1.81	0.680	0.740	0.014
1–6 years	4.26 ± 2.42	6.63 ± 1.63	5.05 ± 1.06	5.97 ± 2.69	8.64 ± 2.45	< 0.001	0.001	< 0.001
6–12 years	3.46 ± 1.31	4.77 ± 0.92	4.07 ± 0.86	4.86 ± 1.63	6.41 ± 1.94	0.007	0.673	< 0.001
> 12 years	4.13 ± 1.94	7.09 ± 1.45	5.10 ± 1.34	6.40 ± 1.41	8.71 ± 1.78	< 0.001	< 0.001	< 0.001
CSF								
< 1 year	1.3 ± 0.74	1.8 ± 0.52	1.32 ± 0.86	1.55 ± 1.28	3.47 ± 1.11	0.012	0.682	0.012
1–6 years	1.25 ± 0.72	1.96 ± 1.39	1.39 ± 0.74	1.67 ± 1.69	2.71 ± 1.27	< 0.001	< 0.001	0.003
6–12 years	0.90 ± 0.49	1.56 ± 1.69	1.16 ± 0.84	1.54 ± 0.92	2.36 ± 1.24	< 0.001	0.801	< 0.001
> 12 years	0.82 ± 1.30	1.70 ± 1.97	1.14 ± 0.96	1.63 ± 1.19	2.74 ± 1.36	< 0.001	0.101	< 0.001
Air								
< 1 year	-209.6 ± 17.84	-242.7 ± 15.69	-278.2 ± 19.21	-359.5 ± 12.04	-525.9 ± 13.97	0.012	0.012	0.012
1–6 years	-151.3 ± 14.82	-191.8 ± 18.24	-167.0 ± 17.64	-222.0 ± 15.07	-270.3 ± 24.21	0.263	0.010	< 0.001
6–12 years	-133.2 ± 23.3	-183.5 ± 17.61	-158.1 ± 21.93	-205.7 ± 24.63	-292.5 ± 13.81	0.037	0.017	< 0.001
> 12 years	-126.8 ± 19.27	-183.0 ± 22.64	-161.0 ± 12.74	-207.7 ± 21.69	-269.0 ± 21.36	< 0.001	0.055	< 0.001

*p value was calculated by comparison between each level of DLIR with that of ASiR-V.
ASiR = adaptive statistical iterative reconstruction, CSF = cerebrospinal fluid, DLIR = deep learning image reconstruction, FBP = filtered back projection, SNR = signal to noise ratio, WM = white matter

QUANTITATIVE ANALYSIS OF IMAGE QUALITY

The quantitative image analysis results are present in Tables 2 and 3. A gradual reduction of noise in WM, cerebrospinal fluid and air was noted among strength levels of DLIR in sequence from low to high. The SNR of high convexity level WM, lateral ventricle level WM, infratentorial level WM and cerebrospinal fluid in all age groups increased among strength levels of DLIR in sequence from low to high. That of air showed reversed pattern due to negative value HU of air. When compared with ASiR-V, high level DLIR showed significantly improved SNR ($p < 0.05$) in all measured areas. The SNR of low level DLIR showed significant lower p value compared with ASiR-V ($p < 0.05$).

The CNR of high convexity level WM, lateral ventricle level WM, infratentorial level WM in all age groups increased among strength levels of DLIR in sequence from low to high. When compared with ASiR-V, high level DLIR showed significantly improved CNR ($p < 0.05$). And the CNR of low level DLIR showed significant lower p value compared with ASiR-V ($p < 0.05$) in all

Table 3. Comparison of the CNR of Image Sets

	FBP	ASiR-V	DLIR-Low	DLIR-Medium	DLIR-High	*p*-Value*		
						DLIR-Low	DLIR-Medium	DLIR-High
CNR high convexity								
< 1 year	0.80 ± 0.34	1.23 ± 0.42	0.98 ± 0.24	1.17 ± 0.93	1.99 ± 0.67	0.035	0.326	0.012
1–6 years	0.99 ± 0.38	1.44 ± 0.46	1.22 ± 0.48	1.42 ± 1.08	2.01 ± 0.71	< 0.001	0.665	< 0.001
6–12 years	0.81 ± 0.47	1.10 ± 0.32	0.89 ± 0.51	1.04 ± 0.91	1.48 ± 0.58	0.001	0.355	< 0.001
> 12 years	0.93 ± 0.32	1.28 ± 0.69	1.11 ± 0.29	1.32 ± 0.84	1.87 ± 0.69	0.013	0.481	< 0.001
CNR lateral ventricle								
< 1 year	0.74 ± 0.29	1.21 ± 0.65	0.93 ± 0.38	1.00 ± 0.56	1.65 ± 0.54	0.040	0.889	0.036
1–6 years	0.59 ± 0.17	0.87 ± 0.36	0.47 ± 0.21	0.62 ± 0.29	1.58 ± 0.61	< 0.001	0.011	0.001
6–12 years	0.83 ± 0.38	1.26 ± 0.44	1.01 ± 0.28	1.15 ± 0.31	1.63 ± 0.79	< 0.001	0.011	0.005
> 12 years	0.57 ± 0.20	0.90 ± 0.36	0.71 ± 0.45	0.87 ± 0.32	1.40 ± 0.58	< 0.001	0.019	< 0.001
CNR infratentorial								
< 1 year	0.99 ± 0.12	1.34 ± 0.46	0.46 ± 0.17	0.58 ± 0.14	0.98 ± 0.36	0.017	0.207	0.012
1–6 years	0.99 ± 0.25	1.50 ± 0.51	1.19 ± 0.39	1.47 ± 0.37	2.07 ± 0.69	< 0.001	0.056	< 0.001
6–12 years	0.61 ± 0.30	1.01 ± 0.16	0.76 ± 0.22	0.89 ± 0.16	1.10 ± 0.31	0.020	0.250	0.045
> 12 years	0.72 ± 0.14	1.01 ± 0.32	0.81 ± 0.21	0.98 ± 0.29	1.28 ± 0.39	< 0.001	0.248	0.002

*p value was calculated by comparison between each level of DLIR with that of ASiR-V.
ASiR = adaptive statistical iterative reconstruction, CNR = contrast-to-noise ratio, DLIR = deep learning image reconstruction, FBP = filtered back projection

levels, like those of SNR.

QUALITATIVE ANALYSIS OF IMAGE QUALITY

The qualitative image analysis results are presented in Table 4. Sequential decreased score, which means reduction of noise, improvement of GM-WM differentiation of supratentorial and infratentorial level, improvement of sharpness, noted among strength levels of DLIR in sequence from low to high level. Most high level DLIR showed similar scoring compared with ASiR-V without significant difference. And in GM-WM differentiation of supratentorial level in age group between 1 to 6 years old, the high level DLIR scored significantly better than that of ASiR-V ($p < 0.05$).

However, artifact and acceptability didn't showed significant difference among the adapted levels of DLIR. Even worsening of artifact was observed in the third age subgroup, when using high level DLIR when compared with ASiR-V group with significant difference. It adversely affect for overall acceptability (Figs. 1, 2). Image texture was noted most unfamiliar on ASiR-V group when compared with FBP set and showed significant improvement on DLIR image sets ($p < 0.05$). The interobserver agreement between the two readers was more than substantial (mean kappa value about 0.68) (Table 5).

DISCUSSION

This study observed that high level DLIR yielded significantly lower image noise and higher SNR and CNR, compared with conventional FBP and ASiR-V. This result, as the first clini-

cal validation of commercialized DLIR, can prove that using DLIR can aid in achieving better image quality in low-dose pediatric head CT scans compared with ASiR-V. Since ASiR-V is an established reconstruction method for pediatric CT studies (5), high-level DLIR can be an al-

Table 4. Comparison of Qualitative Parameters of Image Sets

	FBP	ASiR-V	DLIR-Low	DLIR-Medium	DLIR-High	*p*-Value*		
						DLIR-Low	DLIR-Medium	DLIR-High
Noise								
< 1 year	3.38 ± 1.21	1.38 ± 1.13	2.38 ± 1.22	1.88 ± 1.42	1.00 ± 0.86	0.023	0.045	0.083
1–6 years	2.68 ± 0.92	1.11 ± 1.14	2.21 ± 1.09	1.39 ± 0.89	1.04 ± 0.63	< 0.001	0.021	0.317
6–12 years	2.20 ± 1.02	1.06 ± 1.02	2.19 ± 0.81	1.42 ± 0.49	1.02 ± 0.69	< 0.001	< 0.001	0.322
> 12 years	2.60 ± 1.48	1.14 ± 0.98	2.12 ± 0.65	1.62 ± 0.67	1.05 ± 0.85	< 0.001	< 0.001	0.534
GM-WM differentiation (supratentorial level)								
< 1 year	3.13 ± 0.79	1.50 ± 0.84	2.25 ± 0.19	2.13 ± 0.54	1.13 ± 0.28	0.056	0.059	0.180
1–6 years	2.50 ± 0.46	1.86 ± 0.39	2.26 ± 0.82	2.04 ± 0.75	1.25 ± 0.64	0.001	0.132	< 0.001
6–12 years	2.44 ± 0.71	1.73 ± 0.82	2.44 ± 0.41	2.15 ± 0.46	1.79 ± 0.51	< 0.001	< 0.001	0.001
> 12 years	2.40 ± 1.21	1.79 ± 0.62	2.40 ± 0.74	2.02 ± 0.66	1.21 ± 0.39	< 0.001	0.058	< 0.001
GM-WM differentiation (infratentorial level)								
< 1 year	2.88 ± 0.49	1.63 ± 0.78	2.25 ± 0.87	1.88 ± 0.39	1.13 ± 0.45	0.129	0.157	0.046
1–6 years	2.29 ± 0.42	1.39 ± 0.39	2.00 ± 0.55	1.86 ± 0.34	1.07 ± 0.24	< 0.001	0.007	0.013
6–12 years	2.13 ± 0.57	1.27 ± 0.61	2.13 ± 0.46	1.85 ± 0.46	1.38 ± 0.34	< 0.001	< 0.001	0.302
> 12 years	2.33 ± 0.38	1.56 ± 0.31	2.10 ± 0.45	1.86 ± 0.59	1.07 ± 0.32	< 0.001	< 0.001	0.003
Sharpness								
< 1 year	3.00 ± 0.44	1.75 ± 0.21	2.25 ± 0.24	1.75 ± 0.59	1.15 ± 0.44	0.234	1.000	0.059
1–6 years	1.86 ± 0.67	1.29 ± 0.41	1.39 ± 0.46	1.21 ± 0.43	1.14 ± 0.33	0.026	0.021	0.206
6–12 years	1.48 ± 0.57	1.19 ± 0.30	1.48 ± 0.37	1.42 ± 0.42	1.08 ± 0.37	0.003	0.062	0.096
> 12 years	2.02 ± 0.64	1.24 ± 0.42	1.38 ± 0.45	1.55 ± 0.31	1.10 ± 0.21	0.262	0.014	0.083
Artifact								
< 1 year	2.75 ± 0.17	2.00 ± 0.27	2.75 ± 0.35	2.50 ± 0.39	2.50 ± 0.36	0.014	0.646	0.234
1–6 years	2.32 ± 0.44	1.64 ± 0.38	2.25 ± 0.51	2.29 ± 0.29	2.07 ± 0.41	0.002	< 0.001	0.007
6–12 years	2.33 ± 0.32	1.75 ± 0.25	2.13 ± 0.18	2.15 ± 0.41	2.29 ± 0.27	< 0.001	< 0.001	< 0.001
> 12 years	2.64 ± 0.29	1.67 ± 0.26	2.10 ± 0.29	2.00 ± 0.32	1.90 ± 0.19	< 0.001	0.012	0.049
Acceptability								
< 1 year	3.25 ± 0.31	1.63 ± 0.23	2.25 ± 0.29	2.25 ± 0.35	2.00 ± 0.32	0.131	0.102	0.257
1–6 years	2.07 ± 0.25	1.46 ± 0.21	2.54 ± 0.31	2.07 ± 0.36	2.04 ± 0.35	< 0.001	< 0.001	0.006
6–12 years	1.92 ± 0.46	1.40 ± 0.33	1.92 ± 0.17	2.13 ± 0.51	2.06 ± 0.26	< 0.001	< 0.001	0.000
> 12 years	2.71 ± 0.56	1.52 ± 0.36	1.95 ± 0.36	2.29 ± 0.39	1.83 ± 0.27	0.002	< 0.001	0.057
Image texture								
< 1 year		1.02 ± 0.28†	4.02 ± 0.96†	2.16 ± 0.63†	1.18 ± 0.91†	< 0.001	< 0.001	< 0.001
1–6 years		1.38 ± 0.38†	3.24 ± 0.66†	3.20 ± 0.72†	1.98 ± 0.42†	< 0.001	< 0.001	< 0.001
6–12 years		1.19 ± 0.42†	4.26 ± 0.38†	2.96 ± 0.42†	2.45 ± 0.39†	< 0.001	< 0.001	< 0.001
> 12 years		1.61 ± 0.30†	3.94 ± 0.74†	3.12 ± 0.51†	2.01 ± 0.67†	< 0.001	< 0.001	< 0.001

*p value was calculated by comparison between each level of DLIR with that of ASiR-V.
†Scores were scored image texture compared with the FBP image set.
ASiR = adaptive statistical iterative reconstruction, DLIR = deep learning image reconstruction, FBP = filtered back projection, GM = gray matter, WM = white matter

Fig. 1. A representative case of head CT of a 29-months-old boy.
A-E. It was reconstructed using FBP (A), ASiR (B), and DLIR-low (C), medium (D), and high (E). An increment of white-gray matter differentiation and sharpness when using higher grade DLIR on the supratentorial level are noted.
ASiR = adaptive statistical iterative reconstruction, DLIR = deep learning image reconstruction, FBP = filtered back projection
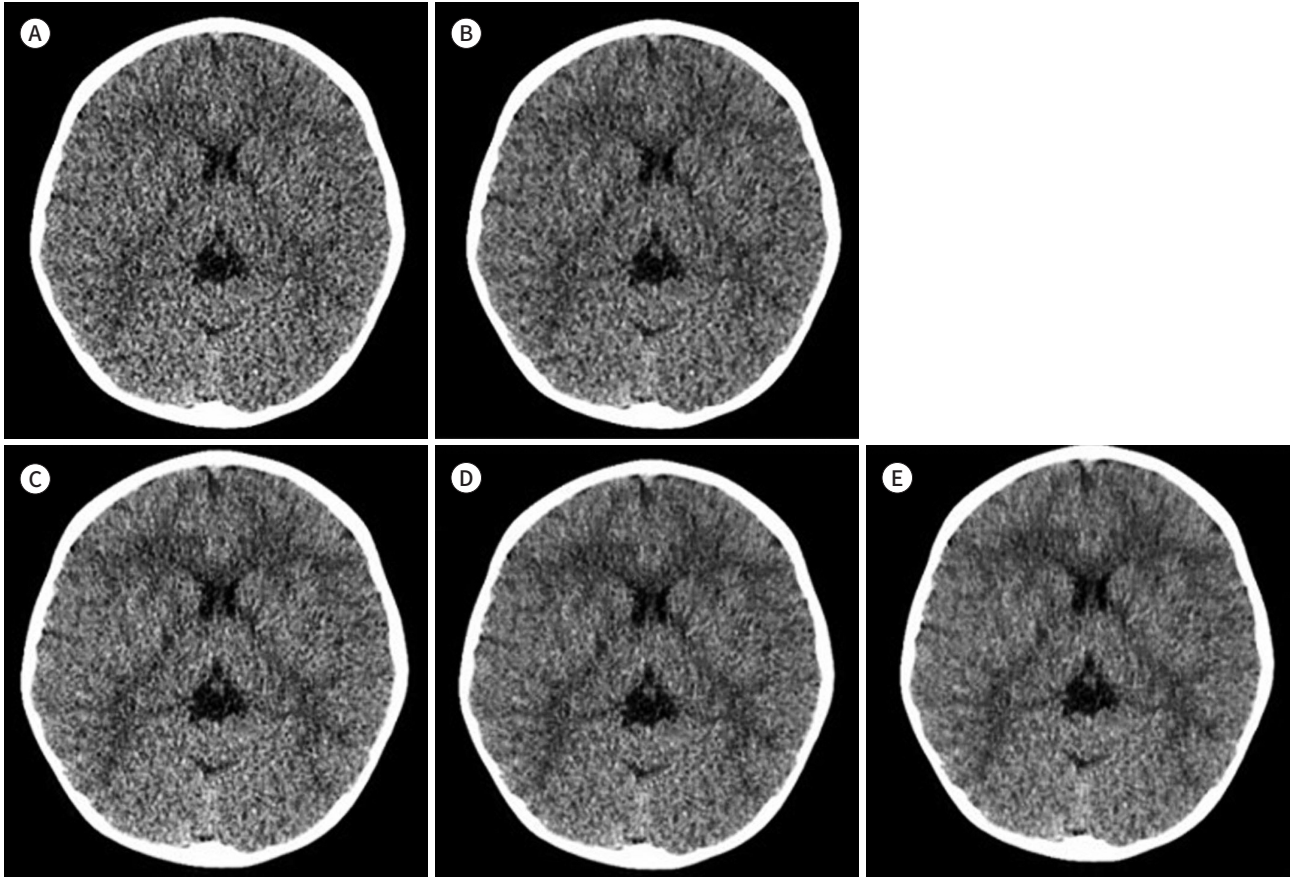
ternative choice for pediatric head CT scans in all age groups.

As the first FDA-granted DLIR platform, TrueFidelity (GE Healthcare) used a deep neural network with massive training sets for both patients and phantoms (6). Training was performed using qualified FBP datasets for the differentiation of noise from signals and suppression of noise without alteration of normal anatomical and pathological structures (6). Therefore, DLIR has no choice but to show better image quality than FBP, but it is meaningful in that it has implemented better image quality than ASiR-V, which uses other reconstruction algorithms compared with FBP.

As TrueFidelity is a recent method for reconstruction, not many reports have been published about its application. There have been only a few published reports about adaptation of this method to phantom study (8) and clinical patients with the abdomen (2) and low-dose chest study (7) for the adult age group. Although these studies were not for the same or similar locations and age groups as our study, they revealed similar trends to our study and proved that noise reduction of DLIR can be applicable. Therefore, a wider adaptation of DLIR for other body parts and age groups is required.

The previous studies about the DLIR method emphasized noting less unfamiliar image tex-

Fig. 2. Infratentorial level of the 29-month-old boy of Fig. 1.
A-G. It was reconstructed using FBP **(A),** ASiR-V **(B)**, and DLIR-low **(C)**, medium **(D)**, and high **(E)**. Increment of artifacts (arrows) of both frontal base and cerebellum when using higher grade DLIR is noted. The magnified images of artifact areas using ASiR-V **(F)** and DLIR-high **(G)** show more prominent artifacts than those using others and an increased density is noted using DLIR-high than that of others.
ASiR = adaptive statistical iterative reconstruction, DLIR = deep learning image reconstruction, FBP = filtered back projection
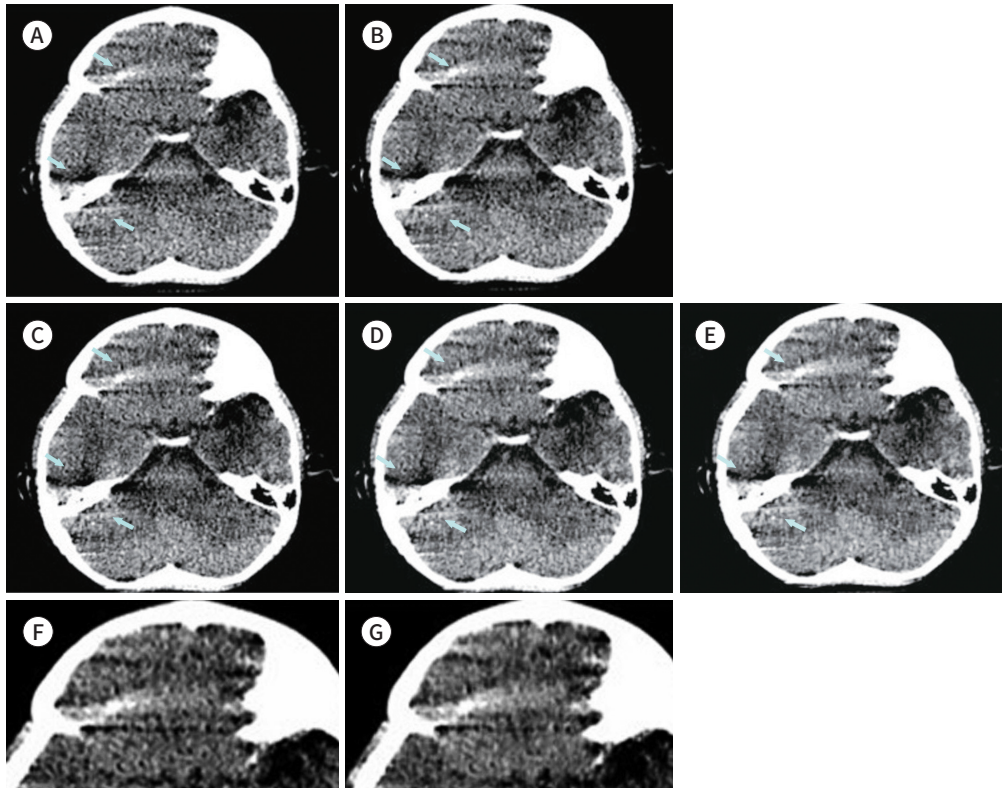
Table 5. Interobserver Agreement of Qualitative Analysis (Weighted Kappa)

	< 1 Year	1–6 Years	6–12 Years	> 12 Years
Noise	0.569	0.846	0.673	0.781
Gray matter-white matter differentiation (supratentorial level)	0.684	0.791	0.591	0.660
Gray matter-white matter differentiation (infratentorial level)	0.614	0.673	0.613	0.703
Sharpness	0.644	0.611	0.839	0.638
Artifact	0.521	0.733	0.640	0.825
Acceptability	0.531	0.642	0.768	0.655
Image texture	0.829	0.796	0.681	0.740

ture compared with conventional IR methods, called as "blotchy" or "plastic-like texture" (12). Previous studies on other vendors (1, 2, 10, 11, 15, 16) or self-developed DLIR (15) also asserted that the image texture of DLIR is more familiar than that of IR methods. Therefore, these DLIR methods can be a solution for the unfamiliar image texture of IR studies. We also concluded that the image texture of DLIR can be more familiar than that of IR methods.

However, among the parameters for qualitative image quality, artifacts revealed no remarkable difference or improvement on higher-level DLIR. It revealed worsening in age

group 3, while the level of DLIR increased. These results adversely affected lowering of the overall acceptability. Other previous study about adaptation of DLIR for pediatric chest and abdomen also mentioned that there is no significant difference in artifact reductions among the reconstruction including FBP, IR and DLIR (17). And this shortness of DLIR may be owing to the limitation of the reconstruction method itself, which differentially remove noises. It can be due to limitation in differentiating artifacts from normal anatomical structures. However, this trend was not noted in previous studies with the same reconstruction modality (2, 7, 8). Therefore, these findings can be specific to the examination of pediatric patients. Artifacts owing to movement are common in pediatric patients; however, the data utilized for training this reconstruction method are mainly adult data (6), and those motion artifacts themselves might not be included in the training set. Thus, it is possible that the differentiation of artifacts might be difficult and requires modification. And there might need more training with pediatric patients image for wider adaptation of DLIR in pediatric examinations with relatively many artifacts than adults.

Our study had some limitations. First, the study population was relatively small, and our investigation was retrospective. Second, the study was carried out at a single institution. Therefore, we consider our findings preliminary. An extremely large sample size may reject the null hypotheses with clinically negligible differences, leading to the possibility that what is insignificant may become significant (18). Third, we divided patients into four age groups; however, the division of age group may be different in other centers with different dose settings for each age group; therefore, a more standardized division is needed.

In summary, high level DLIR improves the image noise, image quality, SNR and CNR of the pediatric head CT scan images with more familiar image texture, compared with ASiR-V. Thus, the use of high level DLIR may be helpful in clinical practice as it yields lower image noise while maintaining overall image quality at equal diagnostic accuracy. Although there needs more training and modification for processing artifacts which can be more common in pediatric age group.

In conclusion, adaptation of DLIR for the pediatric head CT can significantly reduce image noise without degradation of image quality, compared with ASiR-V. For wider adaptation in pediatric CT field, there needs some modification for processing of artifacts.

Author Contributions

Conceptualization, C.H.; data curation, all authors; formal analysis, all authors; investigation, all authors; methodology, C.H.; project administration, C.H.; resources, C.H.; software, all authors; supervision, all authors; validation, L.S.M., Y.S.K.; visualization, C.H.; writing—original draft, L.N.; and writing—review & editing, C.H.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

Funding

REFERENCES

1. Singh R, Digumarthy SR, Muse VV, Kambadakone AR, Blake MA, Tabari A, et al. Image quality and lesion de-

tection on deep learning reconstruction and iterative reconstruction of submillisievert chest and abdominal CT. *AJR Am J Roentgenol* 2020;214:566-573

2. Jensen CT, Liu X, Tamm EP, Chandler AG, Sun J, Morani AC, et al. Image quality assessment of abdominal CT by use of new deep learning image reconstruction: initial experience. *AJR Am J Roentgenol* 2020;215:50-57

3. Kim JH, Kim MJ, Kim HY, Lee MJ. Radiation dose reduction and image quality in pediatric abdominal CT with kVp and mAs modulation and an iterative reconstruction technique. *Clin Imaging* 2014;38:710-714

4. Cho HH, Lee SM, You SK. Pediatric head computed tomography with advanced modeled iterative reconstruction: focus on image quality and reduction of radiation dose. *Pediatr Radiol* 2020;50:242-251

5. Kim HG, Lee HJ, Lee SK, Kim HJ, Kim MJ. Head CT: image quality improvement with ASIR-V using a reduced radiation dose protocol for children. *Eur Radiol* 2017;27:3609-3617

6. Hsieh J, Liu E, Nett B, Tang J, Thibault JB, Sahney S. *A new era of image reconstruction: TrueFidelity™ technical white paper on deep learning image reconstruction.* Milwaukee: GE Healthcare 2019

7. Kim JH, Yoon HJ, Lee E, Kim I, Cha YK, Bak SH. Validation of deep-learning image reconstruction for low-dose chest computed tomography scan: emphasis on image quality and noise. *Korean J Radiol* 2021;22:131-138

8. Greffier J, Hamard A, Pereira F, Barrau C, Pasquier H, Beregi JP, et al. Image quality and dose reduction opportunity of deep learning image reconstruction algorithm for CT: a phantom study. *Eur Radiol* 2020;30:3951-3959

9. Higaki T, Nakamura Y, Zhou J, Yu Z, Nemoto T, Tatsugami F, et al. Deep learning reconstruction at CT: phantom study of the image characteristics. *Acad Radiol* 2020;27:82-87

10. Akagi M, Nakamura Y, Higaki T, Narita K, Honda Y, Zhou J, et al. Deep learning reconstruction improves image quality of abdominal ultra-high-resolution CT. *Eur Radiol* 2019;29:6163-6171

11. Benz DC, Benetos G, Rampidis G, von Felten E, Bakula A, Sustar A, et al. Validation of deep-learning image reconstruction for coronary computed tomography angiography: impact on noise, image quality and diagnostic accuracy. *J Cardiovasc Comput Tomogr* 2020;14:444-451

12. Kaul D, Kahn J, Huizing L, Wiener E, Böning G, Renz DM, et al. Dose reduction in paediatric cranial CT via iterative reconstruction: a clinical study in 78 patients. *Clin Radiol* 2016;71:1168-1177

13. ICRP. ICRP publication 103. The 2007 recommendations of the international commission on radiological protection. Available at. https://www.icrp.org/publication.asp?id=ICRP%20Publication%20103. Published 2007. Accessed August 30, 2020

14. Deak PD, Smal Y, Kalender WA. Multisection CT protocols: sex- and age-specific conversion factors used to determine effective dose from dose-length product. *Radiology* 2010;257:158-166

15. Tatsugami F, Higaki T, Nakamura Y, Yu Z, Zhou J, Lu Y, et al. Deep learning-based image restoration algorithm for coronary CT angiography. *Eur Radiol* 2019;29:5322-5329

16. Liu J, Zhang Y, Zhao Q, Lv T, Wu W, Cai N, et al. Deep iterative reconstruction estimation (DIRE): approximate iterative reconstruction estimation for low dose CT imaging. *Phys Med Biol* 2019;64:135007

17. Yoon H, Kim J, Lim HJ, Lee MJ. Image quality assessment of pediatric chest and abdomen CT by deep learning reconstruction. *BMC Med Imaging* 2021;21:146

18. Faber J, Fonseca LM. How sample size influences research outcomes. *Dental Press J Orthod* 2014;19:27-29

소아용 두부 컴퓨터단층촬영에서 딥러닝 영상 재구성 적용: 영상 품질에 대한 고찰

이 님[1,2] · 조현혜[1]* · 이소미[3] · 유선경[4]

목적 소아 환자에서 두부 컴퓨터단층촬영(이하 CT)에 대한 딥러닝 이미지 재구성(deep learning image reconstruction; 이하 DLIR; TrueFidelity; GE Healthcare, Milwaukee, WI, USA)의 효과를 평가하고자 한다.

대상과 방법 총 126개의 소아 두부 CT 이미지를 수집했으며, adaptive statistical iterative reconstruction (이하 ASiR)-V를 사용한 반복적 재구성 및 세 가지 수준의 DLIR을 사용한 재구성을 시행하였다. 각 이미지 세트 그룹은 환자의 연령에 따라 4개의 그룹으로 구분하였으며 각 연령군의 임상 및 방사선량 관련 데이터를 검토하였다. 양적 매개 변수에는 signal to noise ratio (이하 SNR) 및 contrast to noise ratio (이하 CNR)가 포함되었으며 질적 매개 변수로 영상의 잡음(noise), 회백질의 구분 정도, 선명도, 인공물 및 수용 가능성(acceptability), 영상의 질감이 포함되었고 이에 대한 평가와 비교를 시행하였다.

결과 모든 연령 그룹의 모든 수준의 SNR 및 CNR은 높은 수준의 DLIR 사용 시 증가하였다. ASiR-V와 비교했을 때 높은 수준의 DLIR은 SNR 및 CNR이 개선되었다($p < 0.05$). 그리고 DLIR의 수준이 증가될수록 순차적인 잡음 감소, 회백질 구분 개선, 선명도 개선이 나타났다. 이러한 변수들에서 높은 수준의 DLIR 사용 시 ASiR-V와 유사한 정도의 수치가 측정되었다. 인공물과 수용 가능성의 경우에 적용된 DLIR 수준 간에 큰 차이를 보이지 않았다.

결론 소아 두부 CT에 고수준 DLIR을 적용하면 영상의 노이즈를 줄일 수 있으나 인공물 처리에 대한 개선이 필요하다.

[1]이화여자대학교 의과대학 이화여자대학교 목동병원 영상의학과,
[2]연세대학교 의과대학 세브란스병원 영상의학과,
[3]경북대학교 의과대학 칠곡경북대학교병원 영상의학과,
[4]충남대학교 의과대학 충남대학교병원 영상의학과