Attention-Embedded Decomposed Network with Unpaired CT Images Prior for Metal Artifact Reduction

Binyu Zhao[1], Qianqian Ren[1,*], Jinbao Li[2], Yafeng Zhao[1]

[1] *Department of Computer Science and Technology, Heilongjiang University, Harbin, 150080, China*

[2] *Shandong Artificial Intelligence Institute, Qilu University of Technology (Shandong Academy of Science), Jinan, 250014, China*

* *Email : renqianqian@hlju.edu.cn*

INTRODUCTION

Existing unsupervised methods for metal artifact reduction has two limitations:
1) only use features in image space, which is not enough to restore heavily affected regions;
2) lack the distinction and selection for effective features.

In this study, we use a pair of complementary networks: Content Extraction Network (*CEN*) and Artifact Extraction Network (*AEN*) to decompose the content component and metal artifact component from artifact-affected images, respectively; Image prior is adopted in *CEN* by first inpainting in sinogram space and then refining in image space; Attention mechanism and skip connection are embedded to normal convolutional operation throughout the network to show preference for valid and vital features adaptively.

METHOD

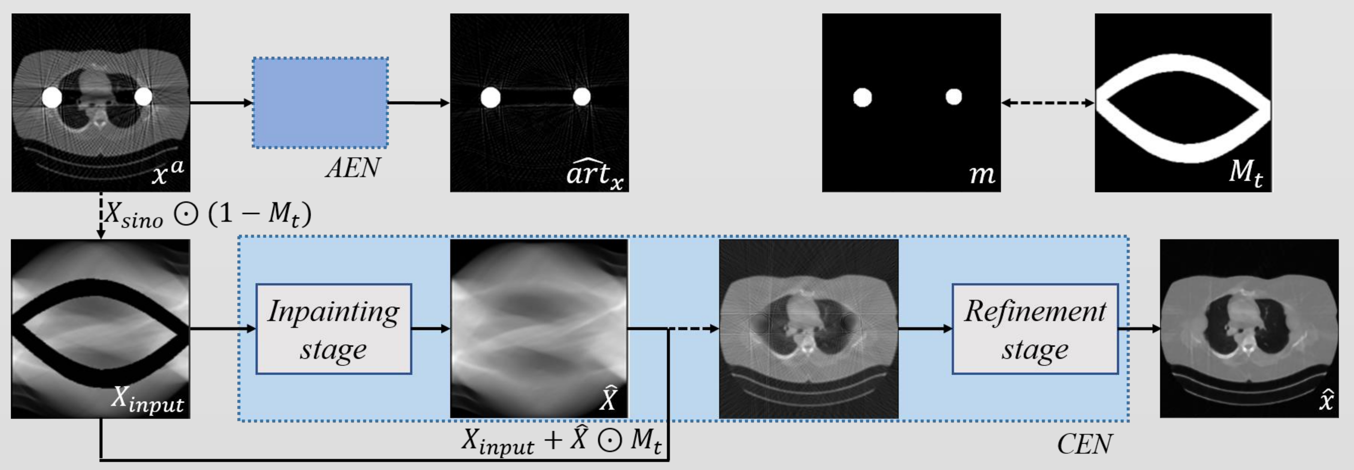

Figure 1: Workflow of *AEDNet*. *AEDNet* has two branches (*AEN* and *CEN*). *AEN* is applied to further correct the output of *CEN*.

$$\hat{X},\hat{x} = G_{CEN}(X_{input})$$
$$\hat{Y},\hat{y} = G_{CEN}(Y_{input})$$
$$\widehat{art}_x = G_{AEN}(x^a)$$
$$\widehat{art}_y = G_{AEN}(y)$$

Total objective function: $\mathcal{L}_{total} = \lambda_{CEN}\mathcal{L}_{CEN} + \lambda_{AEN}\mathcal{L}_{AEN} + \lambda_{reg}\mathcal{L}_{reg}$

$$\mathcal{L}_{reg} = ||\widehat{art}_y||_1$$
$$\mathcal{L}_{AEN} = ||(\widehat{art}_x + \hat{x}) - x^a||_1$$
$$\mathcal{L}_{CEN} = ||\hat{Y} - Y||_1 + ||\mathcal{R}^{-1}(Y_{input} + \hat{Y} \odot M_t) - y||_1 \odot (1-m) + ||\hat{y} - y||_1 \odot (1-m)$$

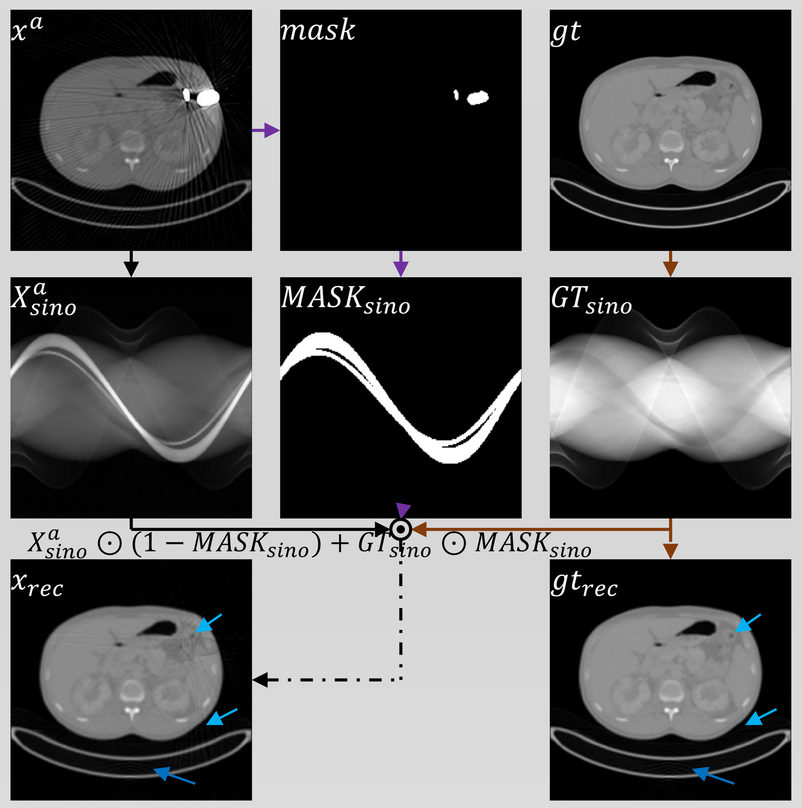

Figure 2: Explanation on problem that residual artifacts exists even if perfectly correction is finished with image prior. This is the reason to use *AEN*.

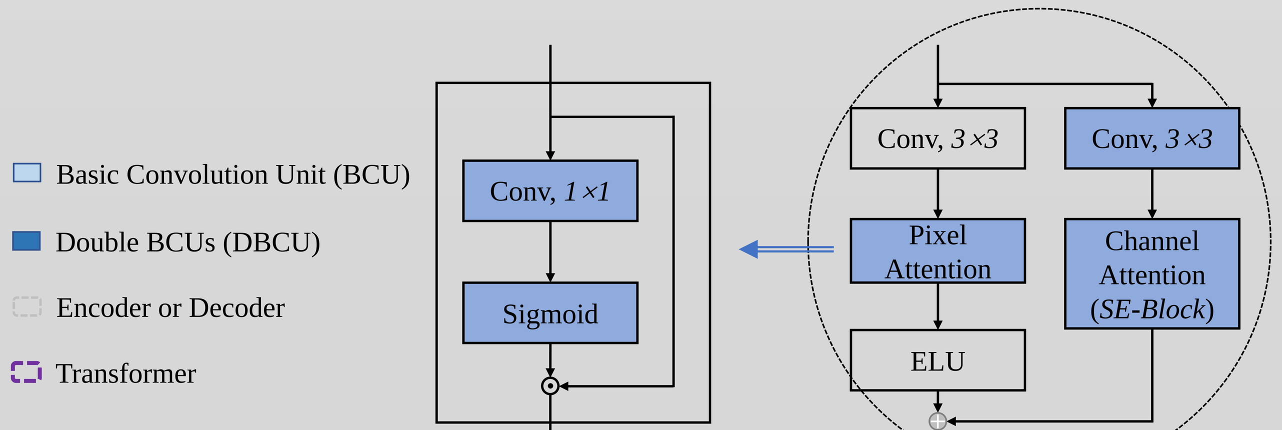

Figure 3(a): Attention and skip connection designs.

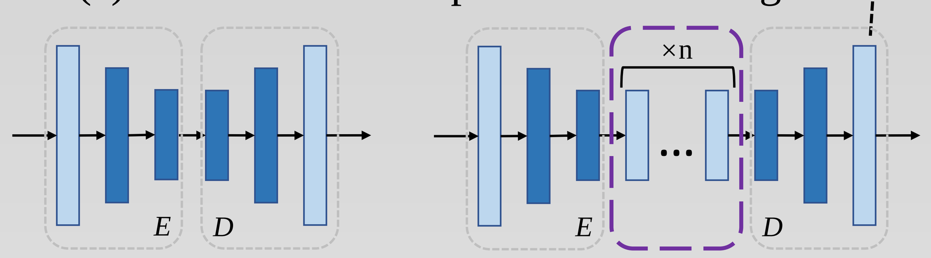

Figure 3(b): Architecture of *AEN (left)*. Figure 3(c): Architecture of Inpainting stage and Refinement stage of *CEN (right)*.

RESULTS

Table 1: Quantitative evaluation on synthesized dataset (DeepLesion).

	Method	PSNR (dB)	SSIM (%)
Supervised	cGANMAR	34.60	92.89
	R²Net	37.05	94.41
	DuDoNet++	38.76	96.96
Unsupervised	ADN	33.88	92.17
	3DGAN	29.16	87.13
	AEDNet (Ours)	**34.53**	**94.42**

Table 2: Ablation Study on synthesized dataset (DeepLesion).

Components	(a)	(b)	(c)	(d)
AEN	×	√	√	√
Skip Connection	√	×	√	√
Attention Mechanism	√	×	×	√
PSNR (dB)	34.36	34.08	34.15	**34.53**
SSIM (%)	94.14	92.75	93.45	**94.42**

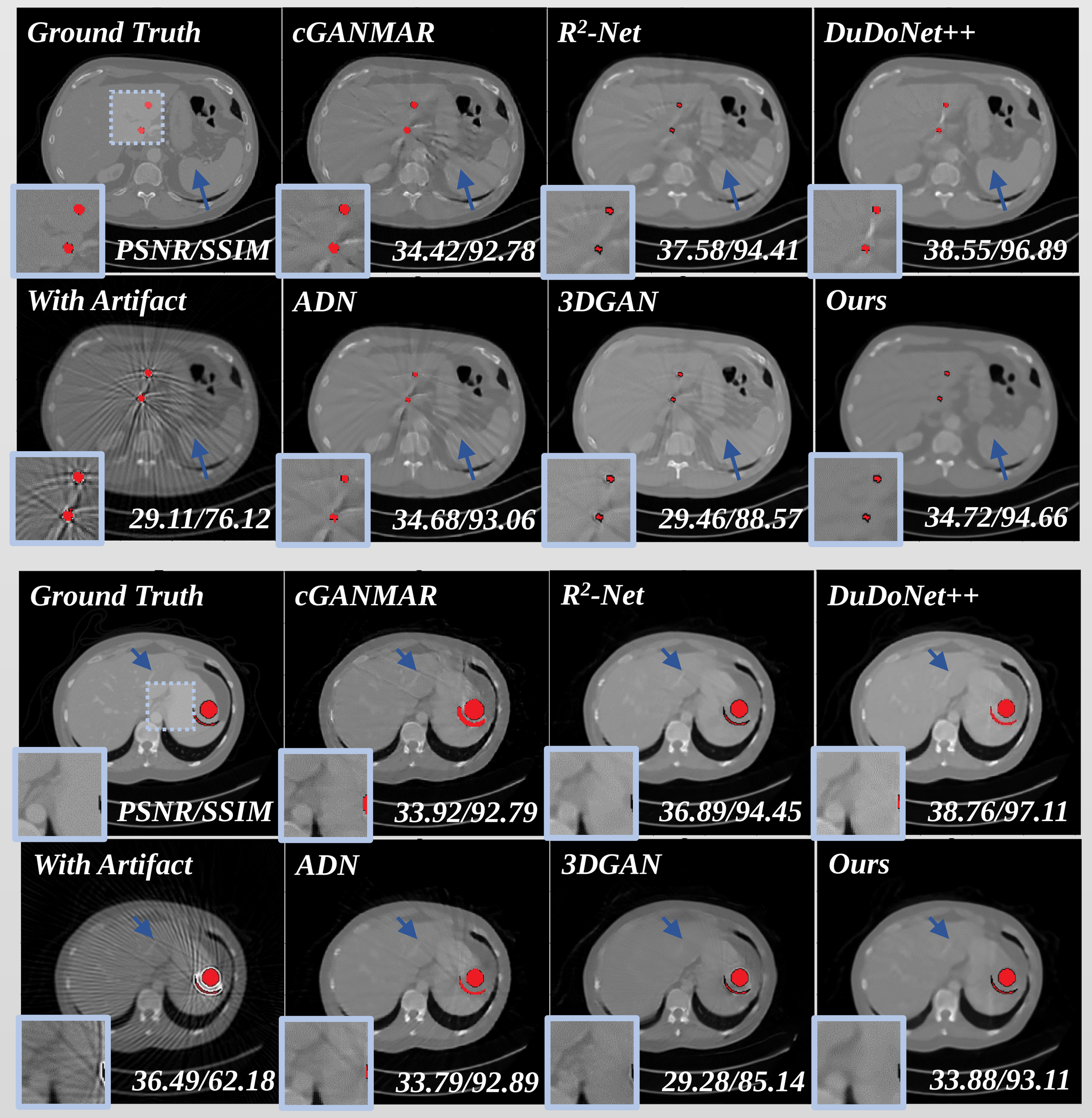

Figure 4: Qualitative comparison with state-of-the-art methods on synthesized dataset (DeepLesion).

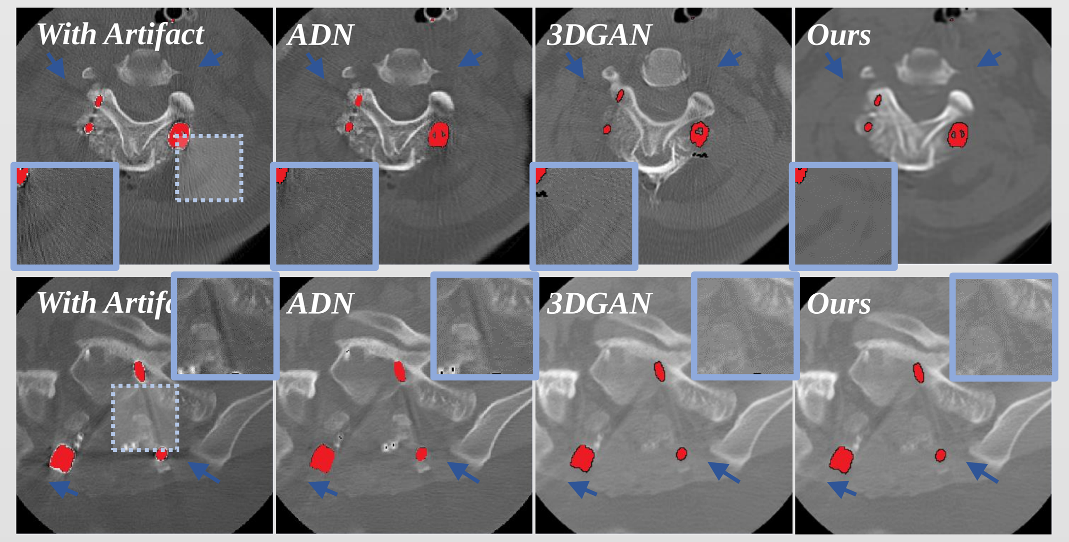

Figure 5: Qualitative comparison with state-of-the-art methods on clinical dataset (vertebrae localization and identification dataset).

Synthesized dataset – DeepLesion: 1) Training set: 1120 synthesized paired images with 90 metal shapes; 2) Test set: 200 synthesized paired images with 10 metal shapes.

Clinical dataset - the vertebrae localization and identification dataset from Spineweb: 1) Training set: 600 images artifact-affected images and 3298 clean images; 2) Test set: 183 artifact-affected images.

Ablation study is conducted on synthetic dataset (DeepLesion) to understand the effectiveness of AEN and the embeddings of BCU.

CONCLUSION

○ The utilization of image prior achieves unsupervised learning for metal artifact reduction;

○ The utilization of decomposed method effectively suppresses residual artifacts;

○ The utilization of attention mechanism and skip connection provides feature fusion and additional flexibility in focusing on important information.

AEDNet successfully alleviates secondary artifact and recovers more details for artifact-affected CT images

FUTURE WORKS

● Blurs appear in corrected images due to information loss in too deep networks. Searching for balance between image deblurring and the effectiveness;

● Consider to validate the robustness of the model and generalize it for more clinical situations.

● …