

Article

Estimating the Volume of Nodules and Masses on Serial Chest Radiography Using a Deep-Learning-Based Automatic Detection Algorithm: A Preliminary Study

Chae Young Lim ¹, Yoon Ki Cha ^{1,*}, Myung Jin Chung ¹, Subin Park ², Soyoung Park ², Jung Han Woo ¹ and Jong Hee Kim ¹

¹ Department of Radiology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, Republic of Korea; star05890@gmail.com (C.Y.L.); mj1.chung@samsung.com (M.J.C.)

² Department of Health Sciences and Technology, SAIHST, Sungkyunkwan University, Suwon 16419, Republic of Korea

* Correspondence: yoonki.cha@samsung.com; Tel.: +82-10-9049-6726

Abstract: Background: The purpose of this study was to assess the volume of the pulmonary nodules and masses on serial chest X-rays (CXRs) from deep-learning-based automatic detection algorithm (DLAD)-based parameters. Methods: In a retrospective single-institutional study, 72 patients, who obtained serial CXRs ($n = 147$) for pulmonary nodules or masses with corresponding chest CT images as the reference standards, were included. A pre-trained DLAD based on a convolutional neural network was developed to detect and localize nodules using 13,710 radiographs and to calculate a localization map and the derived parameters (e.g., the area and mean probability value of pulmonary nodules) for each CXR, including serial follow-ups. For validation, reference 3D CT volumes were measured semi-automatically. Volume prediction models for pulmonary nodules were established through univariable or multivariable, and linear or non-linear regression analyses with the parameters. A polynomial regression analysis was performed as a method of a non-linear regression model. Results: Of the 147 CXRs and 208 nodules of 72 patients, the mean volume of nodules or masses was measured as $9.37 \pm 11.69 \text{ cm}^3$ (mean $\pm$ standard deviation). The area and CT volume demonstrated a linear correlation of moderate strength (i.e., $R = 0.58$, RMSE: $9449.9 \text{ mm}^3 \text{ m}^3$ in a linear regression analysis). The area and mean probability values exhibited a strong linear correlation ($R = 0.73$). The volume prediction performance based on a multivariable regression model was best with a mean probability and unit-adjusted area (i.e., RMSE: 7975.6 mm^3 , the smallest among the other variable parameters). Conclusions: The prediction model with the area and the mean probability based on the DLAD showed a rather accurate quantitative estimation of pulmonary nodule or mass volume and the change in serial CXRs.



Citation: Lim, C.Y.; Cha, Y.K.; Chung, M.J.; Park, S.; Park, S.; Woo, J.H.; Kim, J.H. Estimating the Volume of Nodules and Masses on Serial Chest Radiography Using a Deep-Learning-Based Automatic Detection Algorithm: A Preliminary Study. *Diagnostics* **2023**, *13*, 2060. <https://doi.org/10.3390/diagnostics13122060>

Academic Editor: Thomas Frauenfelder

Received: 12 May 2023

Revised: 9 June 2023

Accepted: 10 June 2023

Published: 14 June 2023



Copyright: © 2023 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/>).

Keywords: pulmonary nodules; chest X-rays; deep-learning; convolutional neural networks; volume estimation

1. Introduction

In the last decade, deep-learning-based automatic detection (DLAD) algorithms have significantly surpassed conventional computer-aided detection (CAD) systems in a variety of tasks related to chest radiographs (CXR) [1–4]. These advanced algorithms have seen an increased adoption in clinical settings for detecting and classifying a wide range of abnormal pulmonary lesions such as nodules, masses, and COVID-19 pneumonia [5–9]. They provide invaluable assistance to physicians, particularly in cases where radiologists are unavailable or overburdened [6,7].

They also enhance the radiologists' performances when used as a support [10–13]. In our study utilizing DLAD, the average sensitivity of radiologists improved from 65.1% to 70.3%, and the number of false-positive findings per radiograph decreased from 0.2 to

0.18 when the radiologists re-reviewed the radiographs with the aid of the DLAD software. Among the 12 radiologists, 104 of 2400 radiographs experienced positive changes (i.e., from a false-negative to a true-positive or from a false-positive to a true-negative) using the DLAD, while 56 of 2400 radiographs showed negative changes.

Despite their impressive capabilities, DLAD algorithms currently lack the ability to compare serial CXRs for assessing lesion size or severity, which is a crucial task performed by radiologists daily.

Convolutional neural networks (CNNs), a specialized form of deep-learning, have piqued interest among the researchers and developers of chest radiography algorithms [14]. Utilizing class activation maps (CAM) in conjunction with CNN models can significantly enhance the image interpretability and localization [15,16]. The CAM computes the probability function for a specific location (x, y) on an input image, generating probability maps that can help pinpoint key areas in the image [16].

The research hypothesis states that the higher the probability of a lesion's existence, the more distinct and denser the lesion appears, potentially indicating a larger projected volume for the lesion. To test this hypothesis, regions of interest (ROIs) with probabilities exceeding a certain threshold were identified as abnormal on the probability map, and the area of the ROI was estimated for each image. By combining this area and probability value, which was calculated using the DLAD classification algorithm, researchers investigated the potential for estimating the volume of the target lesion.

The goal of this research is to apply these findings to the comparison of serial CXRs, allowing for a more comprehensive analysis of the lesion size and severity over time. This advancement could provide physicians with an additional layer of information, thus further enhancing their ability to diagnose and treat various pulmonary conditions. There has been no previous attempt, as shown in this study, to use DLAD for estimating the 3D volume from 2D chest X-rays and to utilize this for assessing changes in lesions. This novel approach may yield academic contributions to the field of deep-learning in medical imaging.

2. Materials and Methods

This study was approved by our institutional review board (IRB No. SMC 2021-07-075-002) and individual consent for this retrospective analysis was waived.

A pilot study was conducted to derive the parameters from the DLAD of the researcher's institution, thereby reflecting the 3 kinds of interval changes (i.e., increased, decreased, and no change) of nodules or masses in the CXRs. The reference standards were the formal readings of each CXR.

To quantitatively analyze the interval change of the nodules or masses, the researchers measured the 3D CT volume of nodules and used this value as a reference standard to perform a linear and non-linear correlation analysis with parameters derived from the pilot study and established prediction models.

2.1. Patients

We retrospectively searched the CXR database of a single, large academic institution (Samsung Medical Center, Seoul, Republic of Korea) from September 2018 to December 2020 using the combination of search terms of "increase-", "decrease-", "grow-", "mass-", "nodule-", and "solitary pulmonary nodule-", which resulted in the identification of 383 patients with interval-changed nodules and masses. Among them, 72 patients with follow-up images with the X-ray scanners of the same manufacturers as the previous images were selected. In the pilot study, when predicting the interval change with the probability map of images taken by scanners from different manufacturers, the accuracy was identified as low.

The ALND has been trained for nodules or masses with a diameter of 4.5 cm or less. Patients with a mass of 4.5 cm or larger were excluded during the initial stage or follow-up scans. Among them, 117 patients with corresponding chest CT scans of both the initial and follow-up radiographs, obtained before or after 7 days from the day of each CXR, were

included. Lastly, we reviewed the PACS images and included 72 patients and a total of 147 chest radiographs with the inclusion and exclusion criteria as described in Figure 1.

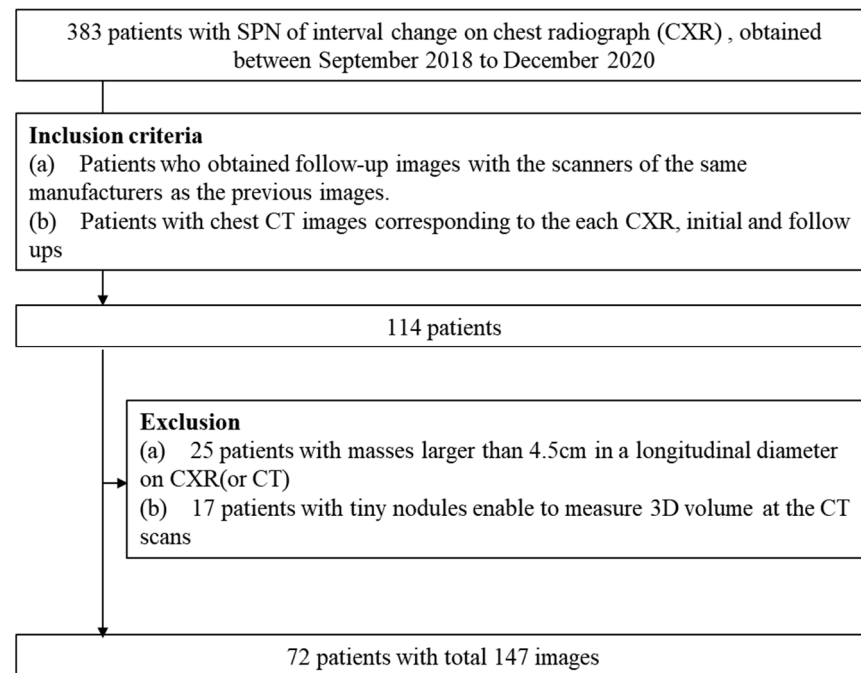


Figure 1. Study flowcharts.

2.2. CAD System for CXR Interpretation

The Samsung Auto Lung Nodule Detection (ALND, version 1.00; Samsung Electronics, Suwon, South Korea) was used in the present study. This deep convolutional neural network (DCNN) was built upon a modified version of ResNet-50 [17], and trained to detect lung nodules on digital radiographs. The training set was 13,710 normal radiographs and 3500 lung nodule (benign or malignant)-containing radiographs, while a test/validation set was 640 normal radiographs and 1480 lung-nodule radiographs (a total of 5700 lung nodules; size: $7 \text{ mm} \leq x \leq 45 \text{ mm}$) [4]. The software was commercially available in Europe and South Korea as of June 2019. For the results, a probability score of 15% was defined as the threshold for a binary classification between the negative and positive results.

The trained ALND generated continuous probability values between 0 and 1 for each chest radiograph, corresponding to the probability that there would be nodules on the chest radiograph. In addition, the DLAD produced continuous activation values ranging from 0 to 1 for each pixel of the images, derived from the class activation function of the neural network [2]. With the cutoff of a 0.2 probability value, we obtained a CAM for each CXR and calculated the preselected parameters from the regions of interest (ROIs), with up to 3 for each image. The five preselected parameters obtained from the pilot study were the area, weighted opacity-area product (WOAP), mean, median, and maximum probability value of each ROI.

2.3. Image Analysis and Definition of Reference Standards with Volume Measurements

Two radiologists (Y.K.C., a thoracic radiologist with 16 years of experience, and M.J.C., a thoracic radiologist with 30 years of experience) reviewed all chest radiographs by consensus. The researchers selected up to three index pulmonary nodules that met the definition of the Fleischner Society glossary of terms and evaluated the interval change of each lesion [18]. At the serial follow-up CXRs, interval changes were classified as an increase or decrease in the extent of nodules and masses. For evaluating the volumes of indexed nodules, we measured the nodule volume semi-automatically with the CT scanner-linked Aquarius iNtuition[®] software (TeraRecon, Houston, TX, USA) for the nodules detected at

the baseline and follow-up CT examinations, based on the region-growing segmentation. In the case of an incomplete segmentation, the radiologist obtained the volume using the manual correction method [19]. Usually, the manual correction method has been adopted for juxta-structural masses or nodules as the volumetry included a large component of the vasculature or pleura [20].

2.4. Statistical Analysis for Establishing a Prediction Model

Spearman correlation coefficient tests were performed to estimate the correlation coefficients between the parameters and reference 3D CT volume, and the in-between parameters. Volume prediction models for the pulmonary nodules were established through univariable or multivariable, and linear or non-linear regression analyses with the parameters. A polynomial regression analysis was performed as a method of the non-linear regression model. The researchers established the degree of regression based on the assumption that the volume would be proportional to the area raised to the power of 1.5 and then adjusted the units between the area and volume (Supplementary Material Figure S1). The root mean square error (RMSE) was measured to compare the performance of the polynomial regression models established at variable degrees.

3. Results

3.1. Demographics and Chest Radiography Features

A total of 147 CXRs from 72 patients (i.e., 39 men [54%] and 33 women [46%]; mean age, 62 years; age range, 33–84 years) with pulmonary nodules were included in the study.

Of the 147 CXRs, a total of 208 nodules were included when we identified up to 3 nodules per image in a descending order of size. The mean volume of the included nodules was 9.4 cm³ (range: 0.11 to 70.4 cm³). Lung cancer was pathologically confirmed in 34 (47%) patients with or without intrapulmonary metastases. Twenty-nine (40%) patients were pathologically or clinically diagnosed with pulmonary metastases of other primary malignancies. Eight (11%) patients were pathologically or clinically diagnosed with benign conditions including pulmonary tuberculosis, granulomatosis polyangiitis, cryptogenic organizing pneumonia, and nontuberculous mycobacteria infection. One patient was pathologically confirmed with primary chondrosarcoma originating from the thoracic cage. Seventy-eight (53%) CXRs were obtained from the scanners manufactured by GE Healthcare and 69 (47%) CXRs were obtained by Samsung Electronics (Table 1).

Table 1. Patient information and nodule characteristics.

Characteristic	Number (Percentage)
Patient information	
No. of chest radiographs	147
Patient with nodules	72
No. of men	39 (54.17%)
No. of women	33 (45.83%)
Mean age	62.00 ± 11.74
Nodule information	
Total no. of nodules	208
No. of nodules per chest radiograph	
One nodule	104
Two nodules	25
Three nodules	18
Disease	
Primary lung cancer	34 (47.2%)
Pulmonary metastasis	29 (40.28%)
Benign	8 (11.11%)
Other primary malignancies in the thorax	1 (1.39%)
Mean nodule volume (cm ³)	9.37 ± 11.69 *
Manufacturer	
GE Healthcare	78 (53.06%)
Samsung Electronics	69 (46.94%)

* mean ± standard deviation.

3.2. Regression Analysis of Volume Prediction of Pulmonary Nodules or Masses

Figure 2 shows the univariable regression analysis to explore the correlation between the area over the probability of 0.2 at the CAM and 3D CT volume of the nodule or mass, the mean probability of the ROIs over the cutoff probability value of 0.2 and the volume, and the in-between area and mean probability.

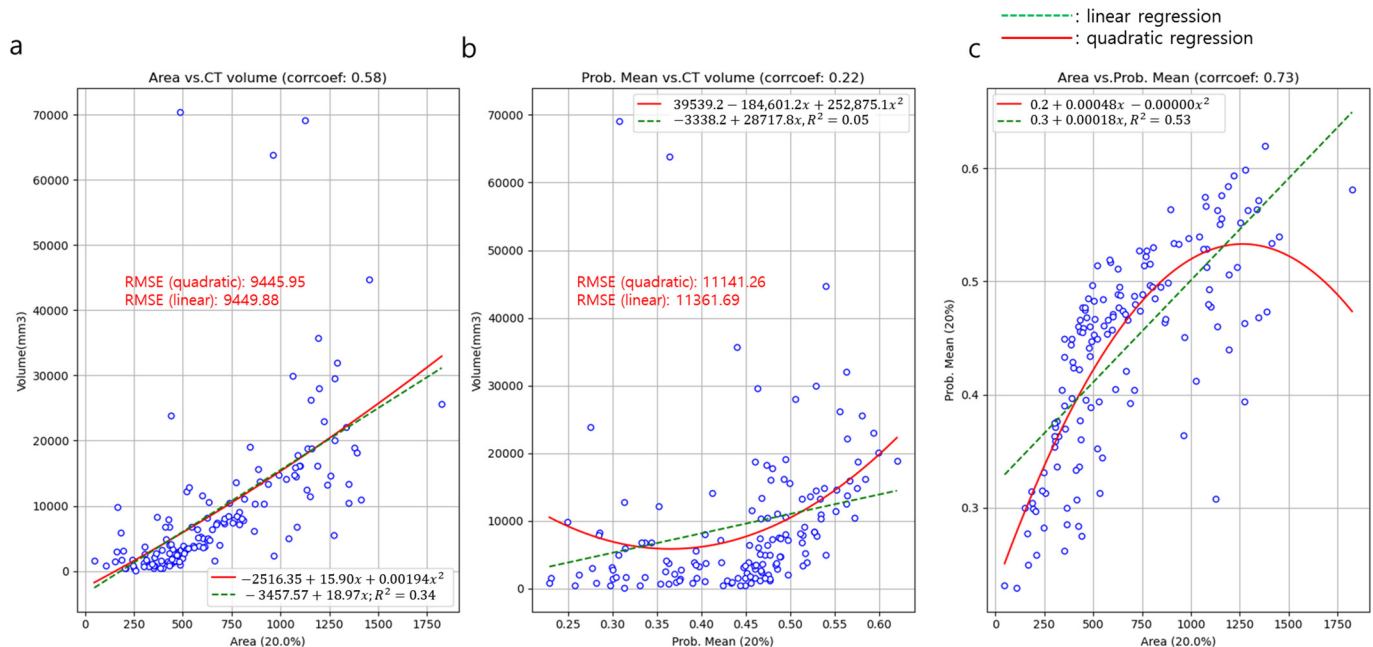


Figure 2. Univariable regression analysis to explore the correlation between the area over the probability of 0.2 (Area) at the CAM and 3D volume of the mass or nodule (a), mean probability (Prob. Mean) of the ROIs over the cutoff probability value of 0.2 and volume (b), and the in-between area and mean probability (c). (Note: regression model was not adjusted for age and sex).

The area and CT volume demonstrated a linear correlation of moderate strength ($R = 0.58$). A quadratic regression analysis among the area and CT volume showed a similar strength of correlation as the linear regression (R -square = 0.34). The area and mean probability of the ROIs exhibited a strong linear correlation ($R = 0.73$). A quadratic regression analysis among the area and probability also showed a similar correlation strength compared to the linear regression (R -square = 0.53). Between the mean probability of ROIs and the CT volume, negligible correlations were noted in both the linear and quadratic regression analysis ($R = 0.22$ and R -square = 0.05).

Figure 3 shows the result of fitting the area and detection probability to the 2D plane using a linear regression equation concerning the CT volume from various angles. The equation for a 2D plane was $14,886.637214 + (29.288656) \times \text{Area} - (57,206.109259) \times \text{Mean probability} = \text{Volume}$. The CT volumes of some masses were relatively large; therefore, there were outliers that rose significantly on the plane (e.g., four cases over 40,000 mm³, annotated by the red circle in Figure 3c). As shown in the univariable regression analysis, there was a strong correlation between the area and detection probability; therefore, in the quadratic regression equation, the coefficients appear as plus and minus signs, respectively.

Figure 4 demonstrates the correlation between the polynomial regression models of the degrees 1 (linear), 1.5 (unit-adjusted), and 2 (quadratic), estimated by the root mean squared error (RMSE). The performance for predicting the 3D CT volume with the area was the best in the unit-adjusted regression model (Figure 4a, RMSE = 9395.73 mm³). The performance for predicting the 3D CT volume with the area and mean probability was also the best at the unit-adjusted multivariable regression model (Figure 4d, RMSE = 7975.55 mm³). The residual plot for the unit-adjusted multivariable regression model is in the Supplementary Material Figure S2. The equation for this model was $= 578,437.473048 + 52,661.249342 \times (\text{Area})^{0.5}$

+ $(-4,508,167.245662) \times (\text{Mean probability}^{0.5}) + 826.73901 \times \text{Area} + (-217,913.487756) \times (\text{Area}^{0.5}) \times (\text{Mean probability}^{0.5}) + 10,744,046.813175 \times \text{Mean probability} + (-3.158884) \times (\text{Area}^{0.5}) + (-804.666435) \times \text{Area} \times (\text{Mean probability}^{0.5}) + 191,918.949126 \times (\text{Area}^{0.5}) \times \text{Mean probability} + (-7,729,391.963889) \times (\text{Mean probability}^{0.5})$. An example of applying this prediction model is demonstrated in Figure 5.

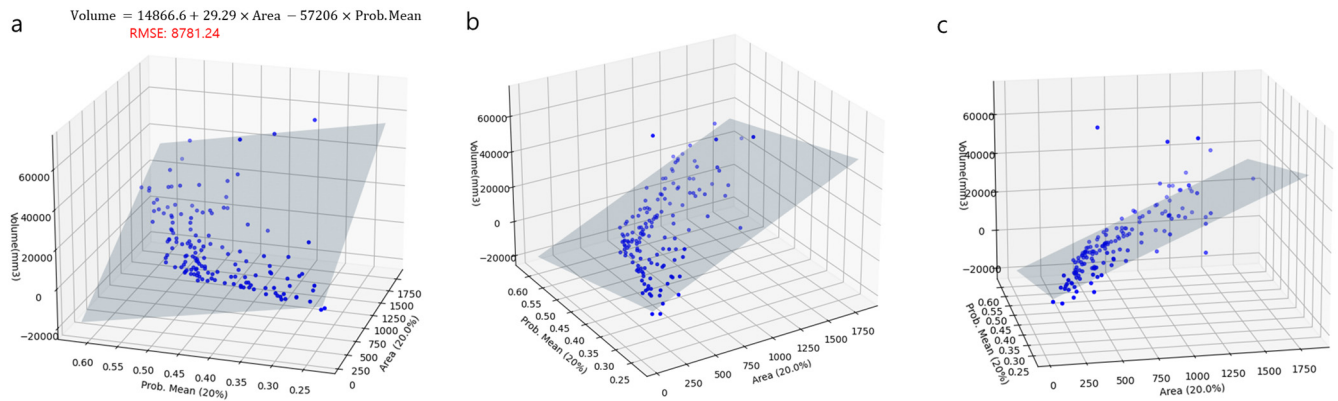


Figure 3. The two-dimensional plane fitted to the three-dimensional linear regression function graphs of area and detection probability for the CT volume at various angles (a–c). Four outliers over the CT volume of 40 cm³ are annotated by the red circle (c). (Note: Area = area over the probability of 0.2 at the class activation map, Prob. Mean = mean probability of the ROIs over the cutoff probability value of 0.2 in the class activation map).

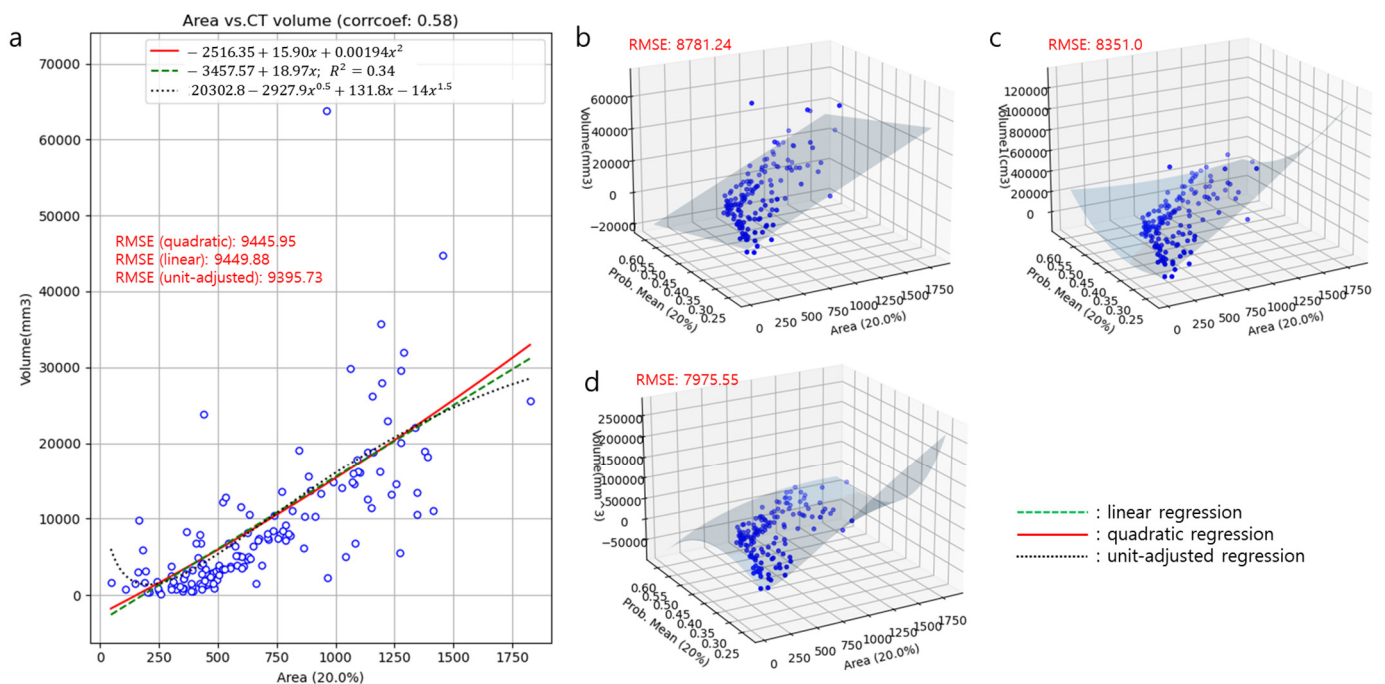


Figure 4. Correlation between the polynomial regression models of degrees 1, 1.5, and 2 estimated by RMSE. Comparison of performances for predicting 3D CT volume with the area among degree 1 (linear), 1.5 (unit-adjusted), and 2 (quadratic) univariate regression model (a). Comparison of performances for predicting 3D CT volume with the area and mean probability among degree 1 (linear), 1.5 (unit-adjusted), and 2 (quadratic) multivariable regression model (b–d). (RMSE = root mean square error).

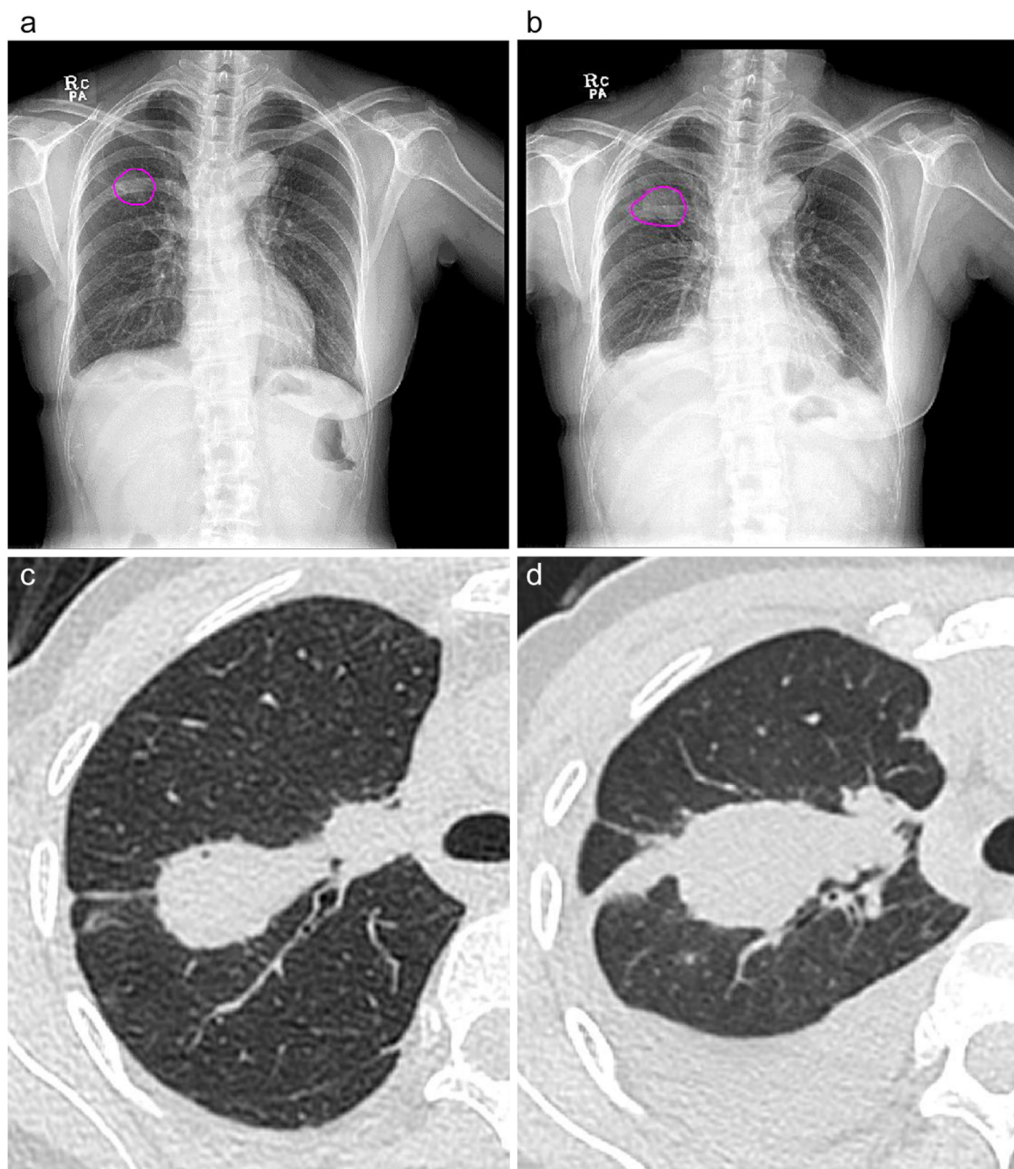


Figure 5. A 49-year-old female with breast cancer (pulmonary metastasis in RUL). In the initial and 2-month follow-up CXR (a,b), ALND-detected and -localized lesion nodule. Pink circles indicate the area, defined as the ROI over the 0.2 probability cutoff at the class activation channel, which is up-sampled and overlapped with the original image. The measured area and mean probability of each image were increased from 761.6 mm² (7.62 cm²) and 0.514 (a) to 1079.6 mm² (10.80 cm²) and 0.529 (b). Corresponding CT scans (c,d) within a week before or after the serial CXRs (a,b), the measured volume of the masses increased from 7.95 cm³ (c) to 14.50 cm³. The predicted volume of the 1.5-degree multivariable regression model was also increased from 8.50 cm³ to 16.35 cm³ (d).

4. Discussion

We analyzed 147 chest radiographs (CXRs) from 72 patients, identifying a total of 208 pulmonary nodules and found a moderate linear correlation between the area of the nodule and its CT volume, and a strong linear correlation between the area and the mean probability of the regions of interest (ROIs). The best performance for predicting the 3D CT volume was achieved with a unit-adjusted regression model, which utilized both the area and mean probability generated by the DLAD.

The choice of a 1.5-degree polynomial regression model was driven by two main considerations. First, we aimed to adjust the units from 2D (mm²) to 3D (mm³), which naturally suggested a 1.5-degree polynomial. Second, we sought to align the dimensions

from the 2D CXRs to the 3D volume of the nodules or masses. The 1.5-degree polynomial regression model demonstrated the best performance in terms of RMSE, indicating that it provided the most accurate estimates of the volume of the pulmonary nodules or masses.

Several previous studies have compared differences in the nodule or mass size in serial chest CT scans using DLAD systems. For example, Xu et al. predicted the treatment response of locally-advanced lung cancer patients using a combined CNN and RNN model, which analyzed the changes in mass on serial chest CT scans obtained during pretreatment and follow-up after radiotherapy [21]. Coreline Ltd. (Seoul, Republic of Korea), a company specializing in DLAD technology, has developed a commercial system capable of automatically matching the same nodule from previous chest CT scans and follow-up scans, facilitating the immediate identification of nodule changes [22]. In our study, 87% of patients had lung cancer or pulmonary metastasis. CXR examinations offer a more cost-effective and accessible method for short-term follow-up compared to chest CT scans, enabling clinicians to quickly assess the disease progression in their patients.

High-performance segmentation models based on U-Net architectures have been reported, providing more precise localization maps [14]. Shimazaki et al. developed a multiclass CXR abnormality detection algorithm using a segmentation method to calculate the exact nodule volume and to enable its comparison in serial images, although the accuracy of comparing the serial images was not validated [17]. We opted not to use U-Net-based architectures for serial image comparison, as our primary goal was to efficiently extract features reflecting the interval changes in nodules rather than achieving high-performance segmentation.

We hypothesize that a probability map can provide information on the rough area of the lesion seen on the CXR, which is not definite; however, a probability map is not made for measuring the area of a lesion but for the probability of the lesion's existence. We thought that since the lesion is more definite, the probability values would get higher and that this might reflect the higher volume of a true lesion. By creating the equations by area and mean probability, our prediction of the lesion volume for nodules and masses was rather well predicted from serial CXRs, as an increasing or decreasing in the size via the reference standard of a measured 3D volume from CT.

Although our ALND model was trained using consecutive CXRs from multiple radiotherapy device manufacturers, our hypothesis did not hold up well when the manufacturer of the serial CXRs changed [4]. Consequently, we developed a prediction model using serial CXRs from the same manufacturer.

Our approach for estimating the volume change using CAM proved effective, even though it was originally developed for multiclass discriminative models, unlike our binary classifiers [16]. It is anticipated that applying this technique to a multiclass discriminative model would yield a robust performance, thus extending its applicability beyond binary classifiers. Other multiclass discriminators have been introduced for CXR abnormalities, such as consolidation, interstitial opacity, pneumothorax, and pleural effusion [23]. Incorporating CAM into these models could lay the groundwork for future studies, including a pneumonia volume estimation for consolidation, ultimately assisting clinicians in making informed therapeutic decisions.

In addition to comparing lesions themselves, accurately identifying lesions across serial imaging follow-ups is a critical aspect of the monitoring process. Li et al. developed the concept of a ResNet-driven Siamese tracker for the visual tracking of images [24,25]. Cai et al. introduced the concept of a deep lesion tracker (DLT) for monitoring lesions in 4D longitudinal CT imaging studies. This method incorporated both image and anatomical considerations through self-supervised learning based on a Siamese neural network. The DLT reported an average Euclidean center error of 7 mm and improved the tumor monitoring workflow in assessing lesion treatment responses while maintaining an accuracy of 85%, which is comparable to manual lesion tracking [26]. Chen et al. introduced a transformer tracking method, known as TransT, which is built upon a Siamese-like feature extraction backbone, an attention-based fusion mechanism, and a classification and regression head. This approach was designed to address the loss of semantic information that occurs during

the correlation operation in traditional tracking algorithms. By incorporating attention mechanisms into the tracking process, TransT effectively enhances the performance of lesion tracking. When tested on large-scale datasets such as LaSOT, TrackingNet, and GOT-10 k, TransT demonstrated a superior performance compared to previously reported lesion tracking algorithms that did not utilize attention mechanisms. Furthermore, Tang et al. proposed a transformer lesion tracker that employed a cross attention-based transformer (CAT) to capture and combine both global and local information for enhanced feature extraction. They also introduced a registration-based anatomical attention module to provide anatomical information to the CAT, enabling it to focus on relevant feature knowledge. A sparse selection strategy (SSS) is presented for selecting the features and reducing the memory footprint in transformer training. This approach improved the average Euclidean center error by at least 14.3% (6 mm vs. 7 mm) compared to the DLT when implemented on a test dataset of CT images, which the DLT also used [27]. Incorporating state-of-the-art tracking methods with the concept of our lesion volume change prediction model could significantly enhance the efficiency and accuracy of radiologists' daily workflows. Furthermore, these techniques may not necessitate extensive training datasets, as the Siamese network for lesion tracking was initially developed for one-shot learning, and our prediction model was also based on a relatively small number of images with a pre-trained DLAD. By combining advanced lesion tracking technology with our volume change prediction model, we can potentially streamline the monitoring process and alleviate the workload for radiologists.

There are several limitations to this study. Firstly, as a single-center study, it included a limited number of patients. Secondly, we did not perform an internal validation of the suggested prediction model, thereby limiting our ability to check the final model's fitness and adjust equation degrees and coefficients. Moreover, the accuracy of CAM for localization remains questionable. Additionally, more recent CAM methods have been proposed, exhibiting an improved performance [28,29]. For instance, Grad-CAM, a popular method for explaining the interpretability of DLAD, generates localization maps without additional training for CAM weights by utilizing the gradients calculated during backpropagation [30]. Specific clinical applications of this prediction model have not been identified. A decision curve analysis with the appropriate setting of clinical outcomes and interventions such as timely CT examination could further be evaluated [31]. Additionally, further studies for a multivariable prediction model of disease progression by adding clinical factors may help with clinicians' decision making. Lastly, unlike detection and segmentation algorithms, which have state-of-the-art methodologies and appropriate metrics including a mean average precision, there are no reported state-of-the-art methodologies to be compared with our study.

5. Conclusions

In conclusion, the prediction model with the area and the mean probability of CAM showed the proper estimation of interval changes in pulmonary nodules or masses on the initial follow-up of serial CXRs with a quantitative volume estimation, based on commercially-available DLAD detecting pulmonary nodules and masses.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/diagnostics13122060/s1>, Figure S1: Relationship between the volume of the 3D object and the area of the 2D projection. Figure S2: Residual plot for the unit-adjusted multivariable regression model for nodule volume prediction.

Author Contributions: C.Y.L.: Data collection, data analysis, interpretation of results, manuscript writing and editing; Y.K.C.: Data collection, data analysis, interpretation of results, manuscript writing, editing, supervision, and final approval of the manuscript; M.J.C.: Study concept, supervision, interpretation of results, critical revision, and final approval of the manuscript; S.P. (Subin Park): Methodology, data analysis, statistical analysis, visualization, and manuscript editing; S.P. (Soyoung Park): Methodology, data analysis, statistical analysis, visualization, and manuscript editing; J.H.W.: Writing—review and editing; J.H.K.: Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of the Samsung Medical Center, Sungkyunkwan University School of Medicine (IRB No. SMC 2021-07-075-002, approved on 20 July 2021).

Informed Consent Statement: Patient consent was waived due to the retrospective nature of this study and the anonymization of patient data.

Data Availability Statement: The datasets generated during this study are not publicly available due to privacy and ethical considerations; however, anonymized data can be provided by the corresponding author upon reasonable request and with the approval of the ethics committee. Researchers interested in accessing the data should contact the corresponding author for further information.

Acknowledgments: We gratefully acknowledge Ha Kyu Jeong for his invaluable assistance in data analysis, model development, and insightful discussions, which greatly contributed to the success of this research. We also express our gratitude to the radiology department staff for their support and cooperation during the data collection process. This work was supported by the Future Medicine 20*30 Project of the Samsung Medical Center under Grant SMX1210791.

Conflicts of Interest: The authors declare no conflict of interest, financial or otherwise, in relation to this study or its publication.

References

1. Hwang, E.J.; Park, S.; Jin, K.N.; Kim, J.I.; Choi, S.Y.; Lee, J.H.; Goo, J.M.; Aum, J.; Yim, J.J.; Park, C.M.; et al. Development and Validation of a Deep Learning-based Automatic Detection Algorithm for Active Pulmonary Tuberculosis on Chest Radiographs. *Clin. Infect. Dis.* **2019**, *69*, 739–747. [[CrossRef](#)] [[PubMed](#)]
2. Nam, J.G.; Park, S.; Hwang, E.J.; Lee, J.H.; Jin, K.N.; Lim, K.Y.; Vu, T.H.; Sohn, J.H.; Hwang, S.; Goo, J.M.; et al. Development and Validation of Deep Learning-based Automatic Detection Algorithm for Malignant Pulmonary Nodules on Chest Radiographs. *Radiology* **2019**, *290*, 218–228. [[CrossRef](#)] [[PubMed](#)]
3. Park, S.; Lee, S.M.; Lee, K.H.; Jung, K.H.; Bae, W.; Choe, J.; Seo, J.B. Deep learning-based detection system for multiclass lesions on chest radiographs: Comparison with observer readings. *Eur. Radiol.* **2020**, *30*, 1359–1368. [[CrossRef](#)] [[PubMed](#)]
4. Sim, Y.; Chung, M.J.; Kotter, E.; Yune, S.; Kim, M.; Do, S.; Han, K.; Kim, H.; Yang, S.; Lee, D.J.; et al. Deep Convolutional Neural Network-based Software Improves Radiologist Detection of Malignant Lung Nodules on Chest Radiographs. *Radiology* **2020**, *294*, 199–209. [[CrossRef](#)]
5. Hwang, E.J.; Park, S.; Jin, K.N.; Kim, J.I.; Choi, S.Y.; Lee, J.H.; Goo, J.M.; Aum, J.; Yim, J.J.; Cohen, J.G.; et al. Development and Validation of a Deep Learning-Based Automated Detection Algorithm for Major Thoracic Diseases on Chest Radiographs. *JAMA Netw. Open* **2019**, *2*, e191095. [[CrossRef](#)]
6. Hong, W.; Hwang, E.J.; Lee, J.H.; Park, J.; Goo, J.M.; Park, C.M. Deep Learning for Detecting Pneumothorax on Chest Radiographs after Needle Biopsy: Clinical Implementation. *Radiology* **2022**, *303*, 433–441. [[CrossRef](#)]
7. Kim, E.Y.; Chung, M.J. Application of artificial intelligence in chest imaging for COVID-19. *J. Korean Med. Assoc.* **2021**, *64*, 664–670. [[CrossRef](#)]
8. Liang, C.H.; Liu, Y.C.; Wu, M.T.; Garcia-Castro, F.; Alberich-Bayarri, A.; Wu, F.Z. Identifying pulmonary nodules or masses on chest radiography using deep learning: External validation and strategies to improve clinical practice. *Clin. Radiol.* **2020**, *75*, 38–45. [[CrossRef](#)]
9. Elhanashi, A.; Lowe Sr, D.; Saponara, S.; Moshfeghi, Y. Deep learning techniques to identify and classify COVID-19 abnormalities on chest X-ray images. In *Real-Time Image Processing and Deep Learning 2022*; SPIE: Bellingham, WA, USA, 2022; pp. 15–24.
10. Ahn, J.S.; Ebrahimian, S.; McDermott, S.; Lee, S.; Naccarato, L.; Di Capua, J.F.; Wu, M.Y.; Zhang, E.W.; Muse, V.; Miller, B.; et al. Association of Artificial Intelligence-Aided Chest Radiograph Interpretation with Reader Performance and Efficiency. *JAMA Netw. Open* **2022**, *5*, e2229289. [[CrossRef](#)]
11. Jang, S.; Song, H.; Shin, Y.J.; Kim, J.; Kim, J.; Lee, K.W.; Lee, S.S.; Lee, W.; Lee, S.; Lee, K.H. Deep Learning-based Automatic Detection Algorithm for Reducing Overlooked Lung Cancers on Chest Radiographs. *Radiology* **2020**, *296*, 652–661. [[CrossRef](#)]
12. Shin, H.J.; Han, K.; Ryu, L.; Kim, E.-K. The impact of artificial intelligence on the reading times of radiologists for chest radiographs. *npj Digit. Med.* **2023**, *6*, 82. [[CrossRef](#)]
13. Sun, J.; Peng, L.; Li, T.; Adila, D.; Zaiman, Z.; Melton-Meaux, G.B.; Ingraham, N.E.; Murray, E.; Boley, D.; Switzer, S.; et al. Performance of a Chest Radiograph AI Diagnostic Tool for COVID-19: A Prospective Observational Study. *Radiol. Artif. Intell.* **2022**, *4*, e210217. [[CrossRef](#)]
14. Hurt, B.; Yen, A.; Kligerman, S.; Hsiao, A. Augmenting Interpretation of Chest Radiographs with Deep Learning Probability Maps. *J. Thorac. Imaging* **2020**, *35*, 285–293. [[CrossRef](#)] [[PubMed](#)]

15. Guo, R.; Passi, K.; Jain, C.K. Tuberculosis Diagnostics and Localization in Chest X-rays via Deep Learning Models. *Front. Artif. Intell.* **2020**, *3*, 583427. [CrossRef] [PubMed]
16. Zhou, B.; Khosla, A.; Lapedriza, A.; Oliva, A.; Torralba, A. Learning Deep Features for Discriminative Localization. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 2921–2929. [CrossRef]
17. Shimazaki, A.; Ueda, D.; Choppin, A.; Yamamoto, A.; Honjo, T.; Shimahara, Y.; Miki, Y. Deep learning-based algorithm for lung cancer detection on chest radiographs using the segmentation method. *Sci. Rep.* **2022**, *12*, 727. [CrossRef] [PubMed]
18. Hansell, D.M.; Bankier, A.A.; MacMahon, H.; McLoud, T.C.; Muller, N.L.; Remy, J. Fleischner Society: Glossary of terms for thoracic imaging. *Radiology* **2008**, *246*, 697–722. [CrossRef] [PubMed]
19. Shin, K.E.; Lee, K.S.; Yi, C.A.; Chung, M.J.; Shin, M.-H.; Choi, Y.-H. Subcentimeter lung nodules stable for 2 years at LDCT: Long-term follow-up using volumetry. *Respirology* **2014**, *19*, 921–928. [CrossRef]
20. Devaraj, A.; Van Ginneken, B.; Nair, A.; Baldwin, D. Use of Volumetry for Lung Nodule Management: Theory and Practice. *Radiology* **2017**, *284*, 630–644. [CrossRef]
21. Xu, Y.; Hosny, A.; Zeleznik, R.; Parmar, C.; Coroller, T.; Franco, I.; Mak, R.H.; Aerts, H. Deep Learning Predicts Lung Cancer Treatment Response from Serial Medical Imaging. *Clin. Cancer Res.* **2019**, *25*, 3266–3275. [CrossRef]
22. CORELINE. Available online: <https://www.corelinesoft.com/en/> (accessed on 1 March 2023).
23. Nam, J.G.; Kim, M.; Park, J.; Hwang, E.J.; Lee, J.H.; Hong, J.H.; Goo, J.M.; Park, C.M. Development and validation of a deep learning algorithm detecting 10 common abnormalities on chest radiographs. *Eur. Respir. J.* **2021**, *57*, 2003061. [CrossRef]
24. Koch, G.; Zemel, R.; Salakhutdinov, R. Siamese neural networks for one-shot image recognition. In Proceedings of the Icm1 Deep Learning Workshop, Lille, France, 6–11 July 2015.
25. Li, B.; Wu, W.; Wang, Q.; Zhang, F.; Xing, J.; Yan, J. Siamrpn++: Evolution of siamese visual tracking with very deep networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 4282–4291.
26. Cai, J.; Tang, Y.; Yan, K.; Harrison, A.P.; Xiao, J.; Lin, G.; Lu, L. Deep lesion tracker: Monitoring lesions in 4D longitudinal imaging studies. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 15159–15169.
27. Tang, W.; Kang, H.; Zhang, H.; Yu, P.; Arnold, C.W.; Zhang, R. Transformer Lesion Tracker. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2022: 25th International Conference, Singapore, 18–22 September 2022; pp. 196–206, Proceedings, Part VI.
28. Jung, H.; Oh, Y. Towards better explanations of class activation mapping. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 1336–1344. [CrossRef]
29. Fu, R.; Hu, Q.; Dong, X.; Guo, Y.; Gao, Y.; Li, B. Axiom-based grad-cam: Towards accurate visualization and explanation of cnns. *arXiv* **2020**, arXiv:2008.02312.
30. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 618–626.
31. Vickers, A.J.; Elkin, E.B. Decision curve analysis: A novel method for evaluating prediction models. *Med. Decis. Making* **2006**, *26*, 565–574. [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.