Deblurring Low-Resolution X-ray Images Using CNN

Seokwon Oh [a], Jinwoo Kim[b], Ho Kyung Kim [a,b,*]
[a] *School of Mechanical Engineering, Pusan Nat'l Univ., Busandaehakro 63beon-gil, Busan 46241*
[b] *Center for Advanced Medical Engineering Research, Pusan Nat'l Univ., Busandaehakro 63beon-gil, Busan 46241*
[*] *Corresponding author: hokyung@pusan.ac.kr*

1. Introduction

Ill-posed problems such as deblurring, denoising, and super-resolution, are key issues in the computer vision field. In general, indirect-conversion x-ray detectors suffer from low spatial resolution due to the x-ray and light spreading in the phosphor layer, which is converting incident x-rays to light photons. This signal spreading makes it difficult to inspect micro-defects in samples such as ball-grid arrays in printed circuit boards (PCBs). While the signal spreading can be reduced by using a thinner phosphor layer, it requires high x-ray exposure, which leads to a high tube load.

In this study, we propose a deep learning method, which is one of the most powerful tools for ill-posed problems to deblur the low-resolution PCB images. Two network models based on the U-Net [1] architecture are proposed and the training and testing image sets are obtained experimentally using x-ray detectors with two different-thick phosphor layers. The performances of proposed networks are quantitatively evaluated in terms of spatial resolution and noise.

2. Methods and Materials

2.1 Datasets and preprocessing

Datasets for training and testing the networks were obtained experimentally by a lab-made x-ray imaging system. An x-ray detector was composed of a 1024 × 512-formatted photodiode array (RadEye 1^{TM}, Teledyne Rad-icon Imaging Corp.) and replaceable phosphor layers. Two gadolinium oxysulfide (Gd2O2S:Tb) phosphor layers with different thicknesses were considered; 90 μm for high-resolution (HR) imaging and 300μm for low-resolution (LR) imaging. X-ray images for 14 different types of PCB samples were acquired using 60 kVp and 90 kVp tungsten-target x-ray spectra and two phosphor layers. The LR and HR images were aligned using the homography method to minimize geometric misalignment due to the replacement of phosphor layers. The number of images was augmented 6 times by rotating and flipping and each image was divided into 10 sub-patches. Therefore, we have 3360 sub-patches in total.

2.2 Network architecture

The signal spreading in the LR images y can be expressed briefly by,

$$y = H * x + v, \quad (1)$$

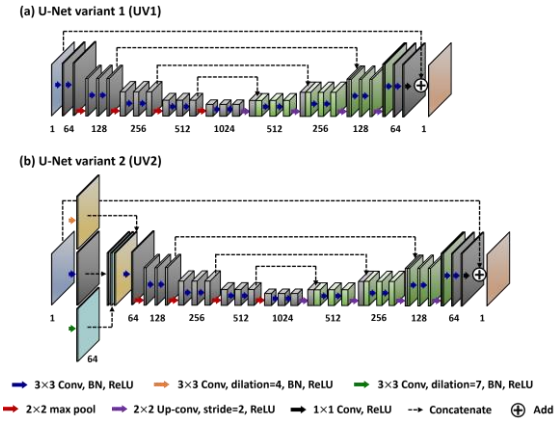

Fig. 1. Architectures of the proposed network models: (a) U-Net variant 1 and (b) U-Net variant 2.

where H denotes the point-spread function of the LR detector, x is the ideal projection image, and v means additional electronic noise. The networks should be trained to predict the ideal projection image x from LR images y, but we set the HR images as a target rather than x since they cannot be acquired experimentally.

Figure 1 shows two U-Net variant networks proposed for the restoration of the LR images. As shown in Fig. 1(a), U-Net variant 1 (UV1) has the same architecture as U-Net but a skip-connection between input and output is employed for the preservation of high-frequency information and stability of the network. U-Net variant 2 (UV2) also has the skip-connection between input and output but it has three different types of convolutional kernels, which can account for the large signal spreading in the LR detector, for its first layer. Conventional 3 × 3 convolutional kernels may not be enough to cover the point-spread function of the LR detector. In this regard, the dilated convolutional kernels with dilation 4 and 7 are used to the first layer of UV2.

2.3 Performance evaluation

The performances of the proposed networks are comparatively evaluated using the peak signal-to-noise ratio (PSNR) [2] and structural similarity (SSIM) [3] index between output and HR images. Furthermore, the modulation-transfer function (MTF) [4] and normalized noise-power spectra (NNPS) [5] are evaluated to investigate the resolution and noise performances in the spatial-frequency domain. The MTF and NNPS are respectively calculated using additional edge-knife and flood-field images, which are deblurred by networks.

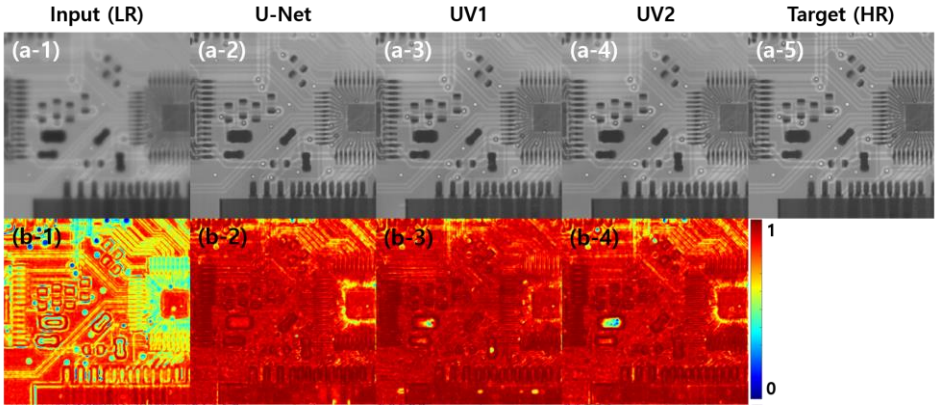

Fig. 2. Comparison of (a) the projection images for a PCB sample and (b) the corresponding SSIM distributions. The projection images are displayed in a range of $\mu \pm 4\sigma$.

The NNPS is multiplied by air kerma K at the detector surface for a fair comparison.

3. Preliminary Results

Tab. I summarizes the PSNR and SSIM values for the test images. the proposed networks show slightly higher performances compared to the U-Net.

Fig. 2 compares the deblurred images and corresponding SSIM distributions. Visual inspection validates that all the networks improved the spatial resolution of a given LR image.

Fig. 3 shows the MTF and $K \cdot NNPS$ performances for images predicted from various network models. U-Net shows the best fit for the target MTF but provides slightly lower $K \cdot NNPS$ performance at the mid-frequency ranges (3~6 mm^{-1}) compared to the proposed networks.

Table I: Summary of the PSNR and SSIM values.

	U-Net	UV1	UV2
PSNR	24.059 ±1.200	24.783±3.293	25.337±3.036
SSIM	0.879 ±0.019	0.892 ±0.028	0.879 ±0.030

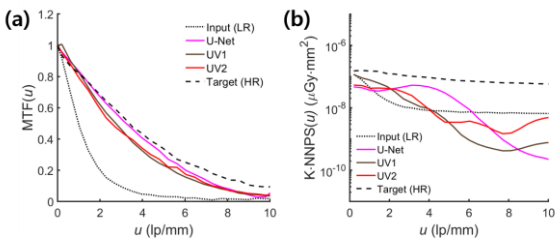

Fig. 3. Comparison of (a) MTF and (b) $K \cdot NNPS$ for various networks.

4. Discussion and Conclusion

Two U-Net-based networks for deblurring low-resolution x-ray images have been investigated. The deblurring performance of U-Net was improved by adding the skip connection and the additional convolutional filters that can cover the extent of the point-spread function of the LR detectors. The spatial-frequency analysis using MTF and NNPS showed good agreement with the conventional PSNR and SSIM assessment. This evaluation method may be helpful to figure out how the neural networks work in a specific frequency.

However, since the networks were trained only with the PCB images, the predicted images may suffer from unexpected artifacts when the edge-knife images, which have a much simpler structure, are input to the networks. These artifacts may distort the MTF curve. The contrast-transfer function measurement using line-pair patterns may be an alternative for assessing spatial-resolution characteristics. More detailed analysis in the spatial-frequency domain and training and testing the networks with various resolutions will be our future studies.

ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2021R1A2C1010161).

REFERENCES

[1] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, Medical Image Computing and Computer-Assisted Intervention (MICCAI), Vol. 9351, pp. 234-241, 2015.
[2] Z. Wang, A. C. Bovik, Mean Squraed Error: Love It or Leave It? A New Look at Signal Fidelity Measures, IEEE Signal Processing Magazine, Vol. 26, No. 1, pp. 98-117, 2009.
[3] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image Quality Assessment: from Error Visibility to Structural Similarity, IEEE Transactions on Image Processing, Vol. 13, No. 4, pp. 600-612, 2004.
[4] E. Samei, J. F. Michael, A. R. David, A Method for Measuring the Presampled MTF of Digital Radiographic

Systems Using an Edge Test Device, Medical Physics, Vol. 25, No. 1, pp. 102-113, 1998.
[5] A. D. A. Maidment, M. Albert, P. C. Bunch, I. A. Cunningham, J. T. Dobbins III, R. M. Gagne, R. M. Nishikawa, R. L. Van Metter, R. F. Wagner, Standardization of NPS Measurement: Interim Report of AAPM TG16, In Medical Imaging 2003: Physics of Medical Imaging, Vol. 5030, pp. 523-532, 2003.