Deep Learning method for denoising CT images

Avantika Kothandaraman, Caiwei Zhang, Long Chen
avantk@bu.edu, magzcw@bu.edu, lchen122@bu.edu

1. Task

The goal of this project is to develop a deep learning model that can accurately estimate the standard deviation (STD) map of noise in CT scans. This model will intake raw CT scan data like those shown in Fig.1, process it through a U-Net architecture, and output a detailed STD map, which reflects the localized noise levels across the scan.

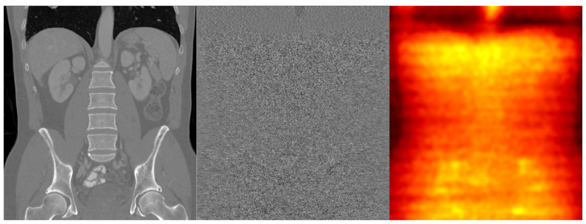

Figure 1. Left: Original CT scan signal; Center: Isolated noise component; Right: Standard deviation map

One of the main challenges of this task is to maintain the precise spatial relationship of the anatomical structures as deep learning models can discern and quantify noise while distorting the medical details. Besides, the model must be robust to variations in noise patterns, which may not be uniform across different scans.

2. Related Work

In [1], the study demonstrates the advantages and limitations of using the Advanced Intelligent Clarity Engine (AiCE) compared with traditional adaptive iterative dose reduction 3D (AIDR 3D) in abdominal CT examinations. Benefits include AiCE's ability to significantly reduce radiation dose by more than 40% while improving image contrast-to-noise ratio (CNR) and reducing image noise without loss of image quality. These improvements are consistent across all body mass index (BMI) categories. Furthermore, AiCE showed potential advantages in identifying hepatocellular carcinoma (HCC), especially in dynamic CT imaging. However, limitations of this study include a small sample size, being restricted to a single institution as a retrospective study, and focusing primarily on two reconstruction algorithms provided by the same manufacturer, which may limit the generalizability of the results.

In [2], According to our findings, low-dose liver CT (LDCT) using deep learning denoising (DLD, ClariCT) was shown to be comparable in overall image quality to standard-dose liver CT (SDCT) using multi-iteration reconstruction technology (MBIR, ADMIRE) Compared to non-inferiority, the dose was reduced by 67%. In the quantitative evaluation, the results of LDCT were significantly better than those of SDCT ($p < 0.05$), while in the qualitative evaluation, the image quality scores were above average in most cases. Performance evaluation metrics for lesion detection did not show statistical differences. On the other hand, disadvantages include that reviewers are aware of liver disease. Only the portal venous phase of liver CT was evaluated and not the arterial phase, which is important in detecting hepatocellular carcinoma. The DLD algorithm used in the study, while vendor-agnostic, was only tested on equipment from one vendor (Siemens Healthineers), and further research is needed to confirm that this method of combining mathematical noise reduction and artificial intelligence will work on other vendors. whether similar results can be provided on the manufacturer's equipment.

This study [3] examines image noise and radiomic features in dual-energy CT images reconstructed using a novel deep learning image reconstruction algorithm and other traditional iterative reconstruction algorithms. The results show that compared with other traditional IR algorithms, the DLIR algorithm reduces image noise while maintaining the spatial resolution of the image. This is achieved through signal-to-noise ratio (SNR), contrast-to-noise ratio (CNR) and noise power spectrum (NPS). value is reflected. However, DECT radiomic signatures have high reproducibility between different scans but limited reproducibility between different reconstruction algorithms and different scanners. The DLIR algorithm changes the DECT radiomic signature, resulting in higher consistency with images reconstructed by other traditional IR algorithms than with images reconstructed by ASIR-V. However, the ability of the DLIR algorithm to improve DECT image quality is still a relatively unknown area and requires further analysis and accumulation of clinical evidence.

Although the overall reproducibility of the features was low across different DECT scanners and reconstruction algorithms, the standard identified nine features as robust across DECT scanners and reconstruction algorithms, which provided the basis for further validation and Clinically relevant analysis is possible.

3. Approach

The method proposed in [4] talks about the use of a Deep Learning framework called SILVER to quantify the noise in CT images that a scanner produces. However, it was seen that the estimation of noise in anatomical regions like the lungs was not quite up to the mark. The idea is to create a Standard Deviation (local pixel-wise variance estimation) map of the noise in subsequent, continuous images produced by the scanner for a set of anatomical regions. To further improve the functionality of the model, we propose certain modifications to the workflow. We first load the CT scan data along with the corresponding noise and STD maps. Then, we apply data augmentation techniques such as rotations, reflections, and adding random noise to increase the diversity of the training data, and eventually split the dataset into training, validation, and test sets (80-10-10). The model is trained using both UNet and RatUNet architectures to compare their performance in estimating the STD map. We then implement uncertainty estimation techniques to quantify the model's uncertainty in its predictions. This will help us identify which patches the model struggles to estimate properly. Frontrunners for uncertainty estimation are Monte Carlo Dropout or Bayesian Networks. This is to be followed by a comparative analysis where we train the model using different loss functions (Mean Squared Error, Average Relative Loss and Huber Loss) and compare their performance.

Once the comparative analysis is done, the better model is chosen for further training, validation, optimization and testing. For optimization, we wish to try the Adam optimizer. The trained model will then be validated. As for the final display of results, we visualize the STD map generated by the model to understand the spatial distribution of noise standard deviation across the CT scans. We will also use uncertainty estimation to identify the regions where the model’s predictions are less confident. This is to be done by highlighting the patches of high doubt. We also wish to present Peak Signal-to-Noise Ratio (PSNR) as a part of the final result.

Our approach is based on performing comparative analyses of different model architectures and loss functions. Uncertainty estimation is used to highlight the regions of doubt (patches where the model is not confident of estimating). This helps in isolating just the low-confidence patches alone to be assessed manually by professionals. Further improvements could be made for Iterative Refinement of the STD estimation of these patches.

4. Dataset and Metric

The data set provided by PI consists of ten different CT images. Each image consists of three independent components, namely CT signal, noise signal and STD Map, all in NRRD format, with the former 2 signals in Hounsfield Units (HU). During augmentation, we intend on including newer data with added noise and altered structures. We will carry out an 80-10-10 train-test-validation split. The MONAI library will also be tried during training. Since the goal is to estimate the level of noise and produce an STD Map of the same, our final result will be to produce an STD Map of the various approaches that we try out. Since we are performing a comparative analysis of different DL architectures, we will assess how these models are able to produce the desired STD maps based on accuracy. We will also be comparing the effect of the different loss functions on the accuracy. We will use confidence intervals as a numeric metric for the Uncertainty Estimation. Finally, we will visualize the generated STD Noise Maps as heatmaps, with the “less confident” regions highlighted on them.

5. Approximate Timeline

Make a plan with approximate deadlines, e.g.

Task	Deadline
Data Preprocessing	03/08/2024
Training Data Preparation and Augmentation	03/18/2024
U-Net architecture ready for training	03/28/2024
Initial training with preliminary results	04/03/2024
Have a validated and tested model	04/14/2024
Prepare report and presentation	04/24/2024

References

1) A. Tamura, E. Mukaida, Y. Ota, I. Nakamura, K. Arakita, and K. Yoshioka, "Deep learning reconstruction allows low-dose imaging while maintaining image quality: comparison of deep learning reconstruction and hybrid iterative reconstruction in contrast-enhanced abdominal CT," arXiv.org, Iwate, Japan, 2023.
2) S. Park, J. H. Yoon, I. Joo, M. H. Yu, J. H. Kim, J. Park, S. W. Kim, S. Han, C. Ahn, J. H. Kim, and J. M. Lee, "Image quality in liver CT: low-dose deep learning vs standard-dose model-based iterative reconstructions," 2021

3) Zhong, J., Xia, Y., Chen, Y., Li, J., Lu, W., Shi, X., Feng, J., Yan, F., Yao, W., & Zhang, H. Deep learning image reconstruction algorithm reduces image noise while alters radiomics features in dual-energy CT in comparison with conventional iterative reconstruction algorithms: a phantom study, 2022.
4) Huber, N. R., Kim, J., Leng, S., McCollough, C. H., & Yu, L, "Deep Learning-Based Image Noise Quantification Framework for Computed Tomography." Journal of computer assisted tomography, 47(4), 603–607.