IEEE *Access*
Multidisciplinary : Rapid Review : Open Access Journal

Visibility Enhancement of Lesion Regions in Chest X-ray Images with Image Fidelity Preservation

Ryoichi Ishikawa[1], (Member, IEEE), Tomohisa Yuzawa[1,2], Taiki Fukiage[3], Masataka Kagesawa[1], Toru Watsuji[2], and Takeshi Oishi[1], (Member, IEEE)
[1]Institute of Industrial Science, The University of Tokyo, Komaba, Meguro-ku, Tokyo, Japan, 153-8505
[2]Air Water Inc., Minami Semba, Chuo-ku, Osaka, Japan, 542-0081
[3]NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation, Morinosato-Wakamiya, Atsugi-shi, Kanagawa-ken, Japan, 243-0198

Corresponding author: Ryoichi Ishikawa (e-mail: ishikawa@cvl.iis.u-tokyo.ac.jp).

ABSTRACT X-ray image enhancement can aid a physician's diagnosis by improving lesion visibility. This study proposes a chest X-ray image enhancement framework for enhancing lesion visibility while preserving image features. Our framework assesses the background signals, whereas conventional methods focus on the visibility of the global image. The proposed method predicts the image processing parameters that enhance the lesion signals via the inference neural network. The framework consists of an X-ray image enhancer and an enhanced model predictor for reference. The enhancer regressively estimates the processing parameters for enhancing the lesions using the inference network and processes the input X-ray image. As the inference network requires training, the model predictor computes the reference parameters that maximize the visibility of the lesions within a tolerable loss of fidelity using image pairs—with and without lesions. We created a synthesized dataset, with and without lesions, from healthy chest and phantom lesion X-ray images. The experiments show that after the proposed method was trained on 2000 images, it improved lesion visibility with an acceptable fidelity loss. We also performed pairwise comparisons and confirmed that trade-offs between fidelity loss and visibility gain were attained. A technique for improving lesion visibility while maintaining the fidelity of X-ray images was developed. This method enabled the enhancement of specific signals in the background. Various image processing methods that require parameters could be incorporated into this framework for many different applications.

INDEX TERMS Chest X-Ray, Human Perception, X-Ray Image Enhancement,

I. INTRODUCTION

Chest X-Ray (CXR) images are commonly used to diagnose respiratory disease [1]–[3]. X-ray imaging offers advantages in terms of cost, accessibility, and sensing speed compared with CT or MRI. These images are well-suited to providing an overview of the lungs and to evaluating lung conditions. However, the lesion signal may be buried in the image because bones, muscle tissue, and blood vessels are all being observed [4].

In medical practice, it is common for physicians, rather than computers, to make the final diagnosis. The recent development of deep learning technology produced many methods for enhancing radiographs [5]–[8]. Although the supervised machine learning approaches automatically detect the presence and locations of lesions in diagnostics [9]–[17], particularly in CXR images [14]–[17], these methods require extensive training on annotated data. It is also difficult to per-

form a diagnosis with perfect accuracy due to large variations in the images, which depend on the patient's physique and the shooting conditions. As complete automatic detection is impractical, doctors must make a final diagnosis from the visual cues in the image.

X-ray enhancement preserves image features and contributes to the diagnosis by improving lesion visibility. Several image enhancement techniques have been discussed in image processing research, including contrast adjustment, histogram optimization, noise reduction [18], [19], and super-resolution [20], [21]. Methods that automatically adjust parameters for image filters are also applicable to X-ray images. However, image filters are not designed to enhance only certain regions, i.e., lesions [22]–[24]. If an image filter emphasizes visual features other than the lesion, the signal to be observed may become buried. Moreover, some filters may hide features per the given parameter. Put differently,

by enhancing the visual features of the lesion, we help the physician greatly during the diagnostic process.

This study proposes a technique for improving the human visibility of lesions using a deep learning framework using a prediction model. Although it is difficult to know the lesion's position in advance, learning its spatial distribution, shape, and frequency characteristics should be possible. In addition, a human visibility prediction model enables us to recognize regions that match these characteristics. The proposed method also requires the preservation of image features based on the fidelity metric to prevent missed lesions.

The contributions of this paper are as follows:

- Proposal of a learning-based X-ray image processing framework for improving visibility while maintaining fidelity.
- Presentation and evaluation of X-ray imaging algorithms by visibility using simulated data and lesion models.
- Implementation of differentiable human perception-based visibility and image fidelity evaluation modules.

The visibility prediction model originally aimed for Image Quality Assessment (IQA) to detect noises perceived by human visual systems in image processing, such as compression. The proposed method considers lesions as noise and estimates their visibility by comparing images with and without lesions. However, capturing such pairs of X-ray images would be impossible. Therefore, we used composite X-ray images of lesion phantoms and healthy lungs to train our model.

The remainder of this manuscript is as follows: Section II provides an overview of the related work. We explain the visibility-based X-ray image enhancer and the reference model predictor in Section III and IV, respectively. The experimental results are presented in Section V. Section VI discusses the results, and we conclude the study in Section VII.

II. RELATED WORK

This section briefly reviews general image enhancement methods, as well as those specific to medicine, and discusses the positions of the proposed methods.

A. GENERAL IMAGE ENHANCEMENT

General image enhancement has been studied extensively. Various methods have been proposed, including intensity-based, retinex-based, frequency-based, and learning-based methods.

1) Intensity-based method

Intensity-based methods adjust the image contrast based on pixel intensity, using techniques like histogram equalization (HE) [25]–[29] and intensity conversion [30]–[32]. HE rearranges the pixel values using the cumulative distribution function so that the output of the intensity histogram becomes equalized. Intensity conversion, including tone mapping, is another option for enhancing images. These methods define mapping functions that assign output intensity values to the input. The most commonly available functions are the gamma function and logarithmic function [33].

2) Retinex-based methods

The retinex theorem [34] models the way humans perceive colors or illuminations. The observed optical information is divided into a reflection component, which is derived from the scene objects, and an illumination component. The image is sharpened by separating and enhancing these components. The retinex method separates reflections using a Gaussian kernel for a single size [35], multiple sizes [36], or multiple channels [37] and applies a contrast enhancement filter to the reflections. This theorem has also been integrated into Deep Neural Network-based methods [38].

3) Frequency-based methods

Frequency-based methods include image processing in the frequency-domain and frequency-based decomposition. Frequency-domain methods use the Fourier transform to create an image and then apply a homomorphic filter to emphasize high-frequency components. They then use the inverse Fourier transform to reconstruct the image [39]. Image decomposition methods decompose and reconstruct images by filtering low-frequency components [40] or by emphasizing high-frequency components such as edges. For example, bilateral filter [41], guided filter [42], and weighted least square filter [43] are popular image decomposition methods.

4) Learning-based methods

Many learning-based methods have been proposed in recent years. These supervised methods must train on a dataset that includes before-and-after enhancement [44]–[46]. The expected input images are generated by downsampling, degrading, or adding noise to achieve super-resolution, edge enhancement, and noise reduction. The Generative Adversarial Network (GAN) is an effective framework for image enhancement. Some supervised approaches use data pairs and employ GANs in their loss functions [47], [48]. However, GAN-based unsupervised approaches do not require the preparation of image pairs. A generator provides an enhanced image from the input, and the discriminator uses reference images to determine whether the image is in the desired domain [49]–[51].

B. X-RAY IMAGE ENHANCEMENT

Based on the abovementioned studies related to general image enhancement methods, we describe methods for X-ray imaging.

1) Retinex-based method

The retinex method is effective for X-ray image enhancement. Chen and Zou [52] used the two-scale retinex method for bright and dark regions in CXR images. YangDay and Zhang [53] proposed a single-scale retinex method with segmentation and weighting schemes to prevent noise amplification and the over-enhancement of artifacts.

2) Decomposition method

Some methods improve X-ray image visibility by decomposing the features of both the target and nontarget features. Huang and Nguyen [54] introduced a technique for improving X-ray images by attenuating their tissue components. A method known as ATACE [24] decomposes the tissue components and details to enhance image contrast. The bone suppression method [55], [56] is another approach for improving the visibility of the region of interest. It removes the bone structure from the original X-ray image. Madmad et al. [57] proposed a CNN approach to decomposing X-ray images into base and fine structures using a synthetic dataset.

3) Frequency-based method and other methods

Frequency-based methods are popular in X-ray image enhancement. Xiao *et al.* [39] combined homomorphic filtering methods and noise reduction for X-ray image enhancement. Kim et al. [58] proposed a CNN-based tone mapping method that involved generating a detailed reference layer through X-ray image synthesis. Madmad and Vleeschouwer [27] emphasized texture and shape components by utilizing a bilateral HE operator. Anand et al. [59] trained a CNN module using high-contrast images obtained from a contrast-limited adaptive HE (CLAHE) algorithm.

Other X-ray image enhancement techniques include denoising [18], [19], [60], and super-resolution [20], [21]. Some studies have explored the application of image enhancement methods as a preprocessing step in lesion detection [7], [8], [61].

C. POSITION OF THE PROPOSED METHOD

Our method is a framework for evaluating and enhancing the visibility of specific regions in the input X-ray image, whereas conventional methods evaluate the whole image. We use a method that quantifies the visibility, based on human visual perception (HVP), of the difference between two images. This method also maintains visual features, including tissues and bones, to prevent misses. Like some image enhancement methods, it requires tuning of the parameters; however, these are intuitive parameters that determine image fidelity.

III. VISIBILITY-BASED X-RAY IMAGE ENHANCER

The proposed method uses X-ray images as input, estimates the parameters that improve lesion visibility, and outputs images processed according to these parameters. Fig. 1 shows an outline of the proposed framework. The image enhancer comprises an image processing module and a parameter predictor network module, as shown in the figure. This section explains each image processing module and the estimation network. The following Section IV covers the process of estimating the chosen parameters to train the predictor network.

We assume that the input X-ray image displays only the lung region, which we extracted to eliminate the differences between the imaging environments and between individual physical dimensions as much as possible, using the segmentation labels supplied with the dataset.

A. IMAGE PROCESSING MODULE

Although various algorithms can be used for the image processing module, our proposed framework uses an algorithm that combines high degrees of freedom intensity adjustment and frequency adjustment. Intensity adjustment employs tone curve correction represented by Radial Basis Functions (RBFs), whereas frequency adjustment employs image reconstruction using Laplacian image sequences.

1) Intensity adjustment

The intensity adjustment defines a tone curve $\hat{\eta} = C(\eta)$ for the input pixel intensity value η of an input image $\mathcal{I}$ and the corresponding output intensity value $\hat{\eta}$ of the output image $\hat{\mathcal{I}}$. Assuming that ϕ_i is the ith RBF and that w_i and c_i are the weight and center of ϕ_i, respectively, we define the tone curve as follows:

$$C(\eta) = \sum_{i=1}^{N_{\mathrm{rbf}}} w_i \phi_i(|\eta - c_i|), \qquad (1)$$

where N_{rbf} denotes the number of RBFs. In this experiment, we set $N_{\mathrm{rbf}} = 4$; w_i and c_i are the parameters to be optimized. Note that the input pixel intensity values are normalized by the maximum intensity value in the dataset. The diagram on the left in Fig. 2 shows a tone curve defined by an RBF.

We adopt a multiple harmonic spline function for the RBFs:

$$\phi_i(x) = \begin{cases} x^i & (i = 1, 3, ...) \\ x^i \ln x & (i = 2, 4, ...) \end{cases} \qquad (2)$$

By changing the values w_i and c_i of ϕ_i, we can modify the tone curve as shown in the image in the upper right of Fig. 11. In other words, we seek the values of w_i and c_i that achieve the optimal intensity change for the input.

2) Per-frequency contrast adjustment

The contrast adjustment uses a Laplacian pyramid [62], which decomposes the image into individual frequency components derived from the Gaussian pyramid. We enhance or attenuate the frequency components by multiplying the coefficient parameters for the Laplacian images during the reconstruction process.

Let $\mathcal{G}_1$ be the input image ($\mathcal{I} \rightarrow \mathcal{G}_1$) and let $\{\mathcal{G}_j\}(j = 1, 2, ..., N_{\mathrm{gn}})$ be the Gaussian pyramid of the input. N_{gn} is the number of pyramid levels, and we set $N_{\mathrm{gn}} = 5$ for the experiment. In addition, let $\{\mathcal{L}_j\}$ be the Laplacian pyramid consisting of $N_{\mathrm{gn}} - 1$ images. We generate $\mathcal{L}_j$ by subtracting from the Gaussian image $\mathcal{G}_j$ the image upscaled by one level above, $\mathcal{G}_{j+1}$:

$$\mathcal{L}_j = \mathcal{G}_j - U(\mathcal{G}_{j+1}), \qquad (3)$$

where $U(\mathcal{G})$ indicates the image up-sampling operation of $\mathcal{G}$ by nearest interpolation.

Per-frequency contrast adjustment multiplies the coefficient parameter μ_j against $\mathcal{L}_j$ when the image is reconstructed,
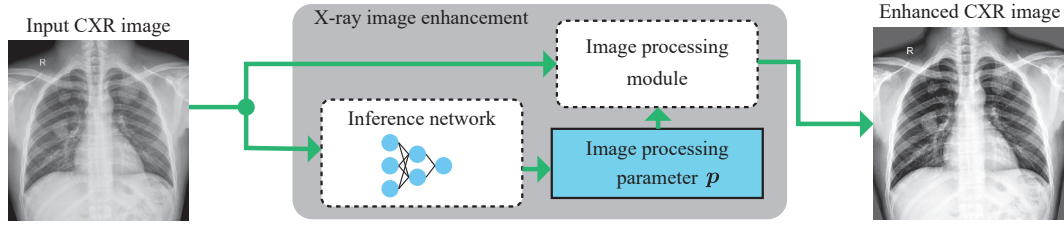
FIGURE 1. X-ray image enhancer estimates the parameters that the image processing module will use to enhance the original CXR image while preserving its features.
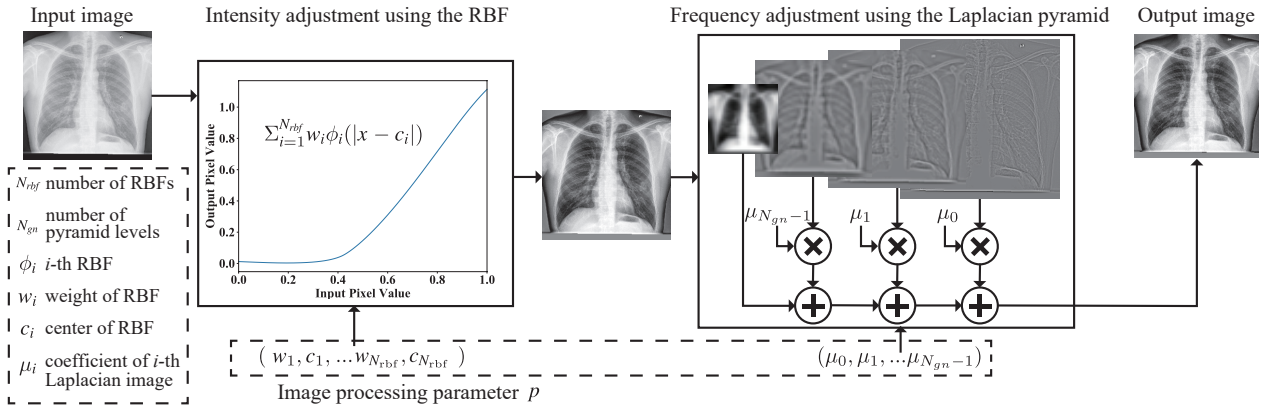

FIGURE 2. Image processing module. The pixel brightness of the input image is converted by the RBF. Next, the converted image is decomposed into a Laplacian pyramid and then reconstructed by multiplication of the frequency components by coefficients.
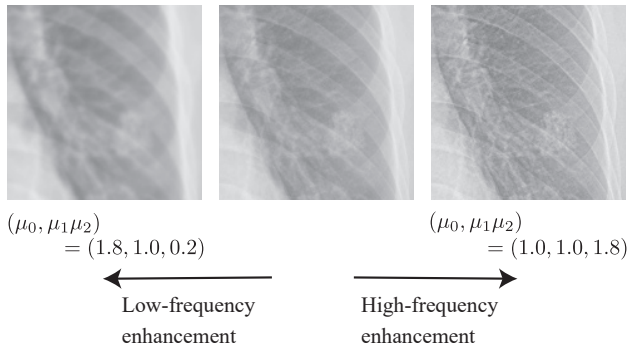

FIGURE 3. Examples of frequency adjustment processing

and this adjustment either enhances or attenuates the corresponding frequency component, as follows:

$$\hat{\mathcal{G}}_j = \mu_j \mathcal{L}_j + U(\mathcal{G}_{j+1}). \qquad (4)$$

Here, μ_j is the parameter to be estimated. We obtain the desired image $\hat{\mathcal{G}}_1$ by reconstructing from the top to the bottom levels of the pyramid. The right middle image in Fig. 2 shows the frequency adjustment process. Adjusting μ_j for each level of the Laplacian image allows us to enhance or suppress specific frequency components. Fig. 3 presents examples that

demonstrate this frequency adjustment.

B. PREDICTOR NETWORK

The estimator is a deep learning model that takes X-ray images as input and regressively estimates the image processing parameters, $\boldsymbol{p} = \{w_1, ..., w_{N_{\mathrm{rbf}}}, c_1, ..., c_{N_{\mathrm{rbf}}}, \mu_1, ..., \mu_{N_{\mathrm{gn}-1}}\}$, that improve lesion visibility while maintaining fidelity with the original image. The training process requires both the X-ray images and the corresponding reference parameters, $\hat{\boldsymbol{p}} = \{\hat{w}_1, ..., \hat{w}_{N_{\mathrm{rbf}}}, \hat{c}_1, ..., \hat{c}_{N_{\mathrm{rbf}}}, \hat{\mu}_1, ..., \hat{\mu}_{N_{\mathrm{gn}-1}}\}$ (see details in Section IV), for image processing, as shown in Fig. 4. We use the Swin-Transformer V2 model [63] as the network architecture.

We calculate the losses for the intensity adjustment l_{int} and frequency adjustment l_{freq} separately during the training process. The total loss is the summation of these losses, where each have equal weight:

$$l_{\mathrm{total}} = l_{\mathrm{int}} + l_{\mathrm{freq}}. \qquad (5)$$

1) Intensity adjustment

We calculate the loss for the intensity adjustment by taking the differences in the estimated and reference tone curves: $C(\eta)$ and $\hat{C}(\eta) = \sum_{i=1}^{N_{\mathrm{rbf}}} \hat{w}_i \phi_i(|\eta - \hat{c}_i|)$. The parameters for the RBF functions w and c derive the tone curves. The loss
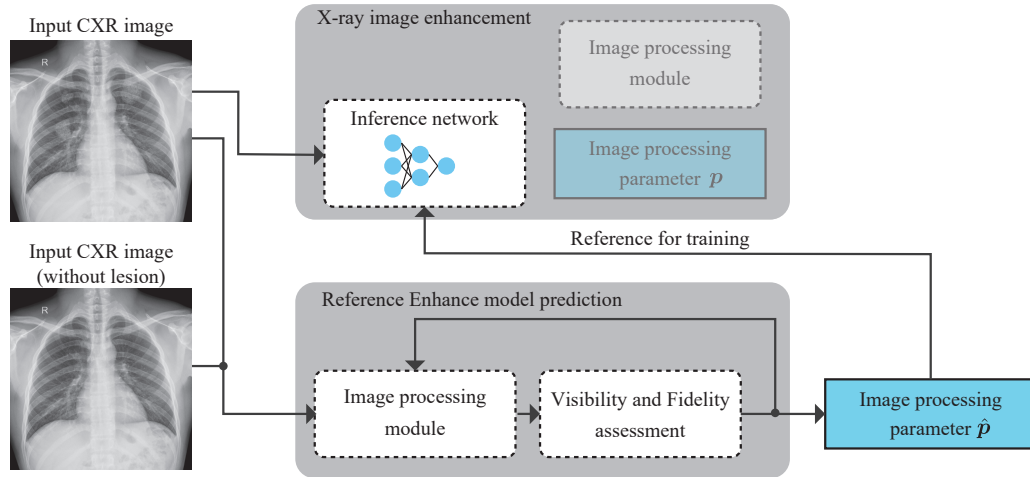
IEEE *Access*

FIGURE 4. Inference network is trained using reference parameters computed with image pairs that have and do not have lesions.

between the curves is calculated as shown below:

$$l_{\text{int}} = \frac{1}{N_{\text{s}}} \sum_{i=0}^{N_{\text{s}}-1} \left(C(\eta_i) - \hat{C}(\eta_i)\right)^2, \tag{6}$$

where N_{s} is the number of samplings. We determine the sampling range and intervals by dividing the range equally between the minimum and maximum intensity values in the lung region. We set $N_{\text{s}} = 12$ for the experiment.

2) Frequency adjustment

In the case of the frequency adjustment parameters, we simply take the mean squared error between the estimated parameter μ_j and the reference parameter $\hat{\mu}_j$:

$$l_{\text{freq}} = \frac{1}{N_{\text{gn}} - 1} \sum_{j=1}^{N_{\text{gn}}-1} (\mu_j - \hat{\mu}_j)^2. \tag{7}$$

IV. REFERENCE ENHANCED MODEL PREDICTOR

We need the image processing reference parameters $\hat{\boldsymbol{p}}$ for the corresponding input image to train the model. The parameters must improve the visibility of the lesion region while maintaining a certain degree of image fidelity to prevent missed information in the X-ray image. If we evaluate the output image with only visibility, the features that provide clues about the lesion may disappear.

A. PROCEDURES

Because we use a visibility predictor model for IQA, which requires a different image (the difference between two images), we need a pair of images with and without lesions. However, obtaining X-ray images without lesions that were taken in the same poses and shooting environment as were those with lesions is almost impossible. Therefore, we applied image augmentation to X-ray images of healthy lungs and lesion phantoms. We explain the details of the image augmentation in Section V-A.

Fig. 5 shows the workflow for the reference parameter prediction. We estimate the reference parameters by repeating the following steps until convergence is reached:

1) First, we apply image processing with the parameter set $\boldsymbol{p}$ to an input without lesions, $\mathcal{I}$, and an image with lesions, $\mathcal{I}^+$, and we obtain the outputs, $\hat{\mathcal{I}}$ and $\hat{\mathcal{I}}^+$.
2) Next, we calculate the visibility gain g_{vis} of the lesions before and after image filtering by taking the differences between $\mathcal{I}$ and $\mathcal{I}^+$ and between $\hat{\mathcal{I}}$ and $\hat{\mathcal{I}}^+$.
3) We also calculate fidelity losses by comparing the input images with lesions before and after filtering, $\mathcal{I}^+$ and $\hat{\mathcal{I}}^+$.
4) We update the parameters $\boldsymbol{p}$ to increase the gain in visibility while maintaining the loss in fidelity to a certain level by network backpropagation.

As the image processing module in Step 1 is explained in Section III-A, the following sections describe other components (visibility assessment, fidelity assessment, and optimization).

B. VISIBILITY ASSESSMENT

Although there are several approaches to visibility prediction, we employed High-Dynamic Range Visible Difference Predictor (HDR-VDP) [64]. The monitors that were used to display the X-ray images had HDR functionality and were compatible with the Grayscale Standard Display Function (GSDF) and were able to show detailed intensity differences. HDR-VDP considers the characteristics of the monitor and allows conversion from image pixel values to display luminance with consideration for HVP. Moreover, HDR-VDP is model-based and more versatile than are learning-based IQA models, which require training on large datasets. Fig. 6 (d) shows an example of the visibility map of the lesions estimated by HDR-VDP.

The visibility predictor Ω takes two images, a reference $\mathcal{I}_{\text{r}}$ and a test $\mathcal{I}_{\text{t}}$, and calculates the predicted visibility map $\mathcal{V}_{\text{t}}$ from the difference between the images: $\mathcal{V}_{\text{t}} = \Omega(\mathcal{I}_{\text{r}}, \mathcal{I}_{\text{t}})$.
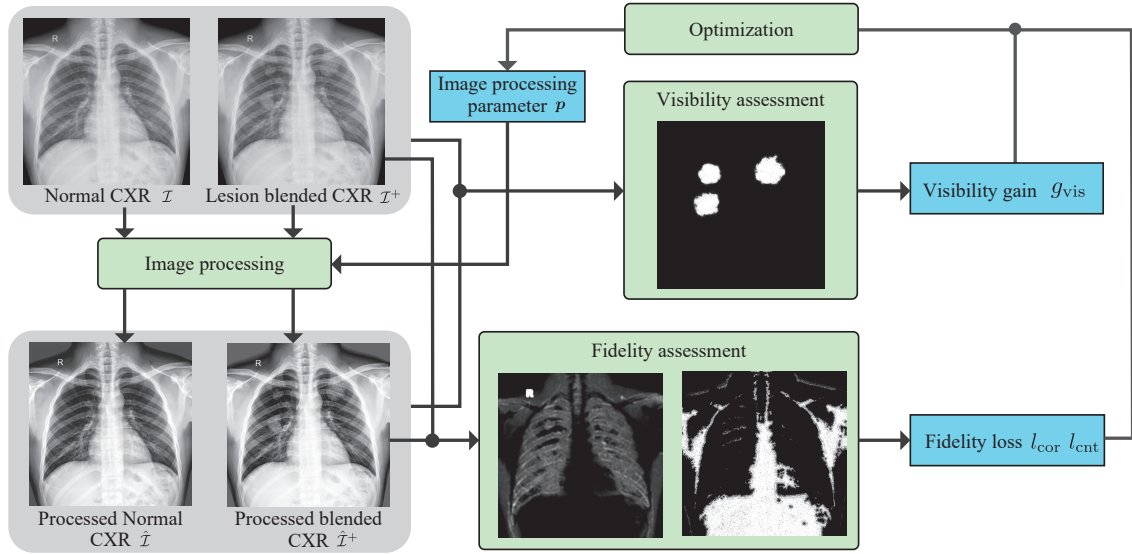
FIGURE 5. Reference parameter prediction. The visibility assessment evaluates, before and after processing, the lesion visibility for each pair of images, which includes images with and without lesions. Fidelity assessment requires blended CXR images before and after processing.

Let $\mathcal{V}^+$ be the predicted visibility map obtained from images with and without lesions before image processing: $\mathcal{V}^+ = \Omega(\mathcal{I}, \mathcal{I}^+)$. Let $\hat{\mathcal{V}}^+$ be the predicted visibility map after image processing: $\hat{\mathcal{V}}^+ = \Omega(\hat{\mathcal{I}}, \hat{\mathcal{I}}^+)$. We define *visibility gain* as an indicator of how much the visibility of the lesion area has improved from the image processing. We derive the visibility gain g_{vis} by computing the ratio of the total of the pixel values of $\mathcal{V}^+$ to that of $\hat{\mathcal{V}}^+$, as follows:

$$g_{\text{vis}} = \frac{sum(\hat{\mathcal{V}}^+)}{sum(\mathcal{V}^+)}, \quad (8)$$

where $sum()$ is a function that sums the values of the pixels in the image. The optimization process maximizes g_{vis} and obtains the image processing parameters.

C. FIDELITY ASSESSMENT

To evaluate the image fidelity between $\mathcal{I}^+$ and $\hat{\mathcal{I}}^+$, we propose a method that combines the *correlation loss* using the correlation between images in the Laplacian pyramid [65] with the *contrast loss*, which indicates the loss of features between images [66]. The former evaluates the correlation of the edge components in the image, whereas the latter detects the loss in contrast based on HVP. If we impose only tolerance limits using correlation loss, we would not be able to guarantee the preservation of image features. However, because image feature preservation is paramount to diagnosis, we use contrast loss simultaneously to prevent the loss of features. In this section, we briefly explain these losses and the optimization method.

1) Correlation loss

The correlation loss l_{cor} between $\mathcal{I}^+$ and $\hat{\mathcal{I}}^+$ is derived from the covariance and standard deviation of the image pixels at each level of the Laplacian pyramid. Let $\{\mathcal{L}_j\}$ and $\{\hat{\mathcal{L}}_j\}$ be the Laplacian pyramids for $\mathcal{I}^+$ and $\hat{\mathcal{I}}^+$, respectively. Let $c(\mathcal{L}, \hat{\mathcal{L}}, \boldsymbol{x})$ be the covariance around pixel $\boldsymbol{x}$ in two Laplacian images, $\mathcal{L}$ and $\hat{\mathcal{L}}$, at the j_{th} level, and let $s(\mathcal{L}, \boldsymbol{x})$ be the standard deviation around $\boldsymbol{x}$ in $\mathcal{L}$:

$$c(\mathcal{L}, \hat{\mathcal{L}}, \boldsymbol{x}) = \Sigma_{\boldsymbol{x}}\left((\mathcal{L}(\boldsymbol{x}) - \overline{\mathcal{L}(\boldsymbol{x})})(\hat{\mathcal{L}}(\boldsymbol{x}) - \overline{\hat{\mathcal{L}}(\boldsymbol{x})})\right), \quad (9)$$

$$s(\mathcal{L}, \boldsymbol{x}) = \sqrt{\Sigma_{\boldsymbol{x}}\left((\mathcal{L}(\boldsymbol{x}) - \overline{\mathcal{L}(\boldsymbol{x})})\right)^2}, \quad (10)$$

where $\overline{\mathcal{L}(\boldsymbol{x})}$ is a weighted average according to the distance from $\boldsymbol{x}$ calculated using a Gaussian kernel. In our experiment, we set the Gaussian kernel size and standard deviation to 15 pixels and 6 pixels, respectively. Next, we calculate the image correlation $f(\mathcal{L}, \hat{\mathcal{L}})$ by summing the covariance values over the entire image and normalizing it by dividing it by the summation of the products of the standard deviations, as follows:

$$f(\mathcal{L}, \hat{\mathcal{L}}) = \frac{\Sigma_{\boldsymbol{x}} c(\mathcal{L}, \hat{\mathcal{L}}, \boldsymbol{x})}{\Sigma_{\boldsymbol{x}}(s(\mathcal{L}, \boldsymbol{x}) s(\hat{\mathcal{L}}, \boldsymbol{x}))}. \quad (11)$$

Then, we obtain the correlation loss l_{cor} by taking the average of the correlation losses of the Laplacian pyramid images: $(1 - f(\mathcal{L}_j, \hat{\mathcal{L}}_j))$ up to $N_{\text{ng}} - 1$ level.

$$l_{\text{cor}} = \frac{1}{N_{\text{ng}} - 1} \sum_{j=1}^{N_{\text{ng}}-1} \left(1 - f(\mathcal{L}_j, \hat{\mathcal{L}}_j)\right) \quad (12)$$

2) Contrast loss

The contrast loss is one of the indicators in the dynamic range-independent IQA [66], and it increases when the features that are visible in the reference image are not visible in the test image(Fig. 6 (e)). The dynamic range-independent IQA [66]

is a metric for distortion of the test image relative to that of the original image based on HVS. This IQA includes three metrics: visible feature loss, invisible feature amplification, and contrast polarity reversal, where the visible feature loss is used as the contrast loss. We omit the explanation for the calculation of contrast loss here; see the details in the original study [66].

Let $\mathcal{M}$ be the contrast loss map [66] calculated from $\mathcal{I}^+$ and $\hat{\mathcal{I}}^+$: $\mathcal{M} = Gamma(\mathcal{I}^+, \hat{\mathcal{I}}^+)$. The loss l_{cnt} is obtained by averaging the pixel values of the obtained map as follows.

$$l_{\mathrm{cnt}} = \frac{1}{N_{\mathrm{img}}} \Sigma_{\boldsymbol{x} \in \mathcal{M}} \mathcal{M}(\boldsymbol{x}). \qquad (13)$$

where N_{img} is the size of the $\mathcal{M}$ image.

D. OPTIMIZATION

The optimal image processing parameter $\hat{\boldsymbol{p}}$ should maximize the visibility gain g_{vis} while maintaining the fidelity losses, which consists of l_{cor} and l_{cnt}, below their chosen threshold values. This is a constrained minimization problem, which is defined as follows:

$$\hat{\boldsymbol{p}} = \underset{\boldsymbol{p}}{\arg\min} \exp\left(-g_{\mathrm{vis}}\right) \qquad (14)$$

$$\text{s.t. } l_{\mathrm{cor}} < \tau_{\mathrm{cor}} \ \wedge \ l_{\mathrm{cnt}} < \tau_{\mathrm{cnt}},$$

where τ_{cor} and τ_{cnt} are threshold values for l_{cor} and l_{cnt}, respectively. We determined the values experimentally, and the users may determine them as they like. We introduce the constraints as penalty terms during the optimization process; Eq. 14 is rewritten as follows:

$$\hat{\boldsymbol{p}} = \underset{\boldsymbol{p}}{\arg\min}(\exp\left(-g_{\mathrm{vis}}\right) + \sigma_1 \Lambda(l_{\mathrm{cor}} - \tau_{\mathrm{cor}})^2$$

$$+ \sigma_2 \Lambda(l_{\mathrm{cnt}} - \tau_{\mathrm{cnt}})^2), \qquad (15)$$

where σ_1 and σ_2 are the weights of the penalty terms, which are gradually increased during optimization, and $\Lambda(x)$ is the ReLU function.

$$\Lambda(x) = \begin{cases} x & x > 0 \\ 0 & x \le 0 \end{cases}, \qquad (16)$$

V. EXPERIMENTS

This section describes the image augmentation approach and the implementation of the datasets used in the experiments. This is followed by the experimental results.

A. DATASET

1) Procedure for the generation of training images

We describe here the procedures for generating a set of X-ray images with and without lesions. The inputs used to generate an image set are an X-ray image of the lungs with no lesions, a segmentation image of the lungs, and an X-ray images of lesions only. For the lung X-ray and segmentation images, we used the Shenzhen CXR Dataset [67], [68]. We imaged

the phantoms that imitated COVID-19 lesions with an X-ray device, as shown in Fig. 7.

We randomly selected a point in the segmented regions of a lung image. The selected point had to be such that the overlaid lesion did not go beyond the boundaries of the lung region; we narrowed the region using a morphological transformation according to the size of the lesion. Then, we overlaid the lesion's image onto the lung's image.

After we overlaid the lesions, the image was a square cutout of the lung region. First, we fitted a minimum bounding box to cover all the lung regions. The bounding box size was then scaled up by 20% to include entire lung regions. Next, we cropped the squared region, which circumscribed the bounding box. Finally, we resized the cropped image to the input size for the network (1024×1024 pixels).

2) Overlaying lesion images

Because the dose during radiography, the sensitivity of the flat panel display, and the thickness of the body are unknown, we need to estimate them to adjust the intensities of the lung and lesion images. We assume that the attenuation coefficients of the body and the lesion are known, whereas the attenuation in the air is small enough to be ignored. Fig. 8 shows an overview of the overlaid lesion images.

We assume that the pixel value $\mathcal{I}(\boldsymbol{x})$ at pixel $\boldsymbol{x}$ in an X-ray image $\mathcal{I}$ is a linear transformation of the incident X-ray intensity $I(\boldsymbol{x})$:

$$\mathcal{I}(\boldsymbol{x}) = \alpha I(\boldsymbol{x}) + \beta. \qquad (17)$$

According to the Beer-Lambert Law, the emitted X-ray with intensity I_0 attenuates while passing through an object with an attenuation coefficient λ and a transmission distance $\delta_{\boldsymbol{x}}$ along the ray incoming toward pixel $\boldsymbol{x}$:

$$I(\boldsymbol{x}) = I_0 \exp\left(-\lambda \delta_{\boldsymbol{x}}\right). \qquad (18)$$

We substitute Eq. 18 into Eq. 17 and obtain the following equation:

$$\mathcal{I}(\boldsymbol{x}) = \hat{\alpha} \exp\left(-\lambda \delta_{\boldsymbol{x}}\right) + \beta \qquad (19)$$

$$\hat{\alpha} = \alpha I_0. \qquad (20)$$

As α and I_0 are not separable, we estimate $\hat{\alpha}$ as their product. Given a lung image $\mathcal{I}_o$ with unknown coefficients α_o and β_o and unknown body thicknesses $\delta_{o,\boldsymbol{x}}$, as well as a lesion image $\mathcal{I}_l$ with α_l, β_l, and $\delta_{l,\boldsymbol{x}}$, we derive the following equations:

$$\mathcal{I}_o(\boldsymbol{x}) = \hat{\alpha}_o \exp\left(-\lambda_o \delta_{o,\boldsymbol{x}}\right) + \beta_o \qquad (21)$$

$$\mathcal{I}_l(\boldsymbol{x}) = \hat{\alpha}_l \exp\left(-\lambda_l \delta_{l,\boldsymbol{x}}\right) + \beta_l. \qquad (22)$$

After $(\hat{\alpha}_o, \beta_o)$ and $(\hat{\alpha}_l, \beta_l)$ are obtained, $\delta_{o,\boldsymbol{x}}$ and $\delta_{l,\boldsymbol{x}}$ can be calculated from these equations to obtain the pixel value $\mathcal{I}^+(\boldsymbol{x})$ of the overlay image $\mathcal{I}^+$, as in the following equation:

$$\mathcal{I}^+(\boldsymbol{x}) = \hat{\alpha}_o \exp\left(-\lambda_o \delta_{o,\boldsymbol{x}} - \lambda_l \delta_{l,\boldsymbol{x}}\right) + \beta_o. \qquad (23)$$

Here, we assume that at least the ranges of $\delta_{o,\boldsymbol{x}}$ and $\delta_{l,\boldsymbol{x}}$ are known as $(\delta_{o,min}, \delta_{o,max})$ and $(\delta_{l,min}, \delta_{l,max})$. Because we can observe the maximum and minimum intensity values
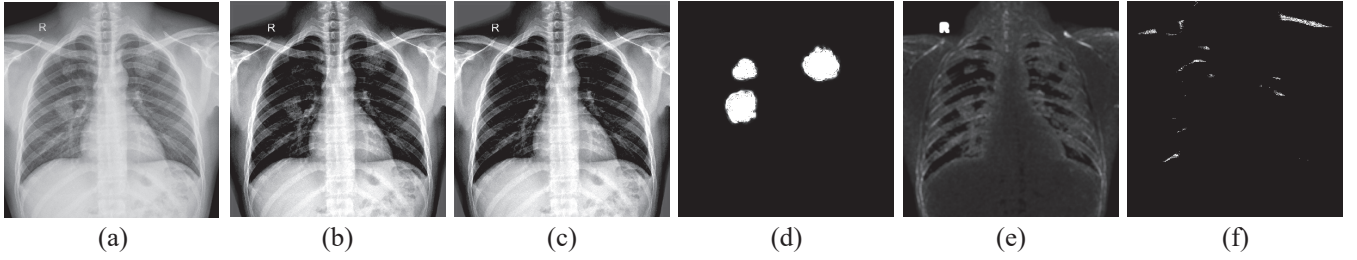
(a) (b) (c) (d) (e) (f)

FIGURE 6. Example of a visibility map and fidelity loss map of (a) an original CXR image, (b) a CXR image with lesions after processing, (c) a CXR image without lesions after processing, (d) a visibility map between (b) and (c), (e) correlation loss of (a) and (b), (f) the contrast loss of (a) and (b)
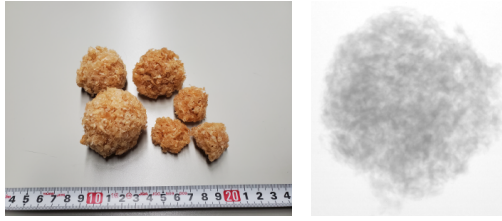

FIGURE 7. Lesion phantom models and an X-ray image overlaid on the CXR image
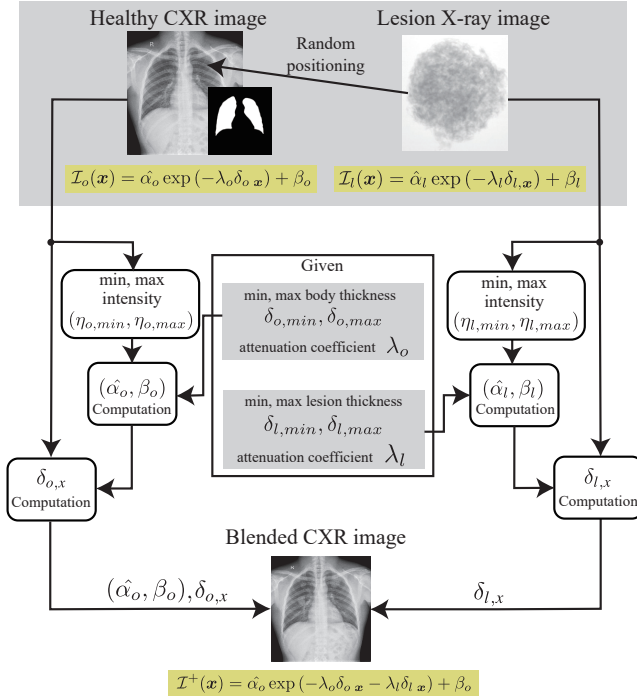

FIGURE 8. Overview of lesion images overlaid onto healthy CXR images

of the lung regions and the lesion regions in the images as $(\eta_{o,min}, \eta_{o,max})$ and $(\eta_{l,min}, \eta_{l,max})$, we obtain $(\hat{\alpha}_o, \beta_o)$ and

$(\hat{\alpha}_l, \beta_l)$ by solving the following simultaneous equations:

$$\eta_{o,min} = \hat{\alpha}_o \exp\left(-\lambda_o \delta_{o,min}\right) + \beta_o, \quad (24)$$

$$\eta_{o,max} = \hat{\alpha}_o \exp\left(-\lambda_o \delta_{o,max}\right) + \beta_o, \quad (25)$$

$$\eta_{l,min} = \hat{\alpha}_l \exp\left(-\lambda_l \delta_{l,min}\right) + \beta_l, \quad (26)$$

$$\eta_{l,max} = \hat{\alpha}_l \exp\left(-\lambda_l \delta_{l,max}\right) + \beta_l. \quad (27)$$

We set the experimental values of the attenuation coefficients to $\lambda_o = 0.07$ and $\lambda_l = 0.09$. We assume that $\delta_{o,max}$ ranges randomly from 24 to $32\,\mathrm{cm}$ considering the thickness of the human body; the random thickness values make it possible to synthesize images in different ratios and to augment the data during training to prevent overfitting to the training data. $\delta_{o,min}$ was set to $10\,\mathrm{cm}$, assuming it corresponded to the thickness of the body without lungs. We measured the lesion phantoms and obtained the maximum transmission distances $\delta_{l,max}$ from 3.3 to $8.0\,\mathrm{cm}$, and we set $\delta_{l,min}$ to $0\,\mathrm{cm}$ outside the lesion.

B. IMPLEMENTATION

1) Reference enhanced model predictor

We implemented the model predictor by LibTorch described in Section IV, which allowed the framework to perform back-propagation. We used the Adam optimizer [69] for optimization, and we set the convergence criteria so that the minimum loss did not update a specified number of times. We also set the threshold values to $\tau_{cor} = 0.005$ and $\tau_{cnt} = 0$. The initial parameters of p should not change the input images after they are processed; we set $w_1 = 1$ and set the other parameters for the intensity adjustment to 0, and all the parameters for frequency adjustment to 1. That is, when $N_{rbf} = 4$ and $N_{gn} = 5$, $w_1 = 1$, $w_{i(=2,3,4)} = 0$, $c_{i(=1,2,3,4)} = 0$, and $\mu_{j(=1,2,3,4)} = 1$.

The visibility gain and contrast loss evaluation assumes the use of medical monitors, which apply the GSDF for brightness correction. Assuming the CXR image is viewed in a dark room, we set the maximum luminance of the monitor to $400\,cd/m^2$ and ambient light to zero. We also set the number of pixels per degree of visual angle to 26.54.

2) Inference model training

We implemented the framework described in Section III-B with PyTorch. We chose the base size network of the Swin

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3528489

IEEE Access·

R. Ishikawa *et al.*: Visibility Enhancement of Lesion Regions in Chest X-ray Images with Image Fidelity Preservation

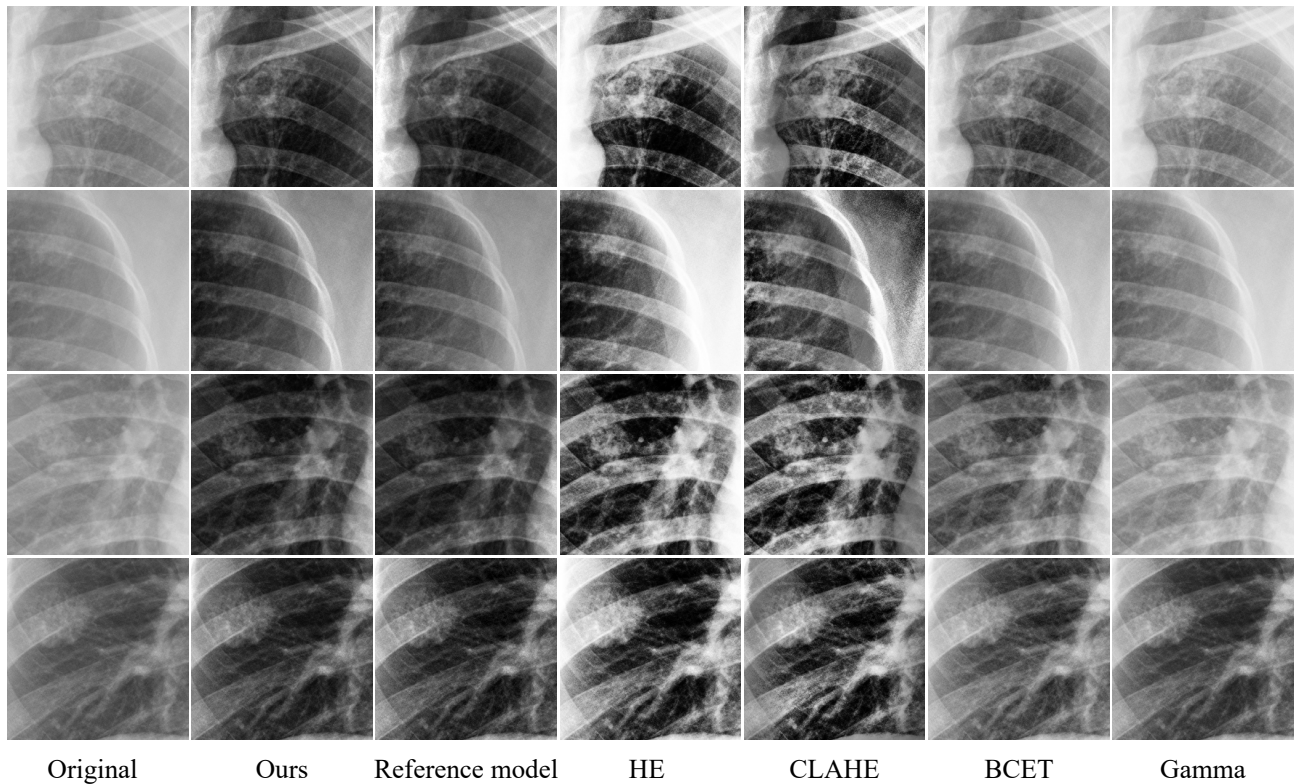| Original | Ours | Reference model | HE | CLAHE | BCET | Gamma |

FIGURE 9. Comparison of the image processing results using each method. Each image is cropped for visualization.

Transformer V2 [63]. The network inputs were the blended and lung segmentation images included in the dataset. The loss function is shown in Eq. 5. We used 2,000 images to train the inference model and 50 images for validation.

C. COMPARATIVE EVALUATION

First, we evaluated the proposed method using visibility gain g_{vis}, correlation loss l_{cor}, and contrast loss l_{cnt} as metrics. The comparative methods were *HE* (histogram equalization), *CLAHE* [25], the Balance Contrast Enhancement Technique (*BCET*) [70], and *Gamma* correction [8]. *HE* flattens the histogram of the overall image brightness. *CLAHE* flattens the histogram in small batches and suppresses the saturation that can occur with *HE*. *BCET* performs a quadratic transformation so that the image brightness is adjusted to the target maximum, minimum, and median values. *Gamma* corrects the brightness by multiplying it by a gamma value.

We applied *HE* and *BCET* to only the lung regions. We calculated the image fidelity losses in only the lung regions. Table 1 shows the mean and variance of the parameters for each method.

Table 2 shows the significance of the proposed method compared with that of the other methods by presenting the results of a one-tailed Wilcoxon signed-rank test [71]. The Wilcoxon signed-rank test was employed because there was no normality in the distribution of any term. The "o" indicates that the proposed method recorded significantly better values,

and the "x" indicates that it recorded inferior values. Note that significant differences exist for all comparisons at the 5% level of significance.

The proposed method achieves larger visibility gains than do the *BCET*, *Gamma*, and *CLAHE* correction with a reasonably small correlation loss. Although *HE* achieves larger visibility gains, it also has larger image correlation and contrast losses. Our results show a trend toward a greater contrast loss than the *Reference model*; however, the loss is not as large as that of *HE*. The reference results estimated by the reference enhanced parameter estimation module show that the results exceed the target threshold ($\tau_{\mathrm{cor}} = 5 \times 10^{-3}$). This is because the constrained optimization is being solved in a penalized manner, which is as expected.

Fig. 9 shows examples of processed images. The images show that the proposed method improves contrast while preserving the features in the image inside of the lungs. Although *HE* improves the contrast inside the lungs, some regions exhibit a loss of features caused by the crushed blacks in the dark areas between the ribs. *CLAHE* excels at enhancing regions that have low contrast; however, because it adjusts the contrast locally, the correlation in the image may become disrupted, which makes it difficult to discern lesions. We did not observe a significant increase in visibility for the *BCET* or *Gamma* correction.

TABLE 1. Mean and standard deviation of the metrics for each method.

	Visibility gain g_{vis} ↑	Contrast Loss ($\times 10^{-3}$) l_{cnt} ↓	Correlation Loss ($\times 10^{-3}$) l_{cor} ↓
Ours	2.96 ± 3.30	1.02 ± 1.43	4.00 ± 0.92
Reference	4.18 ± 3.31	0.16 ± 0.13	5.90 ± 0.36
HE	13.82 ± 36.28	7.65 ± 2.64	23.80 ± 5.54
CLAHE	2.20 ± 1.27	0.23 ± 0.22	22.02 ± 2.19
BCET	1.60 ± 0.55	0.00 ± 0.00	0.85 ± 1.10
Gamma	1.34 ± 0.49	0.02 ± 0.12	2.14 ± 0.74

TABLE 2. Significance for the comparison between our method and the comparison methods. The p values were less than 0.05 in all comparisons, and significant differences existed in all comparisons at the 5% level of significance.

	Visibility gain		Contrast Loss		Correlation Loss	
	Significance	p value	Significance	p value	Significance	p value
HE	x	< 0.001	o	< 0.001	o	< 0.001
CLAHE	o	0.012	x	< 0.001	o	< 0.001
BCET	o	< 0.001	x	< 0.001	x	< 0.001
Gamma	o	< 0.001	x	< 0.001	x	< 0.001

D. SUBJECTIVE EVALUATION

We conducted paired comparisons alongside *BCET*, *CLAHE*, and *HE* to further evaluate the proposed method. This experiment assessed two key criteria: the lesion's visibility and the image fidelity. The visibility assessment required participants to select from a pair of displayed images the image that demonstrated superior visibility of the lesion. For the fidelity assessment, participants were requested to select the image that better preserved the features of the original image. We used 30 CXR images and we conducted comparisons with six pairs of methods. Each participant had to give an answer a total of 180 times for each task. Table 3 provides a comprehensive outline of the instructions for each task. Before the study, we provided an example of missing features and hallucinations that were fabricated manually in Task 2 to standardize the decision criteria as much as possible (Fig. 10). This study was approved by the University-Wide Ethics Review Committee of The University of Tokyo. The review number was 22-450. In addition, we obtained informed consent from the subjects.

This study utilized two vertically positioned monitors in a low-lit setting. The upper monitor displayed a single reference CXR image, whereas the lower monitor displayed two images that were to be compared. The reference image in the upper monitor showed a highlighted lesion for the visibility assessment task, with dashed lines encapsulating the lung regions for the fidelity assessment task. The lower monitor used in the study was "MediCrysta" developed by IO-DATA. Participants were instructed not to bring their faces close to the monitor while completing the designated tasks.

Table 4 shows the measures computed using the Thurstone method [72]. Thirteen participants participated in the visibility assessment task, and the fidelity assessment task involved eight participants. On comparing the results with those shown in Table 1, it is clear that although *CLAHE* demonstrates superior performance in relation to HDR-VDP, this trend is also seen with HDR-VDP with other techniques. Similar to what we see in Table 2, Table 5 reveals that the proposed

method aims to enhance visibility, even at the cost of some reduction in image fidelity.

Table 5 shows the significance of the proposed method compared with the other methods by performing a one-tailed Wilcoxon signed-rank test [71] using the scores obtained using the Thurstone method. An "o" indicates that the proposed method recorded significantly better values, and the "x" indicates that it recorded inferior values. Note that significant differences exist for all comparisons at the 5% level of significance.

VI. DISCUSSION

The objective of this study was to propose a framework that offered an estimation of image processing parameters intended to enhance lesion visibility while preserving the inherent characteristics of the original image. The experimental results demonstrate a successful application of the proposed method for estimating image processing parameters to enhance visibility in comparison with *BCET* and *Gamma*. Furthermore, *Ours* exhibits a lower degree of contrast and correlation loss in the original image than does *HE*, as well as less feature degradation in image correlation than does *CLAHE*.

Table 1 demonstrates that, under a specific correlation loss (set to $\tau_{\mathrm{cor}} = 0.005$), the proposed method surpasses *BCET* and *Gamma* in terms of visibility improvement. However, it should be noted that the proposed method exhibits a larger contrast loss than do *CLAHE*, *BCET*, or *Gamma*, which resulted in noticeable visual impairment in the subject experiments. *HE* grachieves a contrast enhancement that significantly improves visibility in certain images. Nonetheless, *HE* may introduce a wide range of variance, which may cause the lesion signal to remain unenhanced against the background in certain images.

In comparing *Ours* and *Reference*, we see that *Ours* has visibility improvement and lower correlation loss. This observation suggests that the inference network in *Ours* learned the average features of the input images, whereas the reference

TABLE 3. Directions for the tasks

|	Direction of each task
Visibility Assessment	Of the two images displayed on the lower monitor, select the one you think is easier to see the lesion in the annotated area
Fidelity Assessment	Of the two images on the lower monitor, select the one preserves the features of the upper reference image. Please evaluate by only the preservation of the feature, not by its visibility. Please evaluate only in the lung regions (Inside the dashed line).
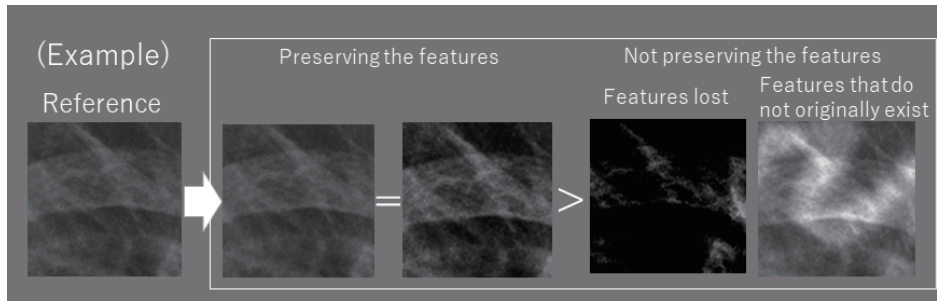
FIGURE 10. Example of feature loss and hallucination with the fidelity assessment task instruction.

TABLE 4. Scores for the Thurstone test

method	Visibility score ↑	Fidelity score ↑
Ours	0.209 ± 0.072	0.455 ± 0.061
HE	0.527 ± 0.087	0.275 ± 0.054
CLAHE	0.606 ± 0.112	0.189 ± 0.049
BCET	0.158 ± 0.066	0.581 ± 0.057

TABLE 5. Significance comparison between our method and comparative methods in the subject experiments. The p values were less than 0.05 for all comparisons and significant differences existed for all comparisons at the 5% level of significance.

	Visibility		Fidelity	
	Significance	p value	Significance	p value
HE	x	< 0.001	o	< 0.001
CLAHE	x	< 0.001	o	< 0.001
BCET	o	< 0.001	x	< 0.001

enhanced model predictor estimated the optimal parameters for images with distinct characteristics.

The contrast loss for *Ours* is larger than that of the *Reference* and the other comparative methods aside from *HE*. The use of RBF in the proposed method introduces a delicate balance between enhancing the contrast and feature loss due to saturation. In other words, the expansion of the brightness range to the maximum possible enhances contrast and improves visibility; however, if this range is exceeded, features are sacrificed due to saturation.

In addition to comparing each method, the subject experiment aims to assess the validity of the quantitative evaluation through HDR-VDP. When the tables are compared, we can observe that HDR-VDP and the subject experiments produced similar results, except in the visibility evaluation of *CLAHE*. The subject experiment evaluated contrast loss and correlation loss simultaneously for the fidelity assessment. In the HDR-VDP evaluator, the fidelity comparison between the proposed method and *CLAHE* revealed that *CLAHE* outperformed in terms of low contrast loss, whereas the proposed

method exhibited superiority in terms of low correlation loss. In the subject experiments, the correlation loss for *CLAHE* is believed to be more pronounced as a fidelity loss.

The comparison between the subject experiment and the HDR-VDP reveals a significant disparity in the visibility results for *CLAHE*. The quantitative assessment conducted using HDR-VDP demonstrated that the proposed method exhibited significantly higher visibility than did *CLAHE*, at a 5% level of significance. Conversely, the results obtained from the subject experiment indicated that *CLAHE* showed greater visibility of the lesion. It is likely that the HDR-VDP evaluated the lower visibility of the lesion in *CLAHE*, influenced by the change in brightness that appeared due to local contrast enhancement. However, considering that the subject experiment included annotations of the lesion's location, it is plausible that the difference in contrast between the lesion and the surrounding area contributed to the lesion's visibility as a signal stimulus.

LIMITATION

One of the limitations of the proposed method is the decrease in performance observed for *Ours* in comparison with *Reference*. The RBF curve and μ parameters output by the reference enhanced model predictor using the training images as input are illustrated on the left side of Fig. 11. Meanwhile, the parameters inferred by the image enhancer using the validation images are presented on the right side of Fig. 11. It is evident that the network in the image enhancer does not adequately learn the characteristic RBF curves, as observed on the *Reference* side. Although the network grasps the general trend of the frequency parameters and contributes to the sharpening of contours, it has limited impact on the parameter estimation for each image.

To address these problems, a learning method that efficiently captures the characteristics of each image should be developed, and an image processing method that enables the network to easily learn the relationship between image features and output parameters should be used. We used RBF because of its high degree of freedom in tone correction. However, we discovered that estimating RBF curves to enhance contrast without causing saturation was challenging. We believe that employing a constrained approach to prevent saturation could enable the estimation of a more effective filter while avoiding black crush.

VII. CONCLUSION

In this study, we proposed a learning-based framework for X-ray image processing that was designed to enhance lesion visibility while preserving image fidelity. To support this, we introduced simulation data and a methodology for evaluating X-ray imaging algorithms based on lesion visibility and models. Furthermore, we implemented differentiable modules for visibility and fidelity evaluation of visibility and fidelity, while leveraging human perception as a guiding principle. These contributions aim to advance the quality and reliability of X-ray imaging.

We used contrast adjustment and frequency adjustment as image processing techniques; however, various image processing methods that require parameter adjustment can be incorporated into this framework. One issue is that reference parameter estimation and inferential training must be repeated when the display environment changes. The display calibration must be performed for each environment, which could be a problem in practical applications. Moreover, because there is a difference between the reference and inference results, improving accuracy is a problem to be addressed in the future.

REFERENCES

[1] K. M. Sunnetci and A. Alkan, "Biphasic majority voting-based comparative covid-19 diagnosis using chest x-ray images," *Expert Systems with Applications*, vol. 216, p. 119430, 2023.

[2] S. Sajed, A. Sanati, J. E. Garcia, H. Rostami, A. Keshavarz, and A. Teixeira, "The effectiveness of deep learning vs. traditional methods for lung disease diagnosis using chest x-ray images: A systematic review," *Applied Soft Computing*, p. 110817, 2023.

[3] E. L. Irede, O. R. Aworinde, O. K. Lekan, O. D. Amienghemhen, T. P. Okonkwo, A. P. Onivefu, and I. H. Ifijen, "Medical imaging: a critical review on x-ray imaging for the detection of infection," *Biomedical Materials & Devices*, pp. 1–45, 2024.

[4] W. B. Gefter, B. A. Post, and H. Hatabu, "Commonly missed findings on chest radiographs: causes and consequences," *Chest*, vol. 163, no. 3, pp. 650–661, 2023.

[5] Y. Guo, X. Hou, Z. Liu, and Y. Zhang, "Ifnet: An image-enhanced cross-modal fusion network for radiology report generation," in *International Symposium on Bioinformatics Research and Applications*. Springer, 2024, pp. 286–297.

[6] P. S. Yadav, B. Gupta, and S. S. Lamba, "A new approach of contrast enhancement for medical images based on entropy curve," *Biomedical Signal Processing and Control*, vol. 88, p. 105625, 2024.

[7] K. Munadi, K. Muchtar, N. Maulina, and B. Pradhan, "Image enhancement for tuberculosis detection using deep learning," *IEEE Access*, vol. 8, pp. 217 897–217 907, 2020.

[8] T. Rahman, A. Khandakar, Y. Qiblawey, A. Tahir, S. Kiranyaz, S. B. A. Kashem, M. T. Islam, S. Al Maadeed, S. M. Zughaier, M. S. Khan, *et al.*, "Exploring the effect of image enhancement techniques on covid-19 detection using chest x-ray images," *Computers in biology and medicine*, vol. 132, p. 104319, 2021.

[9] K. R. Bhatele, A. Jha, D. Tiwari, M. Bhatele, S. Sharma, M. R. Mithora, and S. Singhal, "Covid-19 detection: A systematic review of machine and deep learning-based approaches utilizing chest x-rays and ct scans," *Cognitive Computation*, vol. 16, no. 4, pp. 1889–1926, 2024.

[10] S. Sharma and K. Guleria, "A systematic literature review on deep learning approaches for pneumonia detection using chest x-ray images," *Multimedia Tools and Applications*, vol. 83, no. 8, pp. 24 101–24 151, 2024.

[11] A. U. Ibrahim, M. Ozsoz, S. Serte, F. Al-Turjman, and P. S. Yakoi, "Pneumonia classification using deep learning from chest x-ray images during covid-19," *Cognitive computation*, vol. 16, no. 4, pp. 1589–1601, 2024.

[12] M. A. Al-qaness, J. Zhu, D. AL-Alimi, A. Dahou, S. H. Alsamhi, M. Abd Elaziz, and A. A. Ewees, "Chest x-ray images for lung disease detection using deep learning techniques: A comprehensive survey," *Archives of Computational Methods in Engineering*, pp. 1–35, 2024.

[13] E. Kotei and R. Thirunavukarasu, "A comprehensive review on advancement in deep learning techniques for automatic detection of tuberculosis from chest x-ray images," *Archives of Computational Methods in Engineering*, vol. 31, no. 1, pp. 455–474, 2024.

[14] Z. Z. Qin, S. Ahmed, M. S. Sarker, K. Paul, A. S. S. Adel, T. Naheyan, R. Barrett, S. Banu, and J. Creswell, "Tuberculosis detection from chest x-rays for triaging in a high tuberculosis-burden setting: an evaluation of five artificial intelligence algorithms," *The Lancet Digital Health*, vol. 3, no. 9, pp. e543–e554, 2021.

[15] T. Rahman, A. Khandakar, M. A. Kadir, K. R. Islam, K. F. Islam, R. Mazhar, T. Hamid, M. T. Islam, S. Kashem, Z. B. Mahbub, *et al.*, "Reliable tuberculosis detection using chest x-ray with deep learning, segmentation and visualization," *IEEE Access*, vol. 8, pp. 191 586–191 601, 2020.

[16] D. Haritha, M. K. Pranathi, and M. Reethika, "Covid detection from chest x-rays with deeplearning: Chexnet," in *2020 5th International Conference on Computing, Communication and Security (ICCCS)*. IEEE, 2020, pp. 1–5.

[17] R. Jain, M. Gupta, S. Taneja, and D. J. Hemanth, "Deep learning based detection and analysis of covid-19 on chest x-ray images," *Applied Intelligence*, vol. 51, pp. 1690–1700, 2021.

[18] K. Ren, Y. Gu, M. Luo, H. Chen, and Z. Wang, "Deep-learning-based denoising of x-ray differential phase and dark-field images," *European Journal of Radiology*, p. 110835, 2023.

[19] Q. Wu, H. Tang, H. Liu, and Y. Chen, "Masked joint bilateral filtering via deep image prior for digital x-ray image denoising," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 8, pp. 4008–4019, 2022.

[20] K. Umehara, J. Ota, N. Ishimaru, S. Ohno, K. Okamoto, T. Suzuki, N. Shirai, and T. Ishida, "Super-resolution convolutional neural network for the improvement of the image quality of magnified images in chest radiographs," in *Medical Imaging 2017: Image Processing*, M. A. Styner and E. D. Angelini, Eds., vol. 10133, International Society for Optics and Photonics. SPIE, 2017, p. 101331P. [Online]. Available: https://doi.org/10.1117/12.2249969

[21] M. Greeshma and V. Bindu, "Chest x-ray image super-resolution via deep contrast consistent feature network," in *Recent Trends in Image Processing and Pattern Recognition: 5th International Conference, RTIP2R 2022, Kingsville, TX, USA, December 1-2, 2022, Revised Selected Papers*. Springer, 2023, pp. 78–90.
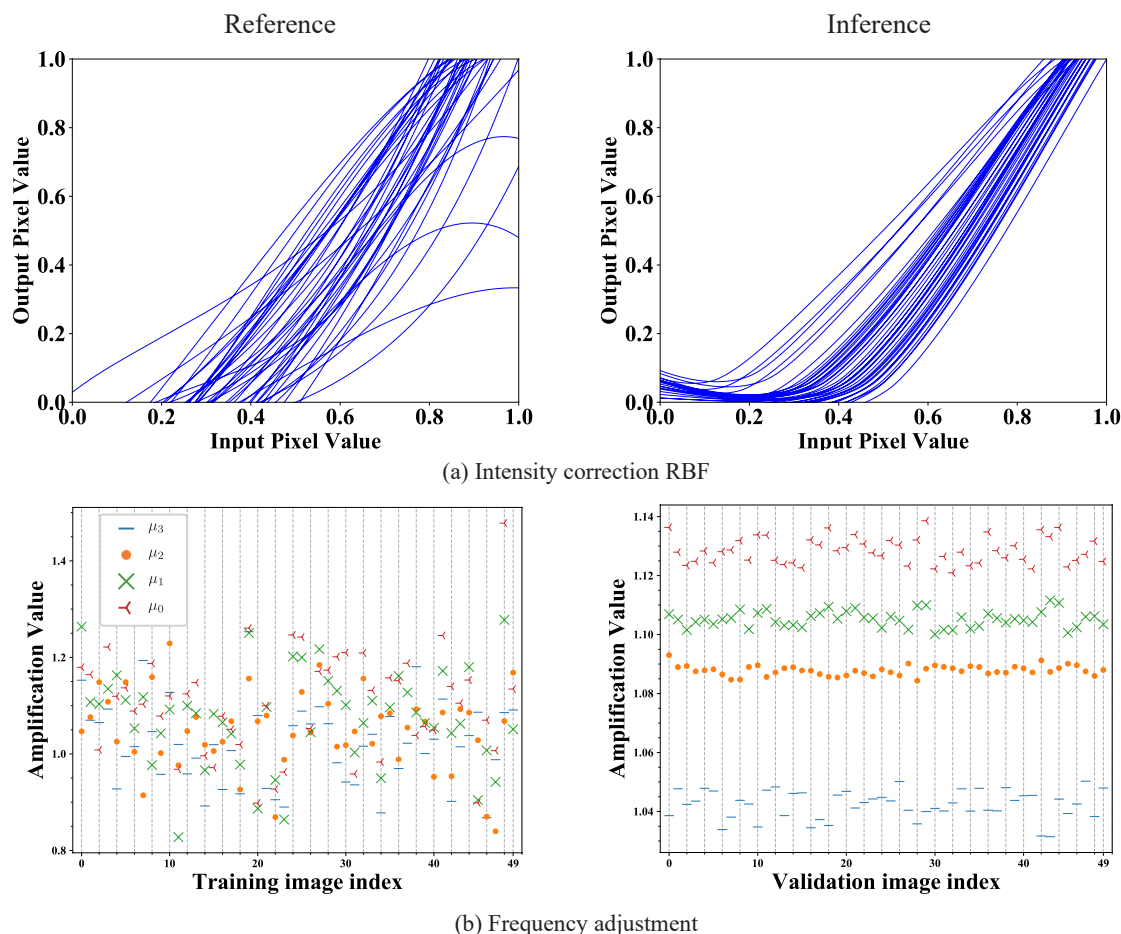
(a) Intensity correction RBF

(b) Frequency adjustment

FIGURE 11. Comparison of reference and inference image processing parameters.

[22] S.-C. Huang, F.-C. Cheng, and Y.-S. Chiu, "Efficient contrast enhancement using adaptive gamma correction with weighting distribution," *IEEE transactions on image processing*, vol. 22, no. 3, pp. 1032–1041, 2012.

[23] X. Guo, Y. Li, and H. Ling, "Lime: Low-light image enhancement via illumination map estimation," *IEEE Transactions on image processing*, vol. 26, no. 2, pp. 982–993, 2016.

[24] S. Kumar and A. K. Bhandari, "Automatic tissue attenuation-based contrast enhancement of low-dynamic x-ray images," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 6, no. 5, pp. 574–582, 2021.

[25] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer vision, graphics, and image processing*, vol. 39, no. 3, pp. 355–368, 1987.

[26] Y. Chang, C. Jung, P. Ke, H. Song, and J. Hwang, "Automatic contrast-limited adaptive histogram equalization with dual gamma correction," *Ieee Access*, vol. 6, pp. 11 782–11 792, 2018.

[27] T. Madmad and C. De Vleeschouwer, "Bilateral histogram equalization for x-ray image tone mapping," in *2019 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2019, pp. 3507–3511.

[28] S. Sahu, A. K. Singh, S. Ghrera, M. Elhoseny, *et al.*, "An approach for denoising and contrast enhancement of retinal fundus image using clahe," *Optics & Laser Technology*, vol. 110, pp. 87–98, 2019.

[29] L. Yang and Z. Zeng, "Medical image enhancement method based on histogram equalization," in *International Conference on Computer Graphics, Artificial Intelligence, and Data Processing (ICCAID 2021)*, F. Wu, J. Liu, and Y. Chen, Eds., vol. 12168, International Society for Optics and Photonics. SPIE, 2022, p. 121680S. [Online]. Available:

https://doi.org/10.1117/12.2631127

[30] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," in *Computer graphics forum*, vol. 22, no. 3. Wiley Online Library, 2003, pp. 419–426.

[31] S. N. Pattanaik, J. A. Ferwerda, M. D. Fairchild, and D. P. Greenberg, "A multiscale model of adaptation and spatial vision for realistic image display," in *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, 1998, pp. 287–298.

[32] N. H. Nguyen, T. Van Vo, and C. Lee, "Human visual system model-based optimized tone mapping of high dynamic range images," *IEEE Access*, vol. 9, pp. 127 343–127 355, 2021.

[33] W. Wang, X. Wu, X. Yuan, and Z. Gao, "An experiment-based review of low-light image enhancement methods," *Ieee Access*, vol. 8, pp. 87 884–87 917, 2020.

[34] E. H. Land and J. J. McCann, "Lightness and retinex theory," *Josa*, vol. 61, no. 1, pp. 1–11, 1971.

[35] D. J. Jobson, Z.-u. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE transactions on image processing*, vol. 6, no. 3, pp. 451–462, 1997.

[36] Z.-u. Rahman, D. J. Jobson, and G. A. Woodell, "Multi-scale retinex for color image enhancement," in *Proceedings of 3rd IEEE international conference on image processing*, vol. 3. IEEE, 1996, pp. 1003–1006.

[37] Z. ur Rahman, D. J. Jobson, and G. A. Woodell, "Retinex processing for automatic image enhancement," *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 100 – 110, 2004. [Online]. Available: https://doi.org/10.1117/1.1636183

[38] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3528489

IEEE *Access*

R. Ishikawa *et al.*: Visibility Enhancement of Lesion Regions in Chest X-ray Images with Image Fidelity Preservation

low-light enhancement," in *British Machine Vision Conference*, 2018.

[39] L. Xiao, C. Li, Z. Wu, and T. Wang, "An enhancement method for x-ray image via fuzzy noise removal and homomorphic filtering," *Neurocomputing*, vol. 195, pp. 56–64, 2016.

[40] S. Agaian and S. A. McClendon, "Novel medical image enhancement algorithms," in *Image Processing: Algorithms and Systems VIII*, J. T. Astola and K. O. Egiazarian, Eds., vol. 7532, International Society for Optics and Photonics. SPIE, 2010, p. 75320W. [Online]. Available: https://doi.org/10.1117/12.839003

[41] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," in *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, 2002, pp. 257–266.

[42] G. Bae, C. Y. Jang, S. I. Cho, and Y. H. Kim, "Non-iterative tone mapping with high efficiency and robustness," *IEEE Access*, vol. 6, pp. 35 720–35 733, 2018.

[43] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," *ACM transactions on graphics (TOG)*, vol. 27, no. 3, pp. 1–10, 2008.

[44] Y. Endo, Y. Kanamori, and J. Mitani, "Deep reverse tone mapping." *ACM Trans. Graph.*, vol. 36, no. 6, pp. 177–1, 2017.

[45] L. Tao, C. Zhu, G. Xiang, Y. Li, H. Jia, and X. Xie, "Llcnn: A convolutional neural network for low-light image enhancement," in *2017 IEEE Visual Communications and Image Processing (VCIP)*. IEEE, 2017, pp. 1–4.

[46] W. Wang, C. Wei, W. Yang, and J. Liu, "Gladnet: Low-light enhancement network with global awareness," in *2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018)*. IEEE, 2018, pp. 751–755.

[47] A. Ignatov, N. Kobyshev, R. Timofte, K. Vanhoey, and L. Van Gool, "Dslr-quality photos on mobile devices with deep convolutional networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 3277–3285.

[48] A. Ignatov, N. Kobyshev, R. Timofte, K. Vanhoey, and L. VanGool, "Wespe: weakly supervised photo enhancer for digital cameras," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 691–700.

[49] R. Wang, B. Jiang, C. Yang, Q. Li, and B. Zhang, "Magan: Unsupervised low-light image enhancement guided by mixed-attention," *Big Data Mining and Analytics*, vol. 5, no. 2, pp. 110–119, 2022.

[50] Y.-S. Chen, Y.-C. Wang, M.-H. Kao, and Y.-Y. Chuang, "Deep photo enhancer: Unpaired learning for image enhancement from photographs with gans," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6306–6314.

[51] A. Rana, P. Singh, G. Valenzise, F. Dufaux, N. Komodakis, and A. Smolic, "Deep tone mapping operator for high dynamic range images," *IEEE Transactions on Image Processing*, vol. 29, pp. 1285–1298, 2019.

[52] S. Chen and L. Zou, "Chest radiographic image enhancement based on multi-scale retinex technique," in *2009 3rd International Conference on Bioinformatics and Biomedical Engineering*. IEEE, 2009, pp. 1–3.

[53] T.-y. YangDai and L. Zhang, "Weighted retinex algorithm based on histogram for dental ct image enhancement," in *2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. IEEE, 2014, pp. 1–4.

[54] C.-C. Huang and M.-H. Nguyen, "X-ray enhancement based on component attenuation, contrast adjustment, and image fusion," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 127–141, 2018.

[55] S. Chen, S. Zhong, L. Yao, Y. Shang, and K. Suzuki, "Enhancement of chest radiographs obtained in the intensive care unit through bone suppression and consistent processing," *Physics in Medicine & Biology*, vol. 61, no. 6, p. 2283, 2016.

[56] M. Eslami, S. Tabarestani, S. Albarqouni, E. Adeli, N. Navab, and M. Adjouadi, "Image-to-images translation for multi-task organ segmentation and bone suppression in chest x-ray radiography," *IEEE transactions on medical imaging*, vol. 39, no. 7, pp. 2553–2565, 2020.

[57] T. Madmad, N. Delinte, and C. De Vleeschouwer, "Cnn-based morphological decomposition of x-ray images for details and defects contrast enhancement," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 2170–2180.

[58] H.-Y. Kim, S. Park, Y.-G. Shin, S.-W. Jung, and S.-J. Ko, "Detail restoration and tone mapping networks for x-ray security inspection," *IEEE Access*, vol. 8, pp. 197 473–197 483, 2020.

[59] S. Anand, R. Roshan, *et al.*, "Chest x ray image enhancement using deep contrast diffusion learning," *Optik*, vol. 279, p. 170751, 2023.

[60] M. Pavoni, Y. Chang, S. ho Park, and Ö. Smedby, "Convolutional neural network-based image enhancement for x-ray percutaneous coronary

[61] A. Matsushima, T.-B. Chen, T. Okamoto, S.-Y. Hsu, N. Itayama, T. Ishibashi, and K. Fukuda, "Basic study for detection of nodules in chest radiography using contrast enhancement and faster R-CNN," in *Medical Imaging 2022: Physics of Medical Imaging*, W. Zhao and L. Yu, Eds., vol. 12031, International Society for Optics and Photonics. SPIE, 2022, p. 1203144. [Online]. Available: https://doi.org/10.1117/12.2611391

intervention," *Journal of Medical Imaging*, vol. 5, no. 2, p. 024006, 2018. [Online]. Available: https://doi.org/10.1117/1.JMI.5.2.024006

[62] P. J. Burt and E. H. Adelson, "The laplacian pyramid as a compact image code," in *Readings in computer vision*. Elsevier, 1987, pp. 671–679.

[63] Z. Liu, H. Hu, Y. Lin, Z. Yao, Z. Xie, Y. Wei, J. Ning, Y. Cao, Z. Zhang, L. Dong, *et al.*, "Swin transformer v2: Scaling up capacity and resolution," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 12 009–12 019.

[64] M. Narwaria, R. K. Mantiuk, M. P. Da Silva, and P. Le Callet, "Hdr-vdp-2.2: a calibrated method for objective quality prediction of high-dynamic range and standard images," *Journal of Electronic Imaging*, vol. 24, no. 1, pp. 010 501–010 501, 2015.

[65] T. Fukiage and T. Oishi, "A content-adaptive visibility predictor for perceptually optimized image blending," *ACM Transactions on Applied Perceptions (TAP)*, 2022.

[66] T. O. Aydin, R. Mantiuk, K. Myszkowski, and H.-P. Seidel, "Dynamic range independent image quality assessment," *ACM Transactions on Graphics (TOG)*, vol. 27, no. 3, pp. 1–10, 2008.

[67] S. Jaeger, A. Karargyris, S. Candemir, L. Folio, J. Siegelman, F. Callaghan, Z. Xue, K. Palaniappan, R. K. Singh, S. Antani, *et al.*, "Automatic tuberculosis screening using chest radiographs," *IEEE transactions on medical imaging*, vol. 33, no. 2, pp. 233–245, 2013.

[68] S. Candemir, S. Jaeger, K. Palaniappan, J. P. Musco, R. K. Singh, Z. Xue, A. Karargyris, S. Antani, G. Thoma, and C. J. McDonald, "Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration," *IEEE transactions on medical imaging*, vol. 33, no. 2, pp. 577–590, 2013.

[69] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[70] L. J. Guo, "Balance contrast enhancement technique and its application in image colour composition," *Remote Sensing*, vol. 12, no. 10, pp. 2133–2151, 1991.

[71] F. Wilcoxon, *Individual comparisons by ranking methods*. Springer, 1992.

[72] L. L. Thurstone and T. G. Thurstone, *Primary mental abilities*. University of Chicago Press Chicago, 1938, vol. 119.

RYOICHI ISHIKAWA received the B.E. degree from the Department of Electrical Engineering, The University of Tokyo, Japan, in 2014, and the M.E. and Ph.D. degrees in Electrical Engineering and Information Systems, The University of Tokyo, Japan, in 2016 and 2019, respectively. From 2019, he is a Project Researcher at the Institute of Industrial Science, The University of Tokyo. His research interests include sensor fusion, calibration, and X-ray vision.

TOMOHISA YUZAWA received the B.E degree from the Department of Biological Production, Tokyo University of Agriculture and Technology, Japan, in 2018, and the M.E. degrees in Department of Integrated Biosciences, The University of Tokyo, Japan.

IEEE *Access*

TAIKI FUKIAGE is a Researcher at Sensory Representation Group of Human Information Science Laboratory in NTT Communication Science Laboratories. He received the Ph.D. in Interdisciplinary Information Studies from the University of Tokyo in 2015. He joined NTT Communication Science Laboratories in 2015, where he studies media technologies based on scientific knowledge about visual perception. He is a member of the Vision Sciences Society and the Vision Society of Japan.

MASATAKA KAGESAWA received the B.S. degree in mathematics from Chiba University, Japan, in 1986 and the M.S. degree in mathematics from Tokyo Metropolitan University, Tokyo, Japan, in 1988. He was a doctor course student at Tokyo Metropolitan University from 1988 to 1990. He received the Ph.D. degree in electric engineering from the University of Tokyo, Tokyo, Japan, in 2002. From 1990 to 1994, he was a Technical Associate at the Institute of Industrial Science, the University of Tokyo. Since 1994, he has been a Research Associate at the same university. His research interests include intelligent traffic systems, digitization of cultural Heritage, virtual reconstruction of ancient objects.

TORU WATSUJI received his B.E. degree from the Department of Agricultural Chemistry, Kobe University, Japan, in 1985, and his M.Sc. and Ph.D. in Science from the Department of Biotechnology, Imperial College London, in 1987 and 1992, respectively. He was a contract researcher at Medisense, Inc. in 1987 and a research assistant in the Department of Haematology at the University of Cambridge in 1988. Subsequently, he worked for Sharp Corporation from 1989 to 2017 and for MTI Ltd. until 2020. Currently, he is employed at Air Water Inc.

TAKESHI OISHI (Member, IEEE) received the B.Eng. degree in electrical engineering from Keio University, in 1999, and the Ph.D. degree in interdisciplinary information studies from The University of Tokyo, in 2005. He is currently an Associate Professor at the Institute of Industrial Science, The University of Tokyo. His research interests include 3D modeling from reality, digital archiving of cultural heritage assets, and mixed/augmented reality.

• • •