



UNIVERSITÀ POLITECNICA DELLE MARCHE  
Repository ISTITUZIONALE

LuCa: A Novel Method for Lung Cancer Delineation

This is the peer reviewed version of the following article:

*Original*

LuCa: A Novel Method for Lung Cancer Delineation / Carletti, M.; Bruschi, G.; Mortada, M. H. D. J.; Burattini, L.; Sbroolini, A.. - In: APPLIED SCIENCES. - ISSN 2076-3417. - 15:22(2025). [10.3390/app152212074]

*Availability:*

This version is available at: 11566/354891 since: 2026-03-30T10:09:29Z

*Publisher:*

*Published*

DOI:10.3390/app152212074

*Terms of use:*

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. The use of copyrighted works requires the consent of the rights' holder (author or publisher). Works made available under a Creative Commons license or a Publisher's custom-made license can be used according to the terms and conditions contained therein. See editor's website for further information and terms and conditions.

This item was downloaded from IRIS Università Politecnica delle Marche (<https://iris.univpm.it>). When citing, please refer to the published version.

(Article begins on next page)

## Article

# LuCa: A Novel Method for Lung Cancer Delineation

Mattia Carletti, Giulia Bruschi, MHD Jafar Mortada , Laura Burattini \*  and Agnese Sbrollini 

Department of Information Engineering, Università Politecnica delle Marche, 60131 Ancona, Italy; matticarletti.2000@gmail.com (M.C.); g.bruschi@pm.univpm.it (G.B.); m.j.mortada@pm.univpm.it (M.J.M.); a.sbrollini@univpm.it (A.S.)

\* Correspondence: l.burattini@univpm.it; Tel.: +39-071-220-4461

## Abstract

Lung cancer remains the leading cause of cancer-related deaths worldwide, with over 2.4 million new diagnoses in 2022. Early diagnosis remains challenging due to the non-specificity of symptoms, often resulting in late-stage detection. Although 2-D and 3-D medical imaging, particularly computed tomography (CT), is widely used for detecting lung cancer, it is associated with manual segmentation, which remains time-consuming and user-dependent. This study proposes LuCa as an innovative 2.5-D deep learning model for lung cancer delineation, which combines the benefits of 2-D segmentation with 3-D volume delineation. The main novelty of LuCa is focused on its pipeline, specifically designed to be of clinical use, in order to guarantee the usability of the method. LuCa employs a U-Net architecture for segmentation, followed by a post-image-processing step for 3-D tumor volume delineation and false-positive correction. The method was trained and evaluated using the “NSCLC-Radiomics” database, comprising CT images of 422 non-small cell lung cancer patients, with clinical manual tumor annotations as ground truth. The model achieved strong performance, with high dice coefficients ( $87 \pm 12\%$ ), intersection over union ( $81 \pm 17\%$ ), sensitivity ( $84 \pm 16\%$ ), and positive predictive value ( $94 \pm 10\%$ ) on the test set. Performance was particularly high for larger tumors, reflecting the ability of the model to delineate more visible lesions accurately. Statistical analysis confirmed the high correlation and minimal error between predicted and ground truth tumor volumes. The results highlight the potential of the 2.5-D approach to improve clinical efficiency by enabling accurate tumor segmentation with reduced computational cost, compared to traditional 3-D methods. Future research will focus on assessing the use of LuCa as real-time clinical decision support, particularly for assessing tumors during treatment.

**Keywords:** deep learning; lung cancer delineation; U-Net; 3-D segmentation



Academic Editors: Pedro Couto and Mehmed M. Kantardzic

Received: 15 September 2025

Revised: 3 November 2025

Accepted: 11 November 2025

Published: 13 November 2025

**Citation:** Carletti, M.; Bruschi, G.; Mortada, M.J.; Burattini, L.; Sbrollini, A. LuCa: A Novel Method for Lung Cancer Delineation. *Appl. Sci.* **2025**, *15*, 12074. <https://doi.org/10.3390/app152212074>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Nowadays, lung cancer is the leading cause of global cancer-related mortality [1]. According to the latest GLOBOCAN estimates 2,480,675 new cases of lung cancer were diagnosed globally in 2022, making it the most frequently diagnosed cancer (12.4% of all cancers globally) [1]. Lung cancers arise through a multi-step process involving the development of multiple genetic and epigenetic alterations [2]. Various factors may contribute to an increase in the risk of developing lung cancer and they could be divided into two macro-categories: non-modifiable risk factors (such as age, gender, ethnicity, and family history) and modifiable risk factors (such as tobacco and cannabis smoking, diet, ionizing radiation, and air pollution) [3,4]. According to the histology of the cells, lung cancer is categorized into two classes, which are non-small cell lung cancer (NSCLC) and

small cell lung cancer (SCLC). NSCLC is the most common type (around 85% of cases) and includes subtypes such as adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. Diagnosis of lung cancer is still not efficient. Over half of lung cancer cases are diagnosed in the last stage. Due to the non-specificity of most of the associated symptoms, indeed, patients may present non-specific systemic symptoms, such as fatigue, anorexia, and weight loss, or direct signs and symptoms caused by the primary tumor, such as chest discomfort, cough, dyspnea, and hemoptysis [5].

The most common diagnostic method for lung cancer is 2-D and 3-D medical imaging, which is the representation of an internal anatomical structure or its functional processes. Computed Tomography (CT) is the imaging technique with the highest sensitivity for the detection of pulmonary nodules [6]. CT assists in finding abnormalities, highlights signs of disease, monitors the response to treatment, and supports the planning of therapy [7]; indeed, increased use of CT has improved the identification of small peripheral nodules [8]. However, clinicians must deal with the reading and interpretation of complex imaging data, where factors such as anatomical structures overlaying the region of interest and poor image quality often obscure the identification of malignant lesions. Then, the segmentation of organs or anatomical structures is still manually performed, causing low annotation quality and high inter-observer variability. This variety does not allow an objective clinical practice, resulting in a high level of false positives and false negatives, especially in the tumor volume characterization. Moreover, this practice is time-consuming and non-objective [9].

Recently, deep learning was introduced as a technique to perform lung cancer segmentation [9–11]. Among the many deep-learning tools, the Convolutional Neural Networks (CNNs) reached good results in image segmentation tasks [12]. Various architectures have been explored to address the complexities of image segmentation, each offering unique solutions while balancing computational costs and segmentation accuracy. Image segmentation through CNN could be performed through 2-D analysis and 3-D analysis. Efficient 2-D architecture has been developed to maintain high segmentation quality with lower resource requirements. The Bi-FPN network integrated bidirectional feature pyramid fusion, achieving accurate results while minimizing computational load [13]. WEU-Net introduced a weight excitation mechanism, emphasizing nodular relevant features and achieving robust contextual learning and better segmentation accuracy [14]. Then, the introduction of dense connections and new loss functions guaranteed the achievement of better segmentation, providing a 1% improvement in comparison to 2-D approaches [15]. Attention and transformer-based architectures have further advanced segmentation by capturing long-range spatial dependencies. A novel framework combined CNNs with transformers, using deformable self-attention to improve tumor segmentation, achieving good results and facilitating real-time radiotherapy workflows [16]. Finally, a two-step approach adapted separate networks for large and small tumors, improving precision across different tumor sizes [17]. In contrast, other works focused on 3-D CNN architecture for accurate volume segmentation. The 3-D densely connected network reduced parameter load while enhancing segmentation detail, outperforming traditional 3-D U-Nets in accuracy [18]. The MF-3D U-Net adopted multiscale feature fusion with trainable down-sampling, which allowed efficient nodule segmentation in complex cases [19]. Retina U-Net improved detection by segmenting based on anatomical region, reducing false positives in PET/CT scans [20]. Multiple attention U-Net further enhanced segmentation quality by incorporating spatial and channel attention mechanisms, achieving high Dice scores on clinical datasets [21]. Two-stage models have also shown potential in balancing segmentation precision and efficiency. A two-stage U-Net effectively refined tumor delineation by first isolating the global tumor area and then refining the region locally, significantly improving the accuracy of PET/CT data [22]. Another approach used a segmented patch-based

method to reduce inference time significantly by processing relevant patches, achieving efficient segmentation in a fraction of the time [23]. The 3-D U-Net combined with graphcut co-segmentation applied dual PET-CT inputs to refine tumor boundaries in lung cancer patients [24]. While these methods demonstrate high accuracy, including the volume 3-D features in the analysis, they face challenges with computational demand, limiting their practical application. Techniques like weighted sampling and customized loss functions were designed to improve learning in challenging nodule regions, such as nodules attached to the chest wall, but remain demanding in resource-limited environments [25]. To reduce computational effort while maintaining a high level of tumor delineation, in the literature, two studies [26,27] presented 2.5-D approaches: these methods constitute a 2-D deep-learning approach for tumor segmentation (aiming to reduce computational effort) and an additional module for the 3-D volume reconstruction (aiming to reach a good tumor delineation). Despite the innovative idea, these methods achieved good performance but still not competitive (dice scores around 80%).

Thus, the present paper aims to assess a new deep-learning method based on an innovative 2.5-D approach, LuCa, which is designed to be competitive with 2-D methods in terms of required resources, and with 3-D methods in terms of delineation accuracy. The main novelty of LuCa is focused on the definition of a pipeline specifically designed to be integrated clinical practice, in order to guarantee the usability of the method.

## 2. Materials and Methods

### 2.1. Data Description

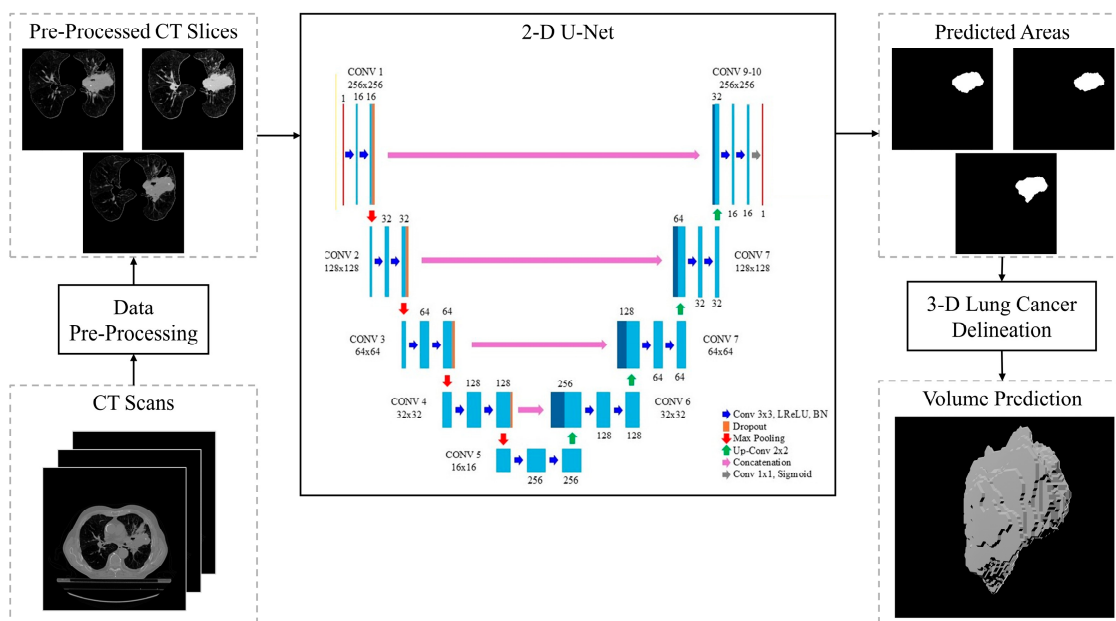
The data used for this project are collected in the “NSCLC-Radiomics dataset”, available online on the “The Cancer Imaging Archive” database [28]. The database contains pretreatment CT scans from 422 NSCLC patients. Patients (290/132 male/female; age:  $68 \pm 10$  years) presented different types of NSCLC at the acquisition time; specifically, 114 patients presented large cell cancer, 152 patients presented squamous cell carcinoma, 51 patients presented adenocarcinoma, 63 patients presented “not otherwise specified” cancer and, finally, 42 patients presented no type of investigation test. Survival time (measured from the start of treatment) ranged between 10 and 4454 days ( $989 \pm 1036$  days). The database also contains the manual annotation of the lungs and of the tumor volume, which was used as ground truth in this study. This delineation was manually performed by a radiation oncologist and consists of the 3-D volume of the primary gross tumor volume (“GTV-1”) and selected anatomical structures (i.e., lung, heart, and esophagus).

Patients with missing files (such as segmentation files) were excluded. Thus, the final population of the study included scans from 391 subjects, with a total of 32,911 pairs of CT slices and masks, which were divided into training, validation, and test datasets. Firstly, the database was split into a training/validation set, which contains CT scans from 312 subjects (80%), and the test set, which contains CT scans from 79 subjects (20%). Then, the training/validation set was split into a training set, which included CT scans from 249 subjects (80% of the training/validation set), and a validation set, which included CT slices from 63 subjects (20% of the training/validation set). The training and validation datasets were used to train the proposed method (the validation dataset was used to implement the early stopping criteria [29]), and the test set was used to evaluate the generalization ability of the entire pipeline. Finally, a data generator was implemented with a batch size of 128, applying a random shuffle of the training and validation datasets.

### 2.2. Proposed Method

The pipeline of LuCa, the proposed 2.5-D deep-learning method for lung cancer delineation, is represented in Figure 1 and described in the following sections. The entire

pipeline has been implemented in Python 3.12.12 Code on the Google Colab environment, using the GPU hardware acceleration and high RAM setups of the PRO version.



**Figure 1.** Block diagram of LuCa, the proposed 2.5-D deep-learning method for lung cancer delineation. The input is the 3-D CT scan, which goes to pre-processing. The pre-processed 2-D CT slices are input to a U-Net, predicting the cancer areas. Finally, an image-processing-based model delineates lung cancer in 3-D, providing the volume prediction.

### 2.2.1. Data Pre-Processing

Data was pre-processed to extract the 2-D slices from each 3-D scan. The pre-processing steps are slice rearrangement, greyscale pixel intensity windowing (range =  $[-1000; +1000]$  HU), lungs extraction by lung segmentation masks, black slices removal, cropping, image resizing to  $256 \times 256$  pixels, JPEG conversion, and CT normalization. Augmentation was not applied.

### 2.2.2. U-Net

Proposed U-Net (Figure 1) consists of an encoder–decoder architecture, proven to be efficient in image processing in capturing information while maintaining good spatial accuracy in a precise segmentation task [30]. The encoder path includes five blocks: each block includes two convolutional layers with an LReLU activation function, followed by a batch normalization block. Convolutional blocks employ a (3, 3) kernel size of (3, 3) with equal padding, spatial dropout set at 0.2 after each block (to help reduce overfitting), and max pooling layers to down-sample the feature maps from  $256 \times 256$  to  $16 \times 16$ . Then, the transposed convolutions mark the beginning of the expansive path, to upsample the feature maps, concatenating them with the corresponding encoder outputs using the skip connections to preserve fine-grained spatial information. The final output layer is a 2-D convolution with a (1, 1) kernel size, followed by a sigmoid activation function. Regarding the training strategy, the Adam optimizer was used with a learning rate of 0.001, and a learning rate scheduler was used to reduce the learning rate by a 0.5 factor in case of stalls in progress. Moreover, the early stopping criterion was used by considering the best validation loss, and restoration of the best weights was implemented to enhance the training performance. Dice Coefficient (DC) was chosen as formulation of the loss function

by considering its sensitivity to class imbalance in binary segmentation. It is defined by Equation (1):

$$DC(PRE, GT) = \frac{2 \cdot |PRE \cap GT|}{|PRE + GT|} \quad (1)$$

where PRE is the predicted set of pixels, and GT is the set of pixels in the ground-truth mask. The loss function was implemented as represented by Equation (2):

$$\text{Loss} = 1 - DC \quad (2)$$

### 2.2.3. Three-Dimensional Lung Cancer Delineation

The 3-D lung cancer delineation was obtained from the 2-D predictions. For each subject, the predictions related to a single CT scan were aligned. Then, a post-processing step was applied to identify and correct the false positive detections. Specifically, this step evaluated the cancer predictions as objects of sequential slices: if a prediction of one slice presented no predictions in the adjacent slices, that prediction was removed; and, if a prediction of one slice had sequential slices with predictions located in a different location, these predictions were considered as not belonging to the same nodule, and the prediction with the smaller area was removed. Finally, all the corrected slices were stacked to construct the 3-D lung cancer, and the volume of the cancer was computed as the sum of the predicted areas multiplied by the thickness between the slices (3 mm in this study).

### 2.3. Statistical Analysis

LuCa has been evaluated by performing a technical evaluation and clinical assessment, focusing on pixel categorization and tumor volume characterization, respectively. Technical evaluation was performed by computing the DC and the intersection over union (IoU) between PRE and GT by using the mathematical formulae defined by Equation (1). Furthermore, sensitivity (SEN) and positive predictive value (PPV) were calculated according to

$$SEN = \frac{TP}{TP + FN} \quad (3)$$

$$PPV = \frac{TP}{TP + FP} \quad (4)$$

where TP are the tumor pixels that were correctly predicted as tumor pixels, FP are the non-tumor pixels that were erroneously predicted as tumor pixels, and FN are the tumor pixels that were erroneously predicted as non-tumor pixels.

Clinical assessment was performed by computing the volumes of tumors of both GT ( $V_{GT}$ ;  $\text{cm}^3$ ) and ( $V_{PRE}$ ;  $\text{cm}^3$ ). Distributions of  $V_{GT}$  and  $V_{PRE}$  volumes were compared by paired *t*-test. Volume similarity was assessed by Bland–Altman analysis:  $\delta(\text{cm}^3)$  and  $\mu(\text{cm}^3)$  are the point-to-point differences and averages between  $V_{GT}$  and  $V_{PRE}$ , respectively. Pearson's correlation analysis assessed volume association:  $\rho$  is the correlation coefficient, and *m* and  $q(\text{cm}^3)$  are the slope and the bias of the linear regression line, respectively. Normal distributions were characterized in terms of mean values (MN) and standard deviation (SD). Statistical significance was set at 0.05.

Both technical evaluation and clinical assessment were performed by considering the entire dataset (overall analysis), the data division sets (training, validation, and test) and the tumor size. Specifically, tumor size was classified as small, medium, or large if its  $V_{GT}$  was lower than  $15 \text{ cm}^3$ , higher than  $15 \text{ cm}^3$  and lower than  $60 \text{ cm}^3$ , or higher than  $60 \text{ cm}^3$ , respectively. Finally, data division and tumor size stratification was merged into subgroups in order to deeply assess the performance of LuCa.

### 3. Results

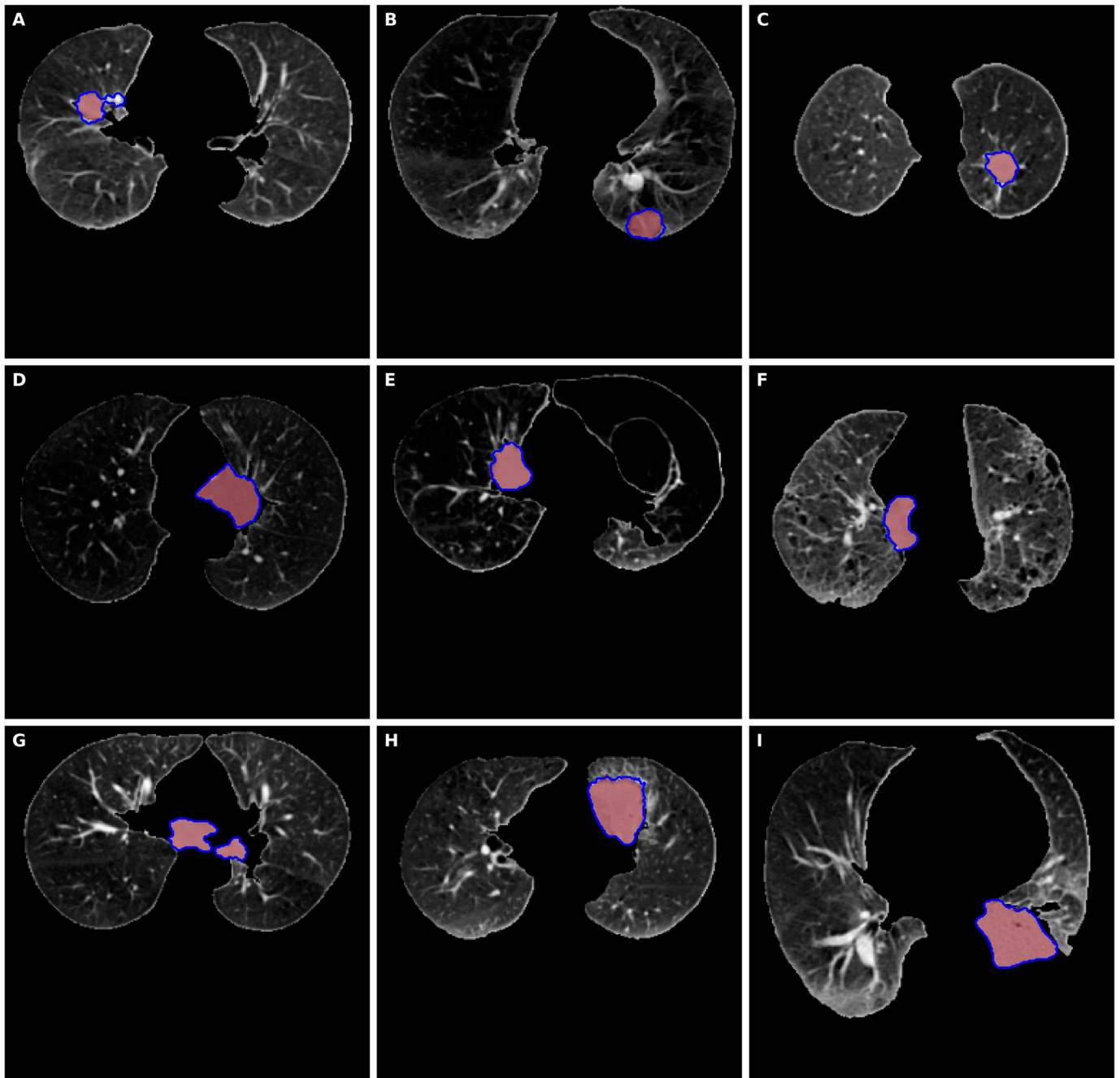
Results of technical evaluation and clinical assessment of LuCa are reported in Table 1.

**Table 1.** Results of the technical evaluation and the clinical assessment of LuCa. Normal distributions are reported in terms of mean value (MN)  $\pm$  standard deviation (SD).

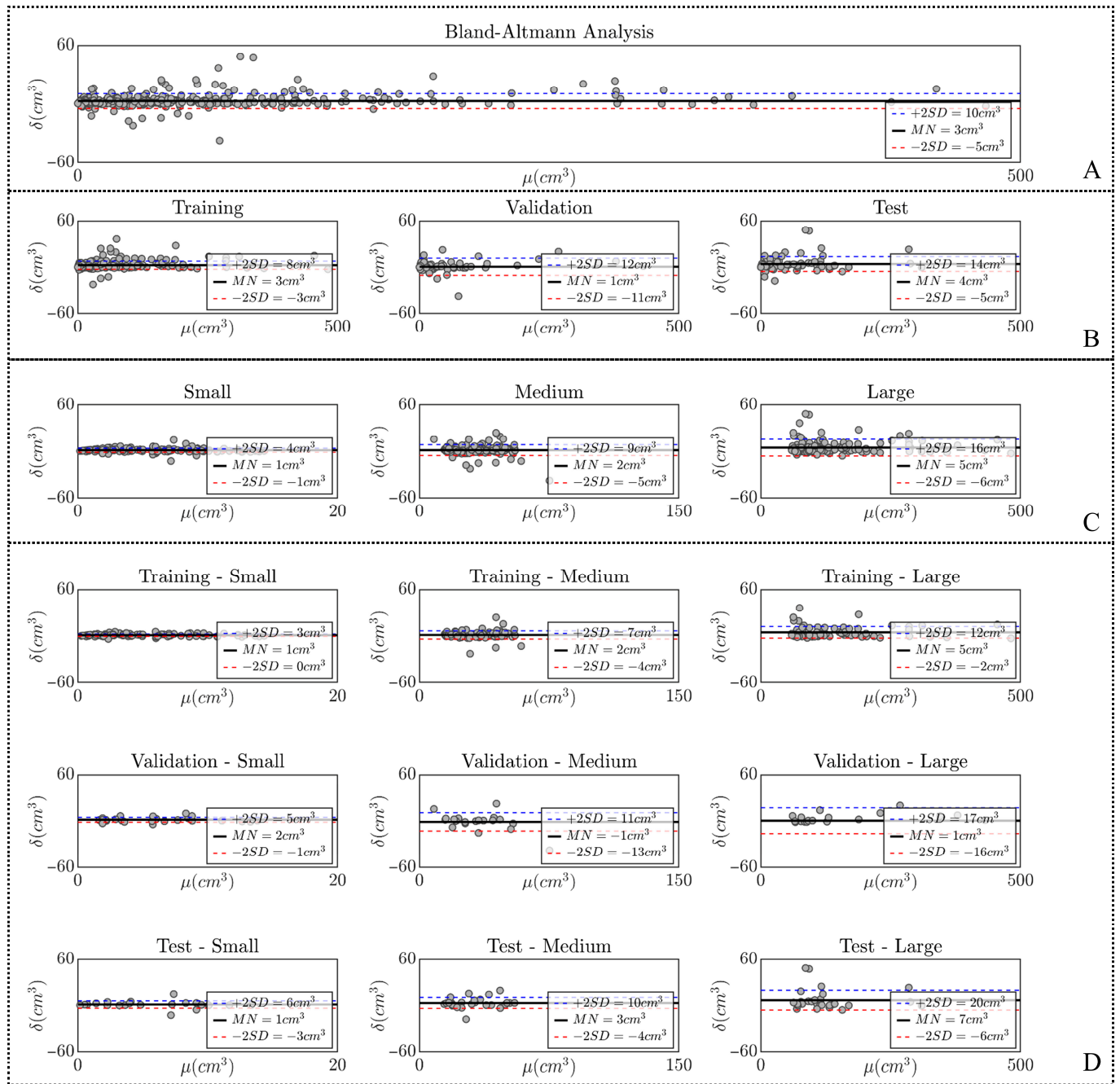
	Stratification	Number	Technical Evaluation				Clinical Assessment		
			DC (%)	IoU (%)	SEN (%)	PPV (%)	V <sub>GT</sub> (cm <sup>3</sup> )	V <sub>PRE</sub> (cm <sup>3</sup> )	
	OVERALL	391	86.8 $\pm$ 14.9	79.3 $\pm$ 19.6	93.7 $\pm$ 9.7	83.6 $\pm$ 18.9	64.2 $\pm$ 79.5	61.5 $\pm$ 78.3	
Data division	TRAINING	249	87.4 $\pm$ 14.2	80.0 $\pm$ 19.0	94.4 $\pm$ 8.0	83.7 $\pm$ 18.7	61.9 $\pm$ 79.4	59.2 $\pm$ 78.2	
	VALIDATION	63	82.6 $\pm$ 19.4	74.2 $\pm$ 23.8	90.2 $\pm$ 14.2	81.2 $\pm$ 23.1	64.6 $\pm$ 93.5	64.1 $\pm$ 92.6	
	TEST	79	88.5 $\pm$ 12.4	81.2 $\pm$ 17.2	94.1 $\pm$ 9.5	85.5 $\pm$ 15.7	71.3 $\pm$ 66.9	66.8 $\pm$ 65.3	
Tumor size	SMALL	135	73.3 $\pm$ 17.5	60.7 $\pm$ 21.1	89.2 $\pm$ 12.0	66.4 $\pm$ 22.9	7.0 $\pm$ 3.7	5.6 $\pm$ 3.9	
	MEDIUM	130	89.4 $\pm$ 10.9	82.1 $\pm$ 14.0	93.3 $\pm$ 9.6	87.6 $\pm$ 12.8	33.8 $\pm$ 13	32.3 $\pm$ 13.8	
	LARGE	126	95.4 $\pm$ 5.7	91.8 $\pm$ 9.0	97.8 $\pm$ 4.6	93.7 $\pm$ 8.0	142.7 $\pm$ 89.9	137.9 $\pm$ 89.6	
Subgroups	TRAIN	SMALL	89	73.5 $\pm$ 17.2	60.9 $\pm$ 20.9	90.3 $\pm$ 10.4	65.7 $\pm$ 22.5	6.6 $\pm$ 3.7	5.2 $\pm$ 3.8
		MEDIUM	85	91.5 $\pm$ 6.7	85.0 $\pm$ 10.3	94.6 $\pm$ 7.4	89.5 $\pm$ 9.2	34.1 $\pm$ 12.6	32.4 $\pm$ 12.4
		LARGE	75	95.8 $\pm$ 4.6	92.4 $\pm$ 7.6	98.2 $\pm$ 1.6	94.0 $\pm$ 7.4	145.0 $\pm$ 92.3	140.1 $\pm$ 92.3
	VALIDATION	SMALL	24	70.2 $\pm$ 18.8	57.1 $\pm$ 22.2	86.4 $\pm$ 15.8	64.8 $\pm$ 25.1	7.5 $\pm$ 4.1	5.9 $\pm$ 4.1
		MEDIUM	20	83.8 $\pm$ 20.3	75.6 $\pm$ 21.7	88.9 $\pm$ 14.1	85.2 $\pm$ 21.6	33.1 $\pm$ 14.2	34.0 $\pm$ 20.1
		LARGE	19	95.2 $\pm$ 8.1	91.6 $\pm$ 11.9	95.7 $\pm$ 11.1	95.2 $\pm$ 4.7	160.8 $\pm$ 117.1	160.2 $\pm$ 113.8
	TEST	SMALL	22	77.0 $\pm$ 17.4	65.3 $\pm$ 21	87.7 $\pm$ 13.0	73.0 $\pm$ 21.6	8.1 $\pm$ 3.2	7.0 $\pm$ 3.9
		MEDIUM	25	86.6 $\pm$ 10.3	77.7 $\pm$ 15.2	92.4 $\pm$ 11.2	83.3 $\pm$ 13.5	33.6 $\pm$ 13.9	30.3 $\pm$ 12.9
		LARGE	32	94.7 $\pm$ 6.5	90.5 $\pm$ 10.3	98.0 $\pm$ 2.3	92.2 $\pm$ 10.5	126.5 $\pm$ 61.9	119.7 $\pm$ 62.5

DC: dice coefficient; GT: ground truth; IoU: intersection over union; LARGE: volumes higher than 60 cm<sup>3</sup>, MEDIUM: volumes higher than 15 cm<sup>3</sup> and lower than 60 cm<sup>3</sup>, PPV: positive predictive value; SEN: sensitivity, SMALL: volumes lower than 15 cm<sup>3</sup>.

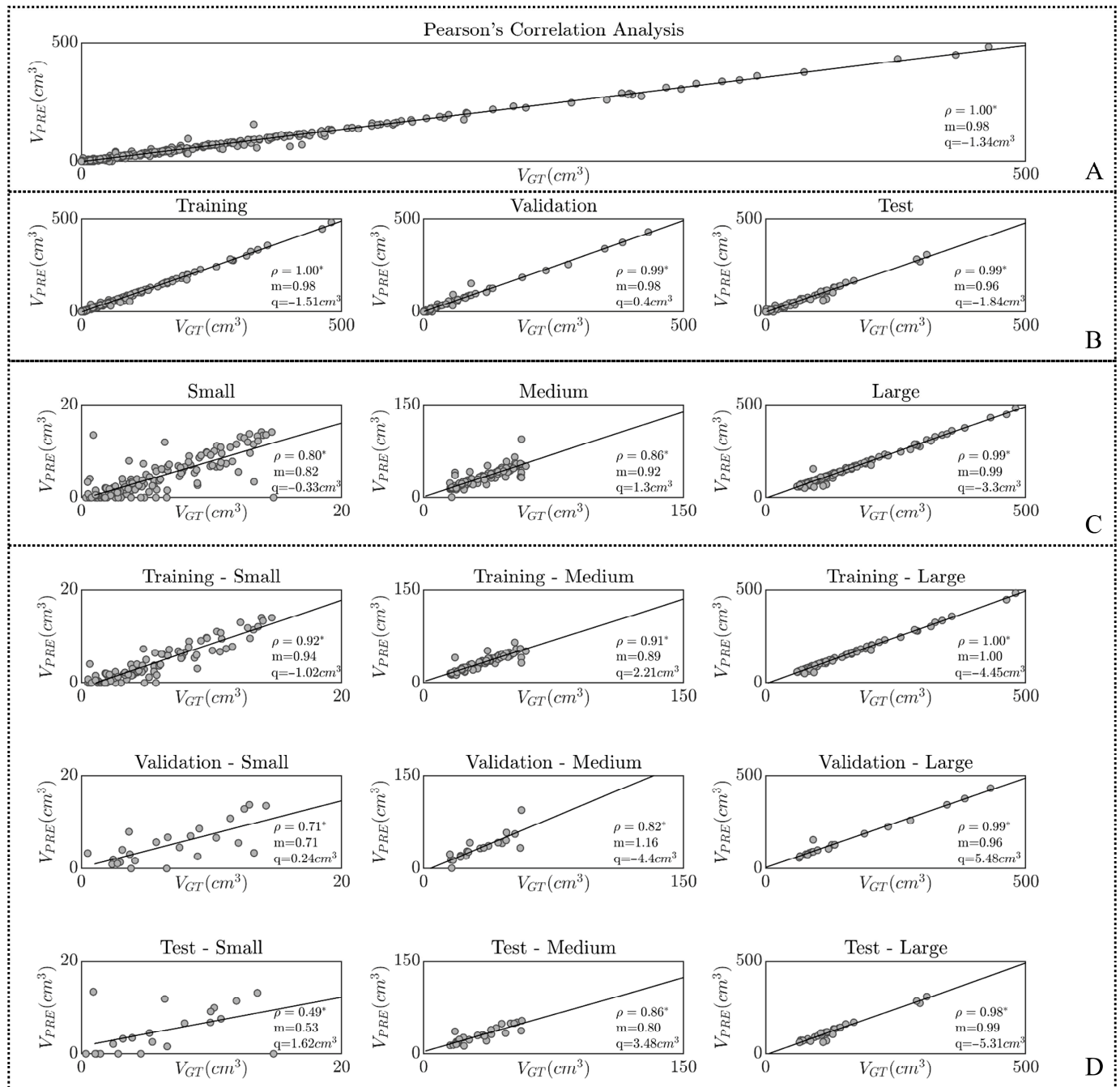
Examples of subgroup-stratified 2D results with ground truth are reported in Figure 2. Technically, testing high values of DC (87  $\pm$  12%), IoU (81  $\pm$  17%), SEN (84  $\pm$  16%), and PPV (94  $\pm$  10%) indicated a very good level of generalization of LuCa; delineation performance increased with tumor size, as expected. Distributions of V<sub>PRE</sub> are not statistically different from V<sub>GT</sub> distributions ( $p$  value higher than 0.85), independently by dataset (training, validation, and test) and by tumor size (small, medium, and large). Results of Bland–Altman and Pearson’s correlation analyses are reported in Figures 3 and 4, showing low errors ( $\delta < 5$  cm<sup>3</sup>) and strong statistically significant agreement ( $\rho > 0.82$ ) between the PRE and GT in all stratifications. Finally, Figure 5 depicts examples of the 3-D volume delineation.



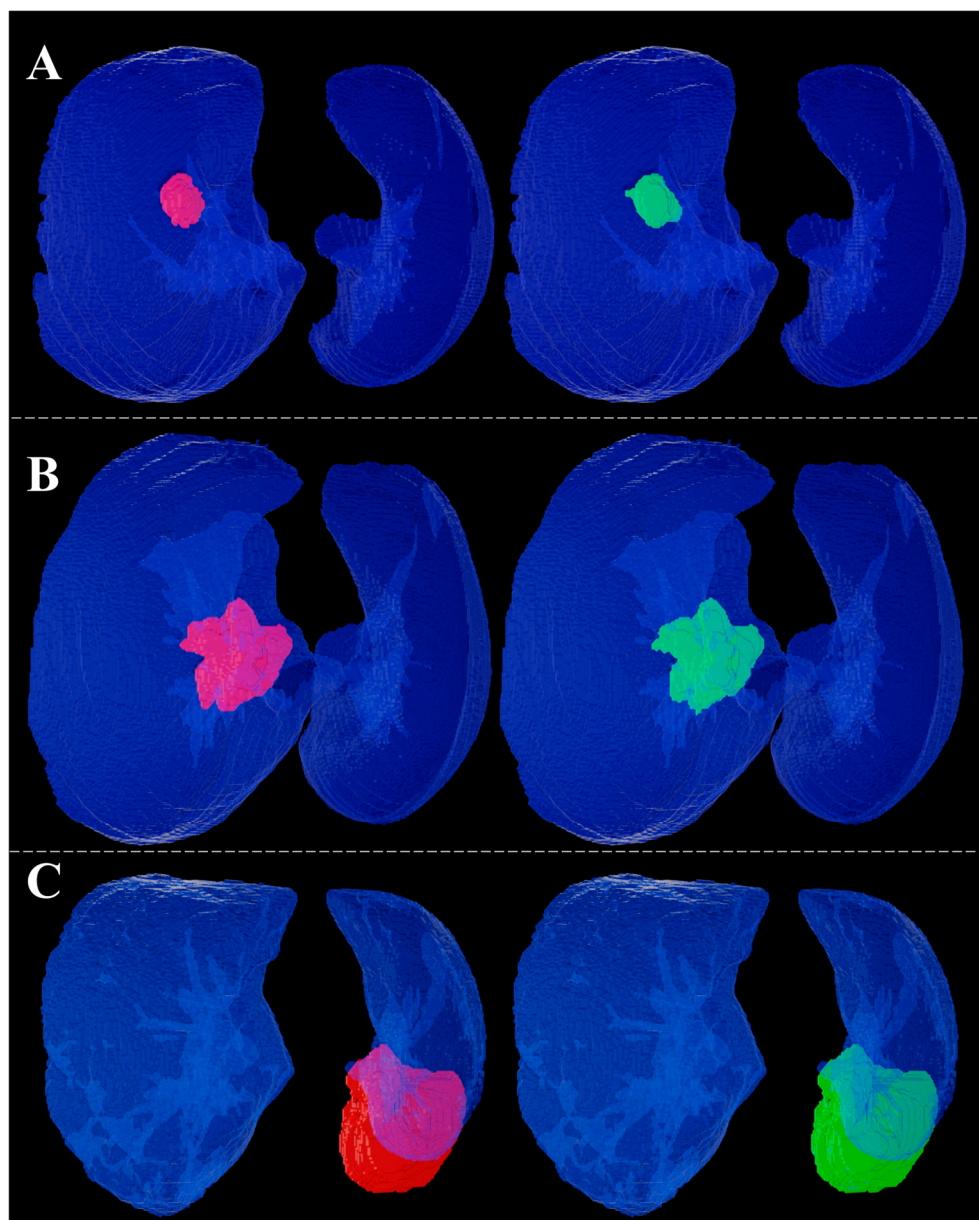
**Figure 2.** Examples of subgroup-stratified 2D prediction (red areas) with ground truth (blue lines). (Panels A–C) represent examples of small tumors obtained from training, validation, and test set, respectively. (Panels D–F) represent examples of medium tumors obtained from training, validation, and test set, respectively. (Panels G–I) represent examples of large tumors obtained from training, validation, and test set, respectively.



**Figure 3.** Bland–Altman analysis performed on the entire dataset (**panel A**), by considering the data division stratification (**panel B**), the tumor size (**panel C**) and the subgroups (**panel D**). The values of mean error (MN) and limits (+2 SD and -2 SD) are reported in the legend.



**Figure 4.** Pearson's correlation analysis performed on the entire dataset (**panel A**) by considering the data division stratification (**panel B**), the tumor size (**panel C**) and the subgroups (**panel D**). The correlation coefficient ( $\rho$ ; \*  $p$ -value lower than 0.05) and the coefficients of the regression line are reported ( $m$ : slope;  $q$ : bias).



**Figure 5.** Examples of the 3-D volume delineation by LuCa: GT 3-D volumes are depicted in red, while the PRE 3-D volumes are depicted in green. (Panels A–C) depict examples of small, medium, and large cancers, respectively.

#### 4. Discussion

The study presents LuCa, a novel deep learning approach for lung cancer segmentation and volume delineation. The main novelty of LuCa is focused on the definition of a pipeline specifically designed to be integrated clinical practice, in order to guarantee the usability of the method. The main concepts for its definition were reliability (i.e., good results in terms of cancer delineation) and usability (i.e., low computational effort to be integrated into the real clinical scenario). The 2.5-D concept is designed to reduce computational effort using a 2-D deep learning model (U-Net) for image segmentation while preserving delineation accuracy through a 3-D image processing tool for cancer volume delineation. U-Net was selected by considering its reliability, transparency, and relatively low computational effort, which is a good tradeoff between computing and clinical use. Finally, the trained model can be exported into Open Neural Network Exchange (ONNX) format, which can be executed

within secure hospital cloud, using runtime available environments (e.g., ONNX Runtime), easily usable in clinical settings, such as PACS system.

As a common technical choice, data manipulation procedures are used to increase the study population and assess the generalizability of the model (e.g., data augmentation). To ensure the population remains representative of the clinical problem, a static data division was applied without pre-processing the population by data augmentation procedures. Despite the fact that in computer science, augmentation is a well-established technique to improve generalization and robustness, in clinical settings this technique can lead to misinterpretation due to the changes in clinical realism of CT data and to the biases in volumetric estimations. Preprocessing and postprocessing pipelines were designed to reduce storage and computational load, essentials in the clinical scenario. Under these conditions, LuCa has been demonstrated effective in lung tumor delineation, providing high statistical metrics (DC = 87%; IoU = 81%; SEN = 84%; PPV = 94%) in the testing dataset, and thus, confirming its ability to generalize the clinical problem of interest.

Clinically, accurate tumor volume delineation is crucial in each stage of oncological patient management, from diagnosis to prognosis, as well as treatment planning. LuCa demonstrated a strong similarity/association between predicted and ground-truth tumor volumes, with no statistically significant differences across datasets (training, validation, and test) or tumor size groups (small, medium, and large). Despite the statistical significance, the performance of LuCa improves with the cancer size. This result is expected according to two main factors. First of all, large tumors are more visible and easily detectable; thus, performance on large anatomical structures is an easier task than detecting small structures in all image processing fields. Secondly, the effect of false positive/negative detection in small structures affects not only automatic analysis but also visual inspection-based procedures, such as those used as ground truth in this study. Considering that the deep-learning methods, such as LuCa, are strongly affected by the gold standard being supervised techniques, the low visibility of small tumors can be inferred by issues related to both visual inspection and automatic tools. Finally, the literature stated that AI-based tools in lung cancer screening may improve sensitivity but increase the number of false-positive results [31]. Thus, our technical method was designed to balance both sensitivity and specificity, dealing with the risk of (not statistically) underestimating small tumors.

In the literature, some other algorithms were proposed to perform lung cancer delineation. Five studies presented 2-D approaches [11–15], providing good results (dice coefficient range: 74–92%) in terms of technical metrics, but the 2-D nature makes them not suitable for lung cancer volumes. On the other hand, eight studies presented 3-D end-to-end approaches [16–23] that were able to reach very high technical results (dice coefficient range: 78–97%) and very good clinical performance in terms of cancer volume delineation; but, these models are not integrable in standard clinical practice considering the need of high computational effort to be applied. Only the other two studies [24,25] presented 2.5-D approaches, similar to LuCa, the one presented in this paper. The results presented by [24,25] confirmed the effectiveness of the 2.5-D hypothesis, but, despite their good technical results (dice coefficient of 75% and 82%, respectively), LuCa provided superior results (dice coefficient of 88%). Thus, the qualitative comparison of LuCa with already published studies in the field confirmed that LuCa is a very good trade-off between clinical feature extraction and computational effort. Unfortunately, a quantitative comparison of LuCa with other methods in the literature is not possible because most of the studies used private datasets, different pipelines, and evaluation criteria, making the comparison not reliable.

Despite its promising performance, the study has several limitations that should be addressed in future work. The multicenter “NSCLC-Radiomics” database is a valid database

for evaluating the performance of a deep-learning model for lung tumor segmentation, guaranteeing reproducibility due to its open access. Despite that, it may not represent the full diversity of clinical cases. For example, most of the data (88%) belongs to dead subjects; thus, an evaluation of the utility of the method as a decision support system cannot be made. Thus, variability in data acquisition settings, image quality, and patient demographics should be validated to assess the generalizability of the model in every clinical scenario. In the future, merging diverse imaging modalities and datasets from multiple centers will help to evaluate the dependency of LuCa on the technical and demographic features of the oncological population. Moreover, LuCa needs ground truth to be trained, such as other supervised procedures. This property remains an important technical issue, considering that it does not allow us to compare the performance of the models with the clinical opinion, which is usually used as ground truth. Thus, future studies will aim to evaluate the clinical use of our deep-learning method by considering several clinical features, different tumor types and by varying its main parameters; moreover, defining unsupervised methodologies independent of manual annotation will be investigated. Finally, despite the superiority of the 2.5-D to 3-D approach in terms of computational effort, LuCa still needs advanced resources to be trained. Thus, exploring hybrid models (including attention mechanisms or transformer-based modules) could reduce the computational cost and, eventually, improve segmentation performance, guaranteeing LuCa accessibility in any kind of environment, especially in low-resource settings.

## 5. Conclusions

The main novelty of LuCa is focused on the definition of a pipeline specifically designed to be integrated clinical practice, in order to guarantee the usability of the method. It proved to be effective in delineating lung cancer volumes by using a 2.5-D approach, offering a computationally efficient and clinically reliable alternative to end-to-end 2-D and 3-D models. Future work will focus on multi-center testing analysis, trying to reproduce the quantitative results published in the literature.

**Author Contributions:** Conceptualization, M.C. and A.S.; methodology, M.C., M.J.M. and A.S.; software, M.C., G.B. and M.J.M.; validation, M.C., M.J.M. and A.S.; formal analysis, M.C. and G.B.; investigation, M.C. and G.B.; resources, L.B.; data curation, G.B.; writing—original draft preparation, M.C. and A.S.; writing—review and editing, G.B., M.J.M. and L.B.; visualization, A.S.; supervision, L.B. and A.S.; project administration, L.B.; funding acquisition, L.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data available in a publicly accessible repository openly available in “NSCLC-Radiomics dataset” on the “The Cancer Imaging Archive” database [<https://www.cancerimagingarchive.net/>] (accessed on 10 November 2025).

**Acknowledgments:** The authors would like to thank Dr. Donatella Di Fabrizio ([https://portale.ospedaliriuniti.marche.it/archivio10\\_personale-clinico-ed-amministrativo\\_0\\_2098.html](https://portale.ospedaliriuniti.marche.it/archivio10_personale-clinico-ed-amministrativo_0_2098.html), accessed on 10 November 2025) for her valuable contribution in reviewing the manuscript. She provided insightful clinical feedback that improved the quality of this work and confirmed the clinical use of LuCa.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Bray, F.; Laversanne, M.; Sung, H.; Ferlay, J.; Siegel, R.L.; Soerjomataram, I.; Jemal, A. Global Cancer Statistics 2022: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2024**, *74*, 229–263. [[CrossRef](#)] [[PubMed](#)]
2. Cooper, W.A.; Lam, D.C.L.; O'Toole, S.A.; Minna, J.D. Molecular Biology of Lung Cancer. *J. Thorac. Dis.* **2013**, *5*, S479–S490. [[CrossRef](#)] [[PubMed](#)]
3. Thandra, K.C.; Barsouk, A.; Saginala, K.; Aluru, J.S.; Barsouk, A. Epidemiology of Lung Cancer. *Contemp. Oncol.* **2021**, *25*, 45–52. [[CrossRef](#)]
4. Malhotra, J.; Malvezzi, M.; Negri, E.; La Vecchia, C.; Boffetta, P. Risk Factors for Lung Cancer Worldwide. *Eur. Respir. J.* **2016**, *48*, 889–902. [[CrossRef](#)]
5. Collins, L.G.; Haines, C.; Perkel, R.; Enck, R.E. Lung Cancer: Diagnosis and Management. *Am. Fam. Physician* **2007**, *75*, 56–63.
6. De Wever, W.; Coolen, J.; Verschakelen, J. Imaging Techniques in Lung Cancer. *ERS J.* **2011**, *7*, 338–346. [[CrossRef](#)]
7. Panunzio, A.; Sartori, P. Lung Cancer and Radiological Imaging. *Curr. Radiopharm.* **2020**, *13*, 238–242. [[CrossRef](#)]
8. *The International Agency for Research on Cancer Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart (IARC WHO Classification of Tumours)*, 1st ed.; World Health Organization: Lyon, France, 2004.
9. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
10. Song, Y.; Liu, Y.; Lin, Z.; Zhou, J.; Li, D.; Zhou, T.; Leung, M.-F. Learning from AI-Generated Annotations for Medical Image Segmentation. *IEEE Trans. Consum. Electron.* **2024**, *71*, 1473–1481. [[CrossRef](#)]
11. Li, Y.; Hao, W.; Zeng, H.; Wang, L.; Xu, J.; Routray, S.; Jhaveri, R.H.; Gadekallu, T.R. Cross-Scale Texture Supplementation for Reference-Based Medical Image Super-Resolution. *IEEE J. Biomed. Health Inform.* **2025**, 1–15. [[CrossRef](#)] [[PubMed](#)]
12. Yamashita, R.; Nishio, M.; Do, R.K.G.; Togashi, K. Convolutional Neural Networks: An Overview and Application in Radiology. *Insights Imaging* **2018**, *9*, 611–629. [[CrossRef](#)] [[PubMed](#)]
13. Annavarapu, C.S.R.; Parisapogu, S.A.B.; Keetha, N.V.; Donta, P.K.; Rajita, G. A Bi-FPN-Based Encoder-Decoder Model for Lung Nodule Image Segmentation. *Diagnostics* **2023**, *13*, 1406. [[CrossRef](#)]
14. Banu, S.F.; Sarker, M.M.K.; Abdel Nasser, M.; Rashwan, H.; Puig, D. WEU-Net: A Weight Excitation U-Net for Lung Nodule Segmentation. *Appl. Sci.* **2021**, *101*, 349–356. [[CrossRef](#)]
15. Lu, D.; Chu, J.; Zhao, R.; Zhang, Y.; Tian, G. A Novel Deep Learning Network and Its Application for Pulmonary Nodule Segmentation. *Comput. Intell. Neurosci.* **2022**, *2022*, 1–6. [[CrossRef](#)]
16. Kunkyab, T.; Bahrami, Z.; Zhang, H.; Liu, Z.; Hyde, D. A Deep Learning-Based Framework (Co-Retr) for Auto-Segmentation of Non-Small Cell Lung Cancer in Computed Tomography Images. *J. Appl. Clin. Med. Phys.* **2024**, *25*, e14297. [[CrossRef](#)]
17. Zhang, F.; Wang, Q.; Fan, E.; Lu, N.; Chen, D.; Jiang, H.; Yu, Y. Enhancing Non-Small Cell Lung Cancer Tumor Segmentation with a Novel Two-Step Deep Learning Approach. *J. Radiat. Res. Appl. Sci.* **2024**, *17*, 100775. [[CrossRef](#)]
18. Zhao, L. 3D Densely Connected Convolution Neural Networks for Pulmonary Parenchyma Segmentation from CT Images. *J. Phys. Conf. Ser.* **2020**, *1631*, 12049. [[CrossRef](#)]
19. Agnes, S.A.; Anitha, J. Efficient Multiscale Fully Convolutional UNet Model for Segmentation of 3D Lung Nodule from CT Image. *J. Med. Imaging* **2022**, *9*, 052402. [[CrossRef](#)]
20. Weikert, T.; Jaeger, P.F.; Yang, S.; Baumgartner, M.; Breit, H.C.; Winkel, D.J.; Sommer, G.; Stieltjes, B.; Thaiss, W.; Bremerich, J.; et al. Automated Lung Cancer Assessment on 18F-PET/CT Using Retina U-Net and Anatomical Region Segmentation. *Eur. Radiol.* **2023**, *33*, 4270–4279. [[CrossRef](#)]
21. Chen, W.; Yang, F.; Zhang, X.; Xu, X.; Qiao, X. MAU-Net: Multiple Attention 3D U-Net for Lung Cancer Segmentation on CT Images. *Procedia Comput. Sci.* **2021**, *192*, 543–552. [[CrossRef](#)]
22. Park, J.; Kang, S.K.; Hwang, D.; Choi, H.; Ha, S.; Seo, J.M.; Eo, J.S.; Lee, J.S. Automatic Lung Cancer Segmentation in [18F]FDG PET/CT Using a Two-Stage Deep Learning Approach. *Nucl. Med. Mol. Imaging* **2023**, *57*, 86–93. [[CrossRef](#)]
23. Shirokikh, B.; Shevtsov, A.; Dalechina, A.; Krivov, E.; Kostjuchenko, V.; Golanov, A.; Gombolevskiy, V.; Morozov, S.; Belyaev, M. Accelerating 3D Medical Image Segmentation by Adaptive Small-Scale Target Localization. *J. Imaging* **2021**, *7*, 35. [[CrossRef](#)]
24. Zhong, Z.; Kim, Y.; Zhou, L.; Plichta, K.; Allen, B.; Buatti, J.; Wu, X. 3D Fully Convolutional Networks for Co-Segmentation of Tumors on PET-CT Images. In Proceedings of the IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; pp. 228–231.
25. Kido, S.; Kidera, S.; Hirano, Y.; Mabu, S.; Kamiya, T.; Tanaka, N.; Suzuki, Y.; Yanagawa, M.; Tomiyama, N. Segmentation of Lung Nodules on CT Images Using a Nested Three-Dimensional Fully Connected Convolutional Network. *Front. Artif. Intell.* **2022**, *5*, 782225. [[CrossRef](#)] [[PubMed](#)]
26. Wang, Y.; Zhou, C.; Chan, H.-P.; Hadjiiski, L.M.; Chughtai, A.; Kazerooni, E.A. Hybrid U-Net-Based Deep Learning Model for Volume Segmentation of Lung Nodules in CT Images. *Med. Phys.* **2022**, *49*, 7287–7302. [[CrossRef](#)] [[PubMed](#)]
27. Wang, S.; Zhou, M.; Liu, Z.; Liu, Z.; Gu, D.; Zang, Y.; Dong, D.; Gevaert, O.; Tian, J. Central Focused Convolutional Neural Networks: Developing a Data-Driven Model for Lung Nodule Segmentation. *Med. Image Anal.* **2017**, *40*, 172–183. [[CrossRef](#)]

28. The Cancer Imaging Archive NSCLC-Radiomics. Available online: <https://www.cancerimagingarchive.net/> (accessed on 10 November 2025).
29. Aggarwal, C.C. *Neural Networks and Deep Learning*; Springer: Cham, Switzerland, 2023.
30. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Springer: Cham, Switzerland, 2015; Volume 9351, pp. 234–241.
31. Geppert, J.; Asgharzadeh, A.; Brown, A.; Stinton, C.; Helm, E.J.; Jayakody, S.; Todkill, D.; Gallacher, D.; Ghiasvand, H.; Patel, M.; et al. Software Using Artificial Intelligence for Nodule and Cancer Detection in CT Lung Cancer Screening: Systematic Review of Test Accuracy Studies. *Thorax* **2024**, *79*, 1040–1049. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.