



This information is current as
of August 24, 2025.

Synthesizing Contrast-Enhanced MR Images from Noncontrast MR Images Using Deep Learning

Gowtham Murugesan, Fang F. Yu, Michael Achilleos, John
DeBevits, Sahil Nalawade, Chandan Ganesh, Ben Wagner,
Ananth J Madhuranthakam and Joseph A. Maldjian

AJNR Am J Neuroradiol 2024, 45 (3) 312-319

doi: <https://doi.org/10.3174/ajnr.A8107>

<http://www.ajnr.org/content/45/3/312>

Synthesizing Contrast-Enhanced MR Images from Noncontrast MR Images Using Deep Learning

Gowtham Murugesan,  Fang F. Yu,  Michael Achilleos,  John DeBevits,  Sahil Nalawade,  Chandan Ganesh,  Ben Wagner,  Ananth J Madhuranthakam, and  Joseph A. Maldjian



ABSTRACT

BACKGROUND AND PURPOSE: Recent developments in deep learning methods offer a potential solution to the need for alternative imaging methods due to concerns about the toxicity of gadolinium-based contrast agents. The purpose of the study was to synthesize virtual gadolinium contrast-enhanced T1-weighted MR images from noncontrast multiparametric MR images in patients with primary brain tumors by using deep learning.

MATERIALS AND METHODS: We trained and validated a deep learning network by using MR images from 335 subjects in the Brain Tumor Segmentation Challenge 2019 training data set. A held out set of 125 subjects from the Brain Tumor Segmentation Challenge 2019 validation data set was used to test the generalization of the model. A residual inception DenseNet network, called T1c-ET, was developed and trained to simultaneously synthesize virtual contrast-enhanced T1-weighted (vT1c) images and segment the enhancing portions of the tumor. Three expert neuroradiologists independently scored the synthesized vT1c images by using a 3-point Likert scale, evaluating image quality and contrast enhancement against ground truth T1c images (1 = poor, 2 = good, 3 = excellent).

RESULTS: The synthesized vT1c images achieved structural similarity index, peak signal-to-noise ratio, and normalized mean square error scores of 0.91, 64.35, and 0.03, respectively. There was moderate interobserver agreement between the 3 raters, regarding the algorithm's performance in predicting contrast enhancement, with a Fleiss kappa value of 0.61. Our model was able to accurately predict contrast enhancement in 88.8% of the cases (scores of 2 to 3 on the 3-point scale).

CONCLUSIONS: We developed a novel deep learning architecture to synthesize virtual postcontrast enhancement by using only conventional noncontrast brain MR images. Our results demonstrate the potential of deep learning methods to reduce the need for gadolinium contrast in the evaluation of primary brain tumors.

ABBREVIATIONS: BraTS = Brain Tumor Segmentation Benchmark; ET = enhancing tumor; GBCA = gadolinium-based contrast agent; MSE = mean squared error; NMSE = normalized mean squared error; PSNR = peak signal-to-noise ratio; RID = Residual Inception DenseNet; SFL = spatial frequency loss; SPL = structural perception loss; SSIM = structural similarity index; T2w = T2-weighted; vT1c = virtual contrast-enhanced T1-weighted; WT = whole tumor

Structural MR imaging offers superior soft tissue contrast compared with other imaging modalities, and plays a crucial role in the evaluation of brain tumors by providing information about lesion location as well as morphologic features such as necrosis, the extent of tumor spread, and the associated mass effect on surrounding brain parenchyma. The administration of

intravenous gadolinium-based contrast agents (GBCAs) shortens T1 relaxation times and increases tissue contrast by accentuating areas where contrast agents have leaked through the blood-brain barrier into the interstitium. This blood-brain barrier breakdown is a feature of certain brain tumors, including high-grade gliomas, and can serve as an important tool for diagnosis and assessment of a treatment response.¹

GBCAs have been used for decades in MR imaging and have historically been considered safe for patients with normal renal function.¹ It is well-known that there is a risk of nephrogenic systemic fibrosis associated with GBCA administration in patients with renal impairment, particularly when linear conjugates of gadolinium are used.² Moreover, recent studies have shown gadolinium deposits in tissues throughout the body, even in the setting of normal renal function, which has raised additional concerns about

Received January 27, 2022; accepted after revision December 1, 2023.

From the Department of Radiology, University of Texas Southwestern Medical Center, Dallas, Texas.

Funding Information: Support for this research was provided by NCI U01CA207091 (A.J.M., J.A.M.) and R01CA260705 (J.A.M.).

Please address correspondence to Fang Frank Yu, MD, UT Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390; e-mail: frankf.yu@utsouthwestern.edu; @FrankFYuMD1



Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A8107>

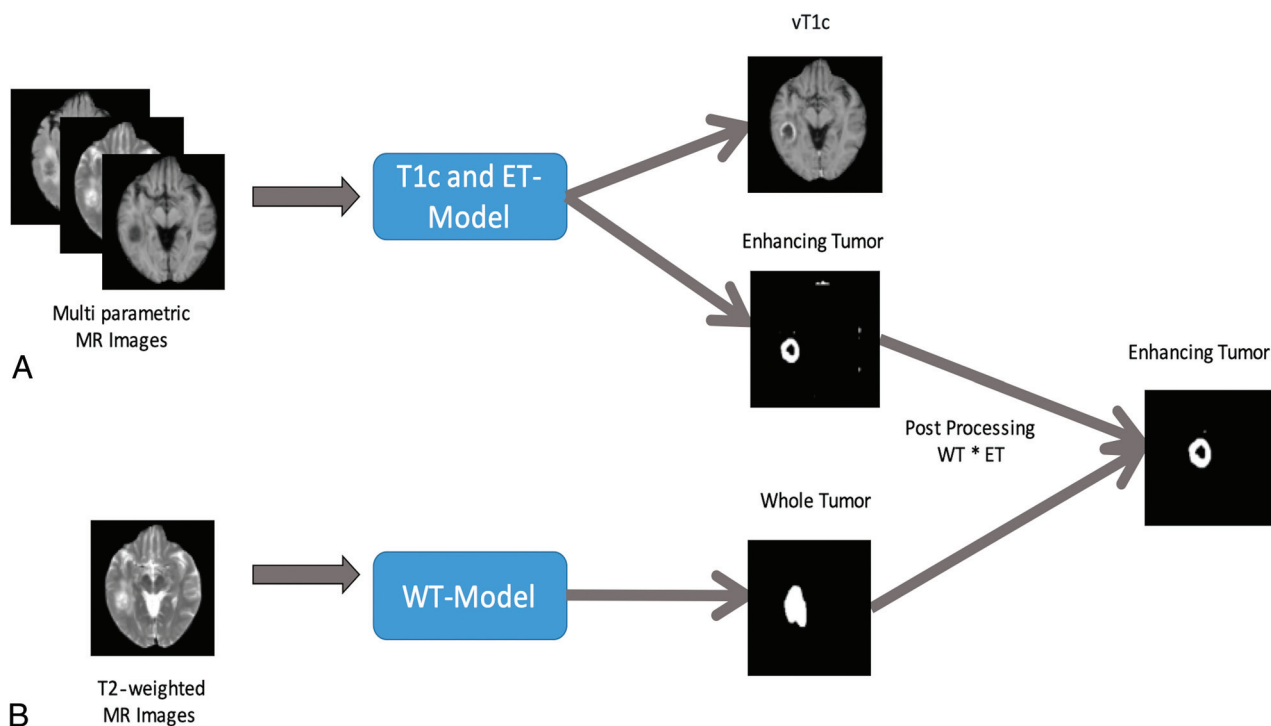


FIG 1. Residual inception DenseNet (RID). A, RID model for virtual contrast enhancement (vT1c prediction) and enhancing tumor (ET) segmentation. B, RID model for whole tumor (WT) segmentation.

the long-term safety of these agents.³ Within the brain, persistent increased signal intensity on T1-weighted (T1w) MR images has been reported within the dentate nucleus and globus pallidus following the prior administration of both linear and macrocyclic GBCAs.

Because of these concerns about the toxicity of gadolinium, there has been growing interest in alternative approaches to contrast-enhanced MR imaging. Examples include manganese-based contrast agents⁴ as well as noncontrast techniques, such as arterial spin-labeling⁵ and chemical exchange saturation transfer.⁶ Recent developments in deep learning algorithms have shown promise in image synthesis and reconstruction. The main goal of this study is to investigate the potential of deep learning methods to simulate contrast enhancement within brain gliomas by using a limited set of standard clinical noncontrast MR images. Our contributions in this work are 3-fold. First, we developed a novel deep learning network to demonstrate the ability of deep learning to synthesize virtual contrast-enhanced T1w images (vT1c) by using only noncontrast FLAIR, T1w, and T2-weighted (T2w) images. Second, we utilized imaging data from different scanners at multiple sites to train the model and evaluated its performance in predicting gadolinium enhancement by using quantitative and qualitative metrics. Third, we analyzed the contribution of each set of input MR images in synthesizing the vT1c image to gain insights regarding the further optimization and streamlining of the MR imaging protocol for clinical application.

MATERIALS AND METHODS

Data and Preprocessing

The multimodal Brain Tumor Segmentation Benchmark (BraTS) data set provides a general platform for developing deep learning

models.⁷ The BraTS 2019 data set used in our study consists of MR imaging data from 460 patients with gliomas, acquired from multiple institutions,⁸ including the University of Pennsylvania; MD Anderson Cancer Center; Washington University School of Medicine in St. Louis; and Tata Memorial Centre in India. The data set has a wide variation in imaging protocols and acquisition parameters. All subjects had precontrast T1w, T2w, and FLAIR as well as postcontrast T1c images. From this set, a single-fold training split of 335 subjects, including 259 high-grade glioma subjects and 76 low-grade glioma subjects, were used for training, while 125 subjects were held out for testing. The training data set was further randomly split into 300 and 35 subjects for the training and in-training validation of the model, respectively.

Data Preprocessing. The standard preprocessing steps performed by BraTS included coregistration to an anatomic brain template,⁹ resampling to isotropic resolution (1 mm³), and skull stripping.¹⁰ Additionally, we performed N4 bias field correction¹¹ to remove radiofrequency inhomogeneity and normalized to zero mean and unit variance.

Network Architecture

A schematic of our proposed network architecture is shown in Fig 1. The residual inception DenseNet (RID) network was first proposed and developed by Khened et al¹² for cardiac segmentation. Our implementation of the RID network incorporated slight modifications in Keras with a TensorFlow backend (Fig 2). In the DenseNet architecture, the GPU memory footprint increases as the feature maps and spatial resolution increases. The skip connections from the down-sampling path to the up-sampling path used elementwise addition, instead of the concatenation operation

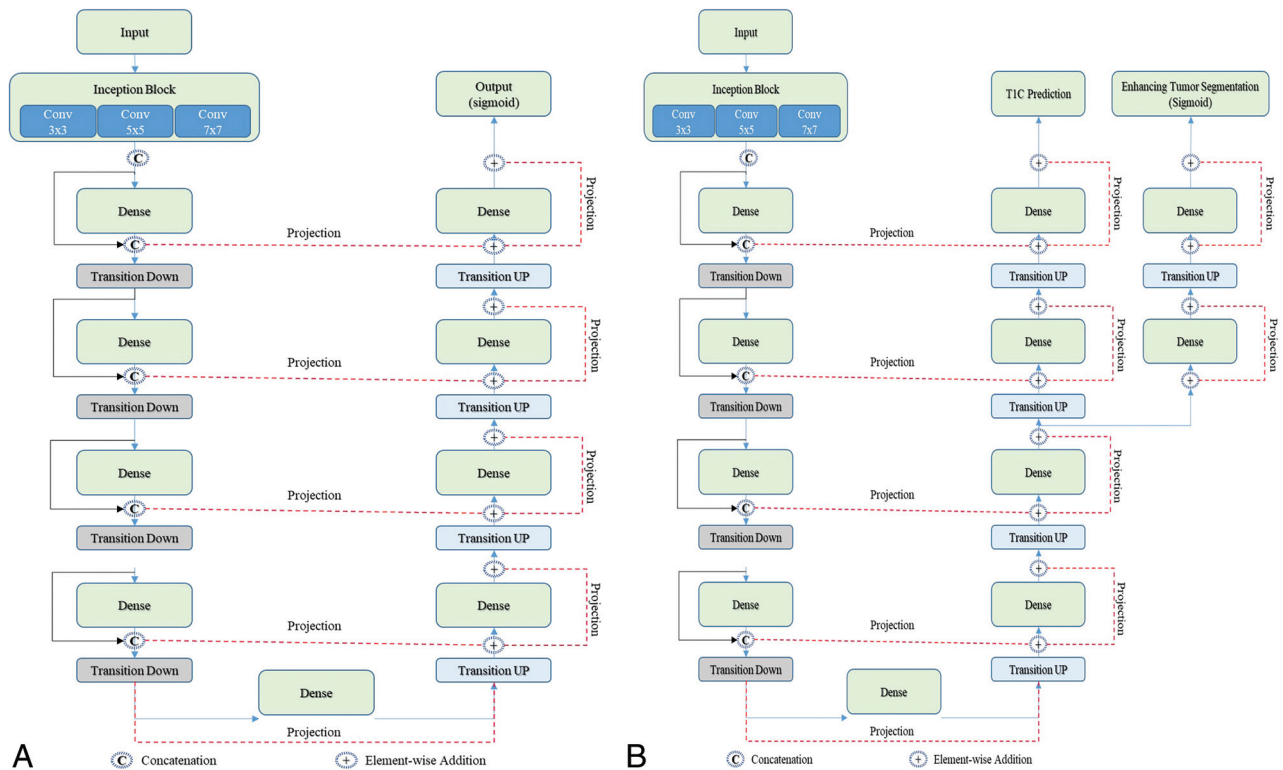


FIG 2. Residual inception DenseNet (RID). A, RID model for whole tumor segmentation. B, RID model for virtual contrast enhancement and enhancing tumor segmentation.

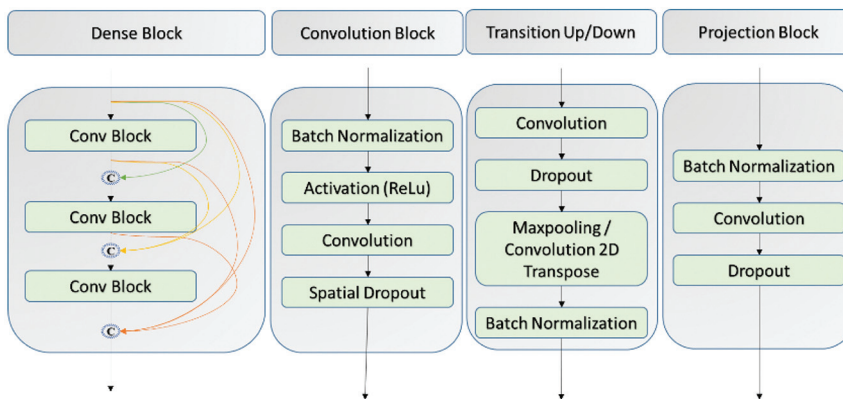


FIG 3. Building blocks of residual inception network. From left to right, dense block, convolution block, transition block, and projection block.

in DenseNet, to mitigate feature map explosion in the up-sampling path. For the skip connections, a projection operation was done by using Batch-Norm-1×1-convolution-drop-out to match the dimensions for element-wise addition (Fig 3). These modifications to the DenseNet architecture helped to reduce the parameter space and the GPU memory footprint without affecting the quality of the segmentation output. In addition to performing dimension reduction, the projection operation helped in learning interactions of cross-channel information¹³ and accelerated convergence. Furthermore, the initial layer of the RID network included parallel convolutional neural network (CNN) branches that were similar to the inception module with

multiple kernels of varying receptive fields, which helped to capture view-point-dependent object variability and learn relationships between image structures at multiple scales.¹⁴

Model Training

The RID model was trained on 2D input patches of size 64×64×3 that were extracted from each image section, with 3 channels (1 for each input image contrast). T1w, T2w, and FLAIR images were concatenated to create the 3 channels of the input. The decoder part of the network was bifurcated to generate 2 outputs: 1) synthesized virtual T1c images (vT1c) and 2) a segmentation

mask of enhancing tumor (ET). Linear and sigmoid activations were utilized for the vT1c generation and ET segmentation, respectively. The mean squared error (L2) loss assumes that the input data set consists of uncorrelated Gaussian signals. This assumption is not always true in real-world data and can result in blurry images. To create sharper output images, we optimized the RID model with the structural perception loss for the vT1c creation and the Dice loss for the ET segmentation. The structural perception loss (SPL), which is further detailed below, is a combination of L2, perception, spatial frequency, and structural similarity loss. Additionally, a separate model, referred to as the whole tumor (WT) model, was trained by using only T2w images to

segment the entire tumor by minimizing Dice loss (Fig 1B). At each stage, the RID model and the WT model were trained until convergence by using Adam optimizers with a learning rate of 0.001 on NVIDIA Tesla P40 GPUs. The tumor grades and manual ground truth annotations for the held out 125 subjects were not made available by BraTS. To facilitate the quantitative analysis and ET segmentations on the held out data set, we derived the annotations by using a model¹⁵ that was trained on the same BraTS 2019 training data set.

Structural Perception Loss

The loss function based on the mean squared error (MSE) between the pixel values of the original images and the reconstructed images is a common choice for learning. However, only using the MSE (L2 loss) results in blurry image reconstruction¹⁶ with a lack of high spatial frequency components that represent edges. Therefore, in addition to the L2 loss, we used the spatial frequency loss (SFL) to emphasize the high-frequency components. Furthermore, a convolutional layer with a Laplacian filter bank as weights was added to the model to emphasize sharp features, such as edges. Perceptual and structural similarity-based (SSIM) losses were also added to improve the model's performance. We used a pretrained VGG-16 network to define perceptual loss functions that measure perceptual differences between predicted images and ground truth images.¹⁷ The VGG loss network remained fixed during the training process. The model was trained to optimize the combination of all of the above losses, which we refer to as structural perception loss (SPL) for simplicity, and can be determined as follows:

$$\text{SPL} = (1 - \alpha) \cdot \text{L2} + \alpha \cdot \text{SSIM} + \beta \cdot \text{SFL} + \gamma \cdot \text{Perceptual},$$

where α , β , and γ represent the normalized contribution of each individual loss. The values were selected to give equal weights for each loss. We combine multiple similarity and error losses to obtain a smooth and realistic virtual contrast synthesis. In this study, we used $\alpha = .5$, $\beta = .5$, and $\gamma = .5$. The combination of multiple loss functions can be interpreted as a form of regularization, as it constrains the search space for possible candidate solutions for the primary task.¹⁸

Evaluation and Statistical Analysis

Quantitative Evaluation. Model performance was evaluated by comparing the model predicted output (vT1c) image to the ground truth T1c image. We computed the SSIM, peak signal-to-noise ratio (PSNR), and normalized mean squared error (NMSE). The PSNR measures the voxelwise difference in signal, the NMSE captures the L2 loss, and the SSIM compares the non-local structural similarity. To evaluate the algorithm's performance for a segmenting enhancing tumor, Dice scores were calculated separately for the whole brain, whole tumor, and enhancing tumor regions by using our previously developed brain tumor segmentation ensemble network.¹⁵ The Dice scores of the ET segmentation were calculated without any correction for the whole brain image after skull stripping (whole brain) but with corrections for whole tumor (after removing predictions outside of the whole tumor segmentation) and ET (after

Table 1: Quantitative evaluation. Analysis of virtual enhancement prediction by using various masks generated by an external model

Mask	SSIM	NMSE	Dice	PSNR
Whole brain	0.91	0.03	0.32	64.35
Whole tumor	0.90	0.01	0.35	48.99
Enhancing tumor	0.90	0.01	0.62	49.93

removing predictions outside of the ET segmentation) to quantify the performance of the model in segmenting ET.

Qualitative Evaluation. To assess the subjective visual quality of the synthesized GBCA enhancement (vT1c), 3 board-certified neuroradiologists (FY [8 years of experience], JD [8 years of experience], and MA [6 years of experience]) rated the synthesized vT1c images by comparing them to the ground truth T1c scans. For each data set, scores were determined by taking into account the general image quality and the degree of visual conformity of the ET region to the ground truth by using a 3-point Likert scale (1 = poor, the algorithm misidentifies the presence or absence of contrast enhancement over the whole tumor volume; 2 = good, the algorithm correctly simulates the signal intensity and the regional extent of enhancement in a portion of the tumor; and 3 = excellent, the algorithm correctly simulates enhancement throughout nearly the full volume of the tumor). The interrater agreement between each rater was computed by using the Fleiss kappa for 3 scale ratings. When discrepancies arose between raters, a consensus rating was obtained through majority voting. The consensus ratings were also dichotomized into low (1) and high (2–3) ratings. The raters also examined the results of the enhancement predictions at a granular level. This included the degree of overestimation or underestimation, the location of enhancement within the tumor (peripheral or lateral), the presence of distant enhancement, the presence of any artifacts, and the overall improvement in image quality.

Importance of the Input MR Sequences for Prediction. To determine the contribution of different input MR imaging sequences on the prediction of the vT1c image, we tested the trained model by iteratively replacing all voxels of each input MR image with zeros while retaining the other 2 input noncontrast MR images.

RESULTS

Quantitative Evaluation

The T1c-ET RID model was tested on 125 held out test subjects. The average PSNR, NMSE, and SSIM for the whole brain were 64.35, 0.03, and 0.91. The whole tumor and ET regions demonstrated lower SSIM and PSNR values compared with the whole brain (Table 1). The Dice coefficients for ET on 125 validation subjects were .32, .35, and .62 for the uncorrected (whole brain), corrected for whole tumor, and corrected for ET cases, respectively. In most cases, the model was able to synthesize T1c images with well-defined enhancing regions, as shown in Fig 4. Out of the 125 subjects tested, only 13 were labeled as low performance after the consensus rating between 3 raters, resulting in an accuracy of 88.8% in synthesizing vT1c. Table 2 summarizes the quality of enhancement, the location of enhancement in the tumor

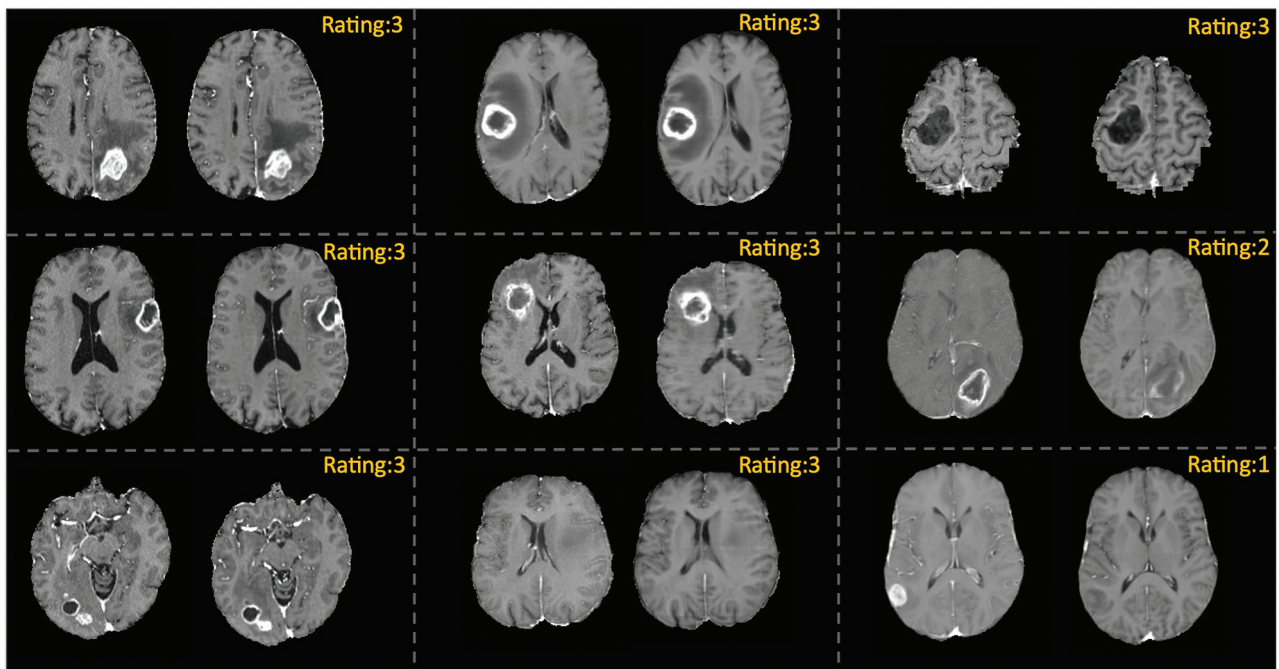


FIG 4. Synthesized virtual contrast enhanced T1w (vT1c) images in 3 different subjects. Ground truth (left column) and synthesized vT1c (right column) image pairs for 9 subjects.

Table 2: Quantitative presence and location of the under/over-estimation of synthetic contrast enhancement, the introduction of artifacts, and the image quality improvement on vT1c

	Reviewer 1 (FY)	Reviewer 2 (MA)	Reviewer 3 (JD)
Overestimate (O)	26	27	20
Underestimate (U)	71	69	58
Both (O and U)	11	8	34
Central	54	35	64
Peripheral	94	71	97
False distant enhancement	10	9	15
Missed distant enhancement	5	4	2
Artifact	17	22	25
Image quality improved	10	NA	3

Note:—NA indicates not applicable.

(peripheral/lateral), the prediction of distant enhancement, the presence of artifacts, and the predicted image quality improvement.

Qualitative Evaluation

Representative images are shown in Fig 4. Comparing the synthesized vT1c images with the ground truth T1c images, 89.6% of the subjective rater scores after consensus were within the good and excellent range (Supplemental Online Data). The intraclass rater reliability of the 3 neuroradiologists was 0.61, indicating moderate interrater agreement by using the 3-point Likert scale. A consensus rating was obtained through majority voting in situations in which the raters had different scores. In cases where the 3 ratings differed, the lowest rating was taken as the consensus. After consensus, a subset of cases (11.2%, 14 cases) was rated as low, in which enhancing regions were not well-captured or were absent, compared with the ground truth T1c data (Fig 5). An example of a low-rated vT1c image is shown in Fig 4.

Importance of the Input MR Sequences for Prediction of Contrast Enhancement

By replacing each input sequence with zeros, we were able to determine which sequences are important in the prediction of specific components of the output vT1c images. The T1w image contributes primarily structural brain information in the predicted vT1c image. The FLAIR and T2w images primarily influence the predicted contrast enhancement (Fig 6).

DISCUSSION

We developed and trained a deep learning model utilizing a diverse multi-institutional data set that was able to synthesize vT1c images for primary brain tumors by using only noncontrast FLAIR, T2w, and T1w images. Qualitative and quantitative evaluations showed the robust performance of the model when predicting tumor enhancement. In most cases, the enhancing and nonenhancing portions of the tumors were correctly predicted.

Gong et al²⁰ developed a deep learning method to predict full-dose T1w postcontrast (T1c) images by using one-tenth of the standard GBCA dose. With respect to this prior work, our study represents an advancement by using only noncontrast sequences to predict T1c images.¹⁹ Narayana et al²⁰ evaluated whether deep learning can predict enhancing demyelinating lesions on MR imaging scans that were obtained without the use of contrast material and demonstrated moderate to high accuracy in patients with multiple sclerosis. Kleesiek et al¹² developed a Bayesian network to predict T1c by using noncontrast T1w, T2w, FLAIR, DWI, and SWI as a 10-channel input.¹² Recently, Calabrese et al²¹ conducted a study to explore the feasibility of dose-free synthesis by training 3D convolutional networks on an internal data set of 400 subjects with 8 noncontrast MR images as input and evaluating

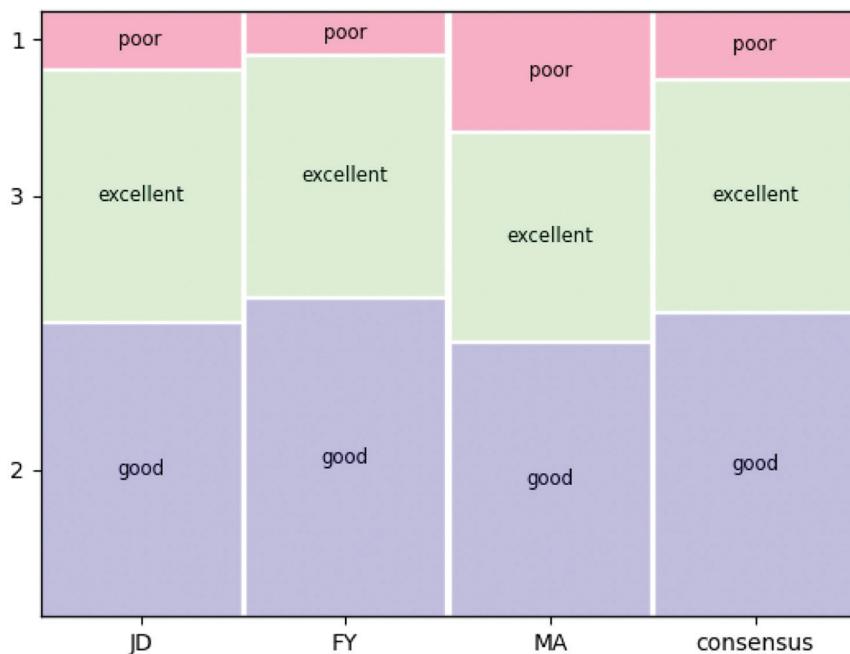


FIG 5. Mosaic plot illustrating the distribution of 3 expert radiologists and their consensus along a 3-point Likert scale.

the model on an external BraTS data set of 200 subjects. For quantitative evaluation, the authors employed an external model that was trained on BraTS data to generate enhancing tumor segmentation by incorporating real (T1c) and virtual contrast (vT1c) in addition to other noncontrast multiparametric MR images (T1, T2, and FLAIR). They reported that the synthesized whole brain postcontrast images exhibited both qualitative and quantitative similarity to the real postcontrast images, as indicated by quantitative metrics such as the Dice coefficients of 0.65 ± 0.25 and 0.62 ± 0.27 for the internal and external BraTS data sets, respectively, for the enhancing tumor compartment. In contrast, our method solely utilizes noncontrast multiparametric MR images (T1, T2, and FLAIR) to predict and segment virtual contrast enhancement, which accounts

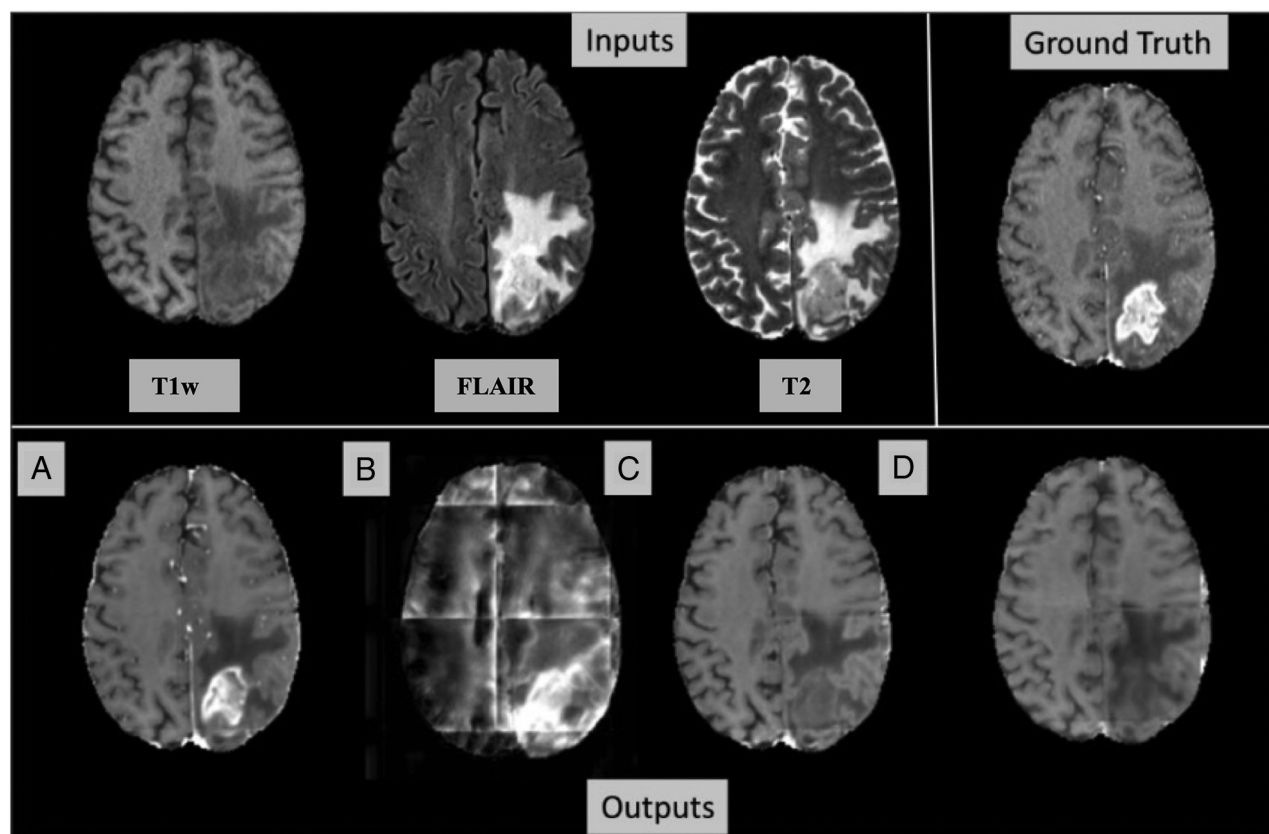


FIG 6. Importance of input sequences example. Top row, input images: T1w, FLAIR, T2, and the ground truth T1c. Bottom row, output images with (A) all inputs (T1w, FLAIR, and T2w) given to the model, (B) T1w replaced with zeros in the input, (C) FLAIR replaced with zeros in the input, and (D) T2 replaced with zeros in the input. The T2 and FLAIR inputs together provide contrast enhancement prediction, whereas T1w input provides primarily anatomic detail.

for the comparatively lower enhancing tumor Dice score observed for the whole brain in our study.

Our results further support this approach by demonstrating the successful prediction of enhancement in almost 90% of the testing data set. Moreover, we were able to achieve comparable results (with superior performance in quantitative metrics, including PSNR and SSIM) while utilizing notably fewer sequences (only T1w, T2w, and FLAIR images) that are standard for clinical brain imaging protocols. Furthermore, the need for fewer sequences also facilitates reduced scan times, which is an important consideration for critically ill, claustrophobic, and cognitively impaired patients.

Another advantage of our strategy included the use of a more diverse data set (BraTS) for both training and testing. Whereas prior studies utilized imaging data from a single institution, BraTS comprises data from multiple sites with variations in acquisition parameters, scanner platforms, and imaging protocols.^{12,19} Introducing more heterogeneity to the training data set enhances the generalizability of the trained networks. Furthermore, testing on a substantially larger data set quantifies the generalizability of the model more accurately. Taken together, these result in a more generalized approach that is robust to differences in imaging hardware and software and is therefore more amenable to clinical translation.

Our analysis of the relative contributions of the input sequences revealed that the FLAIR and T2w images contributed complementary information in predicting enhancement within the tumor. This is consistent with the results presented by Kleesiek et al¹² who noted that T2w images were the most important for predicting contrast. FLAIR and T2w images are generally thought of as having greater contrast-to-noise for the delineation of pathology, compared with T1w images. Tissue changes that are related to disruption of the blood-brain barrier that led to or are seen in association with contrast enhancement, such as necrosis and edema, may be better delineated with these sequences. The T1w images contributed information primarily toward delineating structural details of the brain. T1w images are generally regarded as anatomic images for their ability to capture the fine anatomic details of the brain.

The T1c-ET model failed to predict gadolinium enhancement in subjects for whom 1 or more of the input sequences had a significant motion artifact and for whom the tumor was isointense to normal brain parenchyma on both T2w and FLAIR sequences (Supplemental Online Data). The failure of the model in the latter scenario may be due to an inadequate representation of tumors with these imaging features in the training set. This could be alleviated through the incorporation of additional, larger data sets for training in the future. The deterioration of image quality due to image artifacts, such as motion, could be separately addressed by either preventing them during acquisition or correcting them retrospectively.²² Another potential limitation for implementation is that we used only primary brain tumor cases for the training and testing of the model. The application of the algorithm in cases of sub-centimeter brain metastases and its extension to other body parts represent exciting areas to explore in the future.

An in-depth qualitative review of the synthesized vT1c revealed that, though the enhancement accuracy is satisfactory, there is a

tendency to overestimate or underestimate the enhancement, and there is also a potential for distant enhancement. Regarding this approach, the implications of these observations and the effectiveness of radiologic/surgical decisions and survival predictions based on vT1c, compared with those of real T1c images, must be further investigated before the method can be translated into a clinical tool. Taken together, the results of the current study should be regarded as a proof-of-concept study of clinical feasibility. Future directions to augment the performance of our model include the incorporation of larger data sets and different pathologies as well as the potential acquisition of additional sequences, including rapid low-dose, low-resolution echo-planar gadolinium-enhanced images (as are used for dynamic perfusion MR imaging techniques).¹⁹

CONCLUSIONS

We developed a novel deep learning architecture to synthesize virtual contrast-enhanced T1w images (vT1c) by using only standard clinical noncontrast multiparametric MR images. The model demonstrated good quantitative and qualitative performance in a larger and more heterogeneous data set than those used in prior studies, and showed the feasibility of gadolinium-free predictions of contrast enhancement in gliomas. FLAIR and T2w images were found to provide complementary information for predicting tumor enhancement. Further studies in larger patient data sets with different neurologic diseases are needed to fully assess the clinical applicability of this novel approach.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

1. Villanueva-Meyer JE, Mabray MC, Cha SJN. **Current clinical brain tumor imaging.** *Neurosurgery* 2017;81:397–415 [CrossRef Medline](#)
2. Weinreb JC, Rodby RA, Yee J, et al. **Use of intravenous gadolinium-based contrast media in patients with kidney disease: consensus statements from the American College of Radiology and the National Kidney Foundation.** *Kidney Med* 2021;3:142–50 [CrossRef Medline](#)
3. Hoggard N, Roditi GH. **T₁ hyperintensity on brain imaging subsequent to gadolinium-based contrast agent administration: what do we know about intracranial gadolinium deposition?** *Br J Radiol* 2017;90:20160590 [CrossRef Medline](#)
4. Wang J, Wang H, Ramsay IA, et al. **Manganese-based contrast agents for magnetic resonance imaging of liver tumors: structure-activity relationships and lead candidate evaluation.** *J Med Chem* 2018;61:8811–24 [CrossRef](#)
5. Greer JS, Wang X, Wang Y, et al. **Robust pCASL perfusion imaging using a 3D Cartesian acquisition with spiral profile reordering (CASPR).** *Magn Reson Med* 2019;82:1713–24 [CrossRef Medline](#)
6. Viswanathan S, Kovacs Z, Green KN, et al. **Alternatives to gadolinium-based metal chelates for magnetic resonance imaging.** *Chem Rev* 2010;110:2960–3018 [CrossRef Medline](#)
7. Menze BH, Jakab A, Bauer S, et al. **The multimodal brain tumor image segmentation benchmark (BRATS).** *IEEE Trans Med Imaging* 2014;34:1993–2024 [CrossRef Medline](#)
8. Bakas S, Akbari H, Sotiras A, et al. **Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features.** *Sci Data* 2017;4:170117
9. Rohlfing T, Zahr NM, Sullivan EV, et al. **The SRI24 multichannel atlas of normal adult human brain structure.** *Hum Brain Mapp* 2010;31:798–819 [CrossRef Medline](#)

10. Bakas S, Reyes M, Jakab A, et al. **Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge.** 2018
11. Tustison NJ, Cook PA, Klein A, et al. **Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements.** *Neuroimage* 2014;99:166–79 [CrossRef Medline](#)
12. Kleesiek J, Morshuis JN, Isensee F, et al. **Can virtual contrast enhancement in brain MRI replace gadolinium? A feasibility study.** *Invest Radiology* 2019;54:653–60 [CrossRef Medline](#)
13. Lin M, Chen Q, Yan S. **Network in network.** *arXiv preprint arXiv* 2013;1312.4400
14. Szegedy C, Liu W, Jia Y, et al. **Going deeper with convolutions.** In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Boston, Massachusetts. June 7–12, 2015;1–9
15. Murugesan GK, Nalawade S, Ganesh C, et al. **Multidimensional and multiresolution ensemble networks for brain tumor segmentation.** In: A. Crimi, S. Bakas, eds. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Lecture Notes in Computer Science, vol 12658. Springer; 2021;448–57
16. Ichimura N. **Spatial frequency loss for learning convolutional autoencoders.** *arXiv* 2018;1806.02336
17. Johnson J, Alahi A, Fei-Fei L. **Perceptual losses for real-time style transfer and super-resolution.** In: B. Leibe, J. Matas, N. Sebe, M. Welling, eds. *Computer Vision*. ECCV 2016. Lecture Notes in Computer Science, vol 9906. Springer; 2016;694–711
18. BenTaieb A, Hamarneh G. **Uncertainty driven multi-loss fully convolutional networks for histopathology.** In: M. Cardoso, et al. *Intravascular Imaging and Computer Assisted Stenting, and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis*. Lecture Notes in Computer Science, vol 10552. Springer; 2017;155–63
19. Gong E, Pauly JM, Wintermark M, et al. **Deep learning enables reduced gadolinium dose for contrast-enhanced brain MRI.** *J Magn Reson Imaging* 2018;48:330–40 [CrossRef Medline](#)
20. Narayana PA, Coronado I, Sujit SJ, et al. **Deep learning for predicting enhancing lesions in multiple sclerosis from noncontrast MRI.** *Radiology* 2020;294:398–404 [CrossRef Medline](#)
21. Calabrese E, Rudie JD, Rauschecker AM, et al. **Feasibility of simulated postcontrast MRI of glioblastomas and lower-grade gliomas by using three-dimensional fully convolutional neural networks.** *Radiology Artif Intell* 2021;3:e200276 [CrossRef Medline](#)
22. Duffy BA, Zhao L, Sepehrband F, Alzheimer's Disease Neuroimaging Initiative, et al. **Retrospective motion artifact correction of structural MRI images using deep learning improves the quality of cortical surface reconstructions.** *Neuroimage* 2021;230:117756 [CrossRef Medline](#)