End-to-End Fully Automated Lung Cancer Volume Estimation System

Original Scientific Paper

Pushkar Sathe*
Department of Electronics and Telecommunications Engineering, Mukesh Patel School of Technology Management & Engineering,
SVKM's Narsee Monjee Institute of Management Studies (NMIMS) Deemed-to-University,
Mumbai-400056, India
pushkarmsathe@gmail.com

Alka Mahajan
JK Lakshmipat University, Jaipur-302026, India.
alka.mahajan@jklu.edu.in

*Corresponding author

Deepak Patkar
Department of Radiology, Nanavati Superspeciality Hospital,
Mumbai-400056, India
drdppatkar@gmail.com

Mitusha Verma
Department of Radiology,
Nanavati Superspeciality Hospital,
Mumbai-400056, India
drmitusha@gmail.com

Abstract *– The volume of the tumor plays a very crucial role in deciding the stage of lung cancer which in turn helps in deciding the best treatment and its schedule. Currently used computer-based volume estimation techniques are semi-automatic with limited accuracy. For any automatic lung cancer segmentation system, lung CT scans of hundreds of patients are required along with their corresponding annotated segmentation masks. It is difficult to get accurately annotated data as cancer segmentation of CT scans done by the radiologists, is a time-consuming manual process. Also, it is subjective and prone to intra and inter-observer variability. Further, owing to the irregular shape of the cancerous tumor, accurate volume estimation becomes a challenge with regular convolution models. This paper proposes an end-to-end automatic tumor volume estimation model that estimates volume using the GPR (Gaussian Process Regression) interpolation method. The proposed modified cancer segmentation model uses deformable convolutions. This modification offers a higher segmentation accuracy in terms of IoU (Intersection over Union) and clearly defined nodule boundaries with correct retention of the nodule shape. The research was undertaken in collaboration with Nanavati Hospital, Mumbai, and all the models were validated on a real dataset obtained from the hospital. The proposed model gives a mean segmentation IoU (Intersection over Union) of 0.9035 and a volume estimation accuracy of 93.13% which are almost 5% and 3% higher than 0.8548 and 90.51% which are the corresponding results obtained using a standard U-net++ algorithm.*

Keywords: *Cancer segmentation, Deep learning, Deformable convolution, Lung cancer volume estimation, U-net++*

Received: March 13, 2024; Received in revised form: June 4, 2024; Accepted: July 2, 2024

1. INTRODUCTION

Lung cancer is the most widespread type of cancer and is the second most common cancer after prostate cancer in males and breast cancer in

females. The early detection of lung cancer plays an effective role in diagnosis and leads to an early treatment increasing the likelihood of patient survival rate [1, 2]. Chemotherapy, which is the main treatment for lung cancer requires knowledge of the accurate location of the cancer along with its spread along all the 3 axes. Accurate volume estimation is required for determining the stage of the cancer. Volume estimation can be done using 3D or 2D CT scan images. Though 3D data gives more accurate volume estimation it suffers from a lack of annotated data required for training and higher complexity of the segmentation algorithms. So generally 2D data is used to estimate the slice-wise area of cancer spread and then volume is obtained using these area values. It is very difficult to get accurately annotated data as radiologists usually perform manual segmentation. Human intervention leads to errors arising out of fatigue and subjectivity. To overcome this problem, cancer segmentation models based on deep learning were proposed but they require a large dataset along with corresponding annotations of every CT

scan for training to get acceptable segmentation accuracy that leads to precise area calculation. Also, the majority of cancer segmentation algorithms fail to define the boundaries of cancerous nodules clearly. While many software programs are available for calculating volume automatically, they are semi-automatic as they need demarcations of cancerous portions to be done by radiologists in every or at most on alternate slices. In this paper, a fully automatic tumor volume estimation model is proposed that performs segmentation of cancerous portions by modifying the U-net++ algorithm using deformable convolutions. The volume is calculated using GPR interpolation. The proposed model offers highly accurate volume estimation and requires a lesser number of original CT scans.

The primary contributions of our proposed research are as follows:

1) Designed an improved segmentation algorithm based on a modified convolution that achieved an improved segmentation accuracy in terms of well-defined nodule boundaries and the retention of actual nodule shape.
2) Developed the cancer volume estimation system using interpolation techniques that provided better accuracy of volume estimation.

2. RELATED WORK

In recent times CT scans have been dominantly used for cancer detection and a lot of research on developing AI (Artificial Intelligence) based cancer detection algorithms is reported in the literature. Matt Daykin et al. used a One-Class Support Vector Machine (OCSVM) based model to detect lung abnormality [1]. The method, however, does not use complete images but patches of lung CT scans majorly to increase the size of the data set. Automatic feature extraction or the use of deep learning techniques is also not explored fully. Other reported abnormality detection algorithms [2, 3] majorly dealt with cancers other than lung cancer. Irigoien et al. applied OCC (One Class Classification) to medical data for the detection of various diseases like breast cancer, liver disorders, leukemia, etc. [2]. The authors compared the performance of four different algorithms namely Gaussian, mixtures of Gaussian, Parzen, and typicality approach based on their average AUC (Area Under Curve). The paper reported the best results with an average AUC of 77.4% with the typicality approach. Tarassenko et al. [3] also applied OCC to investigate normality using a large number of available mammograms which do not show any evidence of mass-like structures. The recent advancement in deep learning techniques allows for automatically extracting features from the images, thereby improving the overall performance as compared to conventional CAD (Computer Aided Diagnosis). The use of transfer learning techniques is also reported in literature where fixed feature extraction is done using a pre-trained network. Ardimento et al. [4] applied three transfer learning models viz; VGG (Visual Geometry Group), Xception, and ResNet for feature extraction and combined the results using ensemble architecture to classify the scans as cancerous and noncancerous. Similarly, the state-of-the-art transferable architectures such as VGG-16, VGG-19, GoogLeNet, Inception-V3, ResNet-18, ResNet-50, ResNet-101, InceptionResNet-V2 and 3D multipath VGG like network have also been used for lung cancer feature extraction [5-9] and the performance of SVM (Support Vector Machine) and AdaBoostM2 classifier is analyzed on the deep features extracted from publicly available datasets. In [10] the performance of various machine learning algorithms was evaluated on a lung cancer detection task. The fusion of DenseNet201 with color histogram techniques was used to extract a hybrid feature set. Similarly, an innovative deep-learning model for lung cancer detection by integrating markers from mRNA, miRNA, and DNA methylation was developed [11]. The principal components analysis (PCA) was implemented to streamline features and the synthetic minority over-sampling technique (SMOTE) algorithm was applied to ensure class balance. The PCA-SMOTE model achieved an F1 score of 0.97. Lung nodule segmentation methods can be categorized into traditional approaches and deep learning-based techniques. Traditional methods encompass threshold and region-growing methods, clustering methods, active contour models, and mathematical models [12-20]. On the other hand, deep learning-based methods for segmentation can be further divided into 2D and 3D segmentation networks [21]. While traditional segmentation methods do not necessitate a substantial amount of labeled data for model training, they heavily rely on human intervention and are more focused. These methods primarily rely on shallow image features such as grayscale and texture. In contrast, 2D segmentation networks utilize 2D convolution to extract features from images. The original Fully Convolutional Neural Network (FCN) [22], has found extensive use in various image segmentation fields but it lacks global context information which is very important in the case of medical image segmentation. Ronneberger et al. [23] introduced a U-Net network based on the FCN architecture for medical image segmentation. U-Net incorporates both low-resolution and high-resolution information through skip connections, which is particularly advantageous for segmenting medical images with blurred boundaries. In this approach, low-resolution information is used for target identification, while high-resolution information aids in the localization of the segmentation. Consequently, U-Net has served as a foundation for several improved algorithms. The average IoU of 77.5% ia achieved but the algorithm failed to consider the multiscale information required for accurate segmentation. So, to further improve segmentation results, some researchers have explored the integration of Atrous Spatial Pyramid Pooling (ASPP) [24] instead of the intermediate or output layer of U-Net. This approach enables the extraction of multi-scale image information through different perceptual fields [25-27] but lacks in retention of exact shapes. Though the average IoU of 81.3 was achieved, the algorithm lacked complete recovery

of spatial information. Some 3D segmentation networks utilize 3D convolution to extract features from volumetric images, allowing them to better capture the spatial relationship between nodules and surrounding tissues compared to 2D segmentation networks. Duo et al. [28] introduced a 3D fully convolutional neural network that automatically segments the liver and cardiac great vessels. Milletari et al. [29] incorporated residual connections inspired by ResNet [30] and employed 3D convolution in their V-Net architecture for prostate volume segmentation, departing from the 2D convolution used in U-Net. Deepseed [31] proposed 3D-squeeze-and-excitation (SE) networks, incorporating dynamic scaling with cross-entropy loss [32] to address the sample imbalance problem. Their method was evaluated on the LIDC (Lung Image Database Consortium) [33] and LUNA16 (LUng Nodule Analysis) [34] datasets, demonstrating promising results for lung nodule segmentation. However, it is worth noting that 3D networks suffer from longer training times and may not necessarily outperform 2D networks in terms of evaluation metrics. Zhou et al. [35] proposed U-net++, an extension of U-Net that incorporates dense skip connections to enable multi-scale fusion and feature acquisition at different levels. The model achieves an average IoU of 81.4% but as the algorithm involves non-integral convolution, it fails in maintaining the exact shape of nodules. In [36] the attention module is used with the U-net algorithm having residual connections for fast backpropagation. But model failed to maintain the exact shape of cancerous nodules. Similarly, different deep learning segmentation models were evaluated on different datasets, and the effects of different preprocessing methods were examined [37]. Though the TransUet model achieved the highest segmentation accuracy with an average dice coefficient of 0.81, it failed to maintain the exact shape of the cancerous nodule. Being a transformer-based model, the decision-making process and interpreting predictions was challenging. Also, in [38] a method to integrate feature information through a dual-branch network framework and multi-dimensional fusion module is proposed. By training and validating with multiple data sources and different data qualities, the method demonstrated leading performance on the LUNA16, Multi-thickness Slice Image dataset, LIDC, and UniToChest, with an average dice similarity coefficient of 82%. This method failed to maintain the cancer nodule's exact shape, leading to inaccurate segmentation. Accurate segmentation of cancerous nodules is of high importance as the accuracy of volume estimation and grading is directly depending on the same.

3. METHODOLOGY

Fig. 1 shows the schematic of the proposed system. Raw CT scans are first cleaned and then features are extracted from the isolated lung section. Extracted features are then applied to the cancer detection model as explained in the sections that follow.

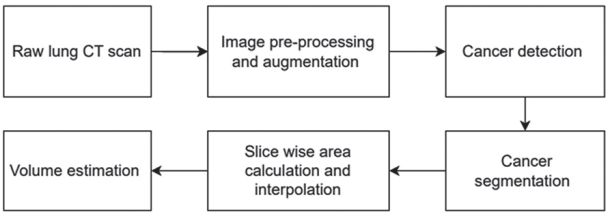


Fig. 1. Schematic of the proposed system

3.1 DATA COLLECTION AND PRE-PROCESSING

Normal and cancerous 2D lung CT scans were made available by the Nanavati Hospital, Mumbai. They were of DICOM (Digital Imaging and Communications in Medicine) file type. Each CT scan is of size 512 X 512. The dataset was checked, classified, and annotated independently by two radiologists and used for training and testing. A total of 872 normal and 146 cancerous scans were collected. Cancerous scans were increased to an optimum of 584 using data augmentation techniques approved by the doctors to balance the available data without the risk of overfitting. All the simulations were carried out using Python running on a server with AMD Ryzen 5 processor, 8GB RAM, and one NVIDIA GeForce GTX 1650 GPU with a compute capability of 7.5. Training of an algorithm took 14 hours and testing of applied image takes just 8 seconds.

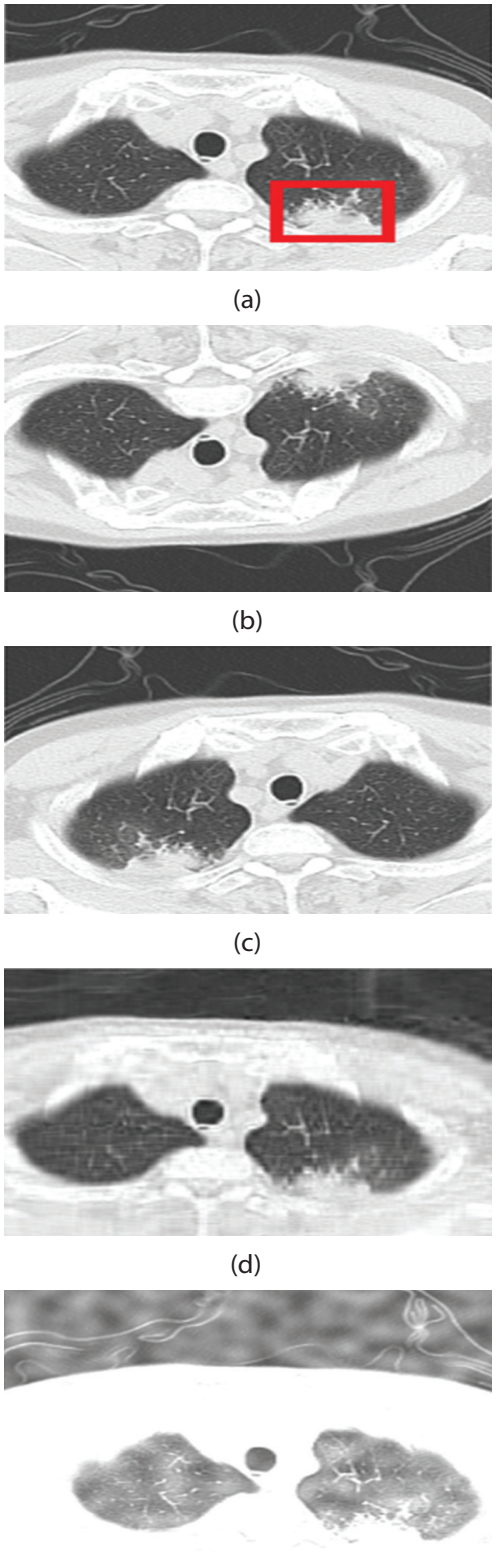
3.2 DATA AUGMENTATION

A significant obstacle in training deep learning models for various tasks is the abundance of data required. This challenge is particularly pronounced in the field of medicine, where limited access to costly imaging resources or a scarcity of study subjects can hinder progress. Additionally, the requirement of annotations for every cancerous scan puts a limit on data collection. Consequently, there is a growing trend of employing data augmentation techniques in research involving small data sets [39, 40].

The demand for extensive data in deep learning networks has spurred the development of various strategies. Only those strategies are used in this work that provide a different visual impact compared to an original image. Vertical flipping, horizontal flipping, PCA (Principal Component Analysis), and image overlay techniques are used in our work for data augmentation.

Fig. 2 shows the images generated using corresponding data augmentation methods for a sample CT scan. Fig. 2(a) shows the original CT scan. The irregularly shaped cancerous nodule is highlighted using the red bounding box. Figs. 2(b) and 2(c) show the results of the flipping operation performed on the original CT scan. After vertical and horizontal flipping the position of the cancerous nodule changed, making it appear as a new image for a deep-learning model. Horizontal flipping is used because objects like cancer exhibit horizontal symmetry as it often looks similar when flipped

horizontally. Vertical flipping is useful in the medical domain where the orientation of objects like cancer varies significantly. Fig. 2(d) shows the result obtained by performing PCA with 20 components. Similarly, Fig. 2(e) shows the new image generated using the image overlay technique.



(a)

(b)

(c)

(d)

(e)

Fig. 2. (**a**) Original CT scan (**b**) Vertical flipping (**c**) Horizontal flipping (**d**) PCA with 20 principal components (**e**) image overlay

3.3 CANCER DETECTION

After data augmentation, the CT scans were applied to the cancer detection model. The cancer detection model was trained using both normal and cancerous CT scans. 872 normal and 584 cancerous CT scans were used in training. The features extracted using VGG were applied to three different classifiers namely SVM, Decision Tree, and Random Forest. The model was validated on a test dataset of 50 normal and 50 cancerous scans.

3.4 CANCER SEGMENTATION

The u-net algorithm proved to be the most significant segmentation algorithm for medical images [23]. Many variations of the basic algorithm have been developed in recent years. The basis of the U-net algorithm is the duplication of feature maps of the encoder to the decoder section to improve the quality of the upsampling feature map. U-net++ further improves the segmentation performance of the U-net algorithm by using the concept of deep supervision and nested skip connections that use convolutional layers [35]. However, in conventional convolutional layers, the receptive field remains unchanged for a given network structure, regardless of the object's size. But use of the same receptive field for objects of different sizes is not ideal. In the field of medical image segmentation, the lesions to be segmented often exhibit irregular shapes and sizes. The use of conventional convolution results in less accurate demarcation of cancer nodules [41, 42].

Fig. 3 shows the schematic of the proposed architecture consisting of 5 layers in the encoder and decoder.

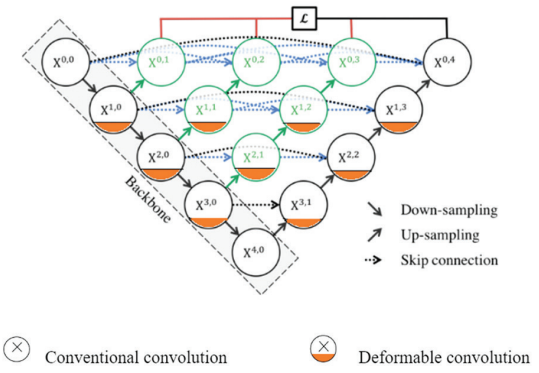


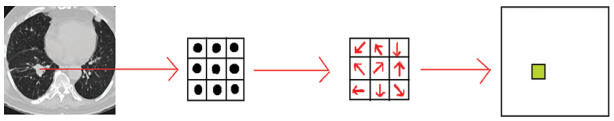
Fig. 3. The schematic architecture of the proposed algorithm

The encoder performs downsampling by capturing the contextual information and reduces spatial dimensions while the decoder performs upsampling and constructs the segmented image based on encoded features. The skip connections concatenate feature maps from the encoder to the decoder at the same spatial resolution. The nested skip pathways combine feature maps from different encoder resolutions to capture multi-scale contextual information. The weighted sum of the feature maps shown by in Fig. 3 improves the aggregation of features from different levels of the encoder.

+To further improve upon the segmentation accuracy, the proposed model uses deformable convolution instead of conventional convolution in the encoder and decoder layers. The increased geometric flexibility results in improved segmentation accuracy by retaining the original shape of the cancerous nodule.

We strategically incorporated deformable convolutions [41] into different parts of the network to identify the most effective place. Considering that the first downsampled convolution block is responsible for extracting the image's basic features, adding deformable convolutions to this block significantly prolongs network training time. Also, after multiple convolution and downsampling operations as the feature map size becomes very small and the feature information becomes abstracted, the benefits of using deformable convolutions are not evident. Taking these factors into account, we included the deformable convolution operation in the basic convolution block of the second, third, and fourth layers of the architecture as shown in Fig. 3.

3.4.1 DEFORMABLE CONVOLUTION



(a) (b) (c) (d)

Fig. 4. (**a**) lung CT scan (**b**) Grid of conventional convolution showing fixed locations of points (**c**) trainable offsets to perform convolution at variable locations instead of fixed places (**d**) features map extracted

Fig. 4(a) shows the lung CT scan. In the case of conventional convolution, the process of convolution is performed using a fixed grid and because of this, the receptive field remains the same. Figure 4(b) shows a fixed convolution grid of size 3 x 3. It is not appropriate to use the constant receptive field in our case as cancer nodules are of varying shape and size. Deformable convolution solves the problem of constant receptive field by using trainable offsets that are added to a fixed grid before performing convolution. Figure 4 (c) shows the trainable offsets. This process allows the network to learn information about nodule boundaries more accurately and it leads to better segmentation accuracy with retention of the exact shape of the cancerous nodule. Figure 4(d) shows the corresponding feature map extracted. Mathematically the deformable convolution can be represented as shown in Eq. (1).

$$y_i(P_0) = \sum_{Pn \in R} w(P_n).\, x(P_0 + P_n + \Delta P_n) \quad (1)$$

Here, the pixel value at P_0 is replaced by weighted addition performed between image pixels and convolution mask at locations $P_n + \Delta P_n$ where

$$P_n \in R \quad (2)$$

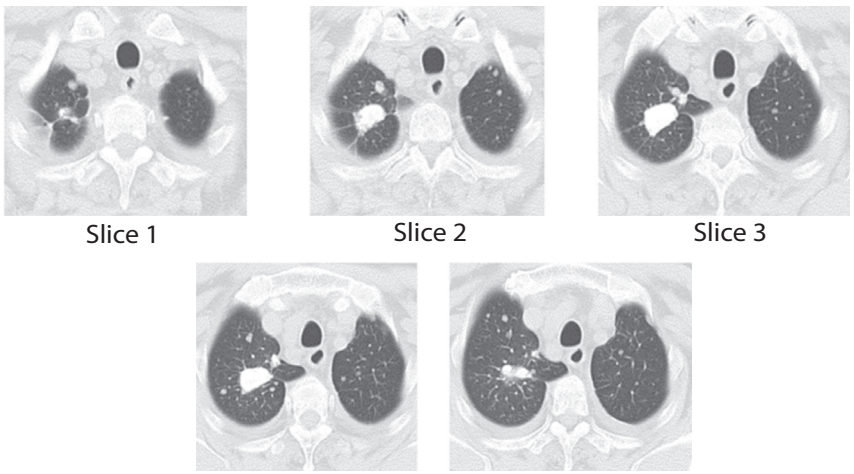
$$R = \{(-1, -1),(-1, 0), \ldots ,(0, 1),(1, 1)\} \quad (3)$$

for a grid of 3 x 3 size and ΔP_n indicates the learnable distance offset for the best location to perform convolution. In equation (1), w indicates the convolution mask and x indicates the image.

3.5. VOLUME ESTIMATION

When a series of CT scans of the patient is applied to the cancer segmentation model, it provides corresponding masks having cancerous portions highlighted for every CT scan. These annotated masks are applied to the volume estimation model that calculates the area of the cancerous portion of every scan and uses the interpolation method to find the volume of the tumor.

From the segmented mask, several white pixels are obtained. The area of the cancerous portion is calculated by multiplying the number of white pixels by the area of each pixel which is obtained from pixel height and pixel width functions of the Python library. Likewise, the area is calculated for all consecutive CT scans and then the volume is obtained from these area values using the interpolation method. Fig. 5 shows the consecutive cancerous scans for one sample patient case. Corresponding area values are 0.435, 2.966, 4.104, 3.8, and 1.755 respectively.



Slice 1 Slice 2 Slice 3

Slice 4 Slice 5

Fig. 5. Consecutive cancerous scans of a sample patient

Each slice is 5 mm thick. This means the real area of cancer spread is known only at intervals of 5 mm each and not in between. Considering the same cancerous area throughout the slice thickness will lead to wrong volume calculation. For better volume estimation, interpolation is used to predict the area values of the cancerous portion 0.01 mm apart slice-wise. Various types of interpolations like linear, bilinear, cubic, Lagrange, and GPR were tried to determine the volume of cancer, and their results are compared with the actual volumes obtained from radiologists. GPR interpolation is found to be the best approach.

3.5.1. GPR INTERPOLATION

The basic idea of GPR is to model the relationship between the input data and the output values as a Gaussian process. A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution. In GPR, we assume that the output values follow a Gaussian process with a mean function $\mu(x)$ and a covariance function $k(x, x')$ that captures the similarity between given data points and interpolation points.

The most commonly used covariance function is the Radial Basis Function (RBF) kernel. This interpolation is the most suitable for irregular data as it can better capture underlying trends and patterns. This property makes GPR the most suitable interpolation method in cancer volume estimation. The key idea of GPR interpolation is to use Bayesian inference to compute the posterior distribution over the unknown value ŷ conditioned on the observed data. The posterior distribution is a Gaussian distribution with a mean $\hat{\mu}$ and a covariance $\hat{\Sigma}$. The mean $\hat{\mu}$ represents the estimated value of $\hat{y}$, and the covariance $\hat{\Sigma}$ quantifies the uncertainty associated with the estimation.

Mathematically GPR interpolation can be explained as follows:

First, the mean function $\mu(x)$ and the covariance function $k(x, x')$ were defined. Then the covariance matrix $K(X, X)$ between the observed input points in X and the covariance vector $k(X, \hat{x})$ between the observed input points in X and the new input point $\hat{x}$ were computed. The covariance scalar $k(\hat{x}, \hat{x})$ between the new input point $\hat{x}$ and itself was then calculated. Finally, the mean vector $\hat{\mu}$ and the covariance matrix $\hat{\Sigma}$ of the posterior distribution were obtained using the formulae [43] given in equations (4) and (5)

$$\hat{\mu} = k(X, \hat{x})^T [K(X, X) + \sigma^2 I]^{-1} y \quad (4)$$

$$\hat{\Sigma} = k(\hat{x}, \hat{x}) - k(X, \hat{x})^T [K(X, X) + \sigma^2 I]^{-1} k(X, \hat{x}) \quad (5)$$

where σ^2 is the noise variance parameter and I is the identity matrix.

The estimated output value $\hat{y}$ at the new input point $\hat{x}$ was obtained by the mean $\hat{\mu}$, and the uncertainty associated with the estimation was quantified by the covariance $\hat{\Sigma}$.

4. 4. RESULTS AND DISCUSSIONS

4.1. CANCER DETECTION

It was found that classification results are optimum for a random forest classifier for the finalized values of hyperparameters when compared to other classifiers. All the cancerous scans were correctly classified by Random Forest, with a recall (or sensitivity) of 100%, accuracy of 99.75%, and precision of 99.5%. Accuracy considers all predictions (both positive and negative) and measures the overall correctness of the model. Whereas, the precision focuses only on the positive predictions and measures how many of those predicted positives are true positives. Also, 100% recall indicates that the model has successfully identified every case who had a disease.

4.2. CANCER SEGMENTATION

In this work, the Intersection over Union (IoU), Dice Similarity Coefficient (DSC), Relative Volume Difference (RVD), Average Surface Distance (ASD), and Hausdorff Distance (HD) were used to evaluate the image segmentation results. IoU is calculated by finding the ratio of the area of overlap between the predicted segmentation and the ground truth segmentation to the area of union between the two. In mathematical terms, it is defined as shown in equation (6)

$$IoU = |X \cup Y| \,/\, |X \cap Y| \quad (6)$$

Here, X and Y represent the sets or regions of pixels that belong to the predicted segmentation and the ground truth segmentation, respectively. A higher IoU value indicates a better segmentation result, with a value of 1 indicating a perfect overlap between the predicted and ground truth segmentations. The Dice coefficient is a similarity measure commonly used in image segmentation to evaluate the overlap between a segmented mask and a ground truth mask. Mathematically it is defined as given in equation (7).

$$DSC=(2*|X \cap Y|)/ (|X| + |Y|) \quad (7)$$

RVD quantifies the relative difference in volume or size between the segmented region and the ground truth region.

$$RVD=(|\text{Volume segmented - ground truth volume}|) / \text{ground truth volume}$$

The RVD value is a measure of how closely the segmented region's volume matches the ground truth volume. It's typically expressed as a percentage, with values closer to zero indicating a more accurate segmentation. If RVD is exactly zero, it means the segmentation perfectly matches the ground truth. A positive RVD indicates an overestimation, meaning the segmented region is larger than the ground truth, while a negative RVD indicates an underestimation, meaning the segmented region is smaller than the ground truth.

ASD on the other hand quantifies the average distance between the surfaces of the segmented region

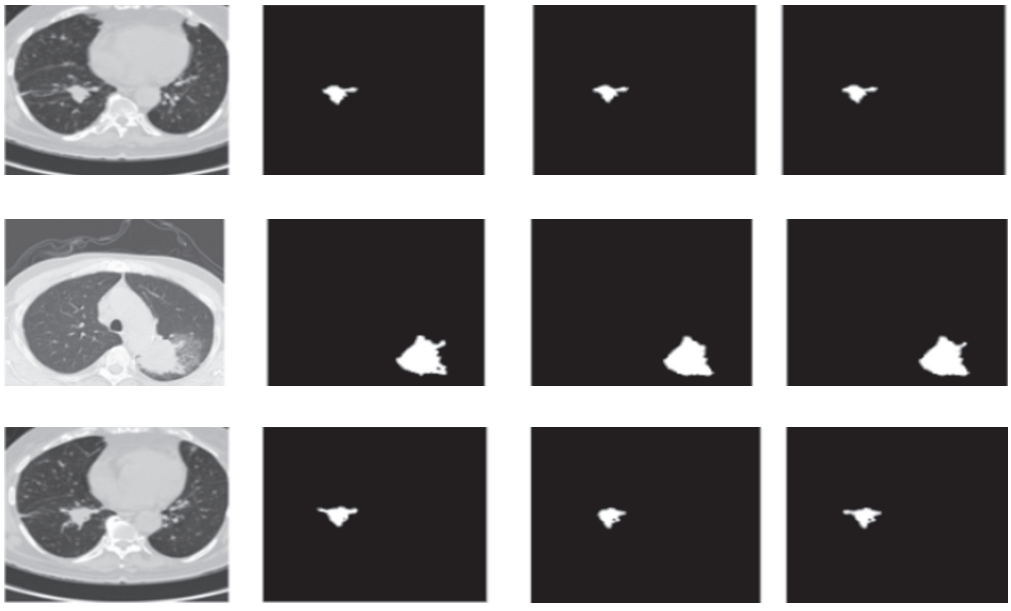
and the corresponding surfaces of the ground truth region. It calculates the distance between each corresponding pair of points, one from the segmented region's surface and one from the ground truth region's surface. Then it calculates the average of these distances. Smaller ASD values indicate a more accurate segmentation, as they imply that the segmented region's surface is closer to the ground truth surface.

HD is also a measure of the dissimilarity between the boundary of a segmented object and a ground truth boundary. It calculates the Euclidean distance between each point in one set and its nearest neighbor in the other set. First, it calculates the Hausdorff distance from the first set to the second set (H1) as the maximum distance among the recorded Euclidean distances from the above step. Similarly, the Hausdorff Distance from the second set to the first set (H2) is calculated. The HD is then defined as the maximum of H1 and H2.

The deformable convolution was used basically to improve the accuracy of the cancerous nodule segmentation. Table 1 shows the performance metrics for the proposed algorithm with regular convolution-based U-net++. The performance metrics for 3 patients CTP1, CTP2, and CTP3 are highlighted in Table 1. It is seen that for all three patients, the RVD value is reduced with the proposed algorithm compared to that with U-net++ which indicates improvement in segmentation. Similarly, there is a reduction in ASD value for three patients, particularly for patients CTP2 and CTP3. Also, the average values of performance parameters for 58 patients indicate that IoU and DCS show considerable improvement of 5.69% and 3.72 % respectively for the proposed model as compared to the U-net++. Similarly, other parameters RVD, ASD, and HD have been reduced for the proposed model by 41.02%, 65.67%, and 65.39% respectively which indeed indicates better segmentation. Fig. 6 shows the results of segmentation obtained using U-net++ and the proposed algorithm for patients. It is seen clearly that the proposed algorithm retains the shape of a nodule by correctly demarcating its boundaries and provides more accurate segmentation compared to U-net++. The retention of the exact shape of the cancerous nodule is of prime concern as the accuracies of the estimation of the highest dimension, volume are directly dependent on it. The spider chart showing the comparison of U-net++ and the proposed algorithm for various segmentation metrics is shown in Fig. 7. The larger difference in the areas covered by the two algorithms in the spider chart indicates the difference in their performance.

Table 1. Comparison of proposed algorithm with U-net++ for various segmentation metrics

Patients	Algorithm/ parameters	IoU	DSC	RVD	ASD	HD
CTP1	U-net++	0.8607	0.925	0.1256	0.1725	3.1623
	Proposed algorithm	0.8887	0.941	0.0874	0.1157	2.00
CTP2	U-net++	0.7723	0.872	0.2375	0.8460	15.6525
	Proposed algorithm	0.8327	0.909	0.1	0.1998	3.6056
CTP3	U-net++	0.9305	0.964	0.0529	0.2115	11.4018
	Proposed algorithm	0.9441	0.971	0.0371	0.0712	3.6056
Average for 58 patients	U-net++	0.8548	0.913	0.156	0.428	10.419
	Proposed algorithm	0.9035	0.947	0.092	0.1469	3.6056



(a) (b) (c) (d)

Fig. 6. (**a**) Original CT scan (**b**) Ground truth (**c**) result of U-net++ (**d**) result of the proposed algorithm

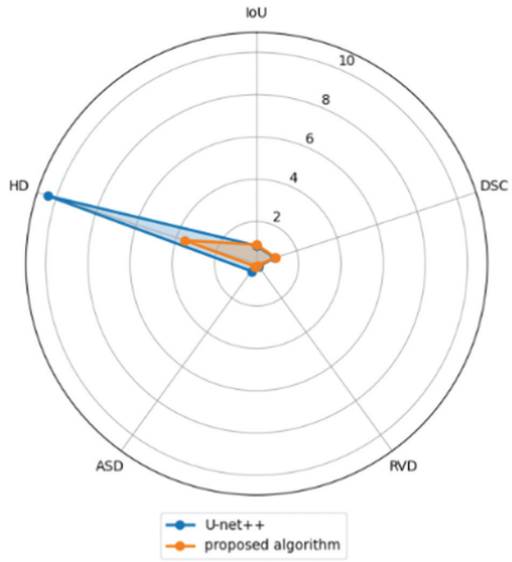


Fig. 7. Spider chart of comparison of U-net++ with the proposed algorithm

4.2. VOLUME ESTIMATION

The results of the cancer segmentation model are applied to the volume estimation model. The

segmentation model generates annotations corresponding to all the cancerous scans. The cancerous portion is demarcated in white and the area corresponding to this cancerous portion is calculated in Python. Interpolation is performed on obtained area values to get area values of unknown places between the slices. Then the volume is obtained by summing up all area values. Various types of interpolations like linear, bilinear, cubic, Lagrange, and GPR (Gaussian Process Regression) are checked in this work and their results are compared. Table 2 shows the results of these interpolation methods applied to CT scans of 3 patients. VP1, VP2, and VP3 are the volumes estimated for 3 patients P1, P2, and P3 respectively. These volumes are compared with ground truth volumes calculated by the radiologists. It is seen that there is a huge difference in calculated volume and ground truth volume when interpolation is not used. In the case of VP3, volume estimation accuracies with the quadratic, cubic, spine, Lagrange, and GPR interpolations are 88.13%, 88.91%, 90.27%, 91.63%, and 94.36% respectively. Similarly, for the other patients, it is found that GPR gives the best volume estimation accuracy with an average value of 93.13%. Fig. 8 indicates the spider chart showing the comparison of various interpolation methods. It is evident from the diagram that the GPR curve is the closest to the ground truth volume.

For comparison purposes, we applied GPR interpolation on the segmentation masks obtained using the U-net++ algorithm. Table 3 indicates volumes obtained using GPR interpolation applied to the Unet++ algorithm and our proposed algorithm for VP1, VP2, and VP3. It is seen that the accuracy of the proposed algorithm is approximately 4% higher than that of U-net++.

The comparison of results obtained is shown in Fig. 9. For all volumes VP1, VP2, and VP3 as well as for average it is seen that the volume estimated using the proposed algorithm is closer to the ground truth compared to that of the U-net++ algorithm.

Volume is one of the important parameters in deciding the stage of the cancer. It also helps radiologists to further decide about the treatment. Though cancer grading depends on multiple factors including the spread of cancer to lymph nodes, spread of cancer to another lung, etc. volume is considered to be the most important factor.

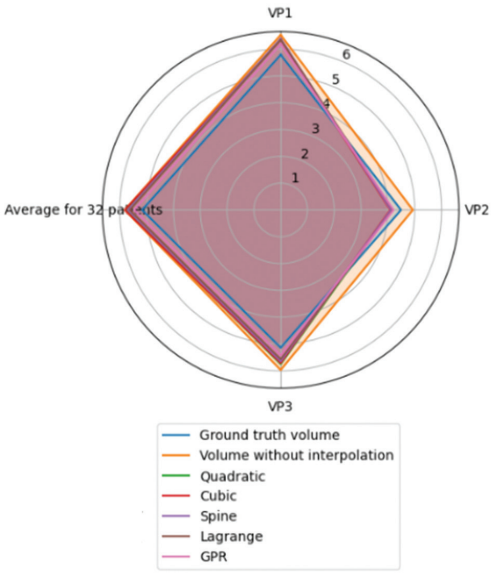


Fig. 8. Spider chart of comparison of various interpolation methods

So if volume estimation goes wrong, it may lead to the inaccurate prediction of the stage of cancer and hence the corresponding dosage and the treatment.

Keeping the aim of a fully automated end-to-end volumetric estimation system, in this work we proposed the idea of deformable convolution and GPR interpolation to achieve better accuracy of cancer nodule segmentation and volume estimation respectively.

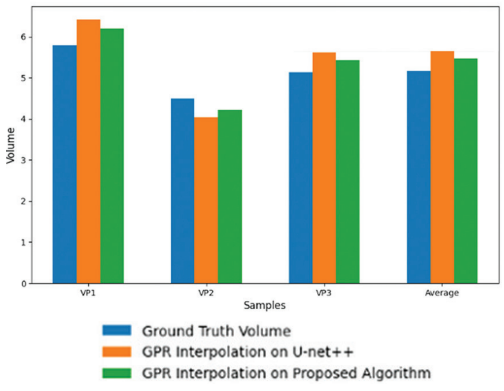


Fig. 9. Bar graph of comparison of volume estimation using U-net++ and proposed algorithm

5. CONCLUSION

This paper proposes a fully automatic system that automatically segments the cancerous portions from the applied CT scans and calculates the volume of the cancerous nodule. The proposed model uses a modified cancer segmentation algorithm which is based on deformable convolution. The modification captures the shape of the nodule more accurately as the convolution was performed at flexible locations instead of fixed locations. These locations were obtained using trainable offsets. This flexibility resulted in improved segmentation accuracy and clearly defined nodule boundaries with correct retention of the nodule shape. The average value of segmentation, an IoU of 0.9035, obtained with the proposed segmentation algorithm is nearly 5% higher than that given by the U-net++ algorithm. The work also uses GPR interpolation resulting in better volume estimation accuracy. It proved to be the best interpolation method for irregularly shaped cancer nodules leading to a final volume estimation accuracy of 93.13%. The fully automatic lung cancer volume estimation system reported here not only removes the need for radiologist intervention and thereby the resulting subjectivity but also improves the overall accuracy of lung cancer detection, segmentation, and volume estimation, leading to an improved cancer grading.

Table 2. Volume estimations (in cubic cm) for various interpolation methods

Nodule volume estimation / Type of interpolation	Ground truth volume	Volume without interpolation	Quadratic	Cubic	Spine	Lagrange	GPR	Volume estimation accuracy in % for GPR
VP1	5.79	6.53	6.37	6.36	6.39	6.35	6.192	93.06
VP2	4.49	4.93	4.13	4.09	4.12	4.18	4.21	93.77
VP3	5.14	5.97	5.75	5.71	5.64	5.57	5.43	94.36
An average of 32 patients	5.17	5.82	5.78	5.81	5.69	5.63	5.52	93.13

Table 3. Comparison of volume estimations (in cubic cm) for U-net++ and proposed algorithm

Nodule volume estimation	Ground truth volume	Volume: GPR interpolation on U-net++	Volume accuracy in % for GPR on Unet++	Volume: GPR interpolation on the proposed algorithm	Volume accuracy in % for GPR on the proposed algorithm
VP1	5.79	6.422	89.08	6.192	93.06
VP2	4.49	4.039	89.95	4.21	93.77
VP3	5.14	5.62	90.66	5.43	94.36
An average of 32 patients	5.17	5.66	90.51	5.52	93.13

6. REFERENCES

[1] M. Daykin, M. Sellathurai, I. Poole, "Comparison of unsupervised abnormality detection methods for interstitial lung disease", Communications in Computer and Information Science, Vol. 894, 2018, pp. 287-298.

[2] I. Irigoien, B. Sierra, C. Arenas, "Towards application of one-class classification methods to medical data", The Scientific World Journal, Vol. 2014, 2014.

[3] L. Tarassenko, P. Hayton, N. Cerneaz, M. Brady, "Novelty detection for the identification of masses in mammograms", Proceedings of the Fourth International Conference on Artificial Neural Networks, Cambridge, UK, 26-28 June 1995 pp. 442-447.

[4] P. Ardimento, L. Aversano, M. L. Bernardi, M. Cimitile, "Deep Neural Networks Ensemble for Lung Nodule Detection on Chest CT Scans", Proceedings of the International Joint Conference on Neural Networks, Shenzhen, China, 18-22 July 2021.

[5] M. Muzammil, I. Ali, I. U. Haq, A. A. Khaliq, S. Abdullah, "Pulmonary nodule classification using feature and ensemble learning-based fusion techniques", IEEE Access, Vol. 9, 2021, pp. 113415-113427.

[6] I. Ali, M. Muzammil, I. U. Haq, A. A. Khaliq, S. Abdullah, "Deep Feature Selection and Decision Level Fusion for Lungs Nodule Classification", IEEE Access, Vol. 9, 2021, pp. 18962-18973.

[7] D. Sethi, K. Arora, S. Susan, "Transfer Learning by Deep Tuning of Pre-trained Networks for Pulmonary Nodule Detection", Proceedings of the IEEE 15th International Conference on Industrial and Information Systems, India, 26-28 November 2020, pp. 168-173.

[8] R. Tekade, K. Rajeswari, "Lung Cancer Detection and Classification Using Deep Learning", Proceedings of the Fourth International Conference on Computing Communication Control and Automation, Pune, India, 16-18 July 2018.

[9] M. S. Rahman, P. C. Shill, Z. Homayra, "A New Method for Lung Nodule Detection Using Deep Neural

Networks for CT Images", Proceedings of the International Conference on Electrical, Computer and Communication Engineering, Cox'sBazar, Bangladesh, 7-9 February 2019.

[10] N. F. Noaman, B. M. Kanber, A. Al Smadi, L. Jiao, M. K. Alsmadi, "Advancing Oncology Diagnostics: AI-Enabled Early Detection of Lung Cancer Through Hybrid Histological Image Analysis", IEEE Access, Vol. 12, 2024, pp. 64396-64415.

[11] T. I. A. Mohamed, A. El-Shamir Ezugwu, "Enhancing Lung Cancer Classification and Prediction with Deep Learning and Multi-Omics Data", IEEE Access, Vol. 12, 2024, pp. 59880-59892.

[12] S. K. Fazilov, O. R. Yusupov, K. S. Abdiyeva, "Mammography image segmentation in breast cancer identification using the OTSU method", Web of Scientist: International Scientific Research Journal, Vol. 3, No. 8, 2022, pp. 196-205.

[13] K. V. Shiny, N. Sugitha, "Automatic brain tumor segmentation on pre-operative MRI using region growing algorithm", Information and Communication Technology for Competitive Strategies, Springer, 2022, pp. 653-665.

[14] C. J. J. Sheela, G. Suganthi, "Morphological edge detection and brain tumor segmentation in magnetic resonance (MR) images based on region growing and performance evaluation of modified fuzzy C-means (FCM) algorithm", Multimedia Tools and Applications, Vol. 79, No. 25-26, 2020, pp. 17483-17496.

[15] H. Deng, J. P. Fitts, C. A. Peters, "Quantifying fracture geometry with X-ray tomography: Technique of iterative local thresholding (TILT) for 3D image segmentation", Computers & Geosciences, Vol. 20, No. 1, 2016, pp. 231-244.

[16] Y. Akbari, H. Hassen, S. Al-Maadeed, S. M. Zughaier, "COVID-19 lesion segmentation using lung CT scan images: Comparative study based on active contour models", Applied Sciences, Vol. 11, No. 17, 2021, p. 8039.

[17] P. M. Bruntha, D. J. Rose, A. T. Shruthi, K. G. Juliet, M. Kanimozhi, "Application of selective region growing algorithm in lung nodule segmentation", Proceedings of the 4th International Conference on Devices, Circuits and Systems, 16-17 March 2018, pp. 319-322.

[18] P. B. Sangamithraa, S. Govindaraju, "Lung tumor detection and classification using EK-mean clustering", Proceedings of the International Conference on Wireless Communications, Signal Processing and Networking, 23-25 March 2016, pp. 2201-2206.

[19] M. Keshani, Z. Azimifar, F. Tajeripour, R. Boostani, "Lung nodule segmentation and recognition using SVM classifier and active contour modeling: A complete intelligent system", Computers in Biology and Medicine, Vol. 43, No. 4, 2013, pp. 287-300.

[20] W. J. Kostis, A. P. Reeves, D. F. Yankelevitz, C. I. Henschke, "Threedimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images", IEEE Transactions on Medical Imaging, Vol. 22, No. 10, 2003, pp. 1259-1274.

[21] X. Hongtao, D. Yang, N. Sun, Z. Chen, Y. Zhang, "Automated pulmonary nodule detection in CT images using deep convolutional neural networks", Pattern Recognition, Vol. 85, 2019, pp. 109-119.

[22] J. Long, E. Shelhamer, T. Darrell, "Fully convolutional networks for semantic segmentation", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7-12 June 2015, pp. 3431-3440.

[23] O. Ronneberger, P. Fischer, T. Brox, "U-Net: Convolutional networks for biomedical image segmentation", Proceedings of the Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5-9 October 2015, pp. 234-241.

[24] L.-C. Chen, G. Papandreou, F. Schroff, H. Adam, "Rethinking atrous convolution for semantic image segmentation", arXiv:1706.05587, 2017.

[25] X. Luo, T. Song, G. Wang, J. Chen, Y. Chen, K. Li,

D. N. Metaxas, S. Zhang, "SCPM-Net: An anchor-free 3D lung nodule detection network using sphere representation and center points matching", Medical Image Analysis, Vol. 75, 2022, p. 102287.

[26] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao, J. Liu, "Ce-Net: Context encoder network for 2D medical image segmentation", IEEE Transactions on Medical Imaging, Vol. 38, No. 10, 2019, pp. 2281-2292.

[27] S. Bose, R. S. Chowdhury, R. Das, U. Maulik, "Dense dilated deep multiscale supervised U-network for biomedical image segmentation", Computers in Biology and Medicine, Vol. 143, 2022, p. 105274.

[28] Q. Dou, L. Yu, H. Chen, Y. Jin, X. Yang, J. Qin, P.-A. Heng, "3D deeply supervised network for automated segmentation of volumetric medical images", Medical Image Analysis, Vol. 41, 2017, pp. 40-54.

[29] F. Milletari, N. Navab, S.-A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation", Proceedings of the Fourth International Conference on 3D Vision, Stanford, CA, USA, 25-28 October 2016, pp. 565-571.

[30] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27-30 June 2016, pp. 770-778.

[31] Y. Li, Y. Fan, "DeepSEED: 3D Squeeze-and-excitation encoder-decoder convolutional neural networks for pulmonary nodule detection", Proceedings of the IEEE 17th International Symposium on Biomedical Imaging, 3-7 April 2020, pp. 1866-1869.

[32] J. Deng, J. Guo, N. Xue, S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15-20 June 2019, pp. 4690-4699.

[33] S. G. Armato III et al. "The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans", Medical Physics, Vol. 38, No. 2, 2011, pp. 915-931.

[34] A. A. A. Setio et al. "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge", Medical Image Analysis, Vol. 42, 2017, pp. 1-13.

[35] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation", Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Springer, 2018, pp. 3-11.

[36] J. Hou, C. Yan, R. Li, Q. Huang, X. Fan, F. Lin, "Lung Nodule Segmentation Algorithm with SMR-Unet", IEEE Access, Vol. 11, 2023, pp. 34319-34331.

[37] W. Chen et al. "CT Lung Nodule Segmentation: A Comparative Study of Data Preprocessing and Deep Learning Models", IEEE Access, Vol. 11, 2023, pp. 94925-94931.

[38] W. Jiang et al. "A Dual-Branch Framework With Prior Knowledge for Precise Segmentation of Lung Nodules in Challenging CT Scans", IEEE Journal of Biomedical and Health Informatics, Vol. 28, 2024, pp. 1540-1551.

[39] F. Liao, M. Liang, Z. Li, X. Hu, S. Song, "Evaluate the Malignancy of Pulmonary Nodules Using the 3-D Deep Leaky Noisy-OR Network", IEEE Transactions on Neural Networks and Learning Systems, Vol. 30, No. 11, 2019, pp. 3484-3495.

[40] C. Shorten, T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning", Journal of Big Data, Vol. 6, No. 1, 2019, p. 60.

[41] J. Dai et al. "Deformable convolutional networks", Microsoft Research Area, June 2017.

[42] X. Huang, B. Pang, T. Zhang, G. Jia, Y. Wang, Y. Li, "Improved Prostate Biparameter Magnetic Resonance Image Segmentation based on Def-Unet", IEEE Access, Vol. 11. 2023, pp. 43089-43100.

[43] C. E. Rasmussen, C. K. I. Williams, "Gaussian Processes for Machine Learning", The MIT Press, 2006.