Seeing the Unseen: Equipping Deep Learning to Enhance Images for Lesion Detection in CT Scans

Nishq Poorav Desai (20320096)
Queen's University, Electrical and Computer Engineering, Kingston, Canada, K7L 2S8

Abstract.
Purpose: The proposed research aims to propose an all-in-one solution for detecting lesions of varying size, severity and occurring locations from Full Body CT scan images with diversified noise and contrast levels. The final goal is to automate enhancing image quality and lesion detection via CNN-based Local and Global Enhancement.
Approach: A ResNet-based Faster RCNN is trained on the CT scan images from DeepLesion Dataset. An additional ResNet18 CNN model is trained to predict parameters to enhance the Local and Global Contrast and remove underlying noise while preserving the image details.
Results: The image enhancement pipeline was able to enhance the performance from 30.77% mAP to 36.19% mAP. Thus, showing a significant improvement in lesion detection without retraining the pre-trained Object-Detector model.
Conclusion: A Deep Learning based pipeline to enhance the local and global contrast of CT Scan images is proposed for the automatic detection of lesions with varying sizes and severity in different body locations. The experiment is validated on Deep Lesion dataset with a significant performance boost with the proposed pipeline. Link to the implementation: `https://github.com/DeSinister/DeepLesionImageEnhancer`

Keywords: Computed Tomography (CT) Scan, Image enhancement, Contrast Limited Adaptive Histogram Equalization (CLAHE), Convolutional Neural Networks (CNN), Object Detection.

*Nishq Poorav Desai, 21npd1@queensu.ca

1 Introduction

Medical imaging is an important aspect in the early detection and accurate diagnosis of various medical conditions, including lesions. CT scans are important for providing visualizations of detailed internal body structures, making them an essential tool for lesion detection. By equipping a deep learning-based image enhancement model to improve the quality and contrast of CT images in the dataset, it is expected to boost the performance of lesion detection models and ultimately improve patient outcomes by enabling earlier and more accurate diagnoses of diseases.

1.1 Literature Review

Object detection can automate this process, reducing the workload and providing more accurate and efficient lesion detection. A lightweight Squeeze-Faster R-CNN was proposed by Alkhaleefa et al.[1] for liver lesion detection from CT images optimized by the reduced depth and decreased number and size of filters. The results in the DeepLesion Dataset improved average precision from 42.7 to 63.8 and a reduced number of parameters, making it beneficial for lesion detection on embedded devices with limited memory . Singh et al.[2] proved that Image Enhancement algorithms could be used when the image has non-uniform illumination, low contrast, or noise to enhance image quality while preserving details for improved object detection in lesion detection in CT scan images.

Image enhancement algorithms could be classified into 2 major categories based on the usage of deep learning. Non-Deep Learning based methods use image processing methods like CLAHE to improve image quality but often fail to adapt to intra-image and inter-image discrepancies. While Deep Learning algorithms like GANs,[3,4] and Image Fusion provide better results at the expense of high amounts of available data and require a lot of time for pre-training.

Traditional Non-Deep Learning-based Image enhancement methods apply domain-specific Image Processing algorithms and are capable of many aspects of noisy images. For example, a modified Contrast Limited Histogram equalization (CLAHE) method was introduced by Gul et al.[5] for Liver Segmentation in 3D CT Volumetric images yielding a state-of-art dice score of 0.97, and further empirically proving the significance of applying CLAHE for CT Scan Images. Sidhu et al.[6] employ PCA and contrast enhancement using CLAHE to highlight retinal blood vessels and extract new blood vessels from the disc and non-disc region, scoring an overall accuracy of 0.9682. Kandhway et al.[7] utilized algorithms such as SSA and KH to clip the limit, based on a proposed fitness function. The proposed optimization-based reformed histogram equalization framework demonstrated superior performance, both qualitatively and quantitatively, compared to all existing methods in terms of edge details, information, contrast, content, and structural similarity.

Recently, deep learning-based image enhancement techniques have shown considerable promise in improving the quality of medical images compared to traditional methods. Zhu et al.[8] proposed a framework that includes image generation and image fusion, for enhancing the contrast and illumination of the MRI image via adaptive attenuation weight matrix and illumination preservation demonstrating state-of-art performance and stability.

Bhutto et al.[9] presented a novel approach to CT and MRI medical image fusion using a CNN-based model with noise-removal and contrast enhancement techniques to pre-process the input images before feature extraction and fusion using CNN. The method outperforms several state-of-the-art methods in CRR and AG metrics and provides improvement in visual quality of the fused images. Sato et al.[10] propose a deep learning-based method for semantic segmenting liver tumours in contrast-enhanced CT scans. The proposed method uses a U-Net architecture with a Hessian-based image enhancer to improve the image quality before segmentation achieving a Dice score of 0.703.

1.2 Problem Statement

Non-Deep Learning based algorithms often do not generalize well with the variety of data, and Deep Learning based algorithms are prone to over-fitting and require huge computation and data for pre-training, there is a lack of an architecture capable of combining the salient features of both. Moreover, applying the same parameters for enhancement can adversely affect the quality because CT scans can have varying levels of noise, artefacts, and non-uniform illumination leading to over-enhancement or under-enhancement of certain features, resulting in an inaccurate representation of the image. Thus, there is a need for an algorithm to consider the intra-image and inter-image contrast and noise-level difference.

1.3 Proposed Solution

To mitigate this issue, image-specific parameters can be used to tailor the algorithm's parameters to each CT scan image. This involves adjusting the parameters of the algorithm to achieve the desired level of enhancement while preserving the image's overall quality. Finally, improving the quality of medical imaging for the benefit of patients and healthcare professionals alike. Secondly, to avoid again pre-training the model, a topped deep learning-based model can be used to support the pre-trained model with enhanced images. Thereby effectively reducing the Time and computation with improved model-specific and image-specific results. Furthermore, one of the largest publicly available medical image datasets called Deep Lesion Dataset is used for evaluating the architecture. The project proposes an all-in-one solution to detect a diverse range of lesion types, sizes, and locations within the body using Full body CT Scan Images. It includes lesions in organs such as the liver, lung, and bone, as well as in soft tissues such as muscle and fat.

The proposed architecture involves a Deep learning model applied for configuring the parameters of Local and global contrast enhancement and Contrast-Limited Adaptive Histogram Equalization (CLAHE) to enhance images fed to Faster RCNN object detection models to detect lesions accurately.

Overall, Medical imaging is essential for detecting and diagnosing medical conditions, including lesions, and CT scans are vital for lesion detection. Traditional and deep learning-based image enhancement algorithms have limitations in generalizing with the variety of data or requiring a lot of computation and data for pre-training.

The proposed solution is to use image-specific parameters to improve image quality via a topped deep learning-based model that can support the pre-trained model with enhanced images, reducing time and computation while improving model-specific and image-specific results. The Deep Lesion Dataset is a diverse and clinically relevant dataset used for training and evaluating lesion detection models. Overall, these advancements can lead to earlier and more accurate diagnoses of diseases, benefiting patients and healthcare professionals.

2 Methods and Materials

The overall architecture pipeline is provided in Fig. 1 The DeepLesion Dataset[11] was used to pre-train the Object Detector containing 32,735 lesions in 32,120 CT slices. The average spacing for the dataset was 0.8 mm in the x-direction and the y-direction, and 3.0 mm between consecutive horizontal slices. The 2D slices of the Dataset are resized to 512px * 512px dimension.

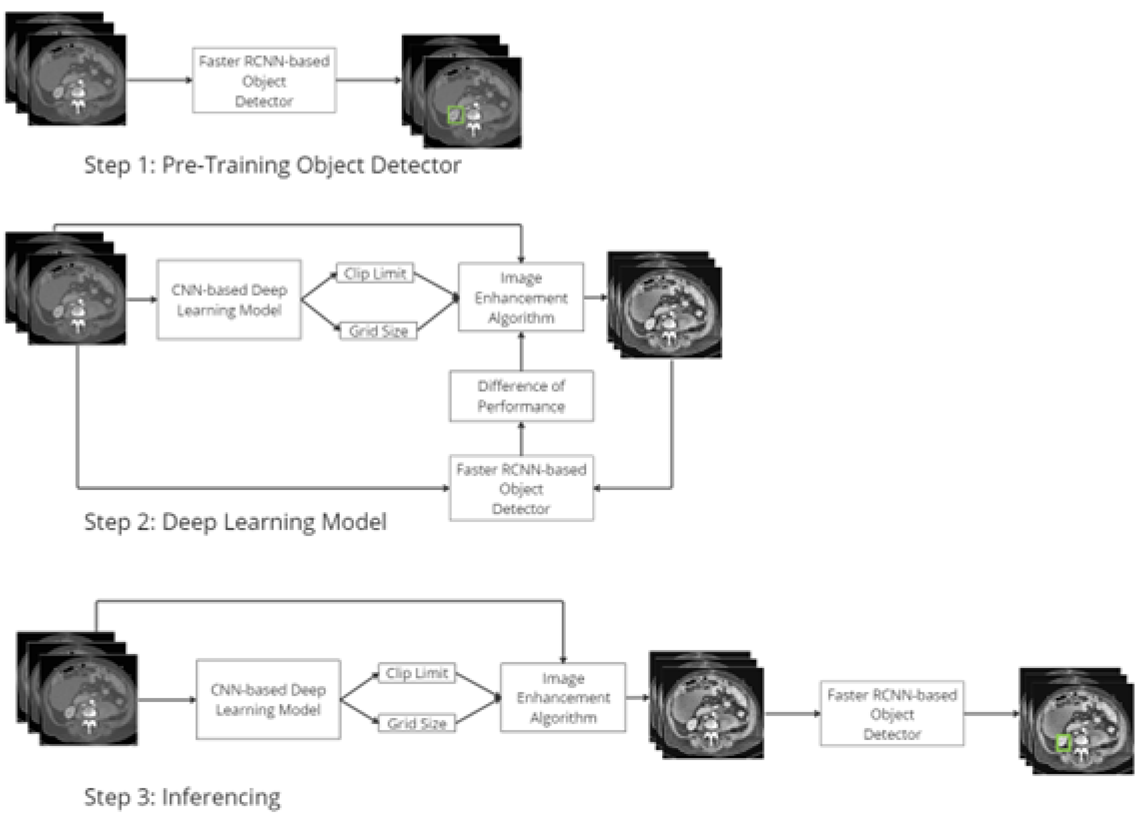


Fig 1 Proposed Architecture Pipeline

Firstly, ResNet50 Feature Pyramid Network-based Faster RCNN Model is pre-trained on the

CT scan images. It is expected that the Object Detector model would provide a decent performance but would not be able to perform with the CT-Scan images having limited contrast.

Table 1 Hyper-parameters for Training Object-Detector

Hyper-parameter	**Value**
Batch Size	26
Number of Epochs	100
Learning Rate	1e-5
Weight Decay	1e-4
Optimizer	Adam
Early Stopping Patience	20

Table 2 Hyper-parameters for Deep-Learning Based Image Enhancer

Hyper-parameter	**Value**
Batch Size	26
Number of Epochs	100
Learning Rate	1.5e-5
Weight Decay	1e-2
Optimizer	Adam
Early Stopping Patience	20

The Local and Global Features are taken into consideration for CLAHE-based image enhancement. The normalized histogram of an image is computed, and the traditional histogram equalization algorithm is described. However, this method often produces unnatural-looking images with visual artefacts.

A general framework for histogram modification[12] is utilized to optimise a weighted sum considering the the original and uniform histograms for optimally modified histogram. The tone distortion measure is introduced to guide the optimization and choose the optimal weighted parameter

λ, producing a mapping function that can adaptively enhance images with less unnatural artefacts.

A combination of hue preservation enhancement framework with the Contrast-Limited Adaptive Histogram Equalization (CLAHE) method is utilized to enhance local contrast and preserve detailed information in an image. This approach replaces the Global Contrast Adaptive method.

Finally, the framework uses pixel-level weights computed based on the contrast and brightness measures of the two images. To ensure a consistent fusion result, the weight maps are normalized, and a Laplacian pyramid is used to prevent the seam problem in the fused image. The result is obtained by applying the weights to the colour channels of the globally and locally enhanced images. The Algorithm utilizes Clip Limit and Grid Size to modulate the maximum amount of contrast enhancement that can be applied to each pixel as well as the size of the local regions used for contrast enhancement.

The parameters of the pre-trained Object Detector are frozen, and a ResNet18-based CNN is used to extract the 2 parameters from the image via multi-objective learning. The input images of size 512px * 512px * 1 channel image are first convoluted to produce 3 channels to input to the ResNet18 Model which is further regressed by following fully connected layers to the image-specific variables Clip Limit and Grid Size, which in turn transforms the input image. The Architecture of the ResNet-based Feature Extractor is given in Fig 2

As there are no ground truth values for the Clip Limit and Grid Size, the loss function is designed to exploit the difference in mean Average Precision between the Original Image, and the Transformed image via the predicted variables. The loss value of 0 indicates no improvement in performance, while the negative value represents the improvement in the model-specific image enhancement. Once the Deep Learning Model is trained, the Inference can be performed with the Input Images first being subjected to the ResNet Model to predict the parameters, which Enhances

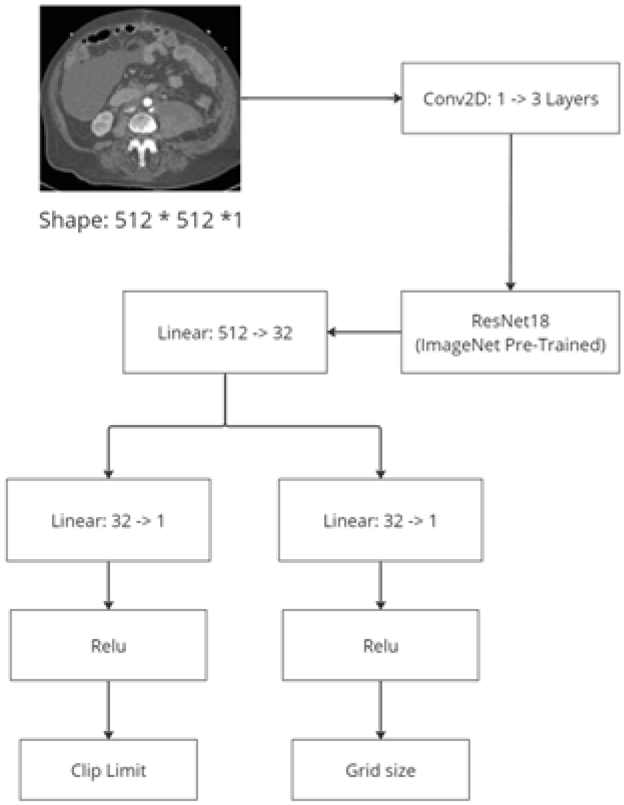


Fig 2 Layered Architecture for Deep Learning Model

the Image, which can be fed to the object detector to output the bounding box predictions.

The proposed pipeline has been validated on the Deep Lesion dataset using the Leave One Out Validation method. About 664 patients' CT Scan images were used as validation sets.

The Metric used to evaluate the proposed architecture is Mean Average Precision. It provides a comprehensive measure of detection accuracy by considering both precision and recall. In medical contexts, the accurate detection of lesions is critical for early diagnosis and treatment planning. False negatives, where a lesion is missed, can lead to delayed diagnosis and treatment, while false positives, where an image region is falsely identified as a lesion, can lead to unnecessary procedures and increased patient anxiety. Mean Average Precision considers both true positives, false positives, and false negatives, giving a more accurate representation of the model's performance. Even after applying the Deep Learning Architecture, the performance is measured in terms of

Mean Average Precision.

3 Results

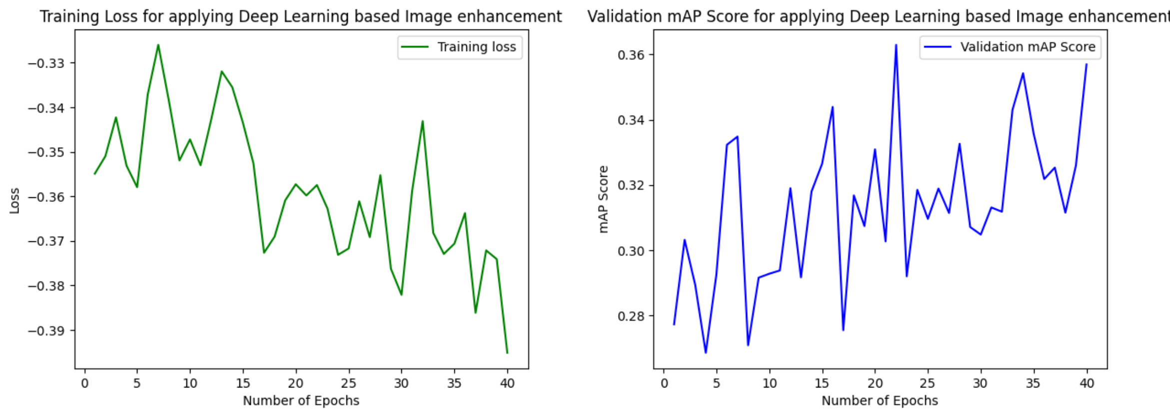
Due to limited Computation resources, out of 4459 Patients, 1587 Patients were used for Training and 664 Patients were used for Validating the pipeline. The size of the Lesion varies from a minimum size of $0.0854mm^2$, to a Mean size of $735.9408mm^2$, and a Maximum size of $50,025.0415mm^2$.

A maximum score of 36.16 mAP was achieved by combining the proposed pipeline, noting a conspicuous increase of 5.42 mAP. Moreover, for the sake of comparison to the Alkhaleefa et al.[1]'s Squeeze-Net, Liver Lesion specific evaluation was conducted, achieving the mAP score of 45.29 and 40.82 by the object detector with and without the image enhancement pipeline. Here, only a small segment of the full dataset was utilised for the training as opposed bt Alkhaleefa et al.,[1] and thus could be the possible reason for limited performance.

Table 3 Performance Comparison

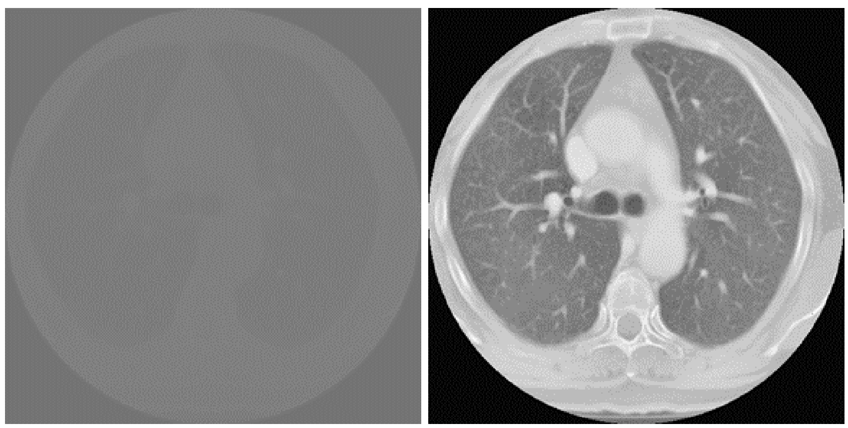
	Algorithm Used	Target Lesion	mAP	Sensitivity
Faster R-CNN based on Optimized Squeeze-Net[1]	Squeeze Net	Liver Lesions	63.8	-
Universal Lesion detector[11]	MX-Net	All Lesion	-	81%
Ours (Object Detector)	ResNet50 FPN Faster RCNN	Liver Lesions	40.82	-
Ours (Object Detector + Deep Learning-based Image Enhancement)	ResNet18 based Local and Global Contrast Enhancement	Liver Lesion	45.29	-
Ours (Object Detector)	ResNet50 FPN Faster RCNN	All Lesions	30.77	-
Ours (Object Detector + Deep Learning-based Image Enhancement)	ResNet18 based Local and Global Contrast Enhancement	All Lesion	36.19	-

The Curve for the loss function for applying the Deep Learning based Image Enhancement Algorithm is provided in the Fig. 3, Early Stopping has been applied after the patience of 25 epochs to avoid overfitting and utilize the optimum model. As observed, the training loss is negative and shows a deviation from the performance without Image Enhancement. More the negative value of the Training Loss, better the improvement in performance. As observed in the figure, the model



(a) (b)

Fig 3 (a) Training Loss for applying Deep Learning based Image enhancement (b) Validation mAP Score for applying Deep Learning based Image enhancement



(a) (b)

Fig 4 (a) Original CT Image with very Low contrast, (b) Enhanced Image with Local and Global Contrast Enhancement

trains to decreased loss up to 22 epochs, and it starts overfitting to the training data, and after certain patience, early stopping was applied.

The Baseline provided by the DeepLesion Dataset encompasses 3 surrounding images, as well as utilizes the full training data of 3,121 patients, and have utilized a deep model for applying Mask-RCNN like MX Net, and have utilized an exhaustive set of anchor scales, and anchor ratios for Object Detection. However, they have only provided the performance to have 5 False Positives per image, which is very vague, and do not interpret about model's efficiency like Mean Average

Precision.

3.1 Ongoing Modules

The dependency of the deep learning model on different Object-Detection modes is currently under testing for the performance of the overall pipeline. The investigation of Grad-Cam[13,14] to investigate the focus of the deep learning model is still being carried on. It is expected that Object Detection models try to generalize on data with low noise levels but try to memorize the data with high noise levels. As a result, Object Detection models might under perform in this scenario,[15–17] However, adding an Image Enhancement mechanism could potentially discourage memorization because of the shift in the content of data, with ameliorated details. Although, while using a Deep Learning Model for Image enhancement on a model it was not used for producing loss function could potentially produce different results. These results can play a crucial role in understanding the patterns the deep learning model is following for enhancing not only image-specific features but also Object-Detection model-specific features. Moreover, the behaviour of the Deep Learning Model can be interpreted based on the Grad-Cam and can be used as a medium to validate the working of the pipeline by observing what part of the image it focuses on.

4 Discussion

The major contributions of the proposed research include:

- An all-in-one solution is proposed for lesion detection with varying sizes, severity, and occurring locations via CT Scan images from different parts of the body.
- An image enhancement Algorithm is added in the pipeline for taking consideration of local and global contrast and reducing noise levels.

- The Image enhancement don't require re-training of the pre-trained model and can be used as a top-up and thus saving computation and time.

- The enhancement algorithm produces significant improvement in terms of performance metrics i.e., by 5.42

The pre-training of the Object-Detector Model is very computationally expensive and was taking a lot more time for training than expected, and the full training data was not utilized. A lot of experiments could be implemented alongside ResNet-based deeper models like DenseNet-121 and Vision Transformers could be investigated for the proposed pipeline. Furthermore, even though Zeng et al.[18] discouraged the use of 3D-based CNN's due to the cube structure, the data exists in 3D, and generally, the surrounding layers can potentially be helpful in making decisions for bounding boxes for lesion detection on a specific layer. Due to the nature of 2D detection, there is a possibility of sub-optimal performance[19,20]

The proposed pipeline points towards a direction for applying image processing for the removal of noise and enhancing Image quality. This post-processing step could be equipped to practically any object-detection model for boosting performance, irrespective of its architecture. Additionally, this pipeline can be extended to other forms of Medical Imaging as well like MRI images,[21,22] and Ultrasound Images[23,24]

After the successful execution of the model-specific investigations and validation of the clinical implications for the focus of the deep learning detector using Grad-Cam on the CT Scan Images, the research project is planned to be published.

4.1 Ablation Study

Other than the inclusion of the Deep-Learning Image Enhancement, the following experiments with different hyper-parameters were undertaken to validate the optimal performance. From Table 4, a dip in performance could be observed while changing the architecture of the model like the number of layers, or reducing hyper-parameters like Batch Size, Learning Rate and Early Stopping Patience.

Table 4 Ablations for Deep-Learning Based Image Enhancer

Ablation	**Results**
Batch Size (26 to 13)	32.57
Number of Fully Connected Layers (2 to 4)	35.68
Learning Rate (1.5e-5 to 1e-4)	31.48
Early Stopping Patience (20 to 10)	36.04

5 Conclusion

The proposed study aims to develop a comprehensive solution that can automatically detect lesions of varying size, severity and occurrence from Full Body CT scan images with varying noise and contrast levels. This will be accomplished by using a combination of CNN-based Local and Global Enhancement techniques to improve image quality and lesion detection. The study uses a ResNet-based Faster RCNN to train on CT scan images from the DeepLesion dataset. Additionally, a ResNet18 CNN model is trained to predict parameters that can enhance the Local and Global Contrast of the images and remove underlying noise while maintaining image details.

The results show that the image enhancement pipeline was able to improve the performance of lesion detection from 30.77% mAP to 36.19% mAP without retraining the pre-trained Object-

Detector model. The experiment was validated on the Deep Lesion dataset and showed a significant performance improvement with the proposed pipeline. Finally, this study points a direction of how applying image processing techniques can be added in the post-processing pipeline to any pre-trained architecture, and discusses and attempts to validate the clinical implications of the proposed pipeline.

References

1 M. Alkhaleefah, T.-H. Tan, V. P. Achhannagari, *et al.*, "Faster r-cnn based on optimized squeezenet for liver lesion detection from deeplesion dataset," in *Proceedings of the 5th International Conference on Graphics and Signal Processing*, 20–26 (2021).

2 R. Singh, S. R. Digumarthy, V. V. Muse, *et al.*, "Image quality and lesion detection on deep learning reconstruction and iterative reconstruction of submillisievert chest and abdominal ct," *American Journal of Roentgenology* **214**(3), 566–573 (2020).

3 M. Hammami, D. Friboulet, and R. Kéchichian, "Cycle gan-based data augmentation for multi-organ detection in ct images via yolo," in *2020 IEEE International Conference on Image Processing (ICIP)*, 390–393, IEEE (2020).

4 T. Mattikalli, T. S. Mathai, and R. M. Summers, "Universal lesion detection in ct scans using neural network ensembles," in *Medical Imaging 2022: Computer-Aided Diagnosis*, **12033**, 864–868, SPIE (2022).

5 S. Gul and M. S. Khan, "On evaluating ct image enhancement techniques for deep learning based 3d liver segmentation," in *2022 Global Conference on Wireless and Optical Technologies (GCWOT)*, 1–6, IEEE (2022).

6 R. Sidhu, J. Sachdeva, and D. Katoch, "Segmentation of retinal blood vessels by a novel hybrid technique-principal component analysis (pca) and contrast limited adaptive histogram equalization (clahe)," *Microvascular Research* **148**, 104477 (2023).

7 P. Kandhway, A. K. Bhandari, and A. Singh, "A novel reformed histogram equalization based medical image contrast enhancement using krill herd optimization," *Biomedical Signal Processing and Control* **56**, 101677 (2020).

8 R. Zhu, X. Li, X. Zhang, *et al.*, "Mri enhancement based on visual-attention by adaptive contrast adjustment and image fusion," *Multimedia Tools and Applications* **80**, 12991–13017 (2021).

9 J. A. Bhutto, L. Tian, Q. Du, *et al.*, "Ct and mri medical image fusion using noise-removal and contrast enhancement scheme with convolutional neural network," *Entropy* **24**(3), 393 (2022).

10 M. Sato, Z. Jin, and K. Suzuki, "Semantic segmentation of liver tumor in contrast-enhanced hepatic ct by using deep learning with hessian-based enhancer with small training dataset size," in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, 34–37, IEEE (2021).

11 K. Yan, X. Wang, L. Lu, *et al.*, "Deeplesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning," *Journal of medical imaging* **5**(3), 036501–036501 (2018).

12 Q.-C. Tian and L. D. Cohen, "Global and local contrast adaptive enhancement for non-uniform illumination color images," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 3023–3030 (2017).

13 M. Takeuchi, T. Seto, M. Hashimoto, *et al.*, "Performance of a deep learning-based identification system for esophageal cancer from ct images," *Esophagus* **18**, 612–620 (2021).

14 S. Sanagavarapu, S. Sridhar, and T. Gopal, "Covid-19 identification in clahe enhanced ct scans with class imbalance using ensembled resnets," in *2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*, 1–7, IEEE (2021).

15 P. Chen, B. B. Liao, G. Chen, *et al.*, "Understanding and utilizing deep neural networks trained with noisy labels," in *International Conference on Machine Learning*, 1062–1070, PMLR (2019).

16 S. Chadwick and P. Newman, "Training object detectors with noisy data," in *2019 IEEE Intelligent Vehicles Symposium (IV)*, 1319–1325, IEEE (2019).

17 K.-Y. Lung, C.-R. Chang, S.-E. Weng, *et al.*, "Rosnet: Robust one-stage network for ct lesion detection," *Pattern Recognition Letters* **144**, 82–88 (2021).

18 N. Zeng, J. Fang, X. Wang, *et al.*, "Factoring 3d convolutions for medical images by depth-wise dependencies-induced adaptive attention," in *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 883–886, IEEE (2022).

19 D. Liang, L. Lin, X. Chen, *et al.*, "Multi-stream scale-insensitive convolutional and recurrent neural networks for liver tumor detection in dynamic ct images," in *2019 IEEE International Conference on Image Processing (ICIP)*, 794–798, IEEE (2019).

20 C. Han, Y. Kitamura, A. Kudo, *et al.*, "Synthesizing diverse lung nodules wherever massively: 3d multi-conditional gan-based ct image augmentation for object detection," in *2019 International Conference on 3D Vision (3DV)*, 729–737, IEEE (2019).

21 J. Kalyani and M. Chakraborty, "Contrast enhancement of mri images using histogram equalization techniques," in *2020 International Conference on Computer, Electrical & Communication Engineering (ICCECE)*, 1–5, IEEE (2020).

22 Z. Ullah, M. U. Farooq, S.-H. Lee, *et al.*, "A hybrid image enhancement based brain mri images classification technique," *Medical hypotheses* **143**, 109922 (2020).

23 S. Sahu, "Comparative analysis of image enhancement techniques for ultrasound liver image," *International Journal of Electrical and Computer Engineering* **2**(6), 792 (2012).

24 P. Singh, R. Mukundan, and R. De Ryke, "Feature enhancement in medical ultrasound videos using contrast-limited adaptive histogram equalization," *Journal of digital imaging* **33**, 273–285 (2020).

List of Figures

List of Tables