RESEARCH ARTICLE

Task-based assessment of resolution properties of CT images with a new index using deep convolutional neural network

Aiko Hayashi[1,2] · Ryohei Fukui[3] · Shogo Kamioka[1] · Kazushi Yokomachi[1] · Chikako Fujioka[1] · Eiji Nishimaru[1] · Masao Kiguchi[1] · Junji Shiraishi[4]

Abstract

In this study, we propose a method for obtaining a new index to evaluate the resolution properties of computed tomography (CT) images in a task-based manner. This method applies a deep convolutional neural network (DCNN) machine learning system trained on CT images with known modulation transfer function (MTF) values to output an index representing the resolution properties of the input CT image [i.e., the resolution property index (RPI)]. Sample CT images were obtained for training and testing of the DCNN by scanning the American Radiological Society phantom. Subsequently, the images were reconstructed using a filtered back projection algorithm with different reconstruction kernels. The circular edge method was used to measure the MTF values, which were used as teacher information for the DCNN. The resolution properties of the sample CT images used to train the DCNN were created by intentionally varying the field of view (FOV). Four FOV settings were considered. The results of adapting this method to the filtered back projection (FBP) and hybrid iterative reconstruction (h-IR) images indicated highly correlated values with the $MTF_{10\%}$ in both cases. Furthermore, we demonstrated that the RPIs could be estimated in the same manner under the same imaging conditions and reconstruction kernels, even for other CT systems, where the DCNN was trained on CT systems produced by the same manufacturer. In conclusion, the RPI, which is a new index that represents the resolution property using the proposed method, can be used to evaluate the resolution of a CT system in a task-based manner.

Keywords Task-based assessment · Resolution property · Modulation transfer function · Computer tomography · Deep convolutional neural network

1 Introduction

The measurement of the modulation transfer function (MTF) is a popular approach to determine the resolution properties of computed tomography (CT) images [1–3]. However, the linearity of the output values with respect to the dose is not

ensured in nonlinear images such as hybrid iterative reconstruction (h-IR) and model-based iterative reconstruction. In addition, contrast and noise with background materials can affect the MTF measurements, making conventional MTF measurement methods, such as edge and wire methods, unsuitable for evaluating the resolution properties under clinical conditions. Recently, Richard et al. [4] proposed a task-based MTF measurement method using the American College of Radiology (ACR) phantom that was closer to the contrast of the human body and could be measured at various contrasts. Multiple-contrast MTFs were obtained using this method with three rods prepared for Hounsfield unit (HU) accuracy measurements. In addition, multiple combinations of MTFs were obtained by varying the dose. However, with this method, it is necessary to add multiple slices to reduce the influence of noise when the MTF is obtained from low-contrast rods. This leads to a problem where a small amount of misalignment of the phantom causes degradation of the

✉ Aiko Hayashi
hayashia@hiroshima-u.ac.jp

1 Department of Radiology, Hiroshima University Hospital, 1-2-3 Kasumi, Minami-ku, Hiroshima 734-8551, Japan

2 Graduate School of Health Sciences, Kumamoto University, 4-24-1 Kuhonji, Kumamoto 862-0976, Japan

3 Department of Radiological Technology, Faculty of Health Sciences, Okayama University, 2-5-1 Shikatacho, Okayama 700-8558, Japan

4 Faculty of Life Sciences, Kumamoto University, 4-24-1 Kuhonji, Chuo-ku, Kumamoto 862-0976, Japan

MTF [5]. Conversely, Takenaga et al. [6] reported a circular edge method without noise effects using a logistic curve fitting technique. They averaged the edge spread function (ESF) within the bin and fitted it with a logistic function to obtain the MTF without requiring additional multi-slice images. However, this method can only be used for reconstructive kernels in low-frequency regions, and there is undershooting at the edges of the high-frequency functions. Therefore, it is difficult to perform preprocessing to obtain a more accurate MTF, owing to the addition of the acquired images and fitting of the ESF.

A reconstruction technique was developed in the late 2010s to remove noise from CT images by using deep convolutional neural networks (DCNNs). This has led to a breakthrough in the tradeoff between resolution properties and noise, which is a drawback of nonlinear images [7]. Recently, DCNN technology has contributed to the improvement in the operational efficiency by reducing exposure, shortening image reconstruction time, and assisting in diagnosis. There are a few reports on the evaluation of noise properties using DCNNs in the assessment of CT image quality [8]. However, to the best of our knowledge, there are no reports on the direct evaluation of the resolution properties of CT images using DCNNs. Thus, in this study, we propose a simpler method to calculate the indices of image resolution properties with an image input using DCNN to solve the problem of measuring the indices of CT resolution properties, which has become more complicated with the advent of nonlinear images. Although the linearity is essential for the application of conventional methods such as MTF, the resolution properties of recent CT images include nonlinearly processed images. However, the DCNN allows comparison of the resolution properties of CT images, regardless whether the images are nonlinear or linear. To examine the consistency of the proposed method, we evaluated the correlations between the MTF values obtained using the conventional method and the estimated indices.

2 Methods

2.1 Experimental materials

An ACR-certified CT phantom (Model 464, Gammex-RMI, Middleton, WI, USA) was imaged using two CT scanners (CT_A and CT_B). CT_A was a 320-row CT system (Aquilion ONE GENESIS edition, Canon Medical Systems Corporation (CMSC), Otawara, Japan) and CT_B was a 160-row CT system (Aquilion precision, CMSC, Otawara, Japan). The CT images were acquired five times under the conditions listed in Table 1, with nonhelical imaging and reconstructed slice thickness of 5 mm. Axial images were reconstructed using two algorithms: filtered back projection (FBP) and h-IR (adaptive iterative dose reduction 3D, CMSC). The ACR phantom contained one cavity and an encapsulated rod composed of bone mimetic, acrylic, and polyethylene. An acrylic rod with a CT value of 120 HU, which was close to that of the organ, was used for the MTF measurements. As shown in Table 1, the sample images for training the DCNN were obtained by varying the exposure conditions and size of the FOV, while the sample images for testing the DCNN were obtained by varying the reconstruction kernel only. In order to train the DCNN to learn the resolution properties, we acquired images by varying the FOV sizes and then created sample images with different resolution properties while maintaining the apparent image size, as detailed in Sect. 2.2.2. On the contrary, for the test images, sample images with different resolution properties were created by varying the reconstruction kernel to evaluate changes in the resolution properties using our proposed method. The frequency enhancement (i.e., image resolution property) of the Canon CT systems can be changed by varying the reconstruction kernel from FC11 to FC15, centered at FC13, which is the standard kernel for the abdominal region. In addition, the larger the reconstruction kernel, the stronger the frequency enhancement. Therefore, in this study, the standard kernel (FC13) was selected as the image for training the DCNN, and the CT images reconstructed by changing the reconstruction kernel from FC11 to FC15 were used as the test images. Because differences in the measured resolution properties between FBP and h-IR

Table 1 Acquisition parameters used for training and testing images obtained from CT_A and CT_B

	Training	Test
Exposure conditions	SD4 (120 kV, 200 mA, 3 s/rot) SD8 (120 kV, 200 mA, 0.5 s/rot) SD12 (120 kV, 100 mA, 0.5 s/rot)	SD8 (120 kV, 200 mA, 0.5 s/ rot)
Reconstruction Kernel	FC13	FC11,12,13,14,15
Reconstruction FOV	FOV 300, 350, 400, 500 mm	FOV 300 mm
Image reconstruction	FBP, h-IR	

could be attributed to noise, the standard deviation (SD) indices (SD4, SD8, and SD12) were used to vary the dose levels in the training images. Three dose levels were considered. However, to compare the difference in the resolution properties estimated by our proposed method for various reconstruction kernels, all test images were obtained at the SD8 dose level.

2.2 Regression learning of DCNN to estimate index of resolution properties

2.2.1 MTF measurement method

In this study, the MTF of the CT images measured using objective methods was used as the teacher signal in the training of the DCNN. Several methods are used to measure the MTF of CT images, such as the wire method [9, 10] and circular edge method [11]. We used the circular edge method recommended by the American Association of Medical Physics to measure the resolution properties of the CT systems. In addition, this method allows task-based evaluation of nonlinear images [11]. Because the circular edge method requires additive averaging with a large number of sample images to eliminate the effect of noise, we chose the method of Takenaga et al. [6] to eliminate this problem by logistic curve fitting. In this method, a region of interest (ROI) containing a rod of acrylic material was selected, and the oversampled ESF was obtained using the distance from the disk center. The obtained ESFs were averaged and rebinned within a bin one-tenth of the pixel size. A logistic curve fitting method was applied to remove noise from the rebinned ESFs, which is expressed as

$$ESF(x) = \frac{a}{1.0 + \exp\{-b(x-c)\}} + d. \tag{1}$$

where a, b, c, and d are the parameters obtained using the iterative nonlinear least squares method. The line spread function was obtained from the denoised ESF after fitting as described above, and the MTF was obtained by performing Fourier transform. The MTFs obtained from the images

were taken three times and averaged to reduce measurement errors. The MTFs were measured for all training and test images scanned from CT_A and CT_B, and the $MTF_{10\%}$ values were obtained for each MTF. It is noted that $MTF_{10\%}$ indicates the value of the spatial frequency (cycles/mm) at which the MTF value becomes 0.1. In general, the higher the $MTF_{10\%}$ value, the higher the resolution. Because the MTF is not affected by noise in the FBP method, the $MTF_{10\%}$ value obtained from the SD4 image was used as the teacher signal in the training, regardless of the SD setting [4].
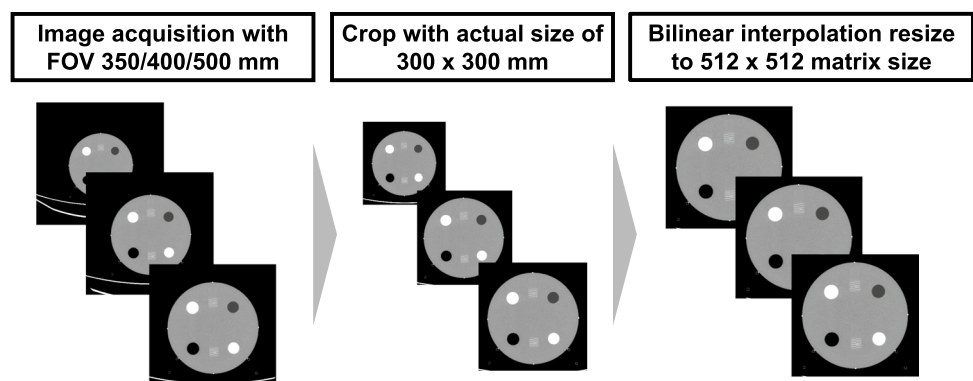
2.2.2 Input sample images for the DCNN

Sample images with the same real and matrix sizes, but different resolution properties, were created to enable the DCNN to learn the resolution property index. These sample images were created using bilinear interpolation from images captured on a CT system at different FOVs. Furthermore, the real MTF of each sample image was obtained using the method described in Sect. 2.2.1. The sample image and the method used to create the input image for the DCNN are described below.

First, the entire phantom was cropped from the images acquired with other FOV sizes with actual dimensions of 300 mm × 300 mm and ROIs with a different matrix size as the 300-mm FOV images. Next, the ROI images with an actual size of 300 mm × 300 mm, cut from the images with an FOV size of 350, 400, and 500 mm, were resized to a matrix size of 512 × 512 pixels using bilinear interpolation (Fig. 1). Thus, the sample images can have the same pixel size as the 300-mm FOV image, but with different resolution properties. Hereinafter, the resolution properties of the sample image acquired with an FOV of 300 mm are defined as the original, and those of the resampled images acquired with an FOV of 350, 400, and 500 mm are defined as high, medium, and low, respectively.

The method used to obtain multiple patch images for the DCNN from a single sample image is shown in Fig. 2. First, the entire 120 HU acrylic rod was cut from the sample image with an ROI (76 × 76 pixels) centered on the

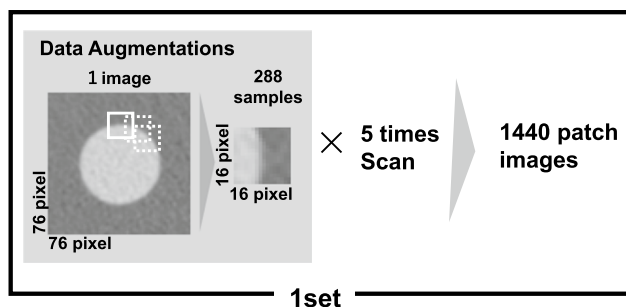Fig. 1 Adjustment with the same pixel size of 300-mm FOV from images acquired with multiple FOVs

| Image acquisition with FOV 350/400/500 mm | Crop with actual size of 300 x 300 mm | Bilinear interpolation resize to 512 x 512 matrix size |

Fig. 2 Multiple patch images created from a single image

rod. The patch images were cut along the circumference of the acrylic rod with a matrix size of 16×16 pixels. Consequently, 288 patch images were obtained in a single shot. Five sample images were acquired for each condition, resulting in $288 \times 5 = 1440$ patch images being prepared for each imaging condition.

2.2.3 DCNN for classification of CT images with different resolution properties

The ability of the DCNN to classify groups of CT images with different resolutions was tested before estimating the CT image resolution property index. A Jupyter notebook was used as the deep learning development tool in Anaconda environment to construct the DCNN. The computer system used comprised an Intel Core i7-11370H processor (Intel, Santa Clara, CA, USA) and NVIDIA GeForce RTX 3050 Ti Laptop GPU (Nvidia, Santa Clara, CA, USA). Windows 10 Home 64-bit with Python 3.7 was used as the operating system, Tensorflow (Google, Mountain View, CA, USA) was used as the framework, and Keras was used as the wrapper. MiniVGG was selected as the original network structure because it has demonstrated high performance in noise classification in previous studies [12]. The DCNN was constructed using a structure search function based on the MiniVGG concept, as shown in Table 2 [13]. The learning parameters were set to 50 epochs, with a learning rate of 0.001 and batch size of 64. Adam was used as the optimization function and categorical cross-entropy was used as the loss function. The images used to construct the DCNN were acquired using CT_A and divided into five sub-sets to apply the holdout method. In the holdout method, one of the five sub-sets was used for testing and the remaining four sub-sets were split in a training-to-validation ratio of 3:1. The FOV sizes of the acquired input images (300, 350, 400, and 500 mm) were used as teacher labels to train the DCNN for classification using input images with different resolutions. Finally, the accuracy of the DCNN was evaluated using the test images.

Table 2 Network structure based on MiniVGG

Layer type	Output size	Filter size/stride
Input image	$16 \times 16 \times 32$	
Conv	$16 \times 16 \times 32$	3×3
ReLu	$16 \times 16 \times 32$	
Conv	$16 \times 16 \times 32$	3×3
Tanh	$16 \times 16 \times 32$	
Bach normalization	$16 \times 16 \times 32$	
Max pooling	$8 \times 8 \times 32$	$2 \times 2/2 \times 2$
Flatten	2048	
Dense	512	
ReLu	512	
Drop out(0.25)	512	
Dense	4	
Softmax	4	

2.2.4 Regression DCNN for estimating indices of the resolution properties of CT images

A part of the DCNN used for classification learning was modified to output an index of the resolution properties of the input CT images. In this modification, the outputs of the last coupling layers of the original DCNN were set to 1 and the last activation function was removed, as shown in Table 3. This resulted in a regression learning DCNN that outputted estimates corresponding to the resolution properties rather than image classification. As in the classification learning case, the learning parameters were set to 50 epochs with a learning rate of 0.001 and batch size of 64. Adam was used as the optimization function and the mean squared error was used to evaluate the learning.

The training images for regression estimation of the DCNN were acquired using CT_A. The input images were the original, high-, medium-, and low-resolution images at

Table 3 Network structure based on MiniVGG for regression

Layer type	Output size	Filter size/stride
Input image	$16 \times 16 \times 32$	
Conv	$16 \times 16 \times 32$	3×3
ReLu	$16 \times 16 \times 32$	
Conv	$16 \times 16 \times 32$	3×3
Tanh	$16 \times 16 \times 32$	
Bach normalization	$16 \times 16 \times 32$	
Max pooling	$8 \times 8 \times 32$	$2 \times 2/2 \times 2$
Flatten	2048	
Dense	512	
ReLu	512	
Drop out(0.25)	512	
Dense	1	

three dose levels (SD4, SD8, and SD12), as shown in Fig. 3. The total number of patch images used to train the DCNN was 17,280 images (1440×4 FOVs $\times 3$ SDs). The training-to-validation ratio was 3:1, and the $MTF_{10\%}$ value measured using the circular edge method was used as the teacher signal. Test images of different resolutions obtained using five reconstruction kernels (FC11–FC15) were used as the input images to evaluate the DCNN trained in this manner. The five outputs of the DCNN [i.e., the resolution property index (RPI)] for each test image were used to evaluate the correlation with the real $MTF_{10\%}$ values.

2.3 Application of the proposed method

2.3.1 Application of the proposed method for nonlinear images

It has been reported that the MTF may exhibit different values in nonlinear images when there are changes in noise [4]. Therefore, it is often difficult to evaluate the resolution of nonlinear images using the MTF. To examine the usefulness of the proposed method for evaluating the resolution properties of nonlinear images, a nonlinear h-IR image was inputted into the DCNN, which was constructed and trained according to the procedure outlined in Sect. 2.2. As in the case of the FBP, the output of the DCNN is the RPI, which represents the resolution properties of the CT image, and this RPI can be used to compare the FBP (which is a linear image) with the h-IR (which is a nonlinear image). In addition, the changes in the DCNN output RPI were compared with the results obtained from the FBP image.

2.3.2 Application of the proposed method for CT images not used for training

In general, the resolution properties of CT systems must be calibrated and evaluated for each system, which is a complicated task for facilities with multiple CT systems. Therefore, the proposed method is highly useful if it uses a DCNN constructed from images obtained from a single CT system to evaluate the resolution properties of other systems. The usefulness of the proposed method for evaluating the resolution properties of images obtained using the CT_B model that was different from that used to construct the DCNN was examined, even if the CT scanners were produced by the same manufacturer. Thus, the output of this DCNN is an RPI that represents the resolution properties of the input image of CT_B, but the DCNN itself was trained on the CT_A image.

3 Results

3.1 MTF measurements

The MTF values for the CT images scanned with CT_A when the resolution was varied at four levels (original, high, medium, and low) are shown in Fig. 4. The worse the resolution properties (larger FOV size), the worse the MTF. The $MTF_{10\%}$ results obtained for each MTF (Fig. 4) are shown in Fig. 5. Based on the $MTF_{10\%}$, it was confirmed that the original showed the highest value, and the results decreased in the following order: high, medium, and low. The edges of the rods were also sharpened in the order of original, high, medium, and low, as shown in Fig. 6. The $MTF_{10\%}$ values obtained from the CT images scanned by varying the reconstruction kernels from FC11 to FC15 are shown in Fig. 7. The higher the number of reconstruction kernels, the more frequencies
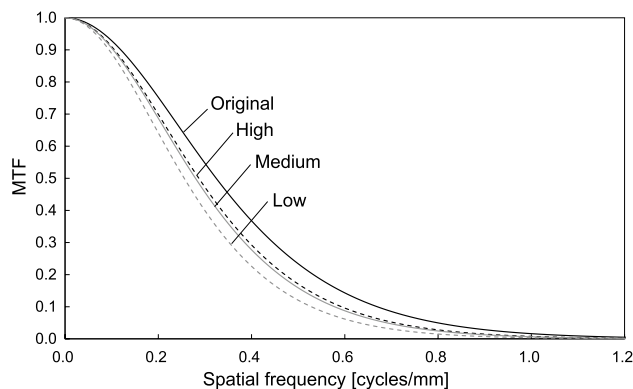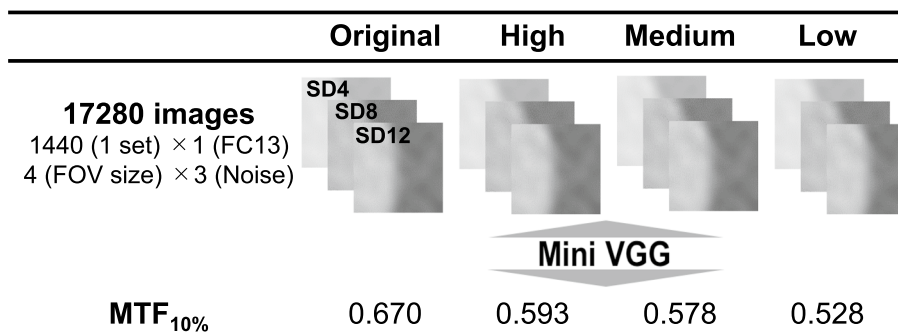

Fig. 4 MTF values obtained for four different FOV settings

Fig. 3 Image dataset for training and $MTF_{10\%}$

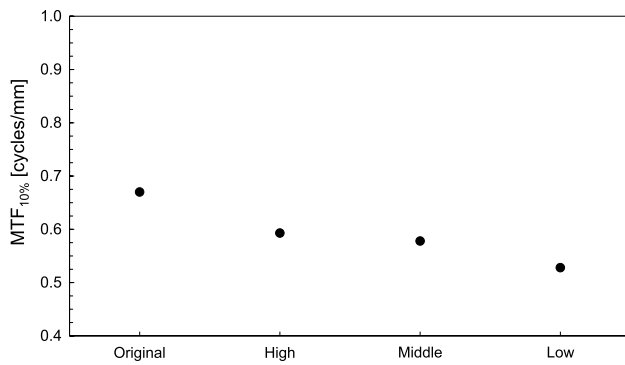|	Original	High	Medium	Low
17280 images 1440 (1 set) $\times 1$ (FC13) 4 (FOV size) $\times 3$ (Noise)	SD4 SD8 SD12			
		Mini VGG		
$MTF_{10\%}$	0.670	0.593	0.578	0.528

Fig. 5 Relationship between the $MTF_{10\%}$ and resolution property level obtained for different FOV settings
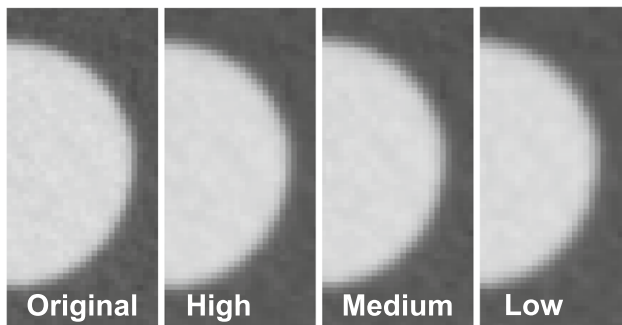

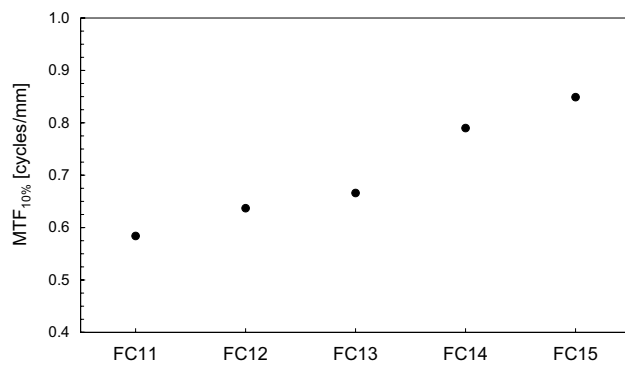**Fig. 6** Rod images acquired for various FOVs (FC = 13)

Fig. 7 Relationship between the $MTF_{10\%}$ and kernel filters

were emphasized, indicating that the $MTF_{10\%}$ tended to be higher. In addition, as shown in Fig. 8, the higher the number of reconstruction kernels, the sharper the edges of the rods.

3.2 Classification of CT images with different resolution properties

The accuracies of the DCNN constructed to classify the CT images at four different resolutions using the confusion
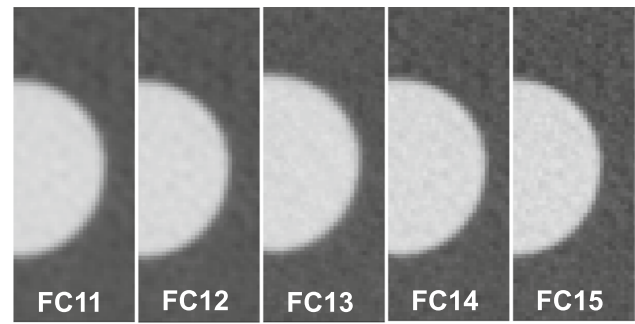
Fig. 8 Rod images acquired for various kernels (FOV = Original)

matrix are listed in Table 4. The rows and columns of the confusion matrix indicate the actual and predicted resolution classes, respectively. Multiclass accuracy, which divides the number of correct predictions by the total number of records, was used as the evaluation index. The classification accuracies were 99.5, 99.8, 99.7, and 100% for the original, high, medium, and low, respectively, with a high accuracy of 99.7% for all labels included.

3.3 Estimation of indices of the resolution properties of CT images

The correlation between the $MTF_{10\%}$ and RPI values estimated from the regression DCNN by varying the reconstruction kernel from FC11–FC15 is shown in Fig. 9. It should be noted that the $MTF_{10\%}$ and RPI values were obtained from the CT images captured under the same imaging conditions, whereas the DCNN was constructed and trained using different image sets. A strong positive correlation ($R^2 = 0.9233$) was observed between the $MTF_{10\%}$ and RPI values.

3.4 Application results using the proposed method

3.4.1 Application results for nonlinear images using the proposed method

The $MTF_{10\%}$ values of the images reconstructed from the SD4, SD8, and SD12 scan data using the FBP and h-IR methods are shown in Fig. 10. Compared with the FBP reconstruction, the h-IR reconstruction tended to decrease the $MTF_{10\%}$ with a decrease in dose (increasing noise). Similar to the FBP reconstruction, the resolution tended to decrease in the order of high, medium, and low, where the original had the highest value.

The $MTF_{10\%}$ values for the images reconstructed from the SD8 scan data using the FBP and h-IR methods with five different reconstruction kernels are shown in Fig. 11. As in the case of the FBP, the higher the number

Table 4 Confusion matrix showing DCNN classification accuracies for four resolution property levels

	Original	High	Medium	Low	Accuracy (%)	Overall accuracy (%)
Original	860	4	0	0	99.5	99.7
High	2	862	0	0	99.8	
Medium	0	0	861	3	99.7	
Low	0	0	0	864	100	
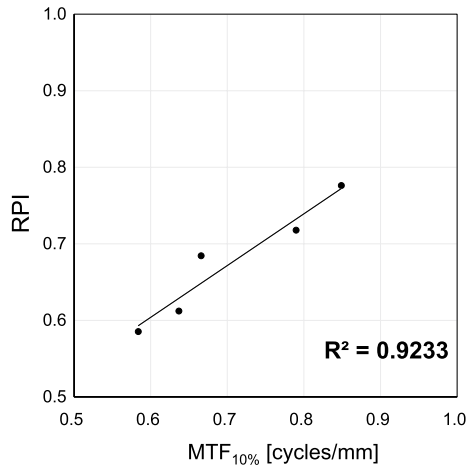
Fig. 9 Comparison between the estimated RPI and real $MTF_{10\%}$

Fig. 11 Relationship between the $MTF_{10\%}$ and kernel filters for FBP and h-IR CT images
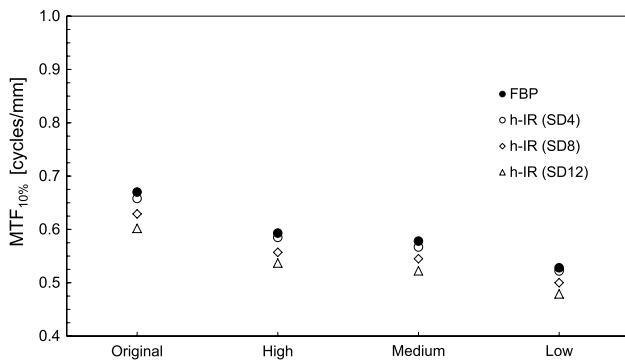
Fig. 10 Relationship between the $MTF_{10\%}$ and resolution property level for FBP and h-IR CT images

Fig. 12 Comparison between the estimated RPI and $MTF_{10\%}$ for h-IR CT images

of reconstruction kernels, the higher the $MTF_{10\%}$ value of the image reconstructed using the h-IR method.

The correlation between the $MTF_{10\%}$ values of the images acquired by changing the reconstruction kernel from FC11 to FC15 using the h-IR method and the RPI estimated from the same images using the proposed method is shown in Fig. 12. The $MTF_{10\%}$ and output RPI values showed a strong positive correlation ($R^2 = 0.9646$) with the h-IR-reconstructed images.
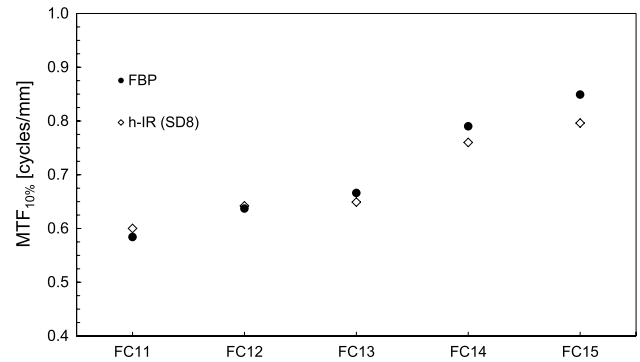
3.4.2 Application results for CT_B images that were not used for training

The $MTF_{10\%}$ values of the CT_B sample images (SD8) reconstructed with five different reconstruction kernels using the FBP and h-IR methods are shown in Fig. 13. It is noted that the CT_B images were not used for DCNN training. Similar to the CT_A, a higher reconstruction kernel value tended to result in higher $MTF_{10\%}$ values. However, the difference in resolution between the FBP and h-IR methods was smaller than that of the CT_A.
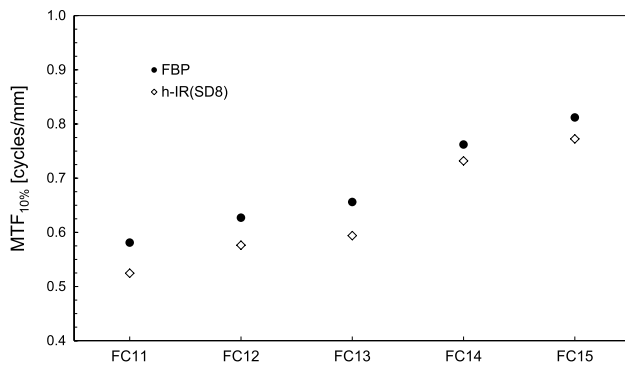
Fig. 13 Relationship between the MTF$_{10\%}$ and kernel filters for FBP and h-IR CT images obtained from a different CT scanner

The correlation between the MTF$_{10\%}$ obtained from the CT_B image and the RPI obtained from the DCNN trained on the input image obtained from CT_A for the same image is shown in Fig. 14. Similar to the results of CT_A, a strong positive correlation was observed between the MTF$_{10\%}$ and output RPI values for the FBP and h-IR reconstruction methods, as shown in Fig. 14a and b, respectively.

4 Discussion

In this study, we used a method to intentionally degrade the resolution at the same matrix size using bilinear interpolation after imaging with various FOV sizes to quantitatively change the resolution of the CT images, as described in Sect. 2.2.2. As a result of obtaining the MTF (using the conventional circular edge method) for the sample image (obtained using the proposed method), we verified that the measured MTF changed with respect changes in the FOV, as predicted and discussed in Sect. 3.1. Therefore, we believe that the images with different resolutions were valid for the purposes of this study.
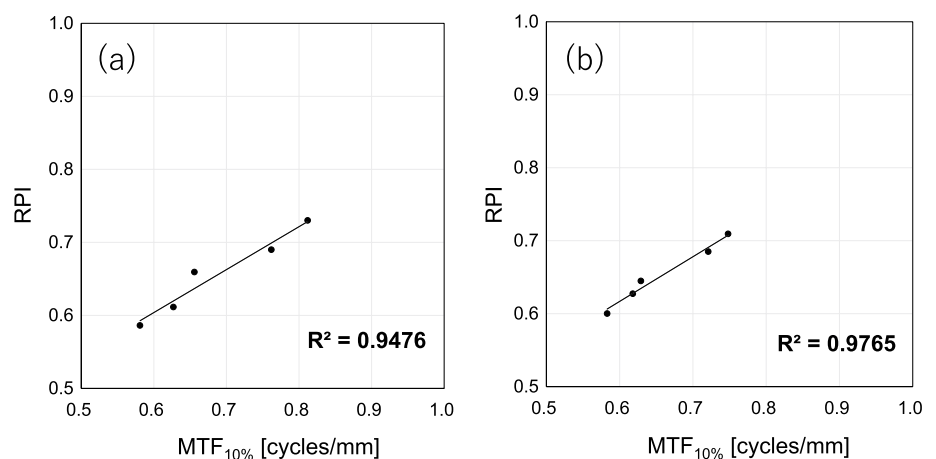
The parameters used to construct the DCNN such as the convolution layer, activation function, and loss function can be changed. Therefore, the structure search function was used to adjust the combination of parameters to vary the performance [14]. In Sect. 3.2, we confirmed that the DCNN optimized using the structural search function could accurately classify the resolution properties by learning the resolution of intentionally degraded images.

The DCNN was trained using sample images whose resolution properties were modified by changing the FOV. In addition, it was confirmed that the DCNN could output resolution property indices for images whose resolution was changed by changing the reconstruction kernel provided by the CT system. One of the problems with deep learning is that the basis for deriving a solution is a black box, in which the two points of concern are the effects of bilinear interpolation and noise on learning. For bilinear interpolation, it is possible that factors other than resolution were learned because images with different resolutions were used during learning. We believe that this concern was eliminated using a test image with a resolution that varied with the frequency intensity of the reconstruction kernel and by outputting an RPI in the same manner. In addition, it is possible that the position dependence of noise was acquired during learning. Sugino et al. [12] reported that images considering the position dependence of noise in the dataset improved the learning results compared with those that did not consider the position dependence of noise. Therefore, the influence of the position dependence of noise was eliminated as much as possible by acquiring and training images with different resolutions by changing the FOV. The fact that the network could estimate the RPI by changing the dose level, even for images that exhibited different noise behaviors, confirmed that the variation in resolution properties with FOVs was the best way for the network to learn the resolution properties of an image.

The proposed method was used to compare the FBP (linear) and h-IR (nonlinear) images, making it possible

Fig. 14 Comparison between the estimated RPI and MTF$_{10\%}$ on (**a**) FBP and (**b**) h-IR CT images obtained from a different CT scanner

to estimate indices that showed a high correlation with the existing MTF values in both cases. The nonlinear images showed slightly lower properties than the linear images when the RPIs were compared. However, we believe that this can be attributed to the decreasing MTF of h-IR with decreasing dose in low-contrast regions such as acrylic rods, as reported by Higaki et al. [7]. Similarly, as shown in Fig. 11, as the reconstruction kernel increased, the difference in the MTF values between FBP and h-IR increased because of the increase in noise caused by frequency enhancement. Thus, as the reconstruction kernel increases, the MTF value of FBP (which is less susceptible to noise) increases and the difference between FBP and h-IR (which is more susceptible to noise) increases.

In general, the resolution properties of nonlinear images vary owing to the noise and contrast. Therefore, measuring the resolution properties using conventional methods requires considerable time and effort. The proposed method is likely to simplify the RPI measurement of nonlinear images, because it can estimate the RPI even if the shooting conditions and functions are changed in a single training session.

The experimental results confirmed that the proposed method estimated RPIs that demonstrated a high correlation with the existing MTF values as well as with the results estimated using the same equipment. This was true even if the equipment used was different from that used for training, provided that the equipment used was produced by the same manufacturer. We believe that this will improve the efficiency of resolution property measurements in hospitals with multiple CT systems. However, the adaptation of this method between CT systems produced by different manufacturers has not been confirmed, and future studies are needed to investigate learning with images that include images from CT systems produced by different manufacturers.

The proposed method offers a high degree of freedom in the shape of the object to be measured. Therefore, we believe that the resolution of CT images can be evaluated using, for example, a human phantom that is more similar to the human body or the clinical image itself. Therefore, optimization of the exposure dose using an evaluation index of the resolution property acquired under more clinically relevant conditions is a future challenge. Next, we would like to use a DCNN trained on phantom images to estimate the resolution properties using sample images extracted directly from clinical images rather than from phantom images. We intend to use this DCNN to evaluate high-resolution clinical images required in orthopedics and other fields.

The limitation of this method is that the higher the resolution of the image input to the DCNN, the more the RPI is underestimated compared with the actual MTF. FC13 was used for all training samples, and therefore, the maximum value of the FC13 resolution property may have affected the maximum estimate. Therefore, it is necessary to use training samples with higher resolution properties to improve the training accuracy.

5 Conclusion

In this study, we proposed a method for obtaining a new index to evaluate the resolution properties of CT images in a task-based manner when the reconstruction method, function, and dose (image noise) were varied. The resolution property index obtained by the proposed method using DCNN was confirmed to be highly correlated with the $MTF_{10\%}$ values obtained by the conventional method. The proposed method is expected to improve the efficiency of measuring the resolution properties of CT images.

Declarations

References

1. Samei E, Flynn MJ, Reimann DA. A method for measuring the presampled MTF of digital radiographic systems using an edge test device. Med Phys. 1998;25(1):102–13.
2. Samei E, Ranger NT, Dobbins JT III, Chen Y. Intercomparison of methods for image quality characterization I. Modulation transfer function. Med Phys. 2006;33(5):1454–65.
3. Judy PF. The line spread function and modulation transfer function of a computed tomographic scanner. Med Phys. 1976;3(4):233–6.
4. Richard S, Husarik DB, Yadava G, Murphy SN, Samei E. Towards task-based assessment of CT performance: system and object MTF across different reconstruction algorithms. Med Phys. 2012;39(7 Part1):4115–22.
5. Friedman SN, Fung GS, Siewerdsen JH, Tsui BM. A simple approach to measure computed tomography (CT) modulation transfer function (MTF) and noise-power spectrum (NPS) using the American College of Radiology (ACR) accreditation phantom. Med Phys. 2013;40(5): 051907.
6. Takenaga T, Katsuragawa S, Goto M, Hatemura M, Uchiyama Y, Shiraishi J. Modulation transfer function measurement of CT images by use of a circular edge method with a logistic curve-fitting technique. Radiol Phys Technol. 2015;8:53–9.
7. Higaki T, Nakamura Y, Zhou J, Yu Z, Nemoto T, Tatsugami F, Awai K. Deep learning reconstruction at CT: phantom study of the image characteristics. Acad Radiol. 2020;27(1):82–7.
8. Hayashi N, Maruyama T, Sato Y, Watanabe H, Ogura T, Ogura A. Evaluating medical images using deep convolutional neural networks: a simulated CT phantom image study. Technol Health Care. 2020;28:113–20.

9. Nickoloff EL. Measurement of the PSF for a CT scanner: appropriate wire diameter and pixel size. Phys Med Biol. 1998;33(1):149–55.

10. Ichikawa K, Hara T, Niwa S, Ohashi K. Method of measure modulation transfer function using metal wire in computed tomography. Nippon Hoshasen Gijutsu Gakkai Zasshi. 2008;64(66):672–80 (**in Japanese**).

11. Samei E, Bakalyar D, Boedeker KL, Brady S, Fan J, Leng S, Myers KJ, Popescu LM, Ramirez Giraldo JC, Ranallo F, et al. Performance evaluation of computed tomography systems: summary of AAPM Task Group 233. Med Phys. 2019;46(11):735–56.

12. Sugino S, Shiraishi J. Dose evaluation of CT images applying image noise classification by convolutional neural network: a fundamental study using phantoms. Nippon Hoshasen Gijutsu Gakkai Zasshi. 2020;76(11):1143–51 (**in Japanese**).

13. Ismail A, Ahmad SA, Soh AC, Hassan K, Harith HH. Improving convolutional neural network (CNN) architecture (MiniVGGNet) with batch normalization and learning rate decay factor for image classification. Int J Integr Eng. 2019. https://doi.org/10.30880/ijie.2019.11.04.006.

14. Zeng X, Luo G. Progressive sampling-based Bayesian optimization for efficient and automatic machine learning model selection. Health Inform Sci Syst. 2017;5:1–21.