Original Research

Improved image quality in contrast-enhanced 3D-T1 weighted sequence by compressed sensing-based deep-learning reconstruction for the evaluation of head and neck

Abstract

Purpose:

To assess the utility of deep learning (DL)-based image reconstruction with the combination of compressed sensing (CS) denoising cycle by comparing images reconstructed by conventional CS-based method without DL in fat-suppressed (Fs)-contrast enhanced (CE) three-dimensional (3D) T1-weighted images (T1WIs) of the head and neck.

Materials and Methods:

We retrospectively analyzed the cases of 39 patients who had undergone head and neck Fs-CE 3D T1WI applying reconstructions based on conventional CS and CS augmented by DL, respectively. In the qualitative assessment, we evaluated overall image quality, visualization of anatomical structures, degree of artifacts, lesion conspicuity, and lesion edge sharpness based on a five-point system. In the quantitative assessment, we calculated the signal-to-noise ratios (SNRs) of the lesion and the posterior neck muscle and the contrast-to-noise ratio (CNR) between the lesion and the adjacent muscle.

Results:

For all items of the qualitative analysis, significantly higher scores were awarded to images with DL-based reconstruction ($p<0.001$). In the quantitative analysis, DL-based reconstruction resulted in significantly higher values for both the SNR of lesions ($p<0.001$) and posterior neck muscles ($p<0.001$). Significantly higher CNRs were also observed in images with DL-based reconstruction ($p<0.001$).

Conclusion:

DL-based image reconstruction integrating into the CS-based denoising cycle offered superior image quality compared to the conventional CS method. This technique will be useful for the assessment of patients with head and neck disease.

Keywords;

Head and neck; contrast-enhanced 3D-T1WI; compressed sensing; deep learning; reconstruction

Introduction

Magnetic resonance imaging (MRI) is one of the useful imaging modalities for the evaluation of the head and neck as it offers clear soft tissue contrast without radiation exposure. With its large scanning range and high spatial resolution, fat-suppressed (Fs) contrast-enhanced (CE) three-dimensional (3D) T1-weighted imaging (T1WI) enables particularly clear visualization of diverse head and neck structures. In most patients with head and neck tumors, both the primary tumor and metastatic lymph nodes can be successfully evaluated for the determination of treatment methods using Fs-CE 3D T1WI as important supportive information [1–4]. However, as is well known, image acquisition of the head and neck is often challenging because of B0 and B1 field inhomogeneity from complicated anatomical shapes and the presence of air or metallic substances [5]. In addition, because the anatomical structures in the head and neck are small with complex arrangements within a narrow space, high-spatial-resolution imaging is needed. Given these factors, the image signal to noise ratio (SNR) for head-and-neck MRI tended to be insufficient for the evaluation. Although longer scanning times with larger number of NEX (i.e., multiple signal averaging) may improve the level of SNR, such long scanning times can cause the patient discomfort and even pain; also, long scanning times increase the likelihood of motion artifacts that decrease image quality.

Parallel imaging is one of the effective techniques for fast scanning and is expected to provide higher SNRs when combined with multiple signal averaging. However, the large number of reduction factors in parallel imaging typically concentrates the image noise at the center of the image, where a high geometry factor (g-factor) is present. In the past decade, a compressed sensing (CS) algorithm has been introduced as a fast acquisition method using random undersampling and iterative image denoising with a wavelet transform filter, and its utility has been demonstrated, especially in 3D sequence design [6]. This technique is considered clinically useful, including in head-and-neck MRI, for fast acquisition that maintains sufficient image quality [4,7,8]. However, since the denoising process by CS is performed basically by an algorithm-based method (i.e., wavelet transform), the variety of noise that can be handled by CS is somewhat limited. More recently, deep learning (DL)-based image reconstruction techniques, particularly using convolutional neural networks (CNNs), have shown promising results for accelerating the data acquisition process of MRI [9,10]. Methods based on deep learning image reconstruction schemes are expected to create superb-quality images from undersampled *k*-space data with shorter acquisition time for head-and-neck Fs-CE 3D-T1WI.

The aim of this study was to assess the utility of DL-based image reconstruction in comparison with that of images reconstructed by the conventional CS-based technique in 3D high-spatial-resolution Fs-CE T1WI. For this purpose, we focused on the advanced DL reconstruction technique of model-based reconstruction using Adaptive-CS-Net [11–14]. In this technique, the DL architecture of Adaptive-CS-Net is embedded in the image-reconstruction process of the CS sensitivity-encoding (SENSE)-based denoising cycle. We hypothesized that using DL to process the large amount of image data would enable more efficient denoising. As a comparison, we used the conventional CS-SENSE-based reconstruction scheme without DL.

Materials and Methods

Patients

The protocol of this retrospective study was approved by our institutional review board, and written informed consent was waived. We selected the cases of 40 patients who were referred to our hospital for head-and-neck evaluation and who underwent MR scanning during the period from April 2022 to January 2023 with the following inclusion criteria: (1) the patient was scanned on a specific MR scanner (see below); (2) MRI sequences including Fs-CE 3D-based T1WI were performed; and (3) image data from both DL- and CS-SENSE-based reconstruction were available. Among them, 1 patient was excluded due to a failure to save the raw data of Fs-CE 3D-T1WI for image reconstruction. Finally, a total of 39 patients were considered eligible for the current study. Twenty-four patients received MR scanning for the initial diagnosis of a head-and-neck tumor. The other 15 patients received MR scanning for evaluation after the treatment of their head-and-neck malignancy as a follow-up examination.

Imaging parameters

All scanning was performed using a 3.0-Tesla MR unit (Ingenia Elition; Philips Healthcare, Best, Netherlands) with a 16-channel neurovascular coil. MRI including the high-spatial-resolution Fs-CE 3D T1WI with a short scanning time was performed to evaluate the whole volume of the head and neck (from the aortic arch to the skull base). The Fs-CE 3D-T1WI data were acquired after an injection of 0.2 mmol gadolinium/kg

(Gadobutrol, Gadovist, Bayer Pharma AG, Berlin, Germany). Image acquisition was performed using a 3D-T1 fast-field echo (T1-FFE) sequence. The MR parameters for the data acquisition were as follows: TR 5.9 ms, TE 1.2 ms, flip angle 8 degree, acquired matrix 300 × 300 (reconstructed matrix, 640 × 640), field of view (FOV) 240 × 240 mm, pixel size 0.375 × 0.375 mm, slice thickness 0.8 mm, slice pitch 0.4 mm (total 500 slices), reduction factor 12, scanning time 1 min 36 s.

Data processing

The DL-based reconstruction used in the present study incorporates technology from vendor prototypes; all image processing was conducted within the MR console. We utilized a reconstruction model incorporating the DL architecture of Adaptive-CS-Net into the CS-based iterative denoising cycle, described as model-based type DL reconstruction. The Adaptive-CS-Network is constructed based-on the U-net shape structure with a soft thresholding function. Specifically, Adaptive-CS-Net replaces the wavelet transform with the sparsifying transform in the CS denoising cycle. Furthermore, datasets associated with image reconstruction, containing domain-specific knowledge such as data consistency, phase behavior, and background information, were individually incorporated into the reconstruction process to mitigate the risk of inappropriate denoising [11]. Significant quality improvement was expected from this advanced denoising method combining the Adaptive-CS-Net-based sparsifying approach with the image reconstruction of a CS denoising cycle.

For the comparison, we used the conventional CS-SENSE technique that combines the CS algorithm and the parallel imaging of SENSE processing [4]. From the same raw-image dataset of Fs-CE 3D T1WI, image reconstruction was performed by DL-based reconstruction and the conventional CS-SENSE model; thus, two image datasets were output from the same original raw image data. Finally, two types of Fs-CE 3D T1WIs processed with DL-based and CS-SENSE reconstruction were output. Axial-based reconstruction images were used for further evaluation.

Image analysis

Qualitative image assessment was performed in all patients. Two board-certified radiologists with 17 and 8 years of experience, respectively, evaluated images of the axial Fs-CE 3D T1WIs with DL-based and CS-SENSE-based reconstruction in blinded fashion; all images were randomized in terms of patient order and between DL-based and CS-SENSE-based reconstruction images in this qualitative image assessment. The overall image quality, visualizations of anatomical structures, and degree of artifacts were visually evaluated in all patient images. In patients with a primary lesion, lesion conspicuity and lesion edge sharpness were also evaluated. Each evaluation was conducted based-on the following five-point (1-5) grading system:

1: very poor image quality, unevaluable for clinical diagnostic use.

2: poor image quality, but possible for diagnostic use.

3: moderate image quality, acceptable for diagnostic use.

4: good image quality, appropriate for diagnostic use.

5: excellent image quality, no limitations for diagnostic use.

Quantitative assessment was performed in patients with the primary tumor lesion. The signal intensity of the primary lesion was measured by placing a polygonal region of interest (ROI) on the axial slice. If the tumor extended into two or more slices, the slice in which the largest area of tumor was depicted was selected. Any area that was suspected to be necrotic, a cystic lesion, or a vessel component was carefully excluded from the ROI. Additionally, a round ROI (1.0 cm dia.) was also placed on the posterior neck muscle, while avoiding the signal intensity of noise or artifact. Both polygonal ROI delineation of the tumor and round ROI placement on the muscle were first performed on CS-SENSE images; thereafter, the ROI was copied onto the DL-based image, with the same ROIs placed at the same location. If bulk noise, which might cause measurement bias, was present in any of the ROIs, the ROI shape and location were modified by referring to both CS-SENSE and DL-based images to exclude the bulk noise. The SNRs of the lesion and the muscle were calculated as the mean signal in the ROI divided by the standard deviation (SD) of the signal in the ROI [4]. For the contrast-to-noise ratio (CNR) an additional ROI was manually drawn on the muscle adjacent to the tumor lesion, and the CNR was calculated as the difference between the mean signal of the tumor lesion and the adjacent muscle ROI divided by the SD of the adjacent muscle's ROI. Quantitative procedures were performed by an experienced radiologist with 17 years of experience.

Statistical analysis

We used kappa statistics to determine interobserver agreement for qualitative analyses (Kappa score=0.00–0.20, poor; Kappa score=0.21–0.40, fair; Kappa score=0.41–0.60, moderate; Kappa score=0.61–0.80, good; Kappa score=0.81–1.00, excellent). Qualitative image scores were compared between the Fs-CE 3D-T1WI with DL-based and CS-SENSE-based reconstruction using the Wilcoxon signed-rank test. SNRs of DL-based and CS-SENSE-based reconstructions of post contrast-enhanced T1W1 were compared for both the lesion and the muscle using a paired t-test after confirmation of normal data distribution by the Shapiro-Wilk test. P values < 0.05 were considered statistically significant. SPSS software (IBM, Armonk, NY) was used for all statistical analyses.

Results

The 39 patients included 29 males and 10 females with a median age of 60 yrs (range 37–78 yrs). Among these patients, 24 patients had pre-treatment primary head and neck tumors as follows: pharyngeal squamous cell carcinoma (SCC) (n=10), oral cavity SCC (n=4), nasal or sinonasal cavity SCC (n=4), and parotid gland tumor (n=6). The other 15 patients were in follow-up status after definitive treatment for their primary lesion.

Interobserver agreements in qualitative analysis between two board-certified radiologists were mostly good (Kappa scores of 0.58–0.65). In all evaluation items in the qualitative analysis, significant differences were observed between images with DL-based and CS-SENSE-based reconstruction. Details of qualitative analysis are presented in Table 1.

In the quantitative analysis, significant improvements in DL-reconstructed images compared to CS-SENSE-based images were observed for the SNRs of both the lesion ($p<0.001$) and the posterior neck muscle ($p<0.001$). A significant difference was also observed in the CNR value ($p<0.001$). Details of the quantitative analyses are presented in Table 2.

Representative cases are presented in Figure 1 and Figure 2.

Discussion

Our results indicated that DL-based image reconstruction embedded into the CS denoising cycle successfully demonstrated improved image quality in both qualitative and quantitative assessments of the Fs-CE 3D-T1 weighted sequence for evaluation of head-and-neck disease. Although the CS technique achieved sufficient or non-inferior image quality with great reduction of scanning time in previous studies focused on the head and neck [4,7,8], the DL-based image reconstruction conducted in the present study showed better image quality than conventional CS-based images. This technique, which integrates the CNN of Adaptive-CS-Net into the CS denoising cycle, may be a useful tool for assessment in patients with head and neck disease for its high-SNR, high-spatial-resolution images within a short scanning time. The short scanning time may contribute directly to the patient's comfort in daily clinical practice.

Generally, images in the head and neck require high spatial resolution due to this region's delicate and complex anatomical structures, as well as high SNR for clear visibility of the wide variety of anatomical tissue and potential lesions. In MR image acquisition, full sampling of the *k*-space is considered to provide the highest SNR; however, acquisition with both high SNR and high spatial resolution requires prohibitive scanning time. The technique of parallel imaging, which is widely available in most scanners for routine clinical MRI, is a popular fast acquisition method to reduce scanning time; however, higher reduction factor acquisition typically concentrates the bulk of the noise in the center of the image, where the g-factor becomes high. CS was introduced to reduce the scanning time while maintaining the image SNR using an effective denoising algorithm of wavelet transformation [6]. Moreover, CS-SENSE has been reported as a powerful algorithm to preserve image quality by effectively combining variable-density *k*-space undersampling, parallel-imaging-based image unfolding, and a CS-based denoising cycle [4,15,16]. However, the quality of the resulting images is not necessarily preserved, due to insufficient denoising with the CS-SENSE algorithm, when a high reduction factor such as 12 was used; in fact, the overall image quality in several CS-SENSE-reconstruction cohorts in the current study was scored as "poor" or "very poor".

More recently, DL-based reconstruction methods for MRI have spread worldwide and have achieved superior image noise reduction [9,10,17,18]. Neural networks in DL algorithms are capable of higher-order processing as denoising, and the use of this technique is likely to increase the range of applications for noise reduction compared to conventional algorithms such as parallel imaging or CS. As a well-known DL-based reconstruction technique, the post-processing type has become rather popular and has been performed in several previous studies in MRI of 3D sequences [19,20]. These studies demonstrated well-preserved image quality using an undersampled dataset with a reduction factor of around 3–6, whereas the present study achieved sufficient image quality in most cases with a higher setting of the reduction factor of 12. The deep learning-based image denoising method utilized in the current study, namely embedding Adaptive-CS-Net into the CS-based denoising cycle, effectively improved the image quality in Fs-CE 3D T1WI of the head and neck. During this denoising process, large amounts of MR signal information from image acquisition and processing could be integrated for more accurate denoising [11]. This denoising method might have the potential to achieve superior denoising effectiveness compared to the post-processing type of DL-based denoising, in which only output images are analyzed. The present study successfully accomplished a sufficient level of denoising even under the excessively high reduction-factor setting of 12.

From a future perspective, it is expected that investigations into DL-based reconstruction will demonstrate its clinical utility for the benefit of patients. Even with a very high reduction factor, DL-based reconstruction can successfully achieve high SNR images with minimal motion artifacts due to the short scanning time; this may result in a more detailed evaluation of head and neck lesions. Specifically, improvements in depicting the anatomical extent of tumor lesions are anticipated to contribute to a more detailed assessment of the resection range. Furthermore, a more detailed evaluation of intra-tumoral characteristics, facilitated by improved image quality through DL-based reconstruction, may lead to an increase in the accuracy of differential diagnoses. The relationship between these advantages and the improvement of patient prognosis should be investigated to underscore the clinical utility of the DL-based reconstruction technique in the near future.

The current study has several limitations. First, the number of patients was small because the current study was conducted at a single institution. Therefore, results of the current study were regarded as preliminary. Second, the current study evaluated a variety of lesions, potentially overlooking any trends specific to lesion location. Further subgroup analysis of a larger total sample divided according to primary site is needed to address this limitation.

Conclusion

The present DL-based image reconstruction strategy, which embeds Adaptive-CS-Net into the CS-based denoising cycle, effectively provided high-spatial-resolution images of sufficient image quality, even within a short scanning time. Image quality provided by this technique demonstrated clear superiority to images derived from the CS-SENSE-based technique. This DL reconstruction technique will be useful for assessment of patients with head-and-neck disease for whom high image quality in a short scanning time is required.

Acknowledgments

This work was supported by a grant from the Japan Society for the Promotion of Science (JSPS) KAKENHI, ID; JP21K07558.

References

[1] Koyfman SA, Ismaila N, Crook D, D'Cruz A, Rodriguez CP, Sher DJ, et al. Management of the Neck in Squamous Cell Carcinoma of the Oral Cavity and Oropharynx: ASCO Clinical Practice Guideline. J Clin Oncol 2019;37:1753–74.

[2] Hiyama T, Kuno H, Nagaki T, Sekiya K, Oda S, Fujii S, et al. Extra-nodal extension in head and neck cancer: how radiologists can help staging and treatment planning. Jpn J Radiol 2020;38:489–506.

[3] Hiyama T, Sekiya K, Kuno H, Oda S, Kusumoto M, Minami M, et al. Imaging of extracranial head and neck lesions in cancer patients: a symptom-based approach. Jpn J Radiol 2019;37:354–70.

[4] Takumi K, Nagano H, Nakanosono R, Kumagae Y, Fukukura Y, Yoshiura T. Combined signal averaging and compressed sensing: impact on quality of contrast-enhanced fat-suppressed 3D turbo field-echo imaging for pharyngolaryngeal squamous cell carcinoma. Neuroradiology 2020;62:1293–9.

[5] Touska P, Connor SEJ. Recent advances in MRI of the head and neck, skull base and cranial nerves: new and evolving sequences, analyses and clinical applications. Br J Radiol 2019;92:20190513.

[6] Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. Magn Reson Med 2007;58:1182–95.

[7] Tomita H, Deguchi Y, Fukuchi H, Fujikawa A, Kurihara Y, Kitsukawa K, et al. Combination of compressed sensing and parallel imaging for T2-weighted imaging of the oral cavity in healthy volunteers: comparison with parallel imaging. Eur Radiol 2021;31:6305–11.

[8] Kami Y, Chikui T, Togao O, Kawano S, Fujii S, Ooga M, et al. Usefulness of reconstructed images of Gd-enhanced 3D gradient echo sequences with compressed sensing for mandibular cancer diagnosis: comparison with CT images and histopathological findings. Eur Radiol 2023;33:845–53.

[9] Higaki T, Nakamura Y, Tatsugami F, Nakaura T, Awai K. Improvement of image quality at CT and MRI using deep learning. Jpn J Radiol 2019;37:73–80.

[10] Lin DJ, Johnson PM, Knoll F, Lui YW. Artificial Intelligence for MR Image Reconstruction: An Overview for Clinicians. J Magn Reson Imaging 2021;53:1015–28.

[11] Pezzotti N, Yousefi S, Elmahdy MS, Van Gemert JHF, Schuelke C, Doneva M, et al. An Adaptive Intelligence Algorithm for Undersampled Knee MRI

Reconstruction. IEEE Access 2020;8:204825–38.

[12] Foreman SC, Neumann J, Han J, Harrasser N, Weiss K, Peeters JM, et al. Deep learning-based acceleration of Compressed Sense MR imaging of the ankle. Eur Radiol 2022;32:8376–85.

[13] Wu X, Tang L, Li W, He S, Yue X, Peng P, et al. Feasibility of accelerated non-contrast-enhanced whole-heart bSSFP coronary MR angiography by deep learning-constrained compressed sensing. Eur Radiol 2023. https://doi.org/10.1007/s00330-023-09740-8.

[14] Yang F, Pan X, Zhu K, Xiao Y, Yue X, Peng P, et al. Accelerated 3D high-resolution T2-weighted breast MRI with deep learning constrained compressed sensing, comparison with conventional T2-weighted sequence on 3.0 T. Eur J Radiol 2022;156:110562.

[15] Vranic JE, Cross NM, Wang Y, Hippe DS, de Weerdt E, Mossa-Basha M. Compressed Sensing-Sensitivity Encoding (CS-SENSE) Accelerated Brain Imaging: Reduced Scan Time without Reduced Image Quality. AJNR Am J Neuroradiol 2019;40:92–8.

[16] Morita K, Nakaura T, Maruyama N, Iyama Y, Oda S, Utsunomiya D, et al. Hybrid of Compressed Sensing and Parallel Imaging Applied to Three-dimensional Isotropic T2-weighted Turbo Spin-echo MR Imaging of the Lumbar Spine. Magn Reson Med Sci 2020;19:48–55.

[17] Zeng G, Guo Y, Zhan J, Wang Z, Lai Z, Du X, et al. A review on deep learning MRI reconstruction without fully sampled k-space. BMC Med Imaging 2021;21:195.

[18] Chen Z, Pawar K, Ekanayake M, Pain C, Zhong S, Egan GF. Deep Learning for

Image Enhancement and Correction in Magnetic Resonance Imaging-State-of-the-Art and Challenges. J Digit Imaging 2023;36:204–30.

[19] Zhou Z, Chen S, Balu N, Chu B, Zhao X, Sun J, et al. Neural network enhanced 3D turbo spin echo for MR intracranial vessel wall imaging. Magn Reson Imaging 2021;78:7–17.

[20] Son JH, Lee Y, Lee H-J, Lee J, Kim H, Lebel MR. LAVA HyperSense and deep-learning reconstruction for near-isotropic (3D) enhanced magnetic resonance enterography in patients with Crohn's disease: utility in noise reduction and image quality improvement. Diagn Interv Radiol 2023;29:437–49.

Table and Figure legends

Table 1. Results of the qualitative assessment by radiologists.

	Reader 1			Reader 2			Kappa-score
	CS-SENSE	DL	p-value	CS-SENSE	DL	p-value	
Overall image quality	2.84±0.92	4.23±0.58	<0.001	2.84±0.59	3.97±0.54	<0.001	0.61
Visualizations of anatomical structures	2.52±0.72	4.07±0.67	<0.001	2.78±0.47	3.89±0.50	<0.001	0.6
Degree of the artifact	2.57±1.03	3.65±0.66	<0.001	2.78±0.57	3.79±0.52	<0.001	0.58
Lesion conspicuity	2.89±1.10	4.14±0.70	<0.001	2.89±0.62	4.10±0.62	<0.001	0.63
Lesion edge sharpness	2.71±1.15	4.00±0.72	<0.001	2.85±0.70	3.96±0.63	<0.001	0.65

Table 1 footnote: Data are mean ± standard deviation, CS-SENSE: compressed-sensing sensitivity-encoding, DL: deep learning.

Table 2. Results of the quantitative SNR and CNR assessments.

		CS-SENSE	DL	p-value
SNR	Lesion	9.7±2.0	13.5±2.1	<0.001
	Muscle	8.1±1.3	10.1±1.5	<0.001
CNR	Lesion to adjacent muscle	6.7±2.5	7.8±2.8	<0.001

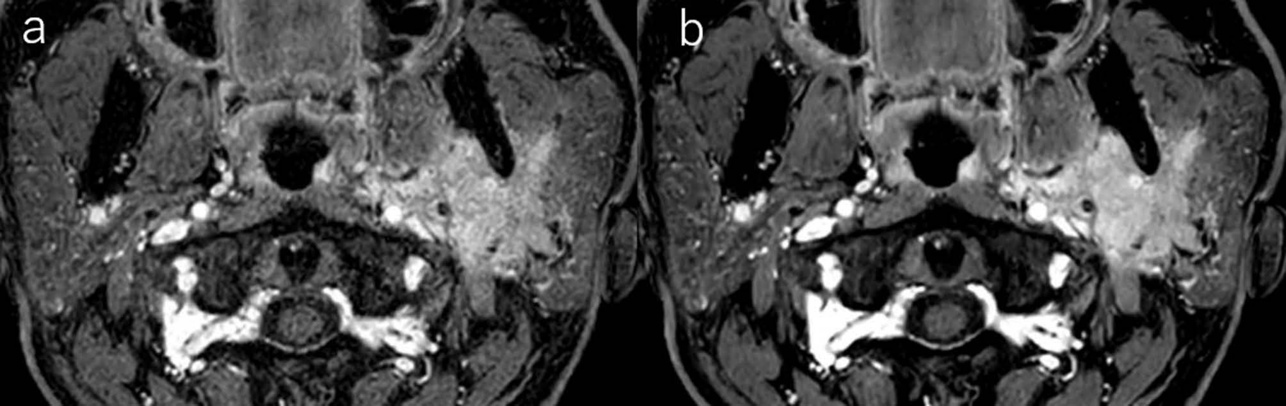
Table 2 footnote: Data are mean ± standard deviation. CS-SENSE: compressed-sensing sensitivity-encoding, DL: deep learning, SNR: signal to noise ratio, CNR: contrast to noise ratio

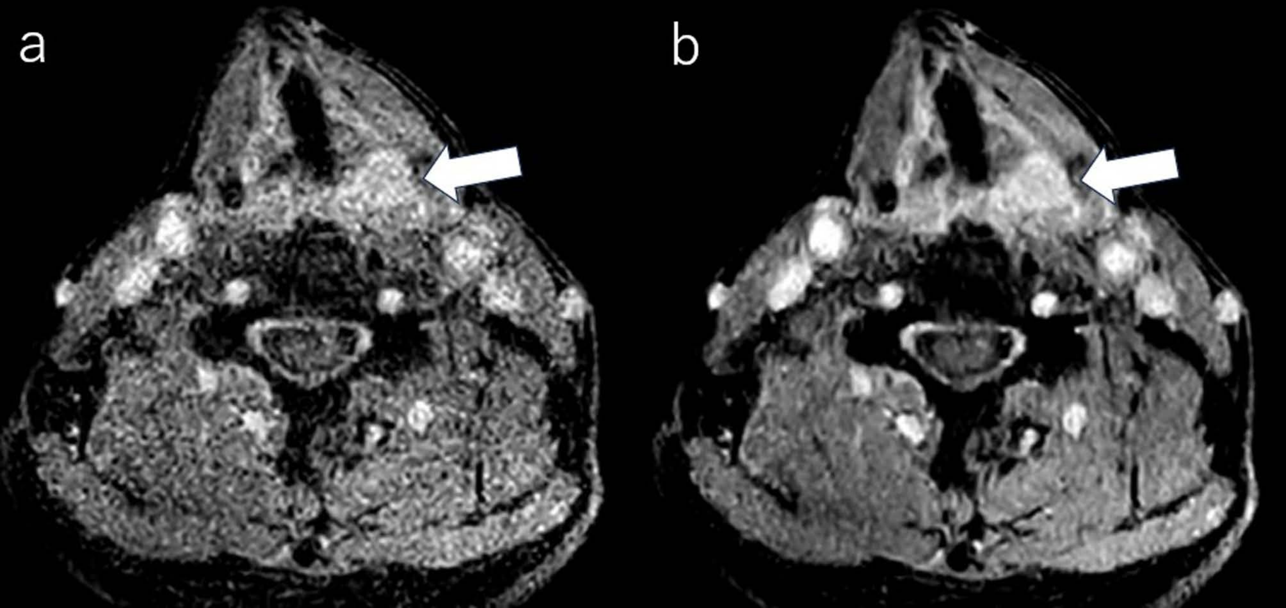
Figure 1. Patient with parotid gland tumor extending into the surrounding tissue.

In Fs-CE 3D T1WI with CS-SENSE-based reconstruction (a), the primary lesion and surrounding tissue was observed to be noisy overall: potentially an obstacle for image interpretation. By contrast, overall clearer image contrast and less noise was observed in Fs-CE 3D T1WI with DL-based reconstruction (b).

Figure 2. Patient with hypopharyngeal SCC.

Strong, distinct noise was observed throughout the image in Fs-CE 3D T1WI with CS-SENSE-based reconstruction (a), making the image inappropriate for evaluation of the tumor (a: arrow) and surrounding tissue. The lesion was clearly observed with less image noise in Fs-CE 3D T1WI with DL-based reconstruction (b: arrow).


a
b


a
b