Deep learning based Max Field-of-View

See More, See Better

Xueli Wang, PhD, Bingjie Zhao; Weiwei Xing; Ximiao Cao; Haifeng Wu; Jiang Hsieh, PhD

Contents

gehealthcare.com

Introduction

Max Field-of-View 2 (MaxFOV 2) is a new deep learning (DL) based innovative option available on the GE computed tomography (CT) scanners that increases the maximum display field-of-view (DFOV) beyond the conventional GE scanner limit. The increase in the maximum DFOV from 50 cm to up to 80 cm allows clinicians to visualize more anatomical information and improve skin line accuracy for optimal radiation therapy simulation and planning.

This white paper will first describe the clinical needs for reconstructing images outside DFOV and examine the shortcomings of existing approaches. It then outlines the approach of a deep learning based MaxFOV 2, including the network selection, training data generation, and the training process. The white paper concludes by presenting quantitative and qualitative evaluation results of clinical and phantom studies.

Clinical Needs

A schematic diagram of a CT system is shown in Figure 1. The fully reconstructable region, depicted by the dotted red circle, is irradiated during the entire CT data acquisition as the X-ray tube and detector rotates around the patient, and is called the scan field-of-view (SFOV). The conventional CT reconstruction is capable of reconstructing an object inside this region. The area between the SFOV and the gantry bore (depicted by the green dotted line) is only "partially sampled" in that X-ray does not irradiate the entire region all the time and therefore measurements are obtained only over a limited range of rotation angles.

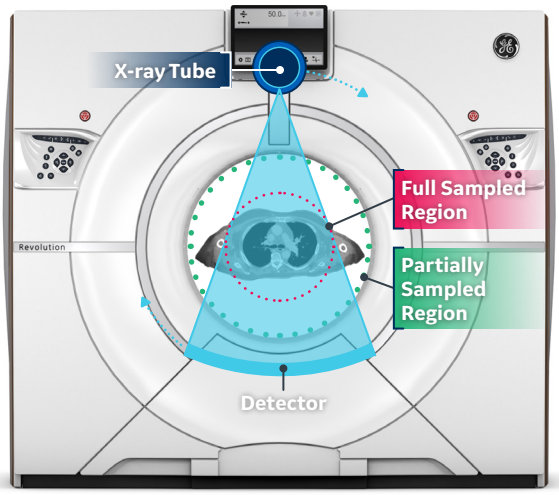
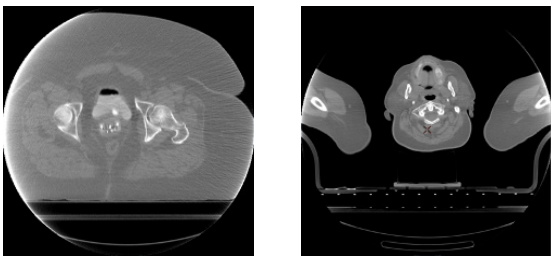


Figure 1: A schematic diagram of a CT system illustrating SFOV. The area inside the red dotted circle is "fully sampled" and defines as the SFOV of the system. Images can be accurately reconstructed inside this region with the conventional CT reconstruction algorithm. The region between the green dotted circle and the SFOV is a "partially sampled region". This region is irradiated by X-ray only over a limited angular range and image cannot be reconstructed by the conventional algorithm in this region.

The diameter of a CT scanner's bore is typically larger than the diameter of the SFOV. In radiation therapy planning, patient's arms or other parts of anatomy sometimes extend outside the imaging area in order to accommodate for the treatment position. Additional accessories are often placed below or around the patients to enable the region of interests (ROI) to be placed near the center of SFOV which results in portions of the anatomy being outside SFOV. For obese patients, anatomy truncation can occur even with the best patient positioning efforts since the width of the patient often exceeds the diameter of SFOV. Patient anatomy that extends outside SFOV cannot be reconstructed with a conventional reconstruction algorithm. In addition, the truncated data results in a peripheral shading artifact which can obscure anatomy even inside the SFOV. Figures 2 (a) and (b) depict two examples of truncation with (a) a cross-sectional pelvis image of an obese patient and (b) a patient with their arms truncated due to additional accessory placed below the patient. In both cases, patient's skin line cannot be accurately estimated, and shading artifacts are observed near the edge of the SFOV.



(a) (b)

Figure 2: Examples of anatomy truncation. (a) an obese patient with size greater than SFOV and (b) a RT patient positioned with additional accessory resulting in arm truncation. Both images were reconstructed with 50 cm FOV.

It is desirable to be able to view the patient anatomy that extends beyond the SFOV and to eliminate truncation artifacts inside the SFOV. For oncology applications, it is important that the patient's entire anatomy can be covered in the CT image in order to accurately estimate the skin line and calculate the dose distribution during radiation therapy planning. When CT images are fused with images from other modalities with a larger scan field-of-view, such as positron emission tomography (PET), it is also useful to be able to cover the entire patient anatomy to enable more accurate attenuation correction.

Challenges with Existing Algorithms

Over the years, various reconstruction algorithms to expand the reconstruction FOV have been proposed. Conventional wide view algorithms utilize the view consistency principles in that the total amount of attenuation integrated over all detector channels remains constant from view to view in parallel beam geometry, and as a result, can precisely locate the truncated views and estimate the amount of truncation.[1] One of the algorithms, the WideView algorithm, takes advantage of this property and estimates the missing portion of the projection by determining a cylindrical water or soft-tissue object that best fits the measurement at the location of truncation.[1,2] Another algorithm, the Max Field-of-View algorithm, estimates the object contour in an iterative fashion by performing multiple forward and backward projections. In addition, a sophisticated transition modification algorithm is utilized to reduce artifact near the edge of the SFOV.[3] Although this algorithm is an improvement,[1] it still relies on filling the truncated regions with constant attenuation similar to soft-tissue.

When the truncated portion of the object is made of uniform soft-tissue material and its contour is close to an ellipsoidal shape, these methods work reasonably well. Since both approaches rely on non-truncated projections, as reference to estimate the amount of truncation, the angular span of the truncated view needs to be limited.

In real clinical situations; however, these simplistic assumptions are often violated. For example, when an air gap is present between the main body and the truncated object as shown in Figure 3(a), such as an arm, the projection measurements at the boundary channels are roughly zero and no information is available to estimate the truncated object location or shape. Another example occurs when multiple truncations are present or when the angle of truncation is larger than 90-degrees, as depicted in Figures 3(b) and (c). In these scenarios, the estimation of the total attenuation of non-truncated projections become less reliable and can lead to increased levels of residual artifacts. When an accessory used to position the patient is outside the SFOV, the soft-tissue assumption of the algorithm is violated, and suboptimal image quality can result as shown in Figure 3(d). The illustrative examples shown in Figure 3 highlight deficiencies of the conventional algorithms resulting from various simplistic assumptions.

In addition to in-plane residual artifacts under stress conditions, conventional algorithms also face challenges in reformatted or volume rendered images. Conventional algorithms try to estimate the missing portion of the projection on a slice by slice basis. Although certain constraints can be placed to ensure image consistency along the z-axis, such an approach has met with limited success due to the complexity of the patient contour. Figure 4 shows a volume rendered image of a patient upper body. Poor consistency in the estimation of patient skin line leads to high-frequency variations and discontinuities in the z-axis.

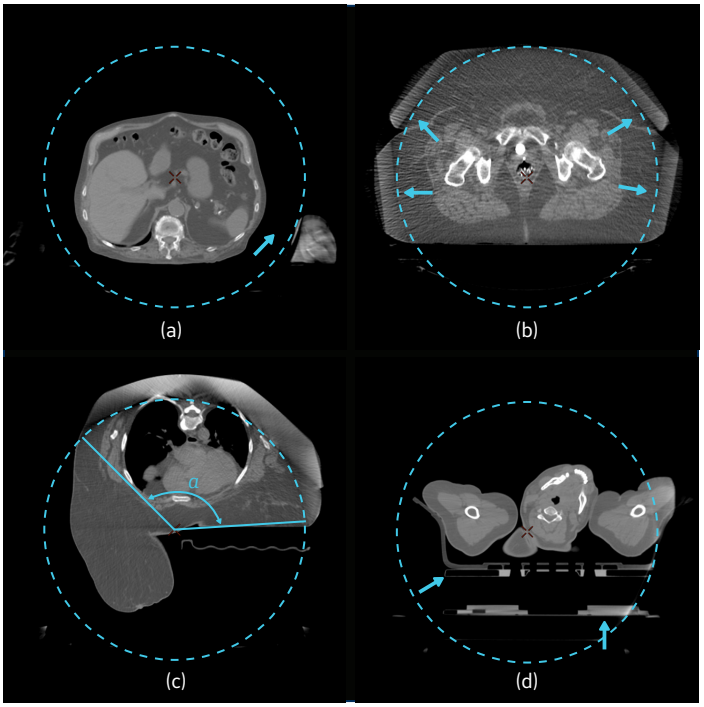


Figure 3: Example of stress cases, (a) an airgap, as shown by the blue arrow, is present between the main body and the truncated portion, (b) multiple truncations take place as shown by the blue arrows, (c) the angle of truncation, (a), is larger than 90-degree; (d) accessory is present beneath the patient as shown by the blue arrows.

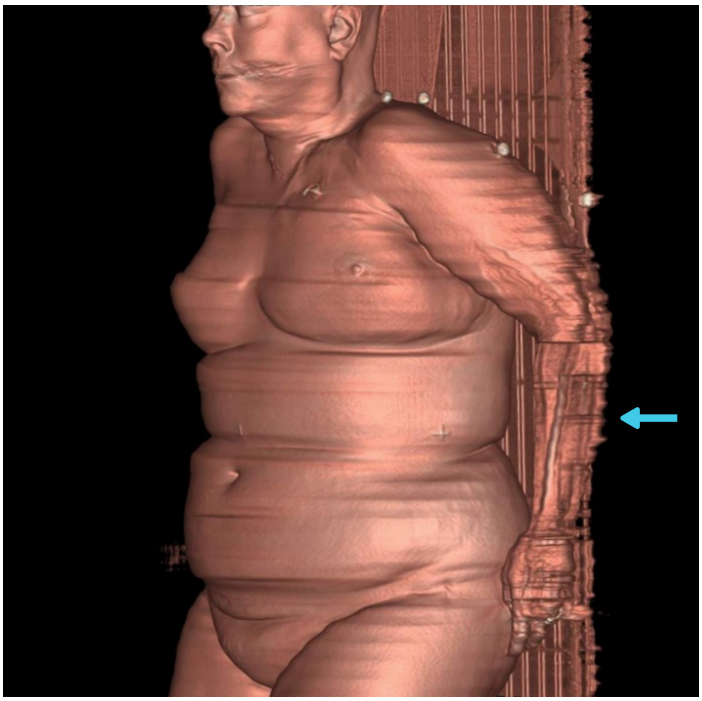


Figure 4: Volume rendered image of a patient body with a conventional algorithm. Because of the algorithm limitation, discontinuities in patient contour is clearly visible.

Deep Learning-Based Solution – MaxFOV 2

With the recent advancement in computing power and development of new technologies, artificial intelligence (AI) has gained significant popularity in recent years and has become available in our daily lives. One of the most popular technologies is deep learning (DL), which has found broad applications including facial recognition, speech recognition, image segmentation, and computer aided detection. Considering the complexity of the clinical situations in RT, it becomes necessary to explore DL's application to handle estimated field-of-view (EFOV) reconstruction.

The fundamental issue with the conventional approaches is the simplified assumptions made in the truncated object models. The simplification is necessary in order to ensure the correction is mathematically trackable. Such constraint; however, is not present in a deep learning-based approach since DL relies on learning from real world examples. At the same time, traditional approaches are based on fundamental properties of X-ray physics and should be effectively leveraged to further refine the DL-based estimation.

GE has developed a new technique called "Max Field-of-View 2 (MaxFOV 2)" to further refine the DFOV extension. MaxFOV 2 can be characterized as a "hybrid" algorithm in that both the conventional and DL-based approaches are combined. The algorithm is divided into three major parts: initial EFOV reconstruction, DL-based image refinement, and final EFOV reconstruction, as illustrated in Figure 5. The first and last step employ classic algorithmic approaches that rely on math and physics, and the middle portion leverages the advantage of the DL approach to handle complicated object shapes and nonhomogeneous attenuations.

Step 1: Initial EFOV Reconstruction

The original sinogram corresponding to 50 cm SFOV (a) is expanded to an EFOV sinogram of up to 80 cm FOV with a traditional EFOV algorithm. An initial EFOV image (b) is reconstructed with a tomographic reconstruction algorithm such as FBP or IR.

Step 2: DL-based Refinement

The DL-based refinement is used to further improve the preliminary EFOV image (b). The goal of this step is to improve upon the initial EFOV image and provide a better estimation of the object being truncated (c).

Step 3: Final EFOV Reconstruction

A synthesized sinogram (d) is generated by forward projecting the DL-based image (c). This sinogram is combined with the original measured sinogram to produce the final sinogram (e), and a tomographic reconstruction algorithm, such as FBP or IR, is then used to produce the final EFOV image (f). Of course, other advanced image reconstruction features, such as metal artifact reduction (MAR), can be used in combination with the proposed EFOV reconstruction algorithm.

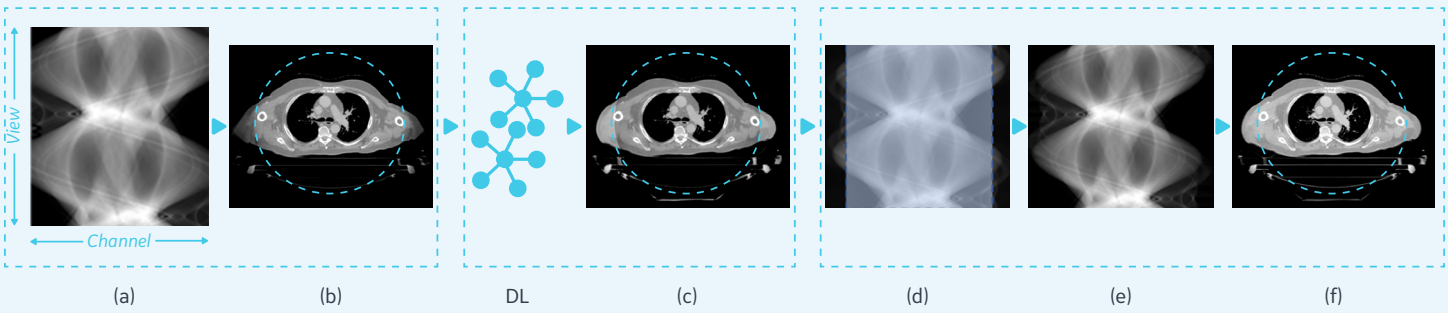


Figure 5: Illustration of MaxFOV 2 process, (a) original sinogram collected on CT scanner with truncation, (b) initial reconstruction of EFOV with a traditional algorithm, (c) refinement of reconstructed image with a trained DL model, (d) synthesized sinogram generated by forward projection, (e) final sinogram generated by combining the original measurement and the synthesized sinogram, (f) final EFOV image.

Deep Learning-Based Solution – MaxFOV 2 *continued*

Before discussing the performance evaluation of MaxFOV 2, a brief description of the DL-based refinement is in order. It is well known that the performance of a DL model depends on the topology of the selected network, the dataset used in training, validation, and testing, and the training process itself. Based on the complexity of the problem, available computational resources, and reconstruction speed considerations, we selected a multi-layer Deep Neural Network (DNN) which contains over 40 convolution layers and over 25 million parameters. Extensive testing and evaluation has shown that such network topology can ensure the quality of the produced images even under the most challenging conditions.

The availability of a high-quality training dataset is equally important. From an algorithm robustness point-of-view, the training dataset needs to cover different data acquisition protocols, such as different kVp, different helical pitches, and different dose levels. In addition, a variety of patient anatomical regions, such as head, chest, abdomen, and pelvis, need to be adequately covered. For RT applications, the presence of a variety of accessories also needs to be included. It is clear that the number of cases used for training is large.

Because our goal is to reconstruct EFOV images of up to 80 cm, it is important to generate perfectly reconstructed images corresponding to the EFOV and use them as the ground truth for DL. Such a requirement is not difficult to obtain for phantom studies since a phantom can be positioned carefully to produce either a truncated image or a non-truncated image. The non-truncated image can be readily used as the ground truth. For clinical studies; however, such flexibility does not exist. Here, we leverage extensively the computer simulation to accomplish the truncated and non-truncated data generation.

Beside the network topology and training dataset, the training process itself also plays a key role in the success of the DL-based algorithm. We chose a supervised training process outlined below:

1. *With truncated data as input, the DL engine generates a refined EFOV data.*
2. *Compare the refined EFOV data with the ground truth to produce a loss function. The loss function is an indicator of the difference between the two. The smaller the loss function value, the smaller the difference between DL output and ground truth.*
3. *Over 25,000,000 parameters of the network are fine-tuned through back-propagation process. This is an iteration processing and the loss function continues to decrease until it reaches a plateau.*
4. *The entire process is repeated on varied training datasets to guarantee the output of the network can accurately match ground truth images in variety of conditions.*

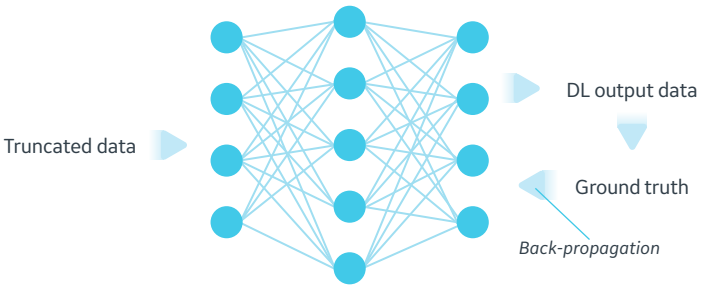


Figure 6: The training process.

The loss function is also useful to ensure the sufficiency condition is met for the training dataset, especially when the size of the training dataset is finite. To ensure the training dataset is sufficiently large, we generated the loss with a number of N, 2N, 3N, and 4N, training datasets as shown in Figure 7. The losses, for the training (aqua curve) and validation loss (dotted aqua), are significantly higher for the N case. When training datasets are larger than or equal to 2N, the losses are quite similar for both the training and validation. This is a confirmation that the selection of 4N training datasets is sufficient.

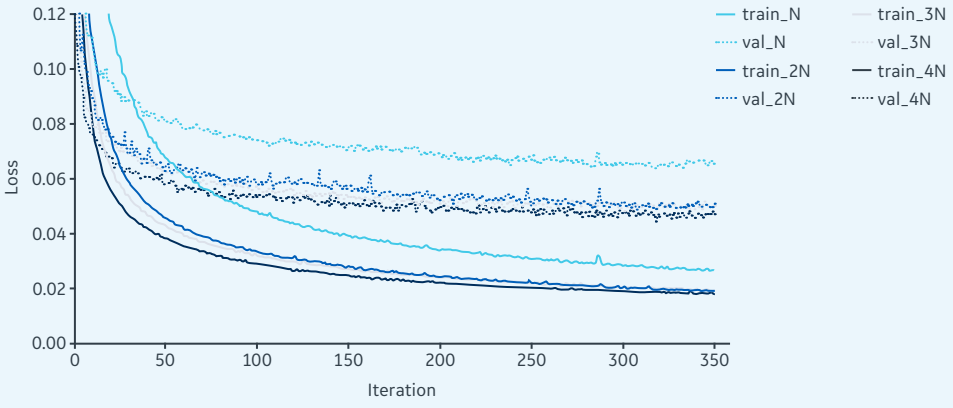


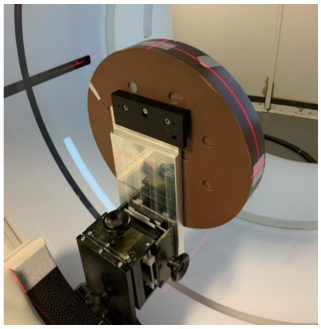
Figure 7: The relationship between the loss function and the training dataset size. The aqua curves, corresponding to dataset size of 1N, plateaus at higher values compared to dark blue (2N dataset), gray (3N dataset), indigo (4N dataset) curves. High similarities among dark blue, gray, and indigo curves indicate that the training dataset size of 4N has met the data sufficiency condition.

MaxFOV 2 Performance

Both quantitative and qualitative evaluations were conducted to assess the performance of MaxFOV 2. Quantitative evaluation was performed with phantom studies and qualitative evaluation was performed with clinical scans.

For quantitative performance assessment, both CT number accuracy of water and skin line accuracy were measured with a Tissue Characterization Phantom (model 467, Gammex, Middleton WI USA). Figure 8 shows the experimental setup with (a) photograph of the phantom and the phantom-holder (b) phantom insert arrangement in the testing scan.

To test different degrees of truncation, the phantom was scanned at different elevation positions relative to the iso-center as show in Figure 9. The ground truth was established by centering the phantom in the SFOV so that no truncation occurred as showed in Figure 9(a). The phantom was then sequentially positioned so that its top edge is close to 60 cm DFOV (b), 70 cm DFOV (c), and 80 cm DFOV (d). In (d), the top of the phantom is positioned near the gantry bore to mimic the worst case truncation. An automated software tool was developed to estimate the CT number accuracy of water and skin line accuracy. Table 1 shows the results obtained from the reconstructed images with the first generation conventional EFOV algorithm (WideView), the second generation conventional EFOV algorithm (MaxFOV), and the DL-based EFOV algorithm (MaxFOV 2), though only 20 HU accuracy improved from MaxFOV to MaxFOV in claimed spec, MaxFOV 2 shows more robustness on varied phantoms, e.g. large body phantom.



(a)

(b)

Figure 8: Phantom setup used in CT number accuracy of water and skin line accuracy test, (a) the tissue Characterization phantom is positioned on a phantom holder to ensure positioning, (b) different material inserts are placed to specify the test condition.

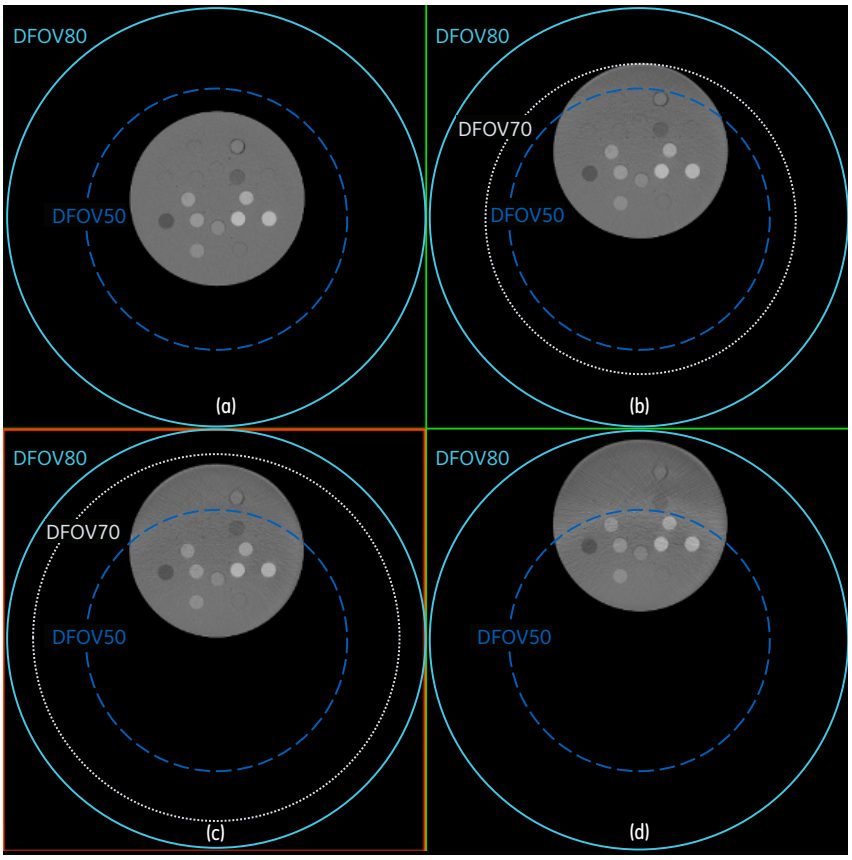


Figure 9: Gammex phantom position used in test.

2004 – 2015 WideView		
Accuracy	**If object extends to 50 – 70 cm DFOV**	**If object extends to 70 – 80 cm DFOV**
CT# of water	No specs	No specs
Skin line	No specs	No specs

2015 – 2020 MaxFOV		
Accuracy	**If object extends to 50 – 70 cm DFOV**	**If object extends to 70 – 80 cm DFOV**
CT# of water	± 40 HU	± 80 HU
Skin line	± 2 mm	± 3 mm

2020 – Future MaxFOV 2 *Developed via deep learning training*		
Accuracy	**If object extends to 50 – 70 cm DFOV**	**If object extends to 70 – 80 cm DFOV**
CT# of water	± 40 HU	± 60 HU
Skin line	± 2 mm	± 3 mm

Table 1: CT number and skin line accuracy comparison of WideView, MaxFOV and MaxFOV 2.

MaxFOV 2 Performance *continued*

To further evaluate performance of MaxFOV 2 in actual clinical scenarios, we collected a large number of clinical cases of MaxFOV 2 and MaxFOV results. In radiotherapy, truncation often occurs in the arm, chest, abdomen, and pelvis of large patients or patients with a variety of accessories. Figure 10 shows examples of challenging truncation scenarios. These scenarios typically lead to unsatisfactory results for conventional EFOV algorithms. Figures from left to right show the presence of air gap (a and e), multiple truncation and high dense object in the truncation region (b and f), over 90-degree truncation (c and g), and accessories in the truncation region (d and h). The blue arrows indicate where MaxFOV algorithm fails to correct for brighter or darker artifacts. In addition, the patient skin line is poorly defined which may impact radiotherapy planning. MaxFOV 2, on the other hand, uses a DL network trained by data collected on multiple CT scanners and successfully overcomes the shortcomings. Due to the robustness of the algorithm and the volumetric nature of the training dataset, MaxFOV 2 provides superior continuity in volume rendered images along the z-axis as compared to MaxFOV, as illustrated in Figures 10 (i) to (p).

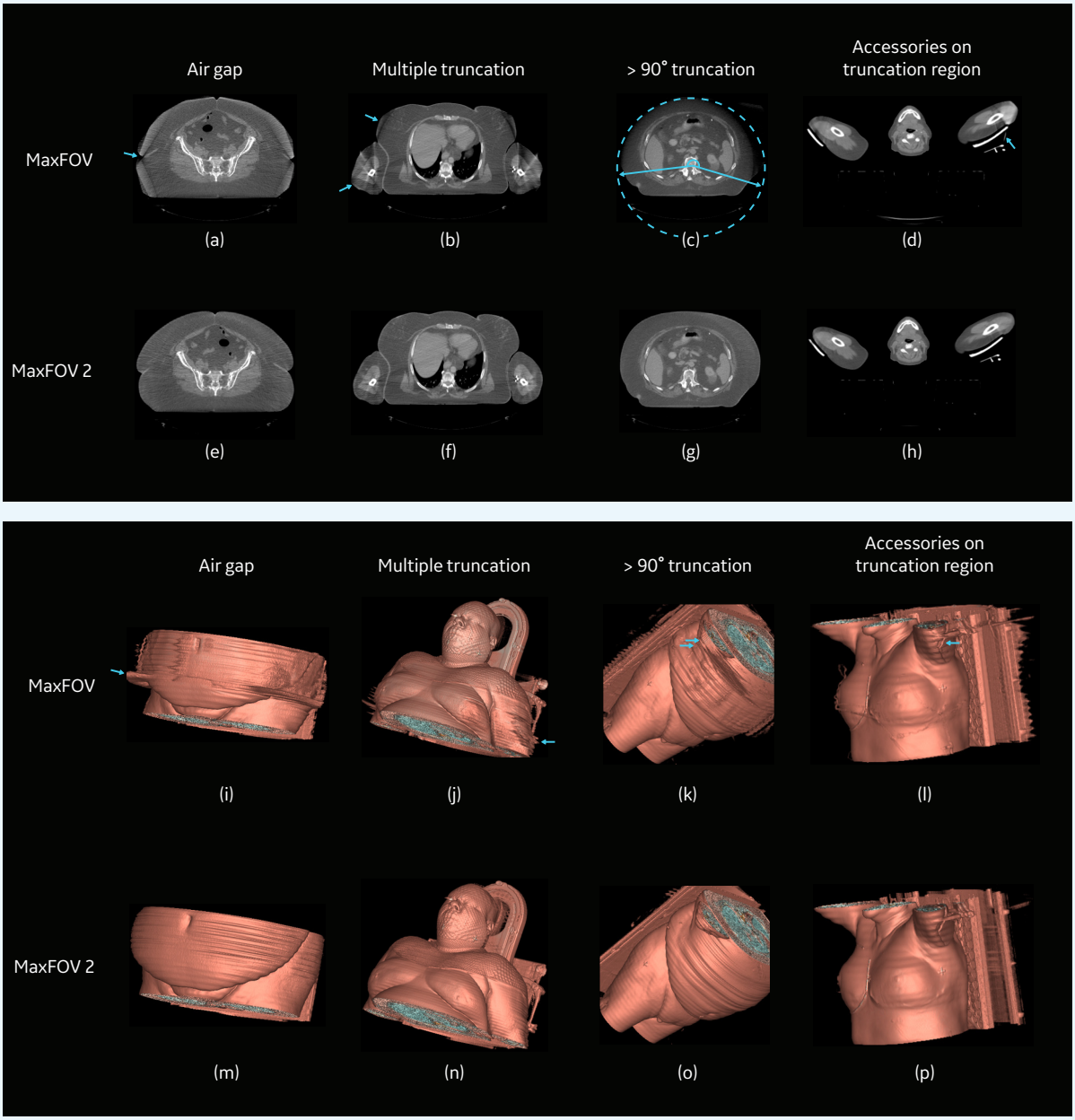


Figure 10: MaxFOV and MaxFOV 2 performance presented in challenging scenarios.

Conclusions

The MaxFOV 2 is a DL-based reconstruction technology with an innovative design and advanced training method. It increases the display FOV beyond the SFOV of the scanner and overcomes shortcomings of conventional algorithms. The MaxFOV 2 improves axial image quality in terms of skin line accuracy and CT number accuracy of water, and provides improved contour in 3D images even under challenging clinical scenarios.

References

[1] Edward H. Chao and Jiang Hsieh, "Extended Field-of-View (WideView) Image Reconstructions", White Paper, GE Healthcare (2004).

[2] J. Hsieh, A. Lonn, and M. Nyka, "Method and apparatus for field-of-view expansion of volumetric CT imaging," US patent 7,756,315 (2010).

[3] Zhihui Sun, Zhi Ye, Ming Yan, and Jiang Hsieh, "White Paper for Max field-of-view CT Reconstruction," GE Healthcare (2015).

June 2021
JB16358XX