

ADAPTIVE CT PREPROCESSING AND ROI STANDARDIZATION FOR PULMONARY NODULE REPRESENTATION

Divya G S^{1*}, Dr. A. Suresh Kumar²

¹Department of Computer Science and Engineering, Bengaluru, India. Mail Id: divyags.siddaraj@gmail.com, Orcid Id: 0009-0009-0723-484X

²JAIN (Deemed-to-be University), Bengaluru, India. Mail Id: sureshkumar.a@jainuniversity.ac.in, Orcid Id: 0000-0001-7145-6337

ABSTRACT

Preprocessing is part of almost every thoracic computed tomography (CT) pipeline for pulmonary nodule analysis, yet its individual steps are rarely measured on their own. This paper evaluates seven preprocessing stages on 30 CT scans from LIDC-IDRI containing 86 radiologist-annotated nodules (37 benign, 26 indeterminate, 23 malignant): Hounsfield unit conversion, isotropic resampling, lung segmentation, intensity standardization, denoising, contrast enhancement, and region-of-interest (ROI) extraction. Two adaptive components are tested, a lung-restricted, noise-adaptive form of contrast-limited adaptive histogram equalization (CLAHE) and a size-adaptive ROI. Selected parameters are chosen by patient-level five-fold cross-validation, and uncertainty is reported as patient-level bootstrap intervals. A fixed lung window kept scans more consistent than per-scan min-max scaling (pairwise histogram distance 0.085 against 0.235). A corrected threshold-based lung mask retained 96.5% of nodules at a mean lung volume of 4.8 L, whereas the mask of the original pipeline covered 86% of the scan. A cross-validated bilateral filter raised structural similarity in the lungs from 0.910 to 0.955 at 25 HU of added noise and raised nodule contrast-to-noise ratio (CNR) by 7.6% without loss of edge sharpness, although at 10 HU no filtering scored higher. CLAHE lowered CNR at every clip limit above 1 (median ratio 0.79 at clip limit 2); the adaptive variant was selected at its weakest setting and did not improve on applying no enhancement (median CNR ratio 0.998, mean 0.919). The size-adaptive ROI reduced the coefficient of variation of nodule occupancy from 2.06 to 0.63 with no truncation. The study shows which steps contribute to a standardized nodule representation and which do not.

KEYWORDS: Pulmonary nodules, thoracic CT, preprocessing, Hounsfield unit standardization, denoising, CLAHE, lung segmentation, ROI extraction, LIDC-IDRI.

INTRODUCTION

Lung cancer accounts for the greatest number of cancer-related deaths as most cases are diagnosed late in their development [1], [2]. Now, only about 15-16% of lung cancer patients survive five years after diagnosis. The use of high-risk groups' low-dose CT scan has resulted in higher detection rate of early-stage cancers and mortality reduction such that the diagnosis of the pulmonary nodule has become a cornerstone of computer-assisted imaging [1]. The literature on artificial intelligence applications to lung cancer imaging has reported a great success rates for the detection of nodules [3], [4], segmentation of images and understanding of malignancy nature [2], with meta-analysis indicating pooled sensitivity and specificity of 0.87 and considerable variability among different studies [5]. Although chest X-ray is exploratory and cheap imaging technique, the capacity to detect nodules is considerably limited due to the anatomy overlaps on the final picture [6]. Therefore, CT continues to be the gold standard imaging technique.

It has been observed that there is a typical process for deep learning systems in relation to CT analysis for lesions found in lungs. The scans are assigned an equivalent in Hounsfield units (HU), segmentation is done for the lungs, an area of interest is cropped, and convolutional networks get applied for classification of data [7], [8]. One model called modified U-Net that combined AlexNet and SVM achieved the accuracy of 97.98% while processing data from LUNA16 dataset [9], while the ensemble model achieved 95% accuracy using a set of three 2D convolutional networks [10]. Researchers also came up with the concept of predicting survivorship based on CT volume assessment in their work [11].

A different set of research predicts the onset of lung cancer without the use of imaging methods. Classic machine learning and stacking ensemble techniques have been employed to process questionnaire/survey information [12], [13], clinical records with survival and treatment response models [14], clinical notes with text classifiers [15], and molecular data with multi-omics networks [16], [17]. One of the pieces of research has a certain relevance for imaging methods: class balancing, the removal of outliers, and scaling of features were enough to significantly change the accuracy of classification on a tabular survey dataset [18]. Thus, preprocessing could not be regarded as simply a neutral preliminary step and its effects should be quantified.

In CT imaging, that particular measurement is mostly absent. The previous studies have mentioned three flaws. First of all, preprocessing has been presented to be a kind of technique, and its performance has been measured only in terms of its ability to classify and fail to indicate the particular contribution of each step. Secondly, the image quality mentioned is calculated for the entire image without any reference, giving little about the actual nodule. Lastly, clips, for example, a clip limit of 2, a Gaussian width of one pixel or a 64-pixel cut are simply taken from conventions without being formalized or divided into training and reporting data.

It has been found that in such cases models lose their efficiency when quality of images and their acquisition differ from the training ones and radiomic indicators are dependent on resolution and noise level [2], [19].

This paper evaluates individually every preprocessing step regarding the annotated nodule. The study employs a selection of samples from the LIDC-IDRI dataset containing radiologist labels defined in DICOM format [20]. It also compares its results with the previous application of the same pipeline, which yielded the results of single-slice processing, and identifies the shortcomings of the previous approach. The study is guided by the following research questions:

RQ1. What method of intensity standardization allows for maintaining a maximal stability of intensities of tissues across several scans?

RQ2. Is the use of a lung mask based on thresholds sufficient for retaining the annotated nodules, and how do different resampling options affect intensity as well as volume of nodules?

RQ3. What denoising filters are able to raise imaging fidelity without degrading edges of nodules, and what is the level of noise when this effect is achieved?

RQ4. Does CLAHE signal with constant or noise-adaptable strength improve the contrast to noise ratio of nodules?

RQ5. Does the size-adaptive ROI allow for more accurate tissue representation in comparison to fixed cropping?

The contributions are as follows.

- A methodological evaluation protocol that evaluates radiologist consensus mask measurements, where parameters are chosen through patient-level cross-validation, and the interval measurement is done using bootstrap intervals.
- Evidence from the stage-by-stage approach with 30 scans and 86 nodules that gives contradictory findings: normalization based on scan shows lower consistency in tissue intensities between scans, while CLAHE parameters help to reduce contrast-to-noise ratio of nodules in all gathered data.
- The two adaptive components with clear definitions and expected outcomes from the use of lung-specific noise-adaptive CLAHE and size-adaptive ROI, where usage of the first improves results at the level of no enhancement, and the usage of the second decreases deviations in nodule occupancy.
- An audit of the original single-slice pipeline that replicates the findings of the study with errors being evaluated: images that measure self-similarity and lung mask covering most of the image, as well as the image center crop that accounted less than 5% of nodules in terms of gathering enough data to support findings.

The structure of the rest of the paper is as follows: The next section focuses on reviewing previous work. The Methodology section explains the processing stages as well as the options that are compared for each stage. The Experimental Setup section describes the inputs used in the experiments and the input measurements processes. The Results and Discussion section has the experimental outcomes as well as uses of the methods in every type of cancer. The last part of the paper gives the conclusions, possibilities for the future work and limitations.

LITERATURE REVIEW

The related work is organized into four sections: nodule detection and prediction, prognosis and radiogenomics using CT images, studies in artificial intelligence that are used in lung cancer imaging, and predictions based on data that doesn't use any image.

CT-based nodule detection and classification.

The task is divided into lobes segmentation, potential nodule detection, and classification for Naseer et al. [9] A modified U-Net is used to segment the lobes in the LUNA16 dataset achieving a Dice score of 90.32% in contrast to the score of 83.40% for standard U-Net. For the task of nodule detection, a second network has been employed and for the classification of 48×48 images the classification is done using AlexNet-SVM perform with 97.98% accuracy. The authors Shah et al. [10] rescale the Hounsfield units for the LUNA16 dataset and build three different two-dimensional convolutional neural networks for fusion to achieve 95% accuracy. Tusher et al. [21] downsize the LIDC-IDRI in size of 224×224 pixels and use Gaussian filters and thresholds in lung segmentation process to combine clinical data with image data in fused network. Ramachandran and others [22] examines ResNet50 and InceptionV3 on 480 CT scans after preprocessing involving scaling and contrast enhancement of the images, respectively. The preprocessing techniques used in all the four studies are shown to be the same except in Naseer et al. study where the segmentation network is evaluated separately.

Prognosis and radiogenomics from CT.

In their work, Tas and Yavuz [11] implement a three-dimensional ResNet-34 to project survival outcomes based on the NSCLC-Radiomics dataset. Their approach converts the DICOM images into 8-bit images using min-max scaling, isolates the tumor area, and limits the number of slices utilized. In a related research study, Tang et al. [19] retrieve 107 radiomic features from gross tumors in the same data set along with clinical features exhibiting the area under the curve as 0.949; the authors acknowledge the impact of the image resolution, noise, and drawing quality on the radiomics process. In their research, Shariaty et al. [23] employ ResNet18 to reveal the presence of EGFR and KRAS mutations based on CT images of nodules and Grad-CAM to prove how the network processes the edges of these nodules, thus showing its ability to focus on that boundary which is classified as the most vulnerable to smoothing and resampling.

Reviews of artificial intelligence in lung cancer imaging.

Yang et al. [2] conducted a survey where they studied detection, segmentation, pathology classification, mutation prediction, prognosis, and enumerated the differences in data quality across hospitals and difficulty in interpretation and the lack of generalization as the major problems. Cellina et al. [1] present a review on the applications of screening, introduce LIDC-IDRI and LUNA16 as the most popular publicly available datasets, and provide examples of areas where the techniques of learning can come into aid including reducing noise and standardization of acquisition process. Javed et al. [24] described the application of deep learning in imaging modalities and stated that preprocessing comes first in the chain of detection, pointing out the risk of information loss because of scaling and cropping the images to make them fit

recommended input size. Kanan et al. [5] conducted a meta-analysis of diagnostic studies and attributed the differences in results to the variability in study parameters. Megat Ramli et al. [6] evaluate the problem of nodule detection with the use of chest x-rays and mention the poor quality of images and data on populations that do not match training datasets negatively influence the results.

Prediction from non-imaging data.

Azimi and Sulaimany [18] demonstrate that through preprocessing, outlier selection, and scaling, the performance of machine learning classifiers is improved, while also applying SHAP and LIME methods to interpret their models. Ishtiaq et al. [13] conduct an experiment in which they compare classifiers using the decision tree, k-nearest neighbours and naive Bayes methods across a risk factors dataset. Tu et al. [12] adopt a stacking ensemble approach to analyze epidemiological questionnaires from 5,421 cases and 10,831 controls. Yuan et al. also apply the machine learning techniques to survival data of 737 patients from clinical records, while Prelaj et al. [14] analyze case studies in immunotherapy effectiveness. Kesiku and Garcia-Zapirain [15] attempt to predict the chances of lung cancer from clinical records using a hybrid model composed of two neural networks - convolutional and recurrent. In addition, Mohamed and Ezugwu [16] propose disease stage classification based on integrated approaches. Kwon et al. [17] collect liquid-biopsy markers. Although these studies do not engage in the analysis of CT images, they are useful to illustrate the point about the changing nature of the models during data processing and how valid the model explanations can be based on the data quality.

Treatment of preprocessing.

Table I presents a comparison of the imaging studies by their preprocessing techniques used as well as by whether they were assessed independently from the model. Table II gives a summary of previous reviews and non-imaging studies and their connection to the current investigation. All imaging studies include intensity scaling, denoising, segmentation, cropping; however, none of them includes the impact of intensity scaling on reproducibility or the impact of filtering on object borders or variation of the amount of cropping the object is occupying. This study touches on these three aspects by performing quantifications on the annotated objects..

TABLE I. Imaging Studies in the Related Work: Data, Model, and Treatment of Preprocessing

Study	Data	Task and model	Preprocessing reported	Evaluated separately
Naseer et al. [9]	LUNA16 CT	Lobe segmentation, candidate extraction, AlexNet-SVM classification	Modified U-Net lobe masks; 48×48 patches	Segmentation only
Shah et al. [10]	LUNA16 CT	Ensemble of three 2D convolutional networks	Rescaling from the Hounsfield scale; candidate patches	No
Tusher et al. [21]	LIDC-IDRI CT and clinical tables	Hybrid convolutional and dense network	Resizing to 224×224 ; scaling to [0, 1]; Gaussian filter; threshold lung mask; augmentation	No
Ramachandran et al. [22]	480 CT images; survey table	ResNet50, CNN, InceptionV3; SVM, random forest, XGBoost	Scaling, contrast adjustment, noise filtering	No
Tas and Yavuz [11]	NSCLC-Radiomics CT	Survival classification, 3D ResNet-34	Per-image min-max to 8 bits; tumour crop; slice selection	No
Tang et al. [19]	NSCLC-Radiomics CT	Radiomic and clinical ensemble for prognosis	Feature extraction from gross tumour volume	No
Shariaty et al. [23]	NSCLC radiogenomic CT	EGFR and KRAS prediction, ResNet18 with Grad-CAM	Nodule image inputs	No
This work	LIDC-IDRI CT, 30 scans, 86 nodules	Evaluation of preprocessing	Seven stages, up to 23 alternatives each	Yes, every stage

TABLE II. Reviews and Non-Imaging Studies in the Related Work and Their Relation to This Study

Study	Data type	Focus	Relation to this study
Yang et al. [2]	Review	AI in detection, diagnosis, treatment and prognosis	Names data quality differences and generalization as open problems
Cellina et al. [1]	Review	AI in CT screening	Describes LIDC-IDRI; lists noise reduction and standardization as roles for AI

Study	Data type	Focus	Relation to this study
Javed et al. [24]	Review	Deep learning for detection across modalities	Places preprocessing first in the detection chain
Kanan et al. [5]	Meta-analysis	Diagnostic accuracy of AI models	Reports heterogeneity between studies
Megat Ramli et al. [6]	Systematic review	Nodule detection on chest radiographs	Image quality limits detection; motivates CT
Azimi and Sulaimany [18]	Survey table	Preprocessing and explainability	Shows that preprocessing alone changes accuracy
Ishtiaq et al. [13]	Risk-factor table	Comparison of classical classifiers	Non-imaging baseline
Tu et al. [12]	Questionnaires	Stacking ensembles for risk prediction	Non-imaging baseline
Yuan et al.	Clinical records	Survival prediction	Non-imaging baseline
Prelaj et al. [14]	Clinical records	Immunotherapy outcome with SHAP	Explanations depend on input quality
Kesiku and Garcia-Zapirain [15]	Clinical notes	Hybrid convolutional and recurrent text model	Non-imaging baseline
Mohamed and Ezugwu [16]	Multi-omics	Stage classification	Molecular alternative to imaging
Kwon et al. [17]	Liquid biopsy	Classification from blood markers	Molecular alternative to imaging

METHODOLOGY

The seven stages of the pipeline are illustrated in Fig. 1, including HU conversion, isotropic resampling, lung segmentation, intensity standardization, denoising, contrast improvement and ROI extraction. At every stage, various options are applied, of which the original pipeline is also one, so that every option can be compared. The alternatives are documented in Table III, and one slice from each stage of the total pipeline is shown in Fig. 2. All processing is done with a 32-bit floating point, but the images are displayed with an 8-bit resolution.

A. Problem Formulation

A scan i is a volume encompassing a set of pixel values p arranged in grid and spread along x -axis and y -axis and z -axis as well. Along with various marked nodules. Every nodule j has a corresponding mask N_j , diameter d_j and a label denoting malignancy of the nodule. In the processing pipeline the scan will be transformed into a set of statically determined size arrays, one per each nodule. There are three requirements for this transformation. Firstly, every tissue in a scan must have the same intensity. Secondly, the nodule in itself and its boundaries must be preserved in the process of transformation. And finally, the nodule of different size should be given in the same manner regardless of its size.

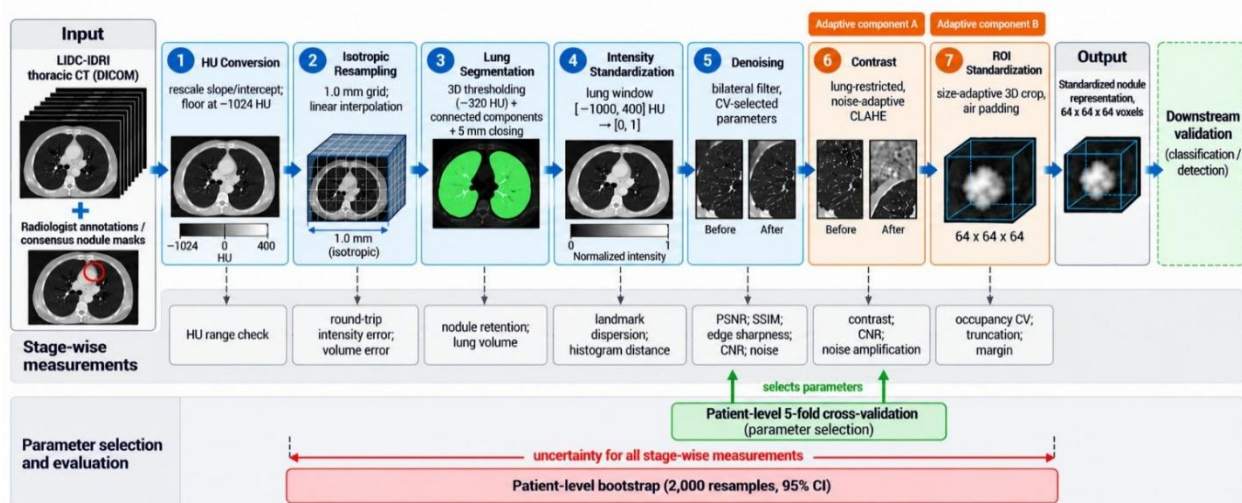


Fig. 1. Processing stages (top row), the alternatives compared at each stage, and the measurements taken (bottom row). Nodule-level measurements use consensus masks derived from the radiologist annotations.

B. HU Conversion and Resampling

Values that are stored are converted into Hounsfield units with the help of the rescale slope s_k and intercept b_k of each slice k . Values outside of the reconstruction area are stored in the form of extremely negative padding codes; they get converted into the value of air to ensure that they do not have any impact on subsequent statistics:

$$H(x) = \max(s_k p(x) + b_k, -1024) \quad (1)$$

Slices are arranged based on their location on the patient axis, and slice spacing is calculated based on the difference between consecutive positions. The volume is resampled into an isotropic grid spacing arrangement Δ^* using separable interpolation along all the axes a .

$$n'_a = \text{round}\left(\frac{n_a \Delta_a}{\Delta^*}\right), a \in \{x, y, z\} \quad (2)$$

Four target spacings (0.5, 1.0, 1.5, and 2.0 mm) and three types of interpolation method (nearest neighbour, linear and cubic convolution) are tested. Masks are linearly resampled and thresholded on 0.5. The entire pipeline uses 1.0 mm with linear interpolation for the process.

C. Lung Segmentation

Lung masks are obtained by thresholding the HU volume at $\tau = -320$ HU and cleaning the result with connected-component and morphological operations:

$$B(x) = [H(x) < \tau], M = \text{close}_r(\text{fill}(\text{LCC}_2(B \setminus B_\partial))) \quad (3)$$

In this case, set B_∂ consists of air mass connected with image boundaries. LCC2 leaves the biggest portions behind (the second must also be at least 10% of the volume of first one), fill removes internal holes, and close is morphological closing with radius r . Four variants are compared. S0 is original pipeline's functionality that thresholds, fills and closes without consideration of air being present outside body. S1 conducts corrected operations slice by slice without execution of closing. S2 conducts two-dimensional closing of the shape with disk of radius of 2, 5 and 10 mm. S3 performs three-dimensional component analysis followed by closing with a ball of 5 mm radius. The body mask is created based on the biggest area of filled part of the body receiving more than -500 HU.

D. Intensity Standardization

Seven mappings from HU to network input are compared. M0 leaves the HU values unchanged. M1 clips to a fixed lung window and rescales linearly:

$$T_{M1}(H) = \frac{\min(\max(H, L), U) - L}{U - L}, L = -1000, U = 400 \quad (4)$$

M2 is per-scan min-max scaling over the whole volume Ω , without clipping:

$$T_{M2}(H) = \frac{H - \min_{\Omega} H}{\max_{\Omega} H - \min_{\Omega} H} \quad (5)$$

M3, M4 and M5 are z-score mappings,

$$T_z(H) = \frac{H' - \mu_R}{\sigma_R} \quad (6)$$

Where H' represents the unclipped volume for M3. And, volumes clipped to $[-1000, 400]$ HU can be used for comparison of M4 and M5. In addition, the mean μ_R and standard deviation σ_R can be computed for the part of M3 and M4 or for the body mask only in M5. ML follows the original scheme used for processing by first clipping to $[-1350, 150]$ HU and applying min-max normalization for each slice. M1 only depends on HU. M2 to M5 is affected by the content of the scan images.

E. Denoising

Filters operate on M1 slices. The Gaussian filter convolves the image with an isotropic kernel of width σ pixels:

$$\hat{I} = I * G_\sigma, G_\sigma(u) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\|u\|^2}{2\sigma^2}\right) \quad (7)$$

The median filter replaces each pixel by the median of its $k \times k$ neighbourhood N_k :

$$\hat{I}(x) = \text{median}\{I(y) : y \in N_k(x)\} \quad (8)$$

The bilateral filter averages neighbours within a window of diameter d , weighting them by spatial distance with width σ_s and by intensity difference with width σ_c , where Z normalizes the weights:

$$\hat{I}(x) = \frac{1}{Z(x)} \sum_{y \in N_d(x)} \exp\left(-\frac{\|x-y\|^2}{2\sigma_s^2}\right) \exp\left(-\frac{(I(x)-I(y))^2}{2\sigma_c^2}\right) I(y) \quad (9)$$

The definition of the non-local means algorithm is based on computing the mean of the pixels with the help of 5×5 surrounding patches that are similar to each other within the search distance of 6 pixels and the strength of filtering h is set equal to 0.8 of 1.0 to the level of the noise. In total, twelve options were examined: without any filtering; Gaussian filtering with $\sigma = 0.5, 1.0$ and 1.5 ; median filtering with $k = 3$ and 5 ; and the bilateral filtering from the original processing ($d = 9, \sigma_s = 0.294, \sigma_c = 75$).

F. Contrast Enhancement

CLAHE divides the image into tiles, equalizes the histogram of each tile and interpolates between tiles. To limit noise amplification, each histogram bin is clipped at a count

$$\beta = \max\left(1, \frac{cN_{\text{tile}}}{B_h}\right) \quad (10)$$

Let us denote c as the clip limit number, N_{tile} as the total pixels in one tile and B_h (which equals 256) as the number of bins to be used. Thus, the excess is being redistributing it throughout all the bins. A clip limit of 1 is equivalent to no changes in the image, while bigger values are resulting in bigger enhancement. In total, five families are being compared. In C0 there is no enhancement at all, while C1 stands for doing global histogram equalization. C2 stands for using CLAHE (contrast limited adaptive histogram equalization) with fixed clip limits equal to 1, 2, 3 or 4 and using a grid of either 8×8 tiles or 16×16 tiles. The clip limit of 2 and tile of 8×8 is what was used originally. C3 deals only with the lungs, that is, it

considers only the lung area and replaces the values of the pixels outside the lung area by the median value of the lung before CLAHE process. C4 applies the same approach as C3 while adjusting the clip limit based on the noise in every slice.

$$c_s = \min\left(\max\left(c_{\text{ref}} \frac{n_{\text{ref}}}{n_s}, 1\right), 4\right) \quad (11)$$

In this instance, n_s represents the calculated noise standard deviation for the image slice, n_{ref} represents the average of n_s across multiple patients used to run the training phase, and c_{ref} is calculated through cross-validation. Since noisier slices have lower limits for clipping, C4 does not take into account any nodule information and therefore can work during inference. CLAHE is applied on the floating-point input image. When using a 16-bit unsigned integer image, the clip count falls below one for most clip limits ranging from one to four in equation (10), which then is set to one making all configurations behave in the same manner. This was seen in the implementation of OpenCV and can be tackled using this method.

G. ROI Extraction

The regions of interest (ROIs) are taken from the 1mm volume and are centered at the centroid of the consensus mask, both in 2D (the center slice) and 3D. R1 is a static crop with a dimension of 32, 48, or 64 mm [8], [25]. R2 is a more dynamic crop where the dimension (s) is determined by the nodule diameter (d) and is resampled at 64 samples per unit, leading to effective sampling spacing v :

$$s = \min(\max(2d + 16, 24), 96) \text{ mm}, v = \frac{s}{64} \text{ mm} \quad (12)$$

All nodules are allowed a margin of $0.5d + 8$ mm on either side, with both lower and upper bounds to maintain the parenchyma of small nodules as well as memory limits for the larger nodules. Cropped images that go outside the volume limits are padded with -1000 HU. The diameter used on the inference step is the one obtained during the detection part; in this algorithm it belongs to the manual label.

Algorithm 1: Evaluated preprocessing pipeline for one scan

Input: CT series; nodule centroids and diameters; c_{ref} and n_{ref} from the training folds

Output: one $64 \times 64 (\times 64)$ ROI per nodule

- 1: convert stored values to HU per slice and floor at -1024 HU
- 2: resample the volume to 1 mm isotropic spacing (linear)
- 3: compute the lung mask M by method S3
- 4: map HU to $[0, 1]$ with the fixed lung window $M1$
- 5: for each axial slice s do
- 6: apply the bilateral filter ($d = 5$, $\sigma_s = 2$, $\sigma_c = 0.05$)
- 7: estimate the noise level n_s and set the clip limit c_s by the adaptive rule
- 8: apply lung-restricted CLAHE with clip limit c_s inside M
- 9: end for
- 10: for each nodule with centroid x and diameter d do
- 11: crop a cube of side $s = \min(\max(2d + 16, 24), 96)$ mm centred on x , padding with air
- 12: resample the cube to 64 samples per side
- 13: end for

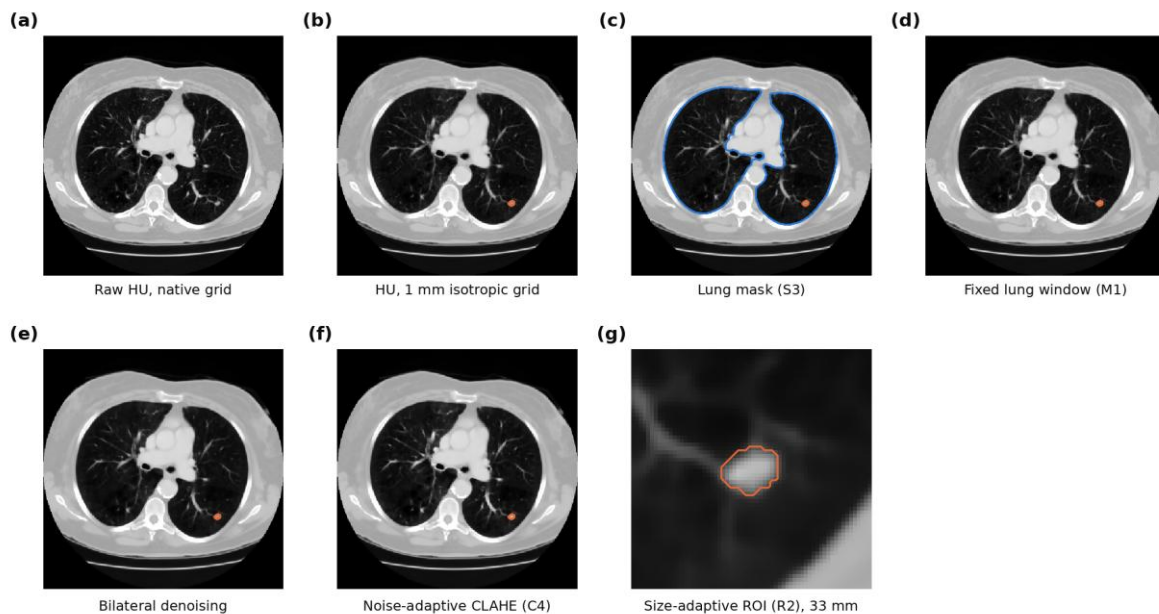


Fig. 2. One slice of scan LIDC-IDRI-0006 at each stage of the full pipeline: (a) raw HU on the native grid, (b) after resampling to 1 mm, (c) lung mask S3, (d) fixed lung window M1, (e) bilateral denoising, (f) noise-adaptive CLAHE with a clip limit of 1.06, (g) size-adaptive ROI of 33 mm. The orange outline is the consensus mask of the nodule.

TABLE III. Alternatives Compared at Each Stage (Settings of the Original Pipeline Are Marked)

Stage	ID	Definition
Resampling	—	Target spacing 0.5, 1.0, 1.5, 2.0 mm $\times$ nearest, linear, cubic interpolation (original: none)
Lung mask	S0	Threshold at -320 HU, fill holes, close (original)
	S1	Per slice: threshold, remove border-connected air, keep two largest components, fill holes
	S2	S1 followed by closing with a disk of radius 2, 5 or 10 mm
	S3	Three-dimensional components, fill, closing with a 5 mm ball
Intensity	M0	HU after the padding floor
	M1	Fixed window $[-1000, 400]$ HU scaled to $[0, 1]$
	M2	Per-scan min–max scaling
	M3 / M4	Per-scan z-score, unclipped / clipped to $[-1000, 400]$ HU
	M5	z-score of the clipped volume with body-mask statistics
	ML	Window $[-1350, 150]$ HU, then per-slice min–max (original)
Denoising	—	None; Gaussian $\sigma = 0.5, 1.0$ (original), 1.5; median 3×3 (original), 5×5 ; bilateral (original and three tuned settings); non-local means $h = 0.8\sigma, 1.0\sigma$
Contrast	C0, C1	No enhancement; global histogram equalization
	C2	CLAHE, clip limit 1 to 4, tiles 8×8 or 16×16 (original: clip 2, 8×8)
	C3	C2 restricted to the lung mask
	C4	C3 with the noise-adaptive clip limit, 8×8 tiles
ROI	R1	Fixed side of 32, 48 or 64 mm (original: 64 pixels at the image centre)
	R2	Size-adaptive side, resampled to 64 samples

Experimental Setup

The questions presented in the research are answered after successfully completing the experiments. The answers to the questions were obtained through the use of measurements that were calculated and defined beforehand, thus being available for every scan or nodule. The answers were summarized with the use of ranges taking into account nodules belonging to different patients. The numbers mentioned below were calculated by the study code based on data.

A. Dataset

The data are the initial 30 patients (LIDC-IDRI-0001 to LIDC-IDRI-0030) of the LIDC-IDRI database, provided by The Cancer Imaging Archive. Each patient has a CT series, and all 30 were employed in full. The annotations were obtained from the standardized DICOM segmentation and structured-report objects derived from the original radiologist markings [20]; this release includes nodules of 3 mm or greater, with each one marked by up to 4 radiologists and assessed in respect of 9 features including malignancy. Annotations regarding a single physical nodule are combined, and a consensus mask is then produced from voxels marked by at least half of the nodule readers. The diameter corresponds to the largest in-plane extent of the consensus mask. The malignancy label is derived from the median value for the ratings of nodule features obtained by the readers according to a five-point scale: above 3 malignancy, below 3 is benign, while exactly 3 indicates that the situation cannot be determined. In total, the experimental dataset comprises of 86 nodules from 29 patients taking the total of 235 annotations – one patient has no nodules of 3 mm or larger and takes part only in the scan level experiments. Characteristics are given in table IV.

All scans have been conducted on machines of the same producer with 120kVp, including 28 scans that have been done with the same reconstruction kernel. The provided data is not really representative of all possible variation in vendors and protocols and this limits the outcome of the intensity experiment.

TABLE IV. Characteristics of the Study Data

Property	Value
Patients and CT scans	30 and 30 (LIDC-IDRI-0001 to LIDC-IDRI-0030)
Scanner models	GE LightSpeed16 (20), LightSpeed Plus (7), LightSpeed VCT (2), LightSpeed QX/i (1)
Tube voltage; reconstruction kernel	120 kVp; STANDARD (28 scans), LUNG (2 scans)
Tube current	130 to 441 mA (median 320 mA)
Slice spacing	1.25 mm (11 scans), 2.5 mm (19 scans)
In-plane spacing; matrix	0.59 to 0.94 mm (median 0.74 mm); 512×512

Property	Value
Slices per scan	114 to 305 (median 140)
Estimated noise σ per scan	Median 7.5 HU (interquartile range 5.1 to 9.3; range 2.5 to 37.9)
Annotations; nodules	235; 86 nodules in 29 patients
Annotations per nodule	One: 22; two: 16; three: 11; four: 37 (48 nodules with three or more)
Nodule diameter	Median 7.6 mm (interquartile range 6.0 to 13.9; range 3.7 to 40.3)
Diameter groups	Under 6 mm: 21; 6 to 10 mm: 37; over 10 mm: 28
Malignancy label, all nodules	Benign 37; indeterminate 26; malignant 23
Malignancy label, three or more readers	Benign 9; indeterminate 20; malignant 19

B. Measurements

Every scan demonstrates consistency where two landmarks are placed in the histogram of HU of the referred body mask, namely, the lung peak ℓ_{lung} , which represents the mode of $[-1000, -500]$ HU region, and the soft tissue peak ℓ_{soft} , which represents the mode of $[-200, 200]$ HU. The landmark mapping $T_{m,i}$ corresponding to method m in the scan i has been applied for both landmarks. In order to find landmarks dispersion for a certain tissue, we compute the standard deviation of the mapped landmark and divide it by the average gap G_m in this case:

$$D_m^t = \frac{SD_i[T_{m,i}(\ell_i^t)]}{G_m}, G_m = \frac{1}{N} \sum_{i=1}^N [T_{m,i}(\ell_i^{\text{soft}}) - T_{m,i}(\ell_i^{\text{lung}})] \quad (13)$$

The next measure is the mean 1-Wasserstein distance W_1 between the body-mask intensity distributions $P_{m,i}$ in pair of scans, normalized by the same gap, one has a fixed random sample of 200,000 voxels per each distribution:

$$\bar{W}_m = \frac{2}{N(N-1)G_m} \sum_{i < j} W_1(P_{m,i}, P_{m,j}) \quad (14)$$

The fidelity of results presented above is obtained by means of adding zero-mean Gaussian noise with the standard deviation of 10, 25, or 50 HU. Thus, the image M_1 slice becomes filtered and compared with I slice with noise.:

$$\text{PSNR} = 10 \log_{10} \frac{R^2}{\frac{1}{|\Omega|} \sum_{x \in \Omega} (I(x) - \hat{I}(x))^2} \quad (15)$$

$$\text{SSIM}(a, b) = \frac{(2\mu_a\mu_b + C_1)(2\sigma_{ab} + C_2)}{(\mu_a^2 + \mu_b^2 + C_1)(\sigma_a^2 + \sigma_b^2 + C_2)} \quad (16)$$

In (16), μ , σ^2 and σ_{ab} are local means, variances and covariance of the two images, and C_1 and C_2 are the usual stabilizing constants. The reference contains the original acquisition noise, so these scores measure the removal of the added noise. Nodular contrast and edges. The average intensity μ_N of nodules in the center slice of each nodule is calculated and assessed against the average intensity μ_B and standard deviation σ_B of the lung mask pixel area surrounding a spherical shell 2-8 mm from the nodules excluding other nodules.

$$C = |\mu_N - \mu_B|, \text{CNR} = \frac{|\mu_N - \mu_B|}{\sigma_B} \quad (17)$$

Edge sharpness is the mean gradient magnitude in a one-pixel band ∂N along the boundary of the consensus mask:

$$E = \frac{1}{|\partial N|} \sum_{x \in \partial N} \|\nabla I(x)\| \quad (18)$$

The level of noise is measured using the wavelet measure of noise standard deviation in a bounding box containing the body. In the reports of denoising and enhancement we provide contrast, CNR, edge sharpness and noise as ratio of value after and before the process completion.

Segmentation and resampling. No reference lung masks are available so that segmentation is evaluated by means of nodules. Fraction of the nodule j left under lung mask M is

$$\rho_j = \frac{|N_j \cap M|}{|N_j|} \quad (19)$$

A nodule is declared retained when ρ_j is equal to or more than 0.5. The data also includes the lung volume and the proportion of the mask area covered by the scan since a mask covering the entire area will retain all nodules. Resampling is observed through the round-trip error, where the volume is resampled from the original area to the target area and back, and the amount of volume change.

$$\text{MAE} = \frac{1}{|M|} \sum_{x \in M} |H(x) - (\mathcal{R}^{-1}\mathcal{R}H)(x)|, \varepsilon_v = \frac{V_r - V}{V} \quad (20)$$

ROI standardization. For nodule j with the ROI of Q_j , the occupancy o_j is defined by the area filled by the nodule, with t_j being the share falling outside of the ROI. Standardization is measured as the coefficient of variation of occupancy for the different nodules.

$$o_j = \frac{|N_j \cap Q_j|}{|Q_j|}, t_j = 1 - \frac{|N_j \cap Q_j|}{|N_j|}, \text{CV}_o = \frac{SD_j(o_j)}{\text{mean}_j(o_j)} \quad (21)$$

The margin (the shortest distance from the nodule to the edge of the ROI in units of diameter) and the background fraction (the share of the ROI that is not covered by the lung mask) are also provided. Table V shows the listed measurements.

TABLE V. Measurements: Definition, Unit of Analysis, and Role

Measurement	Definition	Unit	Preferred
Landmark dispersion	SD of the mapped lung or soft-tissue peak across scans, over the mean landmark gap	Scan	Lower
Histogram distance	Mean pairwise 1-Wasserstein distance of body-mask intensities, over the mean landmark gap	Scan pair	Lower
PSNR, SSIM	Agreement of the filtered noisy slice with the slice before noise was added	Slice	Higher
Contrast, CNR	Nodule-to-ring intensity difference; difference over ring SD	Nodule	Higher
Edge sharpness	Mean gradient magnitude on the nodule boundary	Nodule	Higher
Noise ratio	Estimated noise SD after / before	Slice	Context-dependent
Retention	Share of nodules with at least half their voxels inside the lung mask	Nodule	Higher, at plausible lung volume
Round-trip error	Mean absolute HU difference after resampling to the target grid and back	Scan	Lower
Volume error	Relative change of nodule volume after resampling	Nodule	Closer to 0
Occupancy CV	Coefficient of variation of the nodule share of the ROI	Nodule set	Lower
Truncation	Share of the nodule outside the ROI	Nodule	0

C. Experiments

Eight experiments conducted (see Table VI). Two-dimensional experiments use center slice of every nodule and five other slices during each scan. In total 236 slices are taken from lungs, while E7 applies complete pipeline of the Algorithm 1 and eleven variations of the algorithm where one of the stages is omitted or restored to its original form.

TABLE VI. Experiments, the Question Each Addresses, and the Unit of Analysis

ID	Stage	Question	Units
E0	Inventory	What do the scans and nodules look like?	30 scans, 86 nodules
E1	Intensity	Which mapping keeps tissues most consistent across scans? (RQ1)	30 scans, 435 pairs
E2	Denoising	Which filter restores noisy slices and preserves nodule edges? (RQ3)	236 slices, 86 nodules
E3	Contrast	Does CLAHE raise nodule CNR? (RQ4)	86 nodules, 23 settings
E4	Lung mask	Are annotated nodules retained? (RQ2)	86 nodules, 6 masks
E5	Resampling	What is lost in intensity and nodule volume? (RQ2)	30 scans, 86 nodules
E6	ROI	Does the adaptive ROI standardize occupancy? (RQ5)	86 nodules, 4 ROIs
E7	Ablation	What does each stage contribute to the final ROI?	86 nodules, 12 variants
E8	Cost	How long does each stage take?	30 scans

D. Parameter Selection and Statistical Protocol

Out of the many components included in this research, two specifications were taken into account: denoising property, and cref of adaptive CLAHE. Both were selected by means of five-fold cross-validation of the data concerning patients, so that no patient appeared in both the selection and the outcome. The denoiser specification was that which provided the maximum average lung SSIM at 25 HU noise in the training folds, while cref was the one which gave the maximum median CNR ratio in the training folds, with the case of tie going to the smaller value.

The value of the reference noise n_{ref} is that obtained from the training folds. Measurement for the settings obtained is carried out in the held-out folds and summarized for the whole set of data used in this research. Data for all settings are also provided with descriptive labels.

Results are expressed as either a mean or a median depending on the case, with the 95% percentile bootstrap interval derived from 2,000 resampled patients. The resampling is done at the patient level, meaning that the nodules of a patient will stay with each other and not be mixed with other patients. A standard paired Wilcoxon signed-rank test is used to compare different methods.

E. Implementation and Environment

The pipeline is created in Python utilizing such libraries as NumPy, SciPy, scikit-image, OpenCV, and pydicom. The pipeline automatically runs 56 tests for synthetic volumes including loader tests, annotation reader tests, segmentation variant tests, resampler tests, ROI geometry tests, and metric function tests against known analytical values. The entire process takes only about two hours and can easily be conducted on an ordinary laptop computer without any special graphics hardware.

TABLE VII. Configuration and Environment

Item	Value
HU floor; lung and body thresholds	−1024 HU; −320 HU and −500 HU
Lung window (M1); original window (ML)	[−1000, 400] HU; [−1350, 150] HU
Consensus rule	Voxel marked by at least 50% of the readers of the nodule
Background ring for contrast and CNR	2 mm to 8 mm outside the nodule, inside lung mask S3
Synthetic noise levels; selection level	10, 25, 50 HU; 25 HU
CLAHE	256 bins, floating-point input; tiles 8×8 or 16×16
ROI output size; padding value	64 samples per side; −1000 HU
Cross-validation; bootstrap	5 folds at patient level, seed 0; 2,000 resamples, 95% percentile interval
Processor and memory	Intel Core i7-10610U (4 cores, 8 threads), 31.7 GB; no graphics card used
Operating system; Python	Windows 10; Python 3.11.9
Libraries	NumPy 2.4.6, SciPy 1.17.1, scikit-image 0.26.0, OpenCV 5.0.0, pydicom 3.0.2, pandas 3.0.6, Matplotlib 3.11.2, PyWavelets 1.9.0

RESULTS AND DISCUSSION

Findings are presented according to the stages of processing, with malignancy type breakdown next, pipeline ablation analysis, previous pipeline audit, and cost analysis following. Intervals are provided in brackets, and are expressed in terms of 95% patient bootstrap intervals; values of p are corrected using Holm technique.

A. Intensity Standardization

In Table VIII and Fig. 3, consistency measures are presented for the seven mappings. The landmark dispersions for raw HU (M0) and the fixed window mapping (M1) are equal because M1 is a linear mapping of HU within the defined limitations of 0.033 and 0.008, as determined during a sensitivity analysis for lung and soft-tissue peaks respectively. In the case of HU, the lung peak is found at a value of -868 ± 27 HU and the soft-tissue peak at a value of 53 ± 8 HU across all the thirty scans.

All types of mappings which are based on the scan content lead to increasing the landmark dispersion. Min-max scaling (M2) is the method with the smallest dispersion due to soft tissue, which equals 0.235, with the pairwise histogram distance of M2 equal to 0.235 being even higher than that of M1 (0.085, $p < 0.001$) because this process assumes that the highest and lowest values from the scans are used to create the mapping. The z-scores of the total volume (M3 and M4) produced much higher medical dispersion of the lung (0.076 and 0.075) in comparison to the histogram distance (0.157 and 0.153), as their average and standard deviation depend on the amount of air that surrounds the object being scanned. The restrictions of the statistics to the body (M5) allow the histogram distances to get back the level of M1 (0.082, $p = 0.11$), but at the same time, lung dispersion of M5 was the highest one among all tests (0.125).

There are two qualifications to be considered. Firstly, soft tissue peak was defined earlier as histogram mode ranging from [−200, 200] HU. In 17 out of 30 scans, the mode is the fat peak near −100 HU, while the remaining 13 scans have muscle and blood, thus the soft tissue dispersion is composed of two tissues. The sensitivity analysis was performed in the range [−50, 200] HU with the fat peak being excluded; therefore both results were reported in the same ranking. Secondly, the original mapping ML has a slightly lower histogram distance compared to M1 0.080 due to saturation effect, as the upper limit of 150 HU saturates at least one voxel for 27.9% of nodules while the upper limit of 400 HU gives 11.6%. The lower limit of −1350 HU was never reached. For RQ1, the standard lung window is the most stable mapping since it does not saturate nodules and scan-by-scan normalization does not provide any benefits for calibrated CT data.

TABLE VIII. Consistency Between Scans for Each Intensity Mapping (30 Scans, 435 Scan Pairs; Lower Is Better)

Mapping	Lung-peak dispersion		Soft-tissue dispersion, primary range		Soft-tissue dispersion, sensitivity range		Pairwise histogram distance	
M0: raw HU	0.033	[0.024, 0.040]	0.096	[0.086, 0.098]	0.008	[0.006, 0.010]	0.089	[0.074, 0.104]
M1: fixed lung window	0.033	[0.024, 0.040]	0.096	[0.086, 0.098]	0.008	[0.006, 0.010]	0.085	[0.070, 0.100]
M2: per-scan min-max	0.041	[0.030, 0.050]	0.263	[0.167, 0.336]	0.235	[0.147, 0.300]	0.235	[0.156, 0.317]
M3: z-score, whole volume	0.076	[0.056, 0.092]	0.149	[0.116, 0.177]	0.099	[0.076, 0.117]	0.157	[0.134, 0.181]
M4: z-score, clipped volume	0.075	[0.056, 0.092]	0.149	[0.116, 0.177]	0.097	[0.075, 0.115]	0.153	[0.130, 0.177]
M5: z-score, body statistics	0.125	[0.096, 0.147]	0.100	[0.085, 0.109]	0.023	[0.018, 0.027]	0.082	[0.068, 0.097]
ML: original mapping	0.033	[0.024, 0.040]	0.096	[0.086, 0.098]	0.008	[0.006, 0.010]	0.080	[0.065, 0.094]

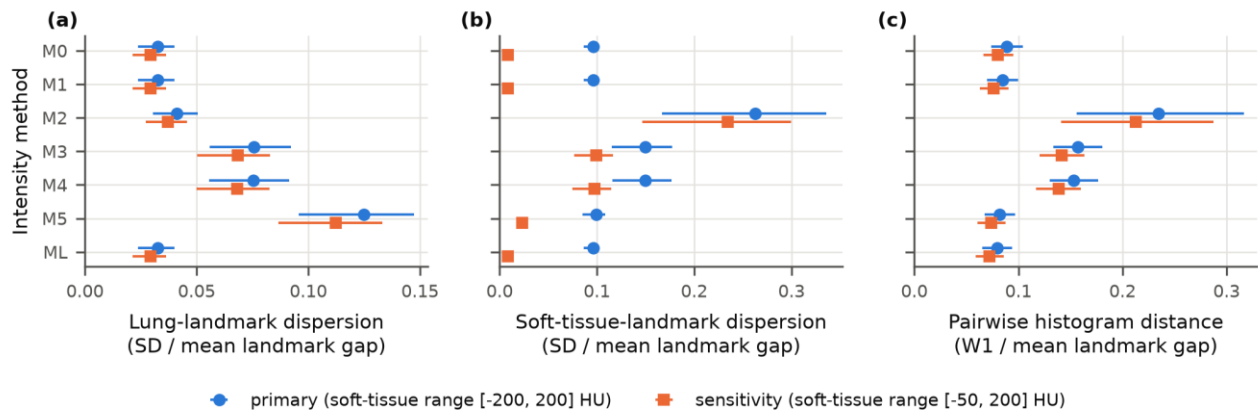


Fig. 3. Consistency measures by intensity mapping with 95% bootstrap intervals: (a) dispersion of the lung peak, (b) dispersion of the soft-tissue peak, (c) pairwise histogram distance. Blue circles use the primary soft-tissue range, orange squares the sensitivity range.

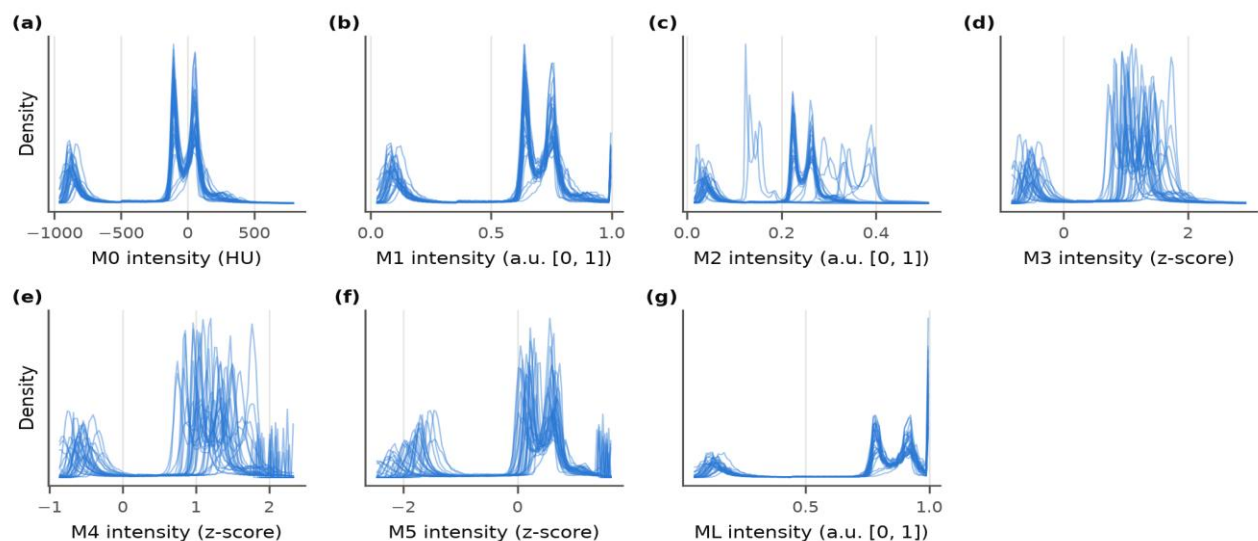


Fig. 4. Intensity distributions inside the body mask for all 30 scans under each mapping: (a) M0, (b) M1, (c) M2, (d) M3, (e) M4, (f) M5, (g) ML. Each curve is one scan; curves that coincide indicate consistent mapping of tissue to intensity.

B. Lung Segmentation

Nodule retention is documented in Table IX and Figure 5. The primitive mask S0 retains the nodules at the rate of 93 percent, yet it also occupies 85.8 percent of the scanned volume, while its mean volume equals 42.2 L, which is approximately nine times higher than the reasonable volume of the lungs. As illustrated in Figure 6, the mask encloses the air surrounding the patient. Therefore, the retention result does not have any practical meaning, since a mask covering everything retains almost all nodules.

The modified masks have an average capacity of 4.7 – 4.8 L and encompass roughly 10% of the scanned image. With the use of the S1 method, nodules sticking to the pleura are identified with an 83.7% retention rate owing to their soft tissue density and the fact that they do not fall within the clean threshold. Use of closing permits a progressive improvement in the results: 89.5% for a 2 mm close, 91.9% for a 5 mm close, and 97.7% for a 10 mm close. When using the 3D S3 method, we classify 96.5% (90.8, 100.0) of all nodules accurately, as well as all 48 nodules marked by more than three readers. According to statistical data, S3 outperforms S1 ($p < 0.001$) and is comparable with S2 associated with a 10 mm close ($p = 0.083$). In regard to RQ2, when using the corrected threshold-based mask, annotated nodules are correctly identified by S3 only if a closing is applied that ranges from 5 to 10 mm. Among the remaining nodules that are not captured by S3, two show malignancy, which is shown in the table below that presents the malignancy distribution.

TABLE IX. Nodule Retention and Mask Size for Each Lung Segmentation Method (86 Nodules, 30 Scans)

Method	Retained, all nodules (%)	Retained, ≥ 3 readers (%)	Lung volume (L)	Share of scan (%)
S0: original	93.0 [88.0, 97.4]	91.7 [84.1, 97.9]	42.23 [37.46, 47.22]	85.8
S1: corrected, per slice	83.7 [76.1, 90.0]	89.6 [81.2, 97.3]	4.74 [4.38, 5.10]	10.3
S2: S1 + 2 mm closing	89.5 [81.8, 95.2]	91.7 [84.1, 97.9]	4.75 [4.39, 5.12]	10.3
S2: S1 + 5 mm closing	91.9 [83.3, 97.2]	95.8 [88.1, 100.0]	4.77 [4.41, 5.14]	10.4
S2: S1 + 10 mm closing	97.7 [93.0, 100.0]	100.0	4.81 [4.45, 5.18]	10.4
S3: three-dimensional, 5 mm closing	96.5 [90.8, 100.0]	100.0	4.79 [4.43, 5.17]	10.4

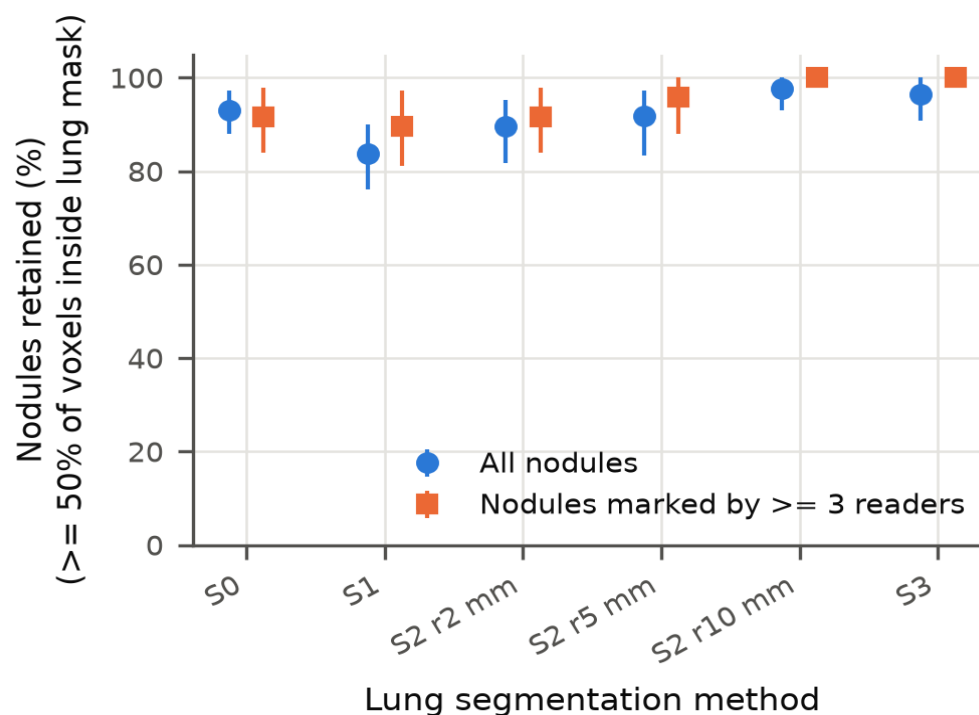


Fig. 5. Share of nodules with at least half of their voxels inside the lung mask, for all nodules and for nodules marked by three or more readers, with 95% bootstrap intervals.

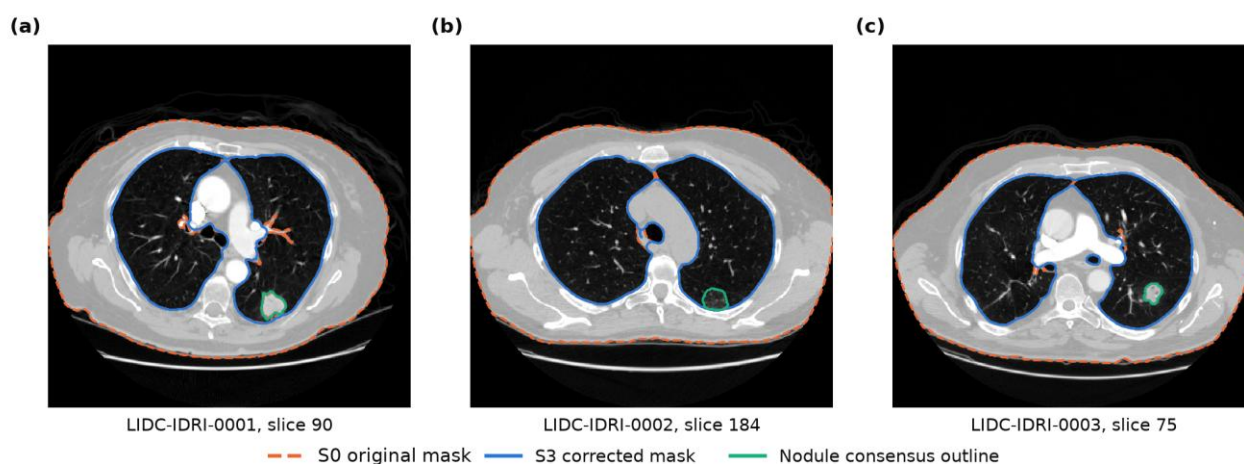


Fig. 6. Original mask S0 (dashed orange), corrected mask S3 (blue) and nodule consensus outline (green) on one slice of three scans. S0 follows the body surface because it includes the surrounding air; S3 follows the lungs and encloses the juxtaleural nodules.

C. Resampling

Table X presents the information regarding the round-trip intensity error and the alteration in the volume of the nodule. At the distance of 1 mm as utilized in this case, cubic interpolation has an error of around half that of linear interpolation (9.0 vs. 19.5 HU, $p < 0.001$), while the performance of linear interpolation is as poor as that of the nearest neighbour (19.2 HU). When moving to 0.5 mm, which is finer than the original mesh, the cubic and linear interpolation errors constitute 1.8 and 9.0 HU respectively; while the fact that the error of the nearest neighbour at this distance is zero means that all the original voxels are accounted for. The errors grow as the grids become coarser, peaking at 38.9 HU for linear interpolation at 2 mm.

At a distance of 1 mm, there is an approximately 5% decrease in nodule volumes (-5.0% $[-6.3, -3.6]$). Figure 7 demonstrates varying effects depending on nodule size; thus, at a distance of 1 mm, the loss forms 6.2% for nodules smaller than 6 mm, 6.7% for nodules sized from 6 to 10 mm, and 2.0% for nodules greater than 10 mm, whereas at a distance of 1.5 mm the loss for nodules smaller than 6 mm is equal to 16.3%. At a distance of 2 mm the mean error is very close to zero, however, the interval is very wide, and also the absolute error is the highest one at all distances (e.g. median equals 8.8% in comparison with 4.0% at 1 mm, $p < 0.001$): due to very small number of voxels errors occur in both directions. For nodules smaller than 6 mm the median number of voxels is equal to 66. The method used in the present study relies on linear interpolation; thus, cubic interpolation should be preferred at distances of 1 mm and smaller; moreover larger distances should be avoided when working with small nodules.

TABLE X. Round-Trip Intensity Error Inside the Lung Mask and Change in Nodule Volume After Resampling

Target spacing (mm)	Nearest (HU)	Linear (HU)	Cubic (HU)	Nodule volume error (%)
0.5	0.00 [0.00, 0.00]	8.98 [7.66, 10.49]	1.82 [1.46, 2.29]	-5.72 [-6.56, -4.51]
1.0	19.21 [16.89, 21.96]	19.54 [17.25, 22.34]	8.98 [7.32, 11.13]	-5.04 [-6.25, -3.56]
1.5	37.96 [34.47, 41.90]	29.37 [26.39, 33.05]	19.38 [16.57, 23.06]	-7.94 [-9.44, -6.31]
2.0	47.02 [42.49, 51.90]	38.87 [35.39, 43.14]	30.63 [26.98, 35.22]	-1.46 [-5.43, 2.44]

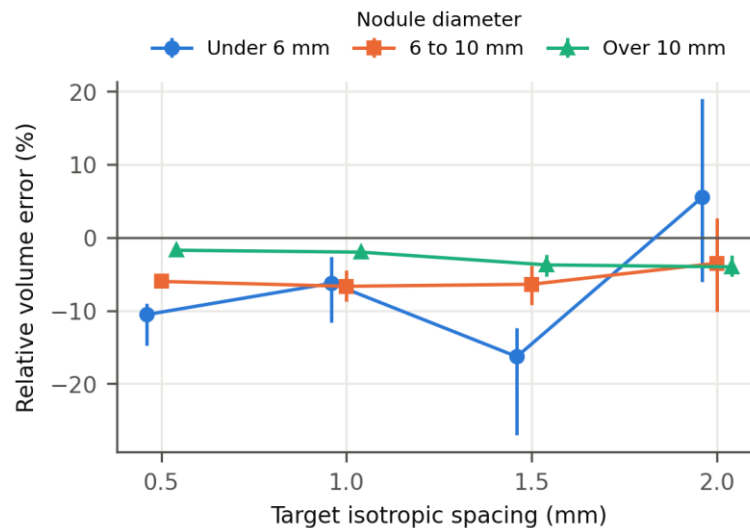


Fig. 7. Relative error in nodule volume after resampling, by target spacing and diameter group (mean with 95% bootstrap interval).

D. Denoising

In all five folds, the bilateral filter with d equal to 5 and σ_c equals 0.05 was the chosen filter through cross-validation, while the Gaussian filter with σ set to 0.5 often came second. Its lung SSIM performance at 25 HU was measured at 0.955 [0.952, 0.956] for held-out folds, compared to 0.910 [0.902, 0.918] without any filtering. Table XI and Fig. 8 contain information on all parameters regarding different noise levels.

The performance of the selected filter varies according to the noise level. In the case of 10 HU, which is close to the actual noise of these scans (7.5 HU median), the performance for the case without filtering reaches 0.983, beating all other filters except for the lightest ones (bilateral with σ_c equal to 0.02 gives 0.985; gaussian with σ equal to 0.5 produces 0.984) while the winning filter score equals 0.964. At 50 HU, the selected bilateral setting (σ_c equal to 0.10) produces the best result of 0.908. It can be concluded that the selected filter for 25 HU is a compromise and it would be better to select the strength of the filter suitable for each level of noise instead of using a single one.

The performance of the original pipeline's configurations is unsatisfactory. The Gaussian filter, with a σ value of 1, achieved a score of 0.906 at 25 HU, while the 3×3 median scored 0.916, neither of which is different from no filtering results ($p = 1.0$ and 0.81). The original bilateral filter has a score of 0.746 at all levels of noise due to its wide range of intensity width of 0.294 or about 410 HU, which is sufficient to make the intensity variations inside the lung negligible, thus producing wide spatial blurring.

The effect of unmodified slices on nodules can be seen in Table XII and Fig. 9. The bilateral filter produces an increase of 7.6% [6.0; 9.3] in CNR while retaining edge sharpness (ratio 1.011). The Gaussian filter has a value of $\sigma = 1$ and makes edges about a quarter less sharp (0.762) without meaningfully improving the CNR; the original bilateral filter on this account loses almost a third of edge sharpness (0.688). The 5×5 median filter has the highest CNR (1.144) at the expense of losing edges (0.803). Non-local means, for the particular settings tested, affect the nodules in a very shallow way. One should keep in mind that the noise values presented in Table XII should be read carefully as the wavelet estimator is sensitive mainly to fine-scale noise and exaggerates the degree of noise reduction. In the case of RQ3, narrow bilateral filtering improves accuracy and CNR of nodules but does not alter edges when the level of noise is considerable, while the results at the usual level of noise of the sample are the other way round.

The sequence of application of the two methods – denoising first and enhancement second – was evaluated using a Gaussian filter with $\sigma = 1$ and CLAHE with a clip limit of 2. Applying the method of denoising first provided a value of CNR equal to 0.795 [0.758, 0.848] while applying the method of enhancement first gave a value of CNR equal to 0.841 [0.803, 0.893] ($p < 0.001$) with a lower level of residual noise. These results do not support the hypothesis that denoising should be applied first before enhancement. It should be noted that both sequences provided a CNR lower than the initial CNR, which is a feature of CLAHE to be discussed next.

TABLE XI. Fidelity Inside the Lung Mask After Filtering Slices with Added Gaussian Noise (236 Slices, 30 Scans; Mean)

Filter	PSNR 10 HU	PSNR 25 HU	PSNR 50 HU	SSIM 10 HU	SSIM 25 HU	SSIM 50 HU
No filtering	42.92	34.97	28.95	0.983	0.910	0.747
Gaussian, $\sigma = 0.5$	40.64	36.91	32.20	0.984	0.950	0.854
Gaussian, $\sigma = 1.0$ (original)	31.41	31.27	30.80	0.913	0.906	0.885
Gaussian, $\sigma = 1.5$	27.44	27.42	27.33	0.817	0.815	0.807
Median, 3×3 (original)	34.99	33.96	31.96	0.933	0.916	0.868

Filter	PSNR 10 HU	PSNR 25 HU	PSNR 50 HU	SSIM 10 HU	SSIM 25 HU	SSIM 50 HU
Median, 5×5	29.13	29.00	28.62	0.810	0.803	0.786
Bilateral, $d = 9$, $\sigma_c = 0.294$ (original)	28.03	28.01	27.94	0.746	0.746	0.746
Bilateral, $d = 5$, $\sigma_c = 0.02$	42.75	36.57	29.69	0.985	0.943	0.777
Bilateral, $d = 5$, $\sigma_c = 0.05$ (selected)	39.09	37.12	32.22	0.964	0.955	0.872
Bilateral, $d = 5$, $\sigma_c = 0.10$	36.43	35.66	33.41	0.942	0.936	0.908
Non-local means, $h = 0.8\sigma$	42.82	37.17	33.52	0.981	0.941	0.878
Non-local means, $h = 1.0\sigma$	42.13	36.65	33.06	0.975	0.928	0.857

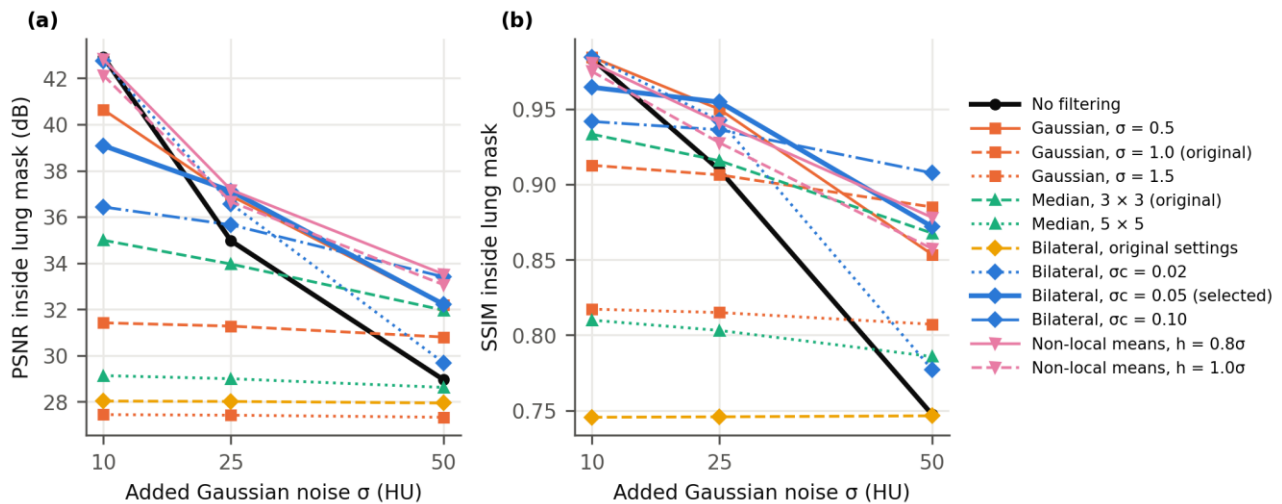


Fig. 8. (a) PSNR and (b) SSIM inside the lung mask against the level of added noise for the twelve filter settings. The thick black line is no filtering and the thick blue line is the cross-validated bilateral filter.

TABLE XII. Effect of Each Filter on the Nodules in Unmodified Slices: Ratio After / Before (86 Nodules; Mean [95% Interval])

Filter	Noise σ ratio	CNR ratio	Edge sharpness ratio
No filtering	1.000 [1.000, 1.000]	1.000 [1.000, 1.000]	1.000 [1.000, 1.000]
Gaussian, $\sigma = 0.5$	0.600 [0.582, 0.617]	1.012 [0.999, 1.027]	0.935 [0.929, 0.941]
Gaussian, $\sigma = 1.0$ (original)	0.144 [0.126, 0.161]	1.016 [0.970, 1.065]	0.762 [0.743, 0.780]
Gaussian, $\sigma = 1.5$	0.049 [0.041, 0.056]	0.979 [0.907, 1.056]	0.601 [0.576, 0.630]
Median, 3×3 (original)	0.568 [0.505, 0.628]	1.092 [1.067, 1.121]	0.934 [0.915, 0.947]
Median, 5×5	0.323 [0.279, 0.366]	1.144 [1.071, 1.219]	0.803 [0.770, 0.829]
Bilateral, $d = 9$, $\sigma_c = 0.294$ (original)	0.083 [0.080, 0.086]	1.112 [1.053, 1.173]	0.688 [0.651, 0.722]
Bilateral, $d = 5$, $\sigma_c = 0.02$	0.406 [0.373, 0.451]	1.028 [1.021, 1.035]	1.005 [1.002, 1.007]
Bilateral, $d = 5$, $\sigma_c = 0.05$ (selected)	0.204 [0.190, 0.224]	1.076 [1.060, 1.093]	1.011 [1.000, 1.019]
Bilateral, $d = 5$, $\sigma_c = 0.10$	0.173 [0.163, 0.182]	1.110 [1.085, 1.136]	0.978 [0.959, 0.993]
Non-local means, $h = 0.8\sigma$	0.663 [0.608, 0.715]	1.004 [1.001, 1.007]	0.999 [0.998, 1.000]
Non-local means, $h = 1.0\sigma$	0.575 [0.516, 0.631]	1.006 [1.002, 1.012]	0.999 [0.997, 1.000]

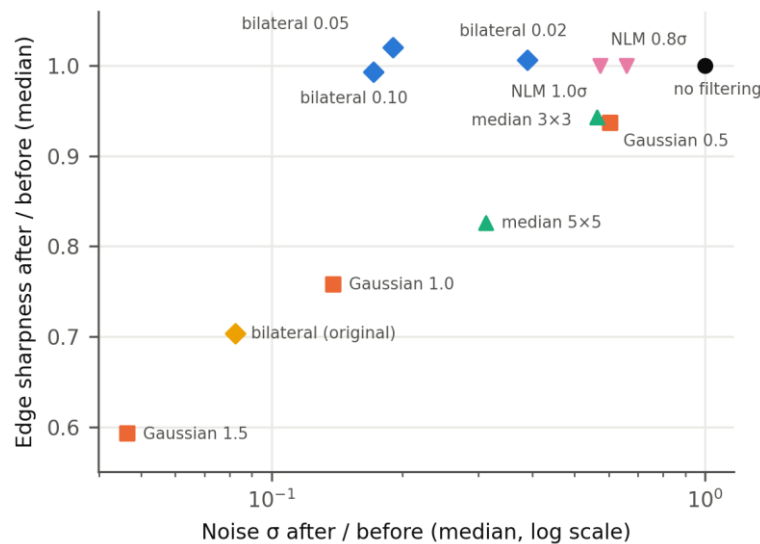


Fig. 9. Noise reduction against retention of edge sharpness on the nodule boundary for each filter setting (medians over 86 nodules). Settings near the top keep edges; settings to the left remove more fine-scale noise.

E. Contrast Enhancement

The present study summarizes the evaluations in Table XIII and Figure 10. According to the results, the Global histogram equalization (C1) minimizes the nodule contrast down to 0.42 from the original value and the CNR to 0.57. Fixed CLAHE (C2) does not change the contrast significantly (remains at 1.01 for the clip limit 2). Nevertheless, noise increases by the factor of 1.69 and thus the median CNR ratio is reduced to 0.790 [0.745, 0.815] for the clip limit 2, to 0.688 for the limit 3 and to 0.614 for the limit 4. The 16×16 tile grid generates the similar results (0.802 for the clip limit 2). When the enhancement is performed only in the lung area (C3), the contrast increases by 22% while suppressing noise growth to 1.19 with the CNR remaining at 0.745 for the clip limit 2. The explanation for this result is that CNR is defined as the ratio with the standard deviation of the tissue around the nodule as the denominator which contains textures and vessels that CLAHE enhances together with the nodule.

In the case of the adaptive variant C4, the choice of $c_{ref} = 1$ was chosen by the cross-validation procedure as the lowest possible clip limit candidate for each fold because CNR depends only on the clip limit in a decreasing manner. In this case, the clip limit is greater than 1 only in the slices quieter than the reference noise levels with a median clip limit equal to 1.08 (1.00-3.06 range). In response, the median CNR is equal to 0.998 [0.934, 1.000] with a mean value of 0.919 [0.871, 0.967]. It means that half of the nodules suffer no changes while the other half suffer losses in terms of CNR.

The trend continues in the various sizes of nodules: a median CNR ratio of 0.79 for 0-6mm nodules occurs at clip limit 2 compared with 0.74 for the 6-10 mm range with a ratio of 0.83 for nodules over 10mm. According to the radiologist subtlety score, CNR ratios lie from 0.74 to 0.81 for rankings between 2 and 5, with the lowest three nodules having a median of 1.33, which is not useful information. For RQ4, CLAHE does not improve the CNR of the evaluated nodules, with the noise-adaptive rule yielding the same result. Although these measures describe the images, it is a different topic whether the improvement in contrast and edges would facilitate classifier work; this is outside the scope of the study.

TABLE XIII. Effect of Contrast Enhancement on the Nodules: Ratio After / Before (86 Nodules; 8×8 Tiles)

Method	Contrast, mean	CNR, median	CNR, mean	Noise σ , mean	Edge sharpness, mean
C0: no enhancement	1.00	1.000	1.000	1.00	1.00
C1: global equalization	0.42 [0.37, 0.48]	0.570 [0.480, 0.622]	0.620 [0.551, 0.705]	1.63 [1.51, 1.74]	0.34 [0.28, 0.40]
C2: CLAHE, clip 1	1.00 [1.00, 1.00]	1.000 [0.999, 1.000]	1.000 [0.999, 1.000]	1.07 [1.04, 1.10]	1.00 [1.00, 1.00]
C2: CLAHE, clip 2 (original)	1.01 [0.97, 1.06]	0.790 [0.745, 0.815]	0.800 [0.769, 0.838]	1.69 [1.64, 1.76]	0.88 [0.85, 0.92]
C2: CLAHE, clip 3	0.99 [0.94, 1.06]	0.688 [0.614, 0.716]	0.702 [0.660, 0.752]	2.23 [2.15, 2.32]	0.80 [0.76, 0.85]
C2: CLAHE, clip 4	0.98 [0.91, 1.06]	0.614 [0.542, 0.644]	0.634 [0.588, 0.690]	2.65 [2.54, 2.79]	0.74 [0.69, 0.80]
C3: lung-restricted, clip 2	1.22 [1.15, 1.30]	0.745 [0.719, 0.788]	0.763 [0.733, 0.804]	1.19 [1.17, 1.21]	1.03 [0.98, 1.10]

Method	Contrast, mean	CNR, median	CNR, mean	Noise σ , mean	Edge sharpness, mean
C3: lung-restricted, clip 4	1.21 [1.12, 1.31]	0.601 [0.552, 0.630]	0.619 [0.579, 0.676]	1.29 [1.25, 1.33]	0.91 [0.85, 0.98]
C4: adaptive, cref = 2	1.25 [1.17, 1.35]	0.763 [0.709, 0.859]	0.772 [0.719, 0.844]	1.21 [1.16, 1.25]	1.06 [0.96, 1.19]
C4: adaptive, cross-validated	1.21 [1.14, 1.27]	0.998 [0.934, 1.000]	0.919 [0.871, 0.967]	1.10 [1.06, 1.13]	1.15 [1.08, 1.21]

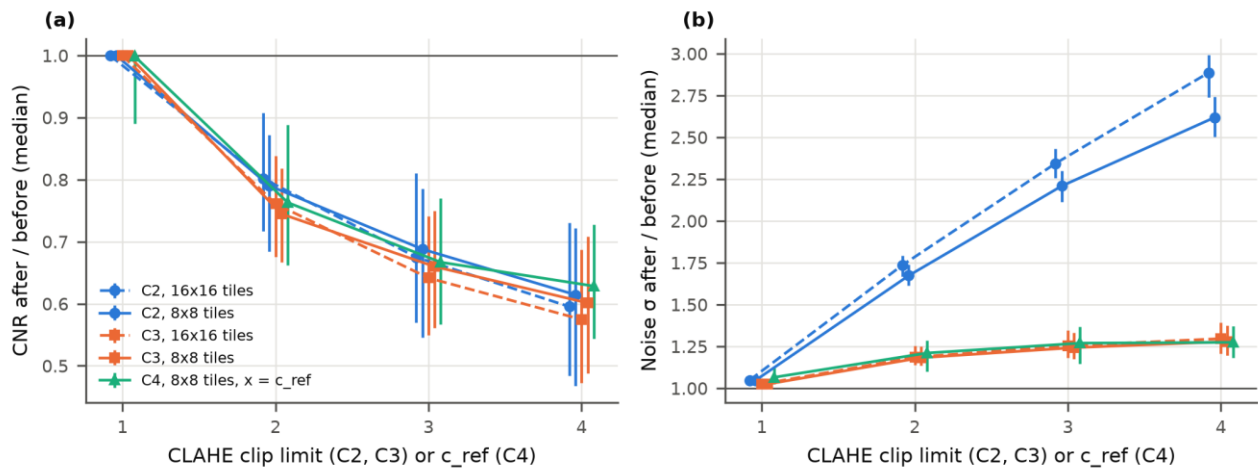


Fig. 10. (a) Median CNR ratio and (b) noise amplification against the clip limit for fixed CLAHE (C2), lung-restricted CLAHE (C3) and noise-adaptive CLAHE (C4, plotted against cref), with 95% bootstrap intervals.

F. ROI Extraction

Table XIV makes a comparison of ROI techniques while Figure 11 illustrates six nodules with both the fixed crop of 64 mm as well as the adaptive crop. With a fixed crop, the nodule occupies different volumes of the array, the coefficient of variation of occupancy is 2.06 in three dimensions at fixed side lengths of the crop as compared to the nodule size without truncation being equal to the coefficient of variation of nodule volume. In three dimensions, the size of the ROI is reduced to 0.63 [0.56, 0.71] and in two dimensions the figure decreases from 1.51 to 0.51 (Figure 12). The use of the smallest fixed crop leads to truncation of 9.3% [3.0, 19.7] of nodules, whereas the use of 48 mm, 64 mm as well as adaptive crop leads to no truncation of the nodules. Moreover, adaptive crop had lesser volume of tissue from outside the lungs as compared to 64 mm crop with background fraction of 0.23 versus 0.38 ($p < 0.001$). The gain is greatest for small nodules. Typically, nodules less than 6 mm account for about 0.02% of a 64 mm cube, and 0.30% of the adaptive cube. However, nodules larger than 10 mm can occupy roughly 1% of either cube type. The downside of this is that the spacing between samples in the adaptive ROI is not consistent across nodules and can range from 0.38 mm to 1.5 mm, with a median of 0.49 mm. As a result, physical size does not correspond to pixel count, so any model requiring this data should receive diameter as another input. In RQ5, the size-adaptive ROI provides a representation of nodules that does not involve discarding size.

TABLE XIV. ROI Methods: Variation of Occupancy, Truncation, Margin, and Background (86 Nodules)

Method	Side (mm)	Occupancy CV, 2D	Occupancy CV, 3D	Nodules truncated, 3D (%)	Median margin ($\times$ diameter)	Background fraction, 3D
R1: fixed, 32 mm	32	1.51 [1.19, 1.78]	2.06 [1.52, 2.61]	9.3 [3.0, 19.7]	1.52	0.21 [0.16, 0.26]
R1: fixed, 48 mm	48	1.51 [1.19, 1.78]	2.06 [1.52, 2.62]	0.0 [0.0, 0.0]	2.53	0.30 [0.25, 0.35]
R1: fixed, 64 mm	64	1.51 [1.19, 1.78]	2.06 [1.52, 2.62]	0.0 [0.0, 0.0]	3.54	0.38 [0.34, 0.43]
R2: size-adaptive	24–96	0.51 [0.45, 0.56]	0.63 [0.56, 0.71]	0.0 [0.0, 0.0]	1.49	0.23 [0.19, 0.28]

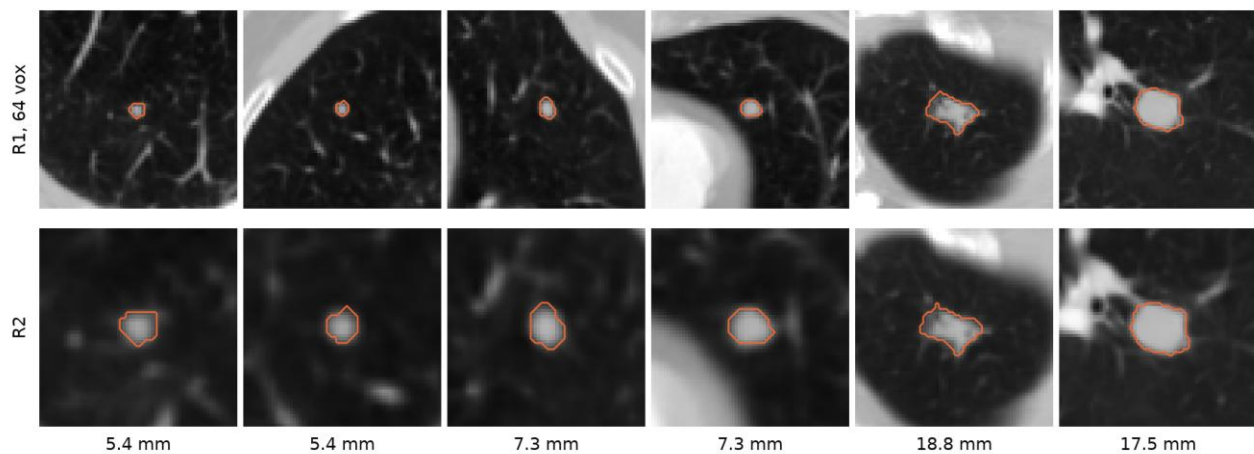


Fig. 11. Centre slices of six nodules (two per diameter group) under the fixed 64 mm ROI (top) and the size-adaptive ROI (bottom), with consensus outlines. Diameters are given below each column.

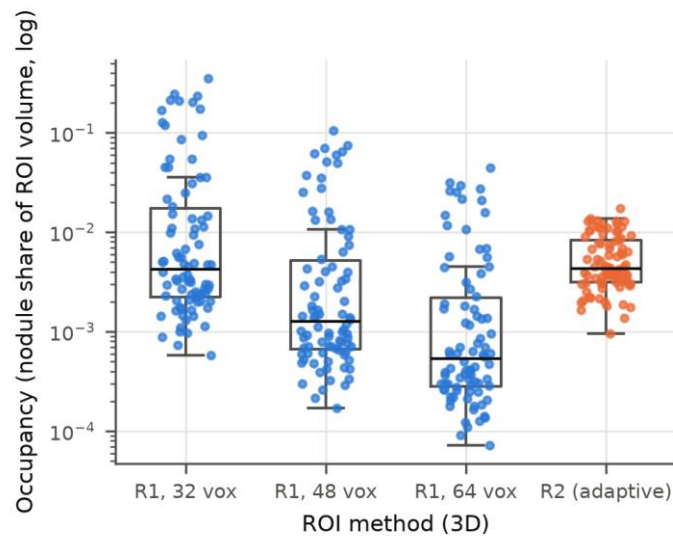


Fig. 12. Occupancy of the three-dimensional ROI by the nodule for the fixed and adaptive methods (log scale; one point per nodule).

G. Results by Malignancy Group

The principal nodule-level findings for the benign, indeterminate, and malignant nodule categories (gathered from the per-nodule data) are presented in Table XV. Malignant nodules are found to be bigger (average diameter of malignant nodules is 18.8 mm while that of non-cancerous nodules is approximately equal to 6.6 mm) and also show lower CNR values (5.32 versus 7.16). Every group shows the positive effect of data processing on CNR. Denoising results in 6-9% CNR increase while CLAHE method with the inundation limit equal to 2 results in its decrease to 18-24%. Also, the adaptive region of interest (ROI) leads to reducing the variation of occupancy by distinctive groups. However, the case of segmentation result is an exception. In case of S3, all indeterminate nodules are detected along with nearly 97.3% of benign nodules, while only 91.3% of malignant ones. It is also a case of a removal of malignant nodules during the classification process that makes it impossible to compensate for further steps; therefore, either dilating (enlarging) the mask processing the nodules is the solution or utilizing a different and more advanced detector capable of serving any additional task.

TABLE XV. Main Nodule-Level Results by Malignancy Group

Group	n	Median diameter (mm)	Median baseline CNR	CNR ratio, selected denoiser (mean)	CNR ratio, clip 2 (median)	Retained by S3 (%)	Occ. CV, R1 64	Occ. CV, R2
Benign	37	6.6	7.16	1.078 [1.065, 1.095]	0.780 [0.735, 0.809]	97.3	3.73	0.61
Indeterminate	26	7.3	6.19	1.091 [1.052, 1.132]	0.762 [0.690, 0.873]	100.0	3.40	0.54
Malignant	23	18.8	5.32	1.056 [1.036, 1.078]	0.823 [0.781, 0.909]	91.3	0.95	0.30

Group	n	Median diameter (mm)	Median baseline CNR	CNR ratio, selected denoiser (mean)	CNR ratio, CLAHE clip 2 (median)	Retained by S3 (%)	Occ. CV, R1 64	Occ. CV, R2
All nodules	86	7.6	6.42	1.076 [1.059, 1.093]	0.790 [0.745, 0.816]	96.5	2.06	0.63

H. Pipeline Ablation

According to Table XVI and Figure 13, the impact on the final ROI can be seen through the removal or restoration of any part of a system. The entire system shows nodule CNR to be equal to 8.09 [6.80, 9.00]. If the denoising system were taken out, the CNR would drop to 7.27 ($p < 0.001$) while using the actual Gaussian filter would give a CNR of 7.11, and edge sharpness 0.228 mm instead of 0.152. The removal of contrast enhancement gives no effect on CNR readings (8.11, $p = 0.74$) but the overall edge sharpness is lower without it (0.188 mm). The displacement of the original CLAHE system leads to maximum loss due to this change (6.08, $p < 0.001$). The reversal to the original intensity mapping will lead the CNR index to 7.75 ($p = 0.002$). The change of the order of processes in comparison with what was presented in the testing gives little impact that follows the previous trend (7.91, $p = 0.016$), which can be explained by the fact that the whole process worked close to none. The removal of resampling and the transition of lung mask filtering do not alter the result noticeably.

The variation in the use of ROI mostly affects standardization rather than contrast. A fixed 64-mm ROI has the same CNR as the adaptive ROI (8.02) but one should note that the variation over the occupancy on the centre slice is three times greater (1.51 vs. 0.51). After cropping, the image center gives the original pipeline adequate sample of the nodule to conduct CNR calculations into only 4 of the 86 nodules.

TABLE XVI. Ablation of the Full Pipeline: Measurements on the Final ROI (86 Nodules; Mean [95% Interval])

Variant	CNR	Edge sharpness (per mm)	Occupancy CV, 2D	p, against full	CNR
Full pipeline	8.09 [6.80, 9.00]	0.2280 [0.1961, 0.2531]	0.513 [0.454, 0.556]	—	
No resampling	7.34 [6.15, 8.21]	0.2629 [0.2155, 0.3001]	0.469 [0.411, 0.508]	0.15	
Original lung mask (S0)	8.05 [6.74, 8.99]	0.2010 [0.1684, 0.2284]	0.513 [0.454, 0.556]	0.74	
No lung mask	8.04 [6.73, 8.99]	0.1936 [0.1623, 0.2213]	0.513 [0.454, 0.556]	0.74	
Original intensity mapping (ML)	7.75 [6.54, 8.63]	0.2410 [0.2073, 0.2663]	0.513 [0.454, 0.556]	0.002	
No denoising	7.27 [6.18, 8.00]	0.2167 [0.1860, 0.2423]	0.513 [0.454, 0.556]	< 0.001	
Original Gaussian filter, $\sigma = 1$	7.11 [5.95, 8.06]	0.1517 [0.1337, 0.1671]	0.513 [0.454, 0.556]	< 0.001	
No contrast enhancement	8.11 [6.79, 9.05]	0.1877 [0.1584, 0.2155]	0.513 [0.454, 0.556]	0.74	
Original CLAHE, clip limit 2	6.08 [5.22, 6.71]	0.1695 [0.1430, 0.1957]	0.513 [0.454, 0.556]	< 0.001	
CLAHE before denoising	7.91 [6.68, 8.77]	0.2217 [0.1901, 0.2477]	0.513 [0.454, 0.556]	0.016	
Fixed 64 mm ROI	8.02 [6.72, 8.89]	0.2026 [0.1800, 0.2195]	1.511 [1.191, 1.780]	0.74	
Image-centre ROI (4 nodules)	2.81 [2.12, 3.71]	0.1230 [0.0956, 0.1422]	5.68	n/a	

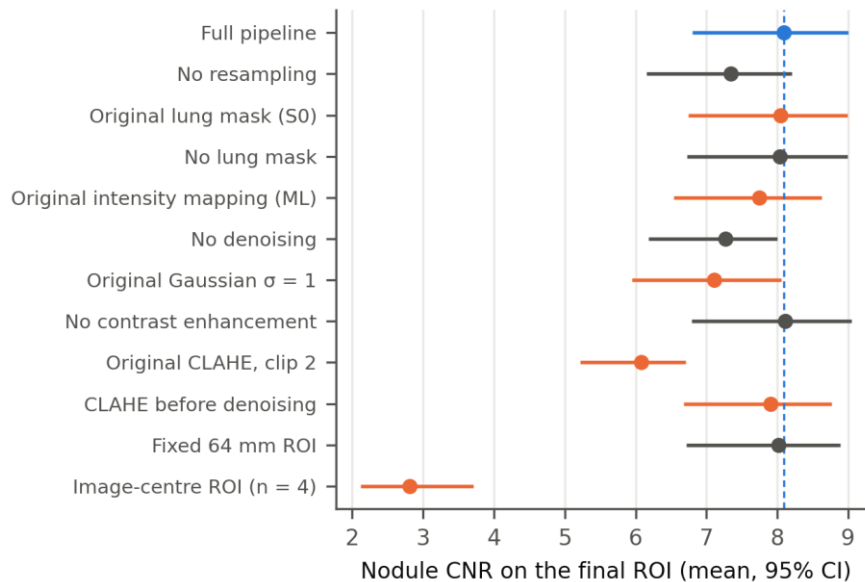


Fig. 13. Nodule CNR on the final ROI for the full pipeline and eleven variants (mean with 95% bootstrap interval). Orange points are variants that restore a setting of the original pipeline; the dashed line marks the full pipeline.

I. Audit of the Original Pipeline

In the initial iteration of the pipeline, the middle slice from a single scan was processed, generating similarity scores based on the investigation of 3 + CLAHE filters. The code could be run again using the LIDC-IDRI-0001 dataset and the entire metric file obtained in this way is identical to the previous one. It implies that the study's conclusions trace back to their sources (Table XVII). As the comparison of similarity scores was made with CLAHE output, thus the filters were ranked depending on the degree of their influence on the image that was obtained using this filter and CLAHE scored the highest value of 1.000. As for the reported signal-to-noise ratio, it was calculated as the mean image divided by its standard deviation, thus the obtained values are almost the same for all methods used within the current study.

TABLE XVII. Audit of the Original Single-Slice Pipeline

Issue	What the original code did	Measured consequence
Self-referenced fidelity	SSIM and PSNR of each filter computed against the CLAHE output	CLAHE: SSIM 1.000, PSNR infinite. Filters ranked by least change (median 0.899, Gaussian 0.896, bilateral 0.787)
Signal-to-noise ratio	Whole-image mean divided by whole-image SD	1.16 to 1.19 for every method; does not respond to noise
Lung mask	Threshold, fill and close without removing outside air	Mask covers 85.8% of the scan (42.2 L against 4.8 L)
ROI location	64-pixel crop at the image centre	Nodule measurable in 4 of 86 cases (4.7%)
Filter settings	Gaussian $\sigma = 1$; bilateral $d = 9$, $\sigma_c = 0.294$	Lung SSIM 0.906 and 0.746 at 25 HU against 0.910 unfiltered; edge sharpness ratios 0.76 and 0.69
Sample	One slice of one scan	No estimate of variation between scans or nodules

J. Computational Cost

As indicated in Table XVIII and Fig. 14, the laptop processor of Table VII has a scanning time of 65.2 ± 18.6 s, with a peak memory usage of 1.78 ± 0.29 GB (maximum 2.46 GB). For the adaptive CLAHE, it accounted for 40.6 s and for the three-dimensional lung mask, the time was 19.3 s. The rest of the stages took no more than 6 s. The contribution of the enhancement stage is zero percent in terms of CNR, and therefore omitting it would make the scanning time be around 25 s per scan.

TABLE XVIII. Wall Time per Scan for Each Stage of the Full Pipeline (30 Scans; Mean $\pm$ SD)

Stage	Time per scan (s)
Loading and HU conversion	2.66 ± 0.94
Resampling to 1 mm	1.38 ± 0.57
Lung mask (S3)	19.33 ± 3.77

Stage	Time per scan (s)
Intensity window (M1)	0.33 ± 0.09
Bilateral denoising	0.49 ± 0.14
Noise-adaptive CLAHE (C4)	40.61 ± 16.38
Size-adaptive ROI (R2)	0.39 ± 0.35
Total	65.17 ± 18.56

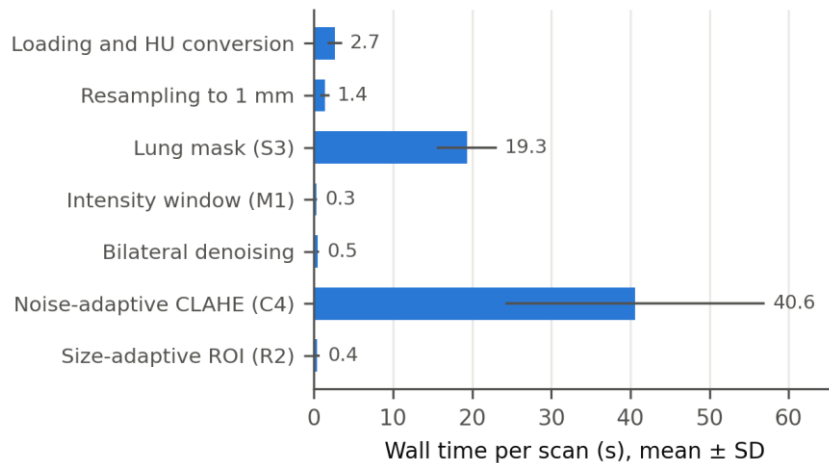


Fig. 14. Wall time per scan for each stage of the full pipeline (mean ± SD over 30 scans).

K. Summary of Findings

Table XIX summarizes responses to the research queries and details which aspects they serve to support. The results correspond to three components of the examined pipeline: the fixed lung window, corrected lung mask with seals, and size-adaptive ROI. One of the components is conditionally supported: bilateral denoising, based on noise level requirements. The CLAHE method does not receive any support.

TABLE XIX. Summary of Findings and the Configuration They Support

Question	Finding	Supported choice
RQ1	A fixed lung window keeps landmark dispersion at the level of raw HU; whole-volume per-scan scaling raises the histogram distance by a factor of 1.8 to 2.8, and body-restricted z-scores raise the lung-peak dispersion	Fixed window [−1000, 400] HU; no per-scan normalization
RQ2	Corrected masks retain 96.5 to 97.7% of nodules at 4.8 L; the original mask covers 86% of the scan. At 1 mm, cubic interpolation halves the intensity error of linear; nodule volume falls by about 5%	Threshold mask with 5 to 10 mm closing; cubic interpolation at 1 mm or finer
RQ3	A narrow bilateral filter raises lung SSIM from 0.910 to 0.955 at 25 HU and nodule CNR by 7.6% with edges intact; at 10 HU no filtering is better	Edge-preserving filter with strength matched to the noise level; none for low-noise scans
RQ4	CLAHE lowers CNR at every clip limit above 1; the adaptive rule is selected at its weakest setting	No contrast enhancement unless a downstream benefit is shown
RQ5	The size-adaptive ROI lowers the occupancy CV from 2.06 to 0.63 with no truncation and less background	Size-adaptive ROI, with the diameter passed separately if needed

DISCUSSION

Well-calibrated data requires less normalization rather than more normalization. CT intensity values are physical measures given on a calibrated scale. The fixed-width window preserves that calibration, because all statistics calculated from the scan, whatever they may be (variance, mean, standard deviation, etc.), replace it with a number depending on a patient size, viewing angle and heavy objects present. As in the Tas and Yavuz [11] study, scaling each image with its own range is common data pre-processing stage. According to the results of the present study, it does not allow to compare scans as intended.

Enhancement and visibility are not the same. CLAHE helps in identifying structures better as it increases the contrast and sharpness of the edges of nodules in the lungs. Nevertheless, for a classifier that is based on the numbers, CLAHE produces worse numbers in terms of detectability since the background texture is amplified more than the nodules. At least in our case, the adaptive rule behaves adequately by reducing the method effectiveness in the noisy slices. Still, the best free

parameter setting is the lowest. Such a negative result of the processing is almost advantageous for classification at the next stage, as this step is removed from the sequence of actions and, consequently, the variability is minimized.

The standardization of the ROI is where the adaptation is advantageous. Grad-CAM analysis shows that classifiers pay attention to the margin of the nodule [23] while radiomic features are calculated from analyzed shapes [19]. In standard cropping, a 5 mm nodule appears as several dozen pixels in a parenchyma, and a 30 mm nodule takes up most of the image. Changing the size of the crop to match the nodule size eliminates the majority of that variation and reduces the share of the chest wall and mediastinum in the input.

Validity threats exist. The sample being small and uniform means that just 30 patients in a row are included, so it is only valid to draw conclusions involving the intensity mapping and noise from scans obtained with other devices and low-radiation protocols. As the identities of the readers were anonymized in the released annotations, counts given correspond to the numbers of the annotations, and nodules less than 3 mm were not counted. There were no reference lung masks to provide evaluation of the segmentation of the detected nodules only. The added noise was found to be white and Gaussian, whereas CT noise was spatially correlated; ranking of filters under real noise might be different. The noise estimator gives an overly optimistic estimate of smoothing. The ROI rule utilizes the diameter of the annotations, whereas it would be assessed based on detector values in real life. All filters and CLAHE operate on the slice basis. A final point that should be pointed out is that each of the measures is related to the image and none to the classification; the Table XIX configuration is shown through images and requires validation through the downstream task.

CONCLUSION

This paper discusses the stages in the CT preprocessing pipeline for lung nodules in a separate manner, on the basis of a patient sample of 30 CT scans and 86 annotated nodules of LIDC-IDRI, by measuring consensus mask, determining patient-specific parameters through cross-validation and calculating bootstrap intervals. Maintaining a fixed lung window resulted in keeping tissue intensities consistent as raw Hounsfield units; however, since there was min-min/max-max and z-scores scaling per scan, the rest of the scans were less consistent. A modified threshold-based lung mask helped to retain nearly all nodules at a realistic lung volume. The cubic method has been proven to retain the intensities better than the linear interpolation method used above; the resampling carried out at 1 mm reduced the volume of the nodule by almost 5%. The narrow bilateral filter helped with better preservation of the image signal compared to the image noise at moderate noise levels without introducing blurriness, however, it produced no advantages when working with the sample that had low noise levels. CLAHE diminished the contrast to noise ratio of nodules regardless of the strength used, and the noise adaptive method was shown to not improve the application of CLAHE technique. The size-adaptive ROI allowed reducing the variability of the number of nodules to about one-third without cutting any of their parts. Out of the two adaptations proposed, one of them was confirmed by the data obtained earlier, while the other one had no backing.

Future Work

The pipeline can be extended in four ways. The first involves evaluating its scale and diversity; the pipeline can be used for an arbitrary number of patients, and repeating the experiment on the total anomaly dataset, with all possible manufacturers and kernels, would test intensity and noise results where they are expected to vary the most. The second is a downstream task where the malignancy classification model is developed using the approaches evaluated in this work, so the relationship between pixel-wise values and classification accuracy would be tested, especially for CLAHE. The third is associated with filtering based on noise according to its estimated value for every scan obtained using a suitable CT noise model. The fourth is the detector; instead of the annotated centroids and diameters of nodules, the results of the nodule detector will determine the propagation of both localization and size errors into the adaptive ROI.

Limitations and Ethical Considerations

This study relies on data from a single vendor, and it does not indicate how images can vary from one vendor or institution to another. It reviews the characteristics of the images but does not claim anything about their diagnostic performance. The malignancy ratings are based on the opinion of radiologists and not on histopathology. The study uses existing data only, which is publicly available and which is labeled as a research dataset according to the provisions of the Creative Commons Attribution license. The researchers made no effort to identify or collect any new data. The techniques introduced do not provide a clinical tool but are meant for use in the research of computer-assisted image analysis.

Data and Code Availability

CT scans form part of the public resource LIDC-IDRI provided by the Cancer Imaging Archive, with its annotations being available through the standardized DICOM representation and copies used herein were downloaded from NCI Imaging Data Commons. The code that produced every table and figure (including the configuration of this code, tests and per-scan and per-nodule result tables), can be obtained from the corresponding author upon a reasonable request.

ACKNOWLEDGMENT

The authors express their gratitude to the National Cancer Institute and the Foundation for the National Institutes of Health for their instrumental role in making the freely available public LIDC-IDRI Database available for the purposes of this study.

REFERENCES

1. M. Cellina et al., "Artificial Intelligence in Lung Cancer Screening: The Future Is Now," *Cancers*, vol. 15, no. 17, p. 4344, Aug. 2023, doi: 10.3390/cancers15174344.
2. D. Yang et al., "Advances in artificial intelligence applications in the field of lung cancer," *Front. Oncol.*, vol. 14, p. 1449068, Sep. 2024, doi: 10.3389/fonc.2024.1449068.

3. J. Ding, A. Li, Z. Hu, and L. Wang, "Accurate Pulmonary Nodule Detection in Computed Tomography Images Using Deep Convolutional Neural Networks," Aug. 29, 2017, arXiv: arXiv:1706.04303. doi: 10.48550/arXiv.1706.04303.
4. H. Peng, H. Sun, and Y. Guo, "3D multi-scale deep convolutional neural networks for pulmonary nodule detection," *PLoS ONE*, vol. 16, no. 1, p. e0244406, Jan. 2021, doi: 10.1371/journal.pone.0244406.
5. M. Kanan et al., "AI-Driven Models for Diagnosing and Predicting Outcomes in Lung Cancer: A Systematic Review and Meta-Analysis," *Cancers*, vol. 16, no. 3, p. 674, Feb. 2024, doi: 10.3390/cancers16030674.
6. P. N. Megat Ramli, A. N. Aizuddin, N. Ahmad, Z. Abdul Hamid, and K. I. Ismail, "A Systematic Review: The Role of Artificial Intelligence in Lung Cancer Screening in Detecting Lung Nodules on Chest X-Rays," *Diagnostics*, vol. 15, no. 3, p. 246, Jan. 2025, doi: 10.3390/diagnostics15030246.
7. H. Jung, B. Kim, I. Lee, J. Lee, and J. Kang, "Classification of lung nodules in CT scans using three-dimensional deep convolutional neural networks with a checkpoint ensemble method," *BMC Med Imaging*, vol. 18, no. 1, p. 48, Dec. 2018, doi: 10.1186/s12880-018-0286-0.
8. Q. Dou, H. Chen, L. Yu, J. Qin, and P.-A. Heng, "Multilevel Contextual 3-D CNNs for False Positive Reduction in Pulmonary Nodule Detection," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 7, pp. 1558–1567, Jul. 2017, doi: 10.1109/TBME.2016.2613502.
9. I. Naseer, S. Akram, T. Masood, M. Rashid, and A. Jaffar, "Lung Cancer Classification Using Modified U-Net Based Lobe Segmentation and Nodule Detection," *IEEE Access*, vol. 11, pp. 60279–60291, 2023, doi: 10.1109/ACCESS.2023.3285821.
10. A. A. Shah, H. A. M. Malik, A. Muhammad, A. Alourani, and Z. A. Butt, "Deep learning ensemble 2D CNN approach towards the detection of lung cancer," *Sci Rep*, vol. 13, no. 1, p. 2987, Feb. 2023, doi: 10.1038/s41598-023-29656-z.
11. M. O. Tas and H. S. Yavuz, "Enhancing Lung Cancer Survival Prediction: 3D CNN Analysis of CT Images Using Novel GTV1-SliceNum Feature and PEN-BCE Loss Function," *Diagnostics*, vol. 14, no. 12, p. 1309, Jun. 2024, doi: 10.3390/diagnostics14121309.
12. H. Tu et al., "Improving Lung Cancer Risk Prediction Using Machine Learning: A Comparative Analysis of Stacking Models and Traditional Approaches," *Cancers*, vol. 17, no. 10, p. 1651, May 2025, doi: 10.3390/cancers17101651.
13. A. Ishtiaq, N. Naseer, Z. Mohy-Ud-Din, A. Akhlaq, H. A. Guesmi, and A. Rashid, "AI-based Lung Cancer Prediction: Development, Comparison, and Evaluation of Machine Learning Models," in *2024 26th International Multi-Topic Conference (INMIC)*, Karachi, Pakistan: IEEE, Dec. 2024, pp. 1–5. doi: 10.1109/INMIC64792.2024.11004407.
14. A. Prelaj et al., "Real-world data to build explainable trustworthy artificial intelligence models for prediction of immunotherapy efficacy in NSCLC patients," *Front. Oncol.*, vol. 12, p. 1078822, Jan. 2023, doi: 10.3389/fonc.2022.1078822.
15. C. Y. Kesiku and B. Garcia-Zapirain, "AI-Enhanced Lung Cancer Prediction: A Hybrid Model's Precision Triumph," *IEEE J. Biomed. Health Inform.*, vol. 29, no. 9, pp. 6287–6300, Sep. 2025, doi: 10.1109/JBHI.2024.3447583.
16. T. I. A. Mohamed and A. E.-S. Ezugwu, "Enhancing Lung Cancer Classification and Prediction With Deep Learning and Multi-Omics Data," *IEEE Access*, vol. 12, pp. 59880–59892, 2024, doi: 10.1109/ACCESS.2024.3394030.
17. H.-J. Kwon et al., "Enhancing Lung Cancer Classification through Integration of Liquid Biopsy Multi-Omics Data with Machine Learning Techniques," *Cancers*, vol. 15, no. 18, p. 4556, Sep. 2023, doi: 10.3390/cancers15184556.
18. P. Azimi and S. Sulaimany, "Effective Preprocessing and Explainability for Better Lung Cancer Prediction," in *2025 29th International Computer Conference, Computer Society of Iran (CSICC)*, Iran, Islamic Republic of: IEEE, Feb. 2025, pp. 1–5. doi: 10.1109/CSICC65765.2025.10967448.
19. F.-H. Tang, Y.-W. Fong, S.-H. Yung, C.-K. Wong, C.-L. Tu, and M.-T. Chan, "Radiomics-Clinical AI Model with Probability Weighted Strategy for Prognosis Prediction in Non-Small Cell Lung Cancer," *Biomedicines*, vol. 11, no. 8, p. 2093, Jul. 2023, doi: 10.3390/biomedicines11082093.
20. A. Fedorov et al., "DICOM re-encoding of volumetrically annotated Lung Imaging Database Consortium (LIDC) nodules," *Medical Physics*, vol. 47, no. 11, pp. 5953–5965, Nov. 2020, doi: 10.1002/mp.14445.
21. M. I. Tusher et al., "Deep Learning Meets Early Diagnosis: A Hybrid CNN-DNN Framework for Lung Cancer Prediction and Clinical Translation," *ijmsphr*, vol. 07, no. 05, pp. 63–72, May 2025, doi: 10.37547/ijmsphr/Volume06Issue05-04.
22. R. Ramachandran, A. A. Jawad, and A. U. J., "Predictive Analysis of Machine Learning and Deep Learning Algorithms for Lung Disease Prediction," in *2025 International Conference on Inventive Computation Technologies (ICICT)*, Kirtipur, Nepal: IEEE, Apr. 2025, pp. 562–568. doi: 10.1109/ICICT64420.2025.11004988.
23. F. Shariaty, V. A. Pavlov, S. V. Fedyashina, and N. A. Serebrennikov, "Integrating Deep Learning and Explainable AI for Non-Invasive Prediction of EGFR and KRAS Mutations in NSCLC: A Novel Radiogenomic Approach," in *2024 V International Conference on Neural Networks and Neurotechnologies (NeuroNT)*, Saint Petersburg, Russian Federation: IEEE, Jun. 2024, pp. 32–35. doi: 10.1109/NeuroNT62606.2024.10585441.
24. R. Javed, T. Abbas, A. H. Khan, A. Daud, A. Bukhari, and R. Alharbey, "Deep learning for lungs cancer detection: a review," *Artif Intell Rev*, vol. 57, no. 8, p. 197, Jul. 2024, doi: 10.1007/s10462-024-10807-1.
25. X. Lu et al., "Multi-level 3D Densenets for False-positive Reduction in Lung Nodule Detection Based on Chest Computed Tomography," *CMIR*, vol. 16, no. 8, pp. 1004–1021, Oct. 2020, doi: 10.2174/1573405615666191113122840.