Low Dose Abdominal CT Image Reconstruction: An Unsupervised Learning Based Approach

Shiba Kuanar, Vassilis Athitsos, K.R. Rao
VLM Lab, University of Texas at Arlington

Dwarikanath Mahapatra
IBM Research, Australia

Zahid Akhtar, Dipankar Dasgupta
University of Memphis, USA

Task: To implement a GAN based auto-encoder network for low dose CT image reconstruction and de-noise the CT images.

Goal: To learn the latent feature maps, achieves more accurate, and visually pleasing Image reconstructions.

Dataset

- 4250 patient abdomen quarter dose CT images collected from publicly available cancer imaging archive (TCIA) and AAPM-Mayo Clinic Low Dose 2018 datasets for training and validation.
- Among them, 3800 selected images are used to train encoder-decoder network and generate image manifolds in latent space. Rest are considered for testing purpose.
- During the training step, 10% of training data are employed as validation to monitor the network performance and tuned the hyper parameters.
- All the CT images have a resolution of 512×512 pixels and a pixel size of 0.875×0.875 mm2.
- Images with a different resolution are resampled to this resolution via bilinear interpolation.

Preprocessing

- Hounsfield unit (HU) calibration in CT DICOM images
- Rescale using linear transformation (Y = info.RescaleSlope.*Y + info.RescaleIntercept)
- Shift (−1000, Air) [to make all HU values in range (0,4000) or more than > 0]
- Voxel range adjustment (DICOM PixelSpacing, imresize)
- 512 × 512 image resize w.r.t to center pixel for Manifold learning
- Normalize Image

Features

- Our network first maps CT images to low dimensional manifolds and then restore the images from its corresponding manifold representations.
- To overcome the dissimilarity, a separate perceptual loss function is included to our network for feature learning
- Our algorithm learns the latent feature maps (latent space) and achieves more accurate image reconstructions.

Our contribution:

- **Problem**: X-ray Computed tomography-based scans expose a high radiation dose and lead to the risk of prostate or abdomen cancers. On the other hand, the low-dose CT scan can reduce radiation exposure to the patient. But the reduced radiation dose degrades image quality for human perception, adversely affects the radiologist's diagnosis. Hence, the need of better reconstruction algorithm.
- Approach: GAN based auto-encoder network to de-noise the CT images**.**
- **Plus:** Significant gains in efficiency.
  - Model provides higher PSNR, SSIM, and better statistical properties of denoised CT images relative to those of normal CT images**.**

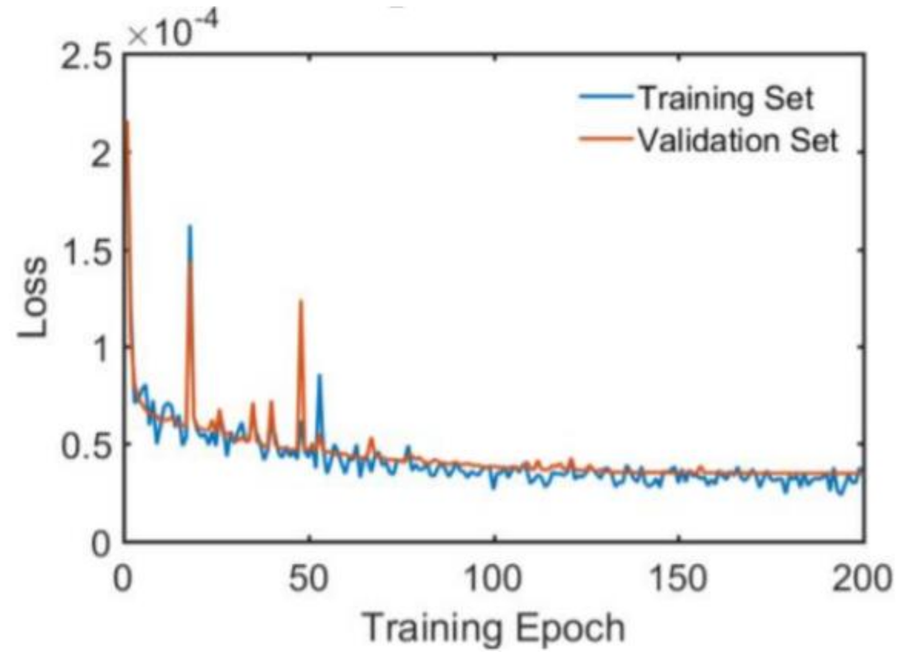


Fig. 1:Loss graphs of our model on dataset

GAN Network Models

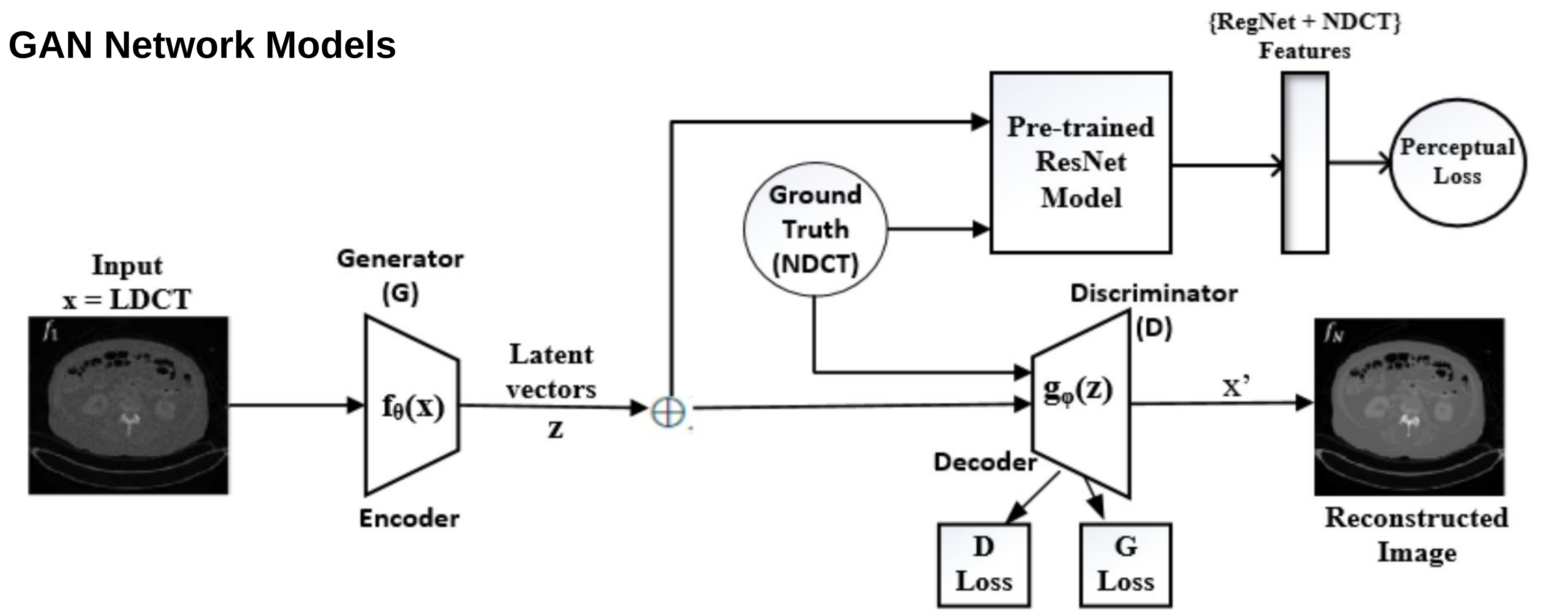


Fig. 2: Framework overview: Auto-encoder architecture: Generator, Discriminator, and RegNet for perceptual loss (i) EEG signals from multiple cortex locations classification.

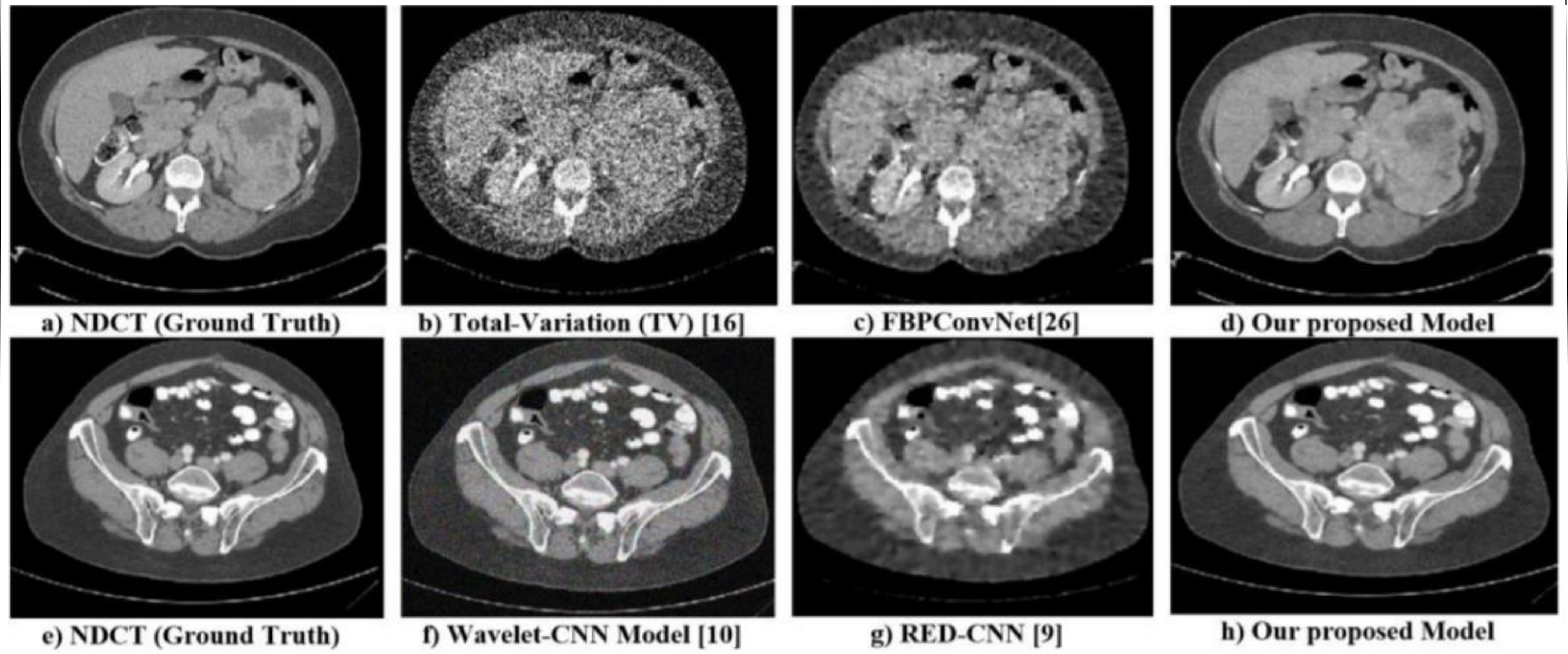


Fig.3: Reconstruction comparison results on test samples. (a), (e) Ground truth; (b) Iterative total-variation; (c) FBPConvNet; (d), (h) Our proposed method; f) Wavelet-CNN method; g) RED-CNN method respectively. The CT images are displayed in a window [-210, 300] HU.

Auto-Encoder Network:

Our model includes a scalable architecture with three components.

- CNN generator includes 6 conv layers. Input CT image to generator passes through a stack of conv layers with various receptive fields, 3×3 kernel followed by 2×2 max-pooling in each layer.
- All hidden layers are equipped with ReLU, max (0, x) and applies thresholding on filter responses.
- Network contains a pre-trained RegNet network (upper half in Fig. 2) and calculates perceptual loss for better image enhancement. The output manifold (z) from the generator and ground truth are fed into the RegNet pre-trained network for respective feature extraction.
- Objective loss function (in Eq. 2) is computed using the extracted features from previous 2 steps
- The reconstruction errors are then back propagated to update the generator weight while keeping the RegNet parameters intact.
- Discriminator D includes 8 convolution layers. After 8 convolution layers, two fully-connected (FC) layers are included, of which first has 1024 outputs and last layer has a single output.
- Network is trained using image patches derived from entire images. After each epoch, we calculated the loss over all image patches for validation.
- it is observed (Fig. 1) that increasing the no of epochs reduces Wasserstein distance (as decay rate becomes smaller). This indicates effectiveness of our RegNet loss introduced in WGAN-RegNet model.

Objective function and Perceptual Loss:

- To overcome the dissimilarity, a separate perceptual loss function is included in the feature space to keep the image details and represented as below:

$$L_{RegNet}(D, G) = E_{x,z}\left[\frac{1}{WHd}\|RegNet(f_\theta(x)) - RegNet(NDCT)\|_F^2\right] \quad (1)$$

  Where RegNet (fθ(x)) is a feature vector extractor, and d, W, and H stand for depth, width.
- Our final loss function is expressed as:

$$\min_G \max_D L_{WGAN}(D, G) + \lambda_1 L_{RegNet}(G) \quad (2)$$

  where λ1 is a regularization term which controls the trade-off between RegNet perceptual loss and WGAN adversarial loss

Results and Analysis

Table I: Statistical Analysis (Mean ± SD) of Image Quality associated with different Models

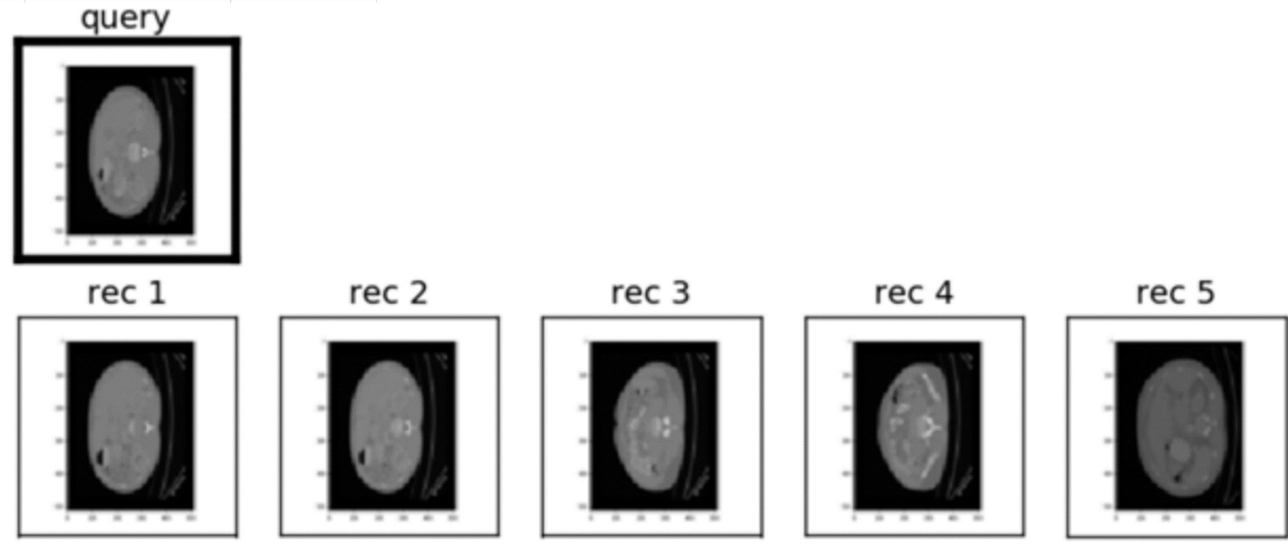
Evaluation Metrics	Reference		TV [15]	FBPConvNet[26]	Wavelet-CNN [10]	Our Model
Artifact Reduction	R1	3.48 ± 0.33	2.16 ± 018	3.22 ± 0.39	3.78 ± 0.01	4.18 ± 0.29
	R2	3.55 ± 0.21	2.21 ± 0.13	3.27 ± 0.44	3.61 ± 0.24	4.06 ± 0.15
Contrast Retention	R1	3.32 ± 0.88	2.44 ± 0.48	3.41 ± 0.11	4.17 ± 0.55	4.22 ± 0.02
	R2	3.74 ± 0.91	2.26 ± 0.11	3.38±0.15	3.82 ± 0.39	4.19 ± 0.16
Noise Suppression	R1	3.49 ± 0.66	2.29 ± 0.72	3.18 ± 0.10	3.86 ± 0.63	3.95 ± 0.72
	R2	3.71 ± 0.14	2.31 ± 0.35	3.26 ± 0.64	3.65 ± 0.13	4.08 ± 0.17
Over all Image Quality	R1	3.83 ± 0.53	2.47 ± 0.12	3.28 ± 0.55	3.92 ± 0.41	4.23 ± 0.59
	R2	3.97 ± 0.47	2.58 ± 0.14	3.05 ± 0.33	3.56 ± 0.27	3.92 ± 0.66

Table II: Quantitative results from different model outputs

Models	PSNR	SSIM	RMSE
TV [15]	24.21	0.562	0.0392
FBPConvNet[26]	29.94	0.865	0.0574
RED-CNN [9]	30.82	0.892	0.0418
Wavelet-CNN [10]	33.51	0.913	0.045
Our Model	37.76	0.944	0.0092

Future Direction:

- We would like to extend our model to find the image similarity search on latent space over huge clinical image datasets.
- This can be helpful in finding better treatment plans and spatial accuracy in dose delivery for diagnosis and prognosis.



Concluding Comments

- Quantitative analysis shows that our auto-encoder model provides higher PSNR, SSIM, and better statistical properties of denoised CT images relative to those of normal CT images.
- Experimental results on the clinical real images show that our proposed model not only removes sharp features effectively but also generates an image with increased contrast.

Acknowledgments

- This work performed during author's summer 2018 internship at Radiation Oncology department at UTSW Medical School and later enhanced by UT Arlington research collaborations.