




# A bi-directional deep learning architecture for lung nodule semantic segmentation

Debnath Bhattacharyya<sup>1</sup> · N. Thirupathi Rao<sup>2</sup> · Eali Stephen Neal Joshua<sup>2</sup> · Yu-Chen Hu<sup>3</sup> 

Accepted: 20 August 2022 / Published online: 8 September 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

## Abstract

Lung nodules are abnormal growths and lesions may exist. Both lungs may have nodules. Most lung nodules are harmless (not cancerous/malignant). Pulmonary nodules are rare in lung cancer. X-rays and CT scans identify the lung nodules. Doctors may term the growth a lung spot, coin lesion, or shadow. It is necessary to obtain properly computed tomography (CT) scans of the lungs to get an accurate diagnosis and a good estimate of the severity of lung cancer. This study aims to design and evaluate a deep learning (DL) algorithm for identifying pulmonary nodules (PNs) using the LUNA-16 dataset and examine the prevalence of PNs using DB-Net. The paper states that a new, resource-efficient deep learning architecture is called for, and it has been given the name of DB-NET. When a physician orders a CT scan, they need to employ an accurate and efficient lung nodule segmentation method because they need to detect lung cancer at an early stage. However, segmentation of lung nodules is a difficult task because of the nodules' characteristics on the CT image as well as the nodules' concealed shape, visual quality, and context. The DB-NET model architecture is presented as a resource-efficient deep learning solution for handling the challenge at hand in this paper. Furthermore, it incorporates the Mish nonlinearity function and the mask class weights to improve segmentation effectiveness. In addition to the LUNA-16 dataset, which contained 1200 lung nodules collected during the LUNA-16 test, the LUNA-16 dataset was extensively used to train and assess the proposed model. The DB-NET architecture surpasses the existing U-NET model by a dice coefficient index of 88.89%, and it also achieves a similar level of accuracy to that of human experts.

**Keywords** Lung cancer · Deep learning · Computer-aided diagnosis · Bidirectional feature extraction · Convolutional neural network

## 1 Introduction

The mortality rate associated with lung cancer is the highest of all cancer types, making it the most dangerous form of the illness [1]. A timely diagnosis has the potential to save many lives. After breast cancer and prostate cancer, the incidence of lung cancer is the third most prevalent kind of disease seen in both men and women [2]. The International Association for the Study of Cancer (IACS) has issued the following forecast on the total number of new cases of lung cancer that will be diagnosed in the USA in the year 2020 [3]:

- In the USA, 235,760 new lung cancer cases are diagnosed each year (119,100 in men and 116,660 in women).
- Lung cancer was the cause of death for about 131,880 persons (69,410 in men and 62,470 in women)

✉ Yu-Chen Hu  
ychu@pu.edu.tw

Debnath Bhattacharyya  
debnathb@kluniversity.in

N. Thirupathi Rao  
nakkathiru@vignaniit.edu.in

Eali Stephen Neal Joshua  
stephen.eali@vignaniit.edu.in

<sup>1</sup> Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur 522 502, India

<sup>2</sup> Department of Computer Science and Engineering, Vignan's Institute of Information Technology (A), Visakhapatnam 530049, AP, India

<sup>3</sup> Department of Computer Science and Information Management, Providence University, 200, Sec. 7, Taiwan Boulevard, Shalu Dist., Taichung City 43301, Taiwan R.O.C.

Pulmonary glands, which are small, spherical lung tumors [4], can potentially develop into lung cancer if they are not detected early enough. For example, a CT scan [5] cannot detect lung cancer in its early stages because the tumors are too tiny and located in the glands. It is not until the illness has progressed to a later stage that symptoms become apparent.

Both CT and magnetic resonance imaging (MRI) [6] are well-known diagnostic tools that assist medical professionals in detecting potential issues at an earlier stage, hence increasing their ability to avert potentially fatal outcomes [7]. In the past, intelligent methods relied on manually designed feature extraction techniques, such as sequential flood feature selection algorithms (SFFSA) [8] or genetic algorithms (GA) [9], which may provide the most accessible features [10]. Deep learning has recently been used in CAD [11] systems to automatically extract image characteristics [12]. As a direct consequence of this, several approaches to processing medical images have been shown to be effective [13].

Small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) [14] are the two forms of lung cancer that are diagnosed most often. Various factors, including the following, have been linked to the development of lung cancer: Smoking creates hazardous particles that may enter the air and can be inhaled [15]. Other factors, including sex, genes, age, and exposure to second-hand smoke, also have a role. People who smoke for extended periods are the most likely to get lung cancer. Signs of lung cancer include yellow fingers, anxiety, long-term illness, tiredness, allergies, wheezing, rumbling, coughing up blood, even in small amounts, hoarseness, shortness of breath, bone pain, headache, trouble swallowing, and chest pain. Lung cancer can be found by looking for signs like these.

Many models have been developed to diagnose early-stage lung cancer, including the improved profuse clustering technique of deep learning with instantaneously trained neural networks (IPCT-DLITNN) and the adaptive hierarchical heuristic mathematical model [16]. There are three different kinds of neural networks: a deep convolutional neural network (DCNN), an artificial neural network (ANN), and an ah-ha hidden Markov model (AHHMM) [17]. The authors [18] discussed this in a piece of writing. The degree of precision and sensitivity that each model has is unique.

To counteract the challenges of effective feature mining and the adaption of such information to the collection of lung nodes, a U-NET model has been proposed [19] in which a weighted bidirectional feature network is utilized. This system applies to a variety of different types of lung nodes. The workflow is with deep learning classifiers with the classification of COVID-19 screening of the pulmonary CT scan infection [20]. The author modified the CT scans to segmented images with the support of CNN architecture [21]. Image segmentation with the computer-aided diagnosis (CADx) on the MRI and CT scans with deep U-NET

segmentation was proposed [22] and worked on the incidence and morbidity of the patients. The authors worked on the 5 K images of the CT scans, collected the samples of 5684 CT scans, and proposed a CNN architecture for accessing the patients' reports [23]. Authors worked on the CT scans of lungs and applied deep learning strategies like CNN and U-NET for the image enhancements and successfully segmented the images. From the studies [20–23], we can observe significant drawbacks that the segmentation of the images was not performed with better dice coefficient due to heavyweight architectures and higher resolution of the CT scans. We tried to outperform this in our proposed DB-NET.

The major contributions of the paper are as follows.

- Nodule volume, determining the position, is essential for lung nodule segmentation.
- Mish activation demonstrated high accuracy when compared with the activation function ReLu.
- Debnath Bhattacharyya-Network (DB-NET) segmentation mask has provided dice coefficient more precisely when compared with traditional U-NET segmentation.
- DB-NET segmentation architecture outperformed when compared with all other segmentation neural networks.

The paper is organized as follows. In Sect. 2, we discuss the background and related works. In Sect. 3, we discuss the proposed methods and architectures. In Sect. 4, we discuss the data and experiments. In Sect. 5, results and discussions are provided. The conclusions and future work are described in Sect. 6.

## 2 Related works

Convolutional neural networks are used to make U-NET. Even though this network only has 23 layers, it is not bad at all. It is not as complicated as networks that have hundreds of layers. When you have a unified network, down-sampling and up-sampling are important parts of it. Use convolutional and pooling layers to get features from the input image during the down-sampling step.

To improve the resolution of the feature map, a method called deconvolution is used. Depending on where you live, this structure is called a decoder (contraction path)–encoder (expansion path). In different ways, convolutional and pooling layers make feature maps with varying amounts of information from the images used, depending on which layer is used. Each of these feature maps is different in how detailed it is. Deconvolution is used to keep the feature map size growing after up-sampling. Then, the down-sampled feature map is merged with the original to get back the less abstract information that was lost and to improve network segmentation.

If we look at a lung CT image, you can see that the U-NET network uses two-dimensional convolution and pooling to get information about nodules. This means that a lot of spatial information has been lost. Down-sampling means that a lot of important information about where things are going is lost. When you up-sample an image, the output is fuzzier and less sensitive to the picture's attributes than the original image output. If you think about all of the problems above, it is important to use an improvised U-NET network to make things even better. Table 1 explains the research gap in the identification of the lung cancer segmentation nodules.

The authors from [24–37] have proposed their work with the same potential of data; as a result, the robustness of the model was missing. If the U-NET was applied to a different variety of data, U-NET was ultimately missing the intersection over union (IOU) and dice coefficient index accuracy. The U-NET architecture enhances the performance of both fully connected and multi-scale converting systems in terms of test results. But the major issues with the earlier models were discussed below:

- The middle strata models learn at a slower rate than the upper strata models; the network may choose to ignore the abstract layers altogether.
- In general, gradients become less noticeable as one moves away from the error calculation and training data output of a network.
- When the object of interest is in a non-standard shape or is located at a specific distance from the image, the U-NET architecture is unable to extract image-derived information.

The advantages of the proposed DB-NET model may also be able to mitigate the negative effects of decreasing gradients in the middle layers of DB-NET models, stated. According to these comparisons, DB-NETs outperform other architectures when it comes to picking up fine details in pictures. Putting together models that have nothing in common or implementing cutting-edge technologies without fully comprehending the effort required can be made quickly and easily.

When making technical decisions about model architectures, we must exercise caution and give equal weight to all model variants, especially when optimizing or disrupting models. Beyond identifying structural details such as heterochromatin concentrations and neuronal synapses, biomedical imaging has a wide range of applications in the field of medicine.

A specific error may need to be detected repeatedly on a small scale to properly configure lighting, orientation, and component sparsity for computer vision algorithms. Convolutional nets, on the other hand, can learn these characteristics without sacrificing information. However, the proposed model DB-NET was applied to a variety of data

and shows impressive results when compared with other benchmark models. Finally, we have tested our model on the diversity of data in the LUNA16 benchmark dataset.

## 3 Materials and methods

### 3.1 Proposed model architecture

#### 3.1.1 U-NET architecture

Additional layers of pooling, including max pooling, ReLU activation, concatenation, and up-sampling, are part of the U-NET model [38]. This passage is about the various ways business is affected by the slow economy. Each section contains four distinct contraction blocks. Before performing a  $2 \times 2$  max pooling, every contraction block takes in an input and applies two  $3 \times 3$  ReLU convolution layers before producing an output. As the pooling layers are stacked, the number of feature maps increases by two. The layer preceding the bottleneck [39] comprises two  $3 \times 3$  convolution layers and two  $2 \times 2$  convolution layers. A vast expanse of circuitry, each block providing input to multiple convolutional layers before sending their combined signals onto two  $3 \times 3$  convolutional layers and a  $2 \times 2$  sampling layer, constitutes the entirety of the expansion section [40, 41]. In addition, the pipeline concatenates the contracting path's feature map with the expanded path's feature map shown in Fig. 1.

#### 3.1.2 Debnath Bhattacharyya-Network (DB-NET) for bidirectional features network

In the realm of medical imaging segmentation, deep learning approaches are showing capable outcomes. U-NET [42], one of the most well-known architectural designs in the world, could be used as a Nodule Candidate Point Generation target for us. Annotated datasets are used to train these networks in this setting. No training data are required for the methods for generating candidate points utilized in the image processing. When we train our DB-NET model, we use the LUNA16 dataset. The presence of nodule sites and their radius, as well as the CT scan value used to generate the binary mask for each scan in the dataset, is all included in LUNA16. For the first topic, we would want to discuss the LUNA16 dataset's pre-processing [43]. CT scans are saved in '.mhd' files, and SimpleITK is used to import the scan image into memory. We have defined three functions for me: Each CT image in the LUNA16 dataset is labeled with nodule spots and the radius of the nodule, which are used in the binary mask generation procedure. To get things started, let us speak about how the LUNA16 dataset was pre-processed. SimpleITK is used to read the CT scans, which are saved in

**Table 1** Related research and gap identification

References	Research objective	Segmentation technique	Research Gap	Split (%)	Accuracy (%)
[24]	Dilated multi-residual blocks network based on U-NET for biomedical image segmentation	The segmented image of DC-U-NET is closer to the ground truth than Otsu, and the region is growing	The position of the CT scan is not accurate. The findings of this study hold up to minimal data	70:30	85.97
[25]	Automatic detect lung node with deep learning in segmentation and imbalance data labeling	Semantic Segmentation	The position of the CT scan is not accurate and the label goes away from the point of intersection. The findings of this study hold up to minimal data	60:40	85.57
[26]	A deep learning model to automate skeletal muscle area measurement on computed tomography images	Ensemble Learning with Semantic Segmentation	This model failed to work on the high-quality digital scans. When high-quality scans were given as input failed to give better accuracy	70:30	84.23
[27]	Multi-level Seg-Unet model with global and patch-based X-ray images for knee bone tumor detection	Patch-based X-ray images for knee bone tumor detection	The mean accuracy of the model was not benchmarked. Thus, we cannot rely on the model performance	70:30	84.81
[28]	Lung cancer and granuloma identification using a deep learning model to extract 3-dimensional radionics features in CT imaging	Tumor segmentation and radiomics feature extraction of the region of interest using gradient boosting	The prediction of the model was not standard. Working efficiently when less amount of data is fed to the classifier	60:40	83.22
[29]	ResBCDU-NET: A deep learning framework for lung CT image segmentation	Bidirectional Convolutional Long Short-term Memory is used as an advanced integrator module	This model failed in the identification of similar image densities. As a result, the model performance was not up to the mark	70:30	83.74
[30]	Automated lung segmentation on chest computed tomography images with extensive lung parenchymal abnormalities using a deep neural network	2D-UNET with three-dimension feature extraction	This model failed when the image had a group of voxels when the data were large	60:40	84.56
[31]	Conventional filtering versus U-NET-based models for pulmonary nodule segmentation in CT images	Semantic segmentation with Seg-U-NET	This model failed in time computational cost and memory computation	70:30	83.65
[32]	Efficacy evaluation of 2D, 3D U-NET semantic segmentation and atlas-based segmentation of normal lungs excluding the trachea and main bronchi	Semantic segmentation	This model requires high-performance computation equipment	70:30	82.63
[33]	Semantic segmentation and detection of mediastinal lymph nodes and anatomical structures in CT data for lung cancer staging	Pixelwise Bio-Medical Image semantic segmentation and Instance Segmentation	This model failed to identify the instances detected and represented by a 3D pixelwise mask, bounding volume, and centroid position	70:30	81.54
[34]	Fully automated lung lobe segmentation in volumetric chest ct with 3D U-NET: Validation with intra- and extra-datasets	Deep convolutional neural network for the image segmentation	The model failed to meet gold standards when it came to performance the accuracy	70:30	82.65
[35]	3D ResNetWork-I for automatic detection of Lung Nodules in CT scans	Deep convolutional neural network for the image segmentation	Global and local features in the CT scan images could not segment properly	70:30	82.36
[36]	Medical Image segmentation using Encoding and Decoding using the Deep learning approaches	X-Shape convolutional neural network	In encoding and decoding model was weakened during the insufficient data	70:30	81.67
[37]	Segmentation and detection on the unsupervised data	Image synthesis and image anomalies using U-NET	They drastically failed to identify the abnormal samples	65:35	85.69

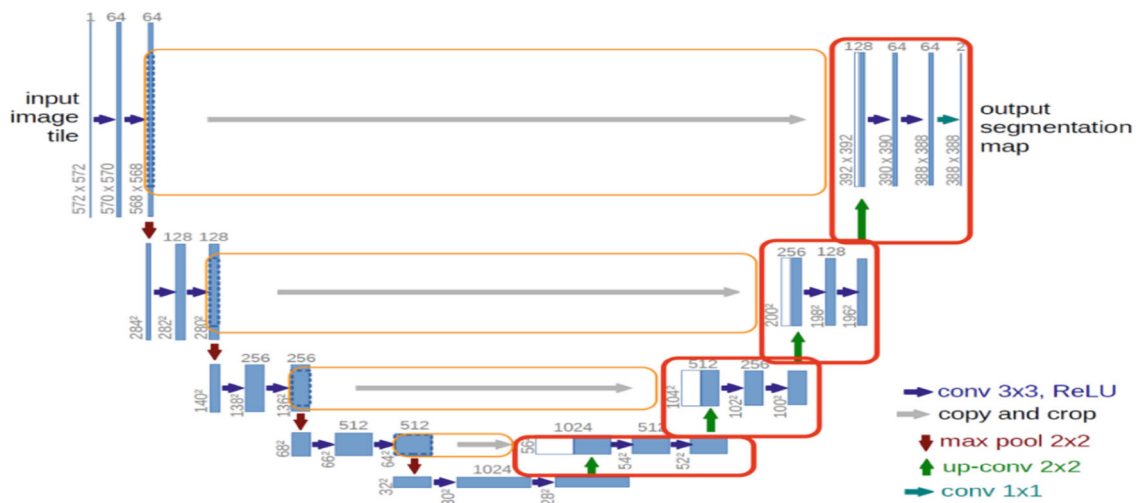


Fig. 1 The basic U-NET architecture [31]

‘.mhd’ files. The following functions are defined and used in this study.

- load\_itk—Used to read a CT\_Scan for the ‘.mhd’ file.
- world\_2\_voxel—Convert world coordinates to voxel coordinates.
- voxel\_2\_world—Convert voxel coordinates to world coordinates.

### 3.1.3 Mish activation function

Although neural networks can take advantage of the nonlinearities [44] that neurons use, neurons can use the activation function built into neural networks. Deep neural networks are effectively trained and evaluated using the capabilities they provide. The strategy implemented by this firm involves utilizing a state-of-the-art activation function known as Mish to assist in their business activities. ReLU and Swish are considered the best activation functions for datasets that are hard, even for the most challenging datasets, but this one is even better. A network based on the Mish programming language is easy to implement in neural networks, making it a particularly good network for neural networks. Figure 2 shows the nonlinearity of the Mish activation.

The self-gate has a mechanism that ensures the gate’s output will be zero if the input falls below a certain threshold. The role of self-gating in helping to prevent the overuse of ReLU-based activation functions (point-wise functions). In this instance, the gating function does not need to change the network parameters because the input to the gate is a scalar value. Mish is similar to the properties of ReLU and Swish, with a range from (0.31 to ∞).

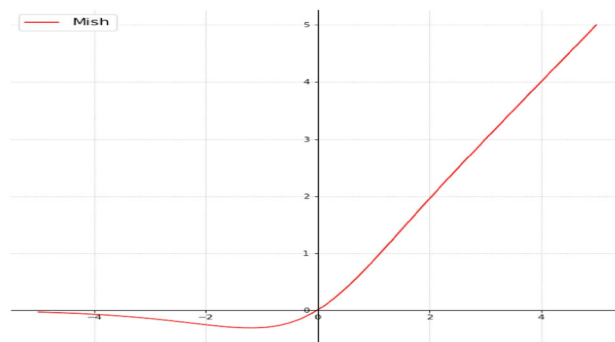


Fig. 2 The graphical representation of Mish activation

Mish activation function is smooth and non-monotonic that can be well defined as:

$$f(X) = X \cdot \tanh(\partial(x)). \tag{1}$$

$$\partial(x) = (1 + e^x). \tag{2}$$

$$f(x) = X \cdot \text{sigmoid}(x). \tag{3}$$

It combines identity, hyperbolic tangent, and softplus. We should remember the tanh and softplus functions at this point.

$$y = x \cdot \tanh(\text{softplus}(x)), \tag{4}$$

where  $\tanh(x) = (e^x - e^{-x}) / (e^x + e^{-x})$ , and  $\text{softplus}(x) = \ln(1 + e^x)$ .

Combining the above two functions, the following equation can be derived.

$$\text{mish}(x) = x \cdot (e^{\ln(1+e^x)} - e^{-\ln(1+e^x)}) / (e^{\ln(1+e^x)} + e^{-\ln(1+e^x)}). \tag{5}$$



Figure 3 shows the proposed architecture for lung cancer segmentation. Each section of the table is separated into four columns. In the first column, you can see the number of layers; in the second, you can see the parameters; in the third, you can see the activation; and in the last, you can see the output of the layers. During the investigation, U-NET identified many additional mistrustful areas than definite nodes, and so we positioned the top 8 node intrants ( $32 \times 32 \times 32$  volume) by descending a gap completing information and saving the positions of the eight most triggered (largest L2 norm) segments. Therefore, to thwart the brighter areas in the image from exclusively serving as an area for the wealthiest business interests, we decided to divide the eight sectors we identified into two non-overlapping subregion groups. When we first divided the  $64 \times 64 \times 64$  image into two distinct segments, we did so into two sections: The first section contained all relevant data, which was necessary to train our classifier; the second section contained all superfluous data, which was necessary to serve as the raw input for our classifier (cancer or not cancer). While in theory, the results of U-NET should return the precise locations of all nodules, thus enabling us to say that images that contain nodules distinguished by U-NET are optimistic for lung tumors, and images that do not contain any nodes spotted by U-NET are negative for lung malignancy, these results should not be interpreted this way in practice because for the results to be accurate, we must ensure that we perform another step between the U-NET analysis and image processing. Table 2 shows the various parameters of the proposed algorithm.

### 3.2 Data augmentation

The data augmentation process consists of three stages.

Stage 1: We started with a concept that did not have any augmented images.

Stage 2: We applied a simple color normalization augmentation.

Stage 3: We rotated 30% of the CT scan images.

The proposed model uses CT images with a size of  $512 \times 512$  pixels; a data augmentation technique was used in the place of a sample strategy to improve the proposed model's generality possible and sturdiness, size, turn, move, rotate, and elastic deformations are data augmentation methods used in the proposed network. We built a model using complex augmentations such as zooming, rotating, and cropping images. An adequate amount of training data must be available to train a DB-NET. Overfitting will occur if only a limited amount of training information is used [45]. Due to the small number of metaphors, the training data were supplemented with image editing to avoid overfitting. The images generated by the microscope are direction invariant

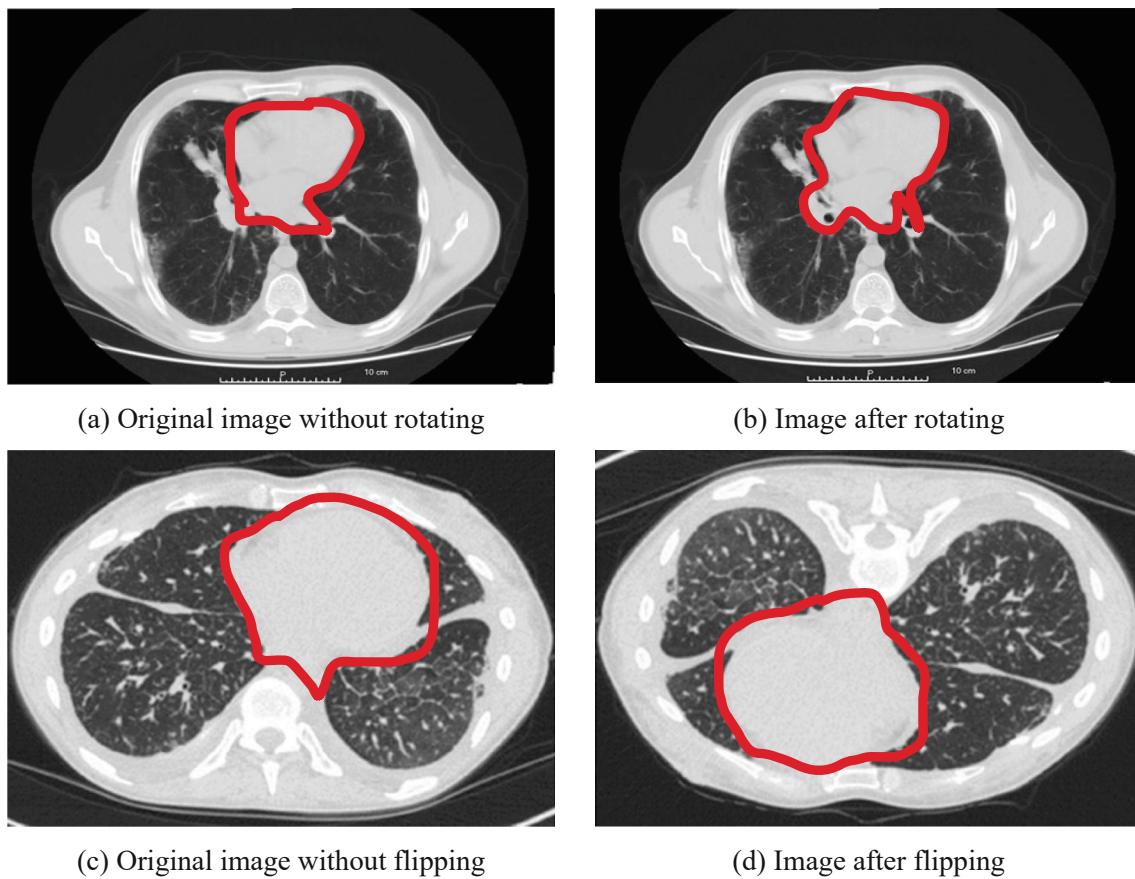
**Table 2** The proposed architecture layers and their respective parameters, activations, and output

Layers	Parameters	Activation	Output
Convolution 1A	$3 \times 3 \times 3$	Mish	$256 \times 256 \times 1$
Convolution 1B	$3 \times 3 \times 3$	Mish	$256 \times 256 \times 32$
Max Pool	$2 \times 2$ , stride 3		$256 \times 256 \times 32$
Convolution 2A	$3 \times 3 \times 3$	Mish	$128 \times 128 \times 32$
Convolution 2B	$3 \times 3 \times 3$	Mish	$128 \times 128 \times 80$
Max Pool	$2 \times 2$ , stride 3		$64 \times 64 \times 80$
Convolution 3A	$3 \times 3 \times 3$	Mish	$64 \times 64 \times 160$
Convolution 3B	$3 \times 3 \times 3$	Mish	$64 \times 64 \times 160$
Max Pool	$2 \times 2$ , stride 3		$32 \times 32 \times 160$
Bi-direction	$2D \times 5$	ReLU	$1.25 \times 10^5$
Convolution 4A	$3 \times 3 \times 3$	ReLU	$32 \times 32 \times 320$
Convolution 4B	$3 \times 3 \times 3$	ReLU	$32 \times 32 \times 320$
Up Convolution 4B	$2 \times 2$		$64 \times 64 \times 320$
Concat	Conv4b, Conv3b		$64 \times 64 \times 480$
Convolution 5A	$3 \times 3 \times 3$	ReLU	$64 \times 64 \times 160$
Conv5B	$3 \times 3 \times 3$	ReLU	$64 \times 64 \times 160$
Up Convolution 5B	$2 \times 2$		$128 \times 128 \times 160$
Concat	Conv5b, Conv2b		$128 \times 128 \times 240$
Convolution 6A	$3 \times 3 \times 3$	Mish	$128 \times 128 \times 80$
Convolution 6B	$3 \times 3 \times 3$	Mish	$128 \times 128 \times 80$
Up Convolution 6B	$2 \times 2$		$256 \times 256 \times 80$
Concat	Conv6B, Conv1B		$256 \times 256 \times 112$
Convolution 6A	$3 \times 3 \times 3$	Mish	$256 \times 256 \times 32$
Convolution 6B	$3 \times 3 \times 3$	Mish	$256 \times 256 \times 32$
Convolution 7	$3 \times 3 \times 3$		$256 \times 256 \times 2$

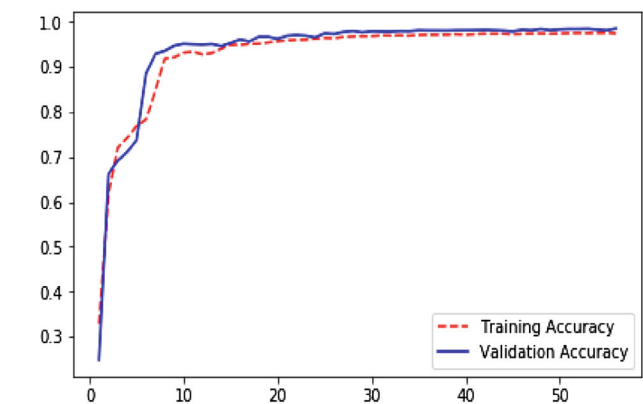
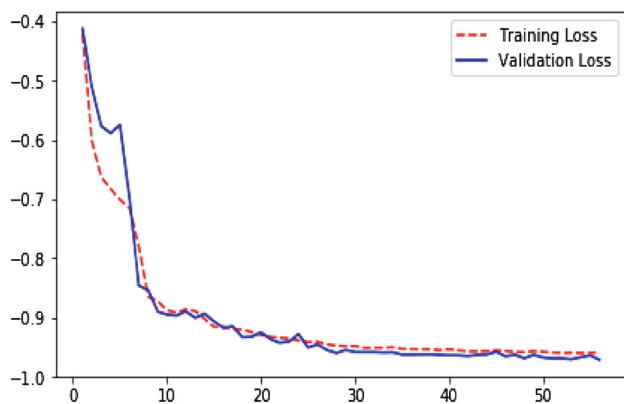
and the perceptiveness of the marked cell in each image differs depending on the conditions. Figure 4 depicts an example of image augmentation after flipping and rotating.

### 3.3 Training and post-processing

The training methods we utilized were the tenfold cross-validation to attain the precise measure capability simplifying the proposed DB-NET model. To deal with the increased training computer, tomography images generator has been employed for image augmentation of the input images and simplifying the capability of the true ground tables. During the model training and optimization, binary-weighted cross-entropy handles the imbalanced data problem where the positive classes were weighted by the negative class in the semantic segmentation training and validation. Finally



**Fig. 4** An example of image augmentation after flipping and rotating



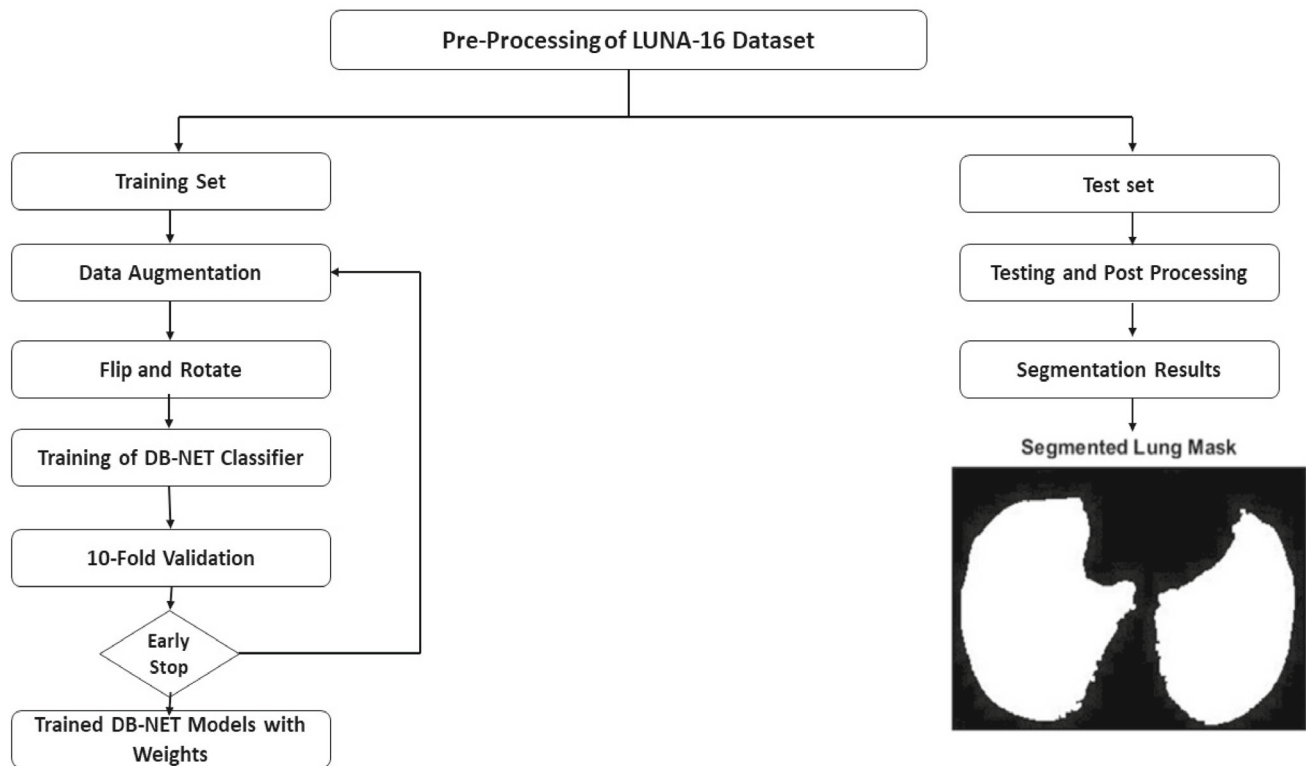
**Fig. 5** Training accuracy vs. validation accuracy on LUNA16 trail set

shown in Fig. 5, in the projected model, we have employed an optimization algorithm—“Adam”—which was used for the following restrictions: The preliminary rate of learning is 0.001, Beta1 = 0.98, Beta2 = 0.988, and the rate of decay is  $1e-7$ . Moreover, two separate batch samples were used to train the present proposed model. Additionally, a unique strategy called the early stopping strategy mechanism was

used to prevent the model from overfitting during the model training.

In the final phase, the post-processing of the proposed DB-NET model has been done. The masks were obtained after every task in raw segmentation metal format (.MHD), which is the best way to store the velocity of data such as CT scans in the system. The testing has been designed to show the best possible images, showing the segmentation results on the





**Fig. 6** Flowchart of the proposed DB-NET model

input CT scan images for the ground truth. Figure 6 depicts the detailed processes of the research progress carried out during the DB-NET model. In the starting phase, we divided the LUNA-16 dataset into 60:40 percentages and allotted 60% for training and 40% for testing the DB-NET. In the next phase, we employed the data augmentation technique to rotate and flip the CT images for better training accuracy. Once we are done with the process, the training set is given as input to our classifiers. We deployed with the tenfold cross-validation.

## 4 Data and experiments

### 4.1 Dataset

For the experiments testing of the DB-NET proposed model, the approach we utilized is the benchmark dataset available on LUNA16 (Lung Nodule Analysis 2016) [46] grand challenge. LUNA16 is resulting from the “Lung Image Database Consortiums Images Collection (LIDC/IDRI).” Input folders have three main things; one is for the sample CT scan images with sample\_1\_images. The stage\_labels folder contains the ground truth of the satge1 training set of images, and stage\_submission shows the format of the submission for stage\_1.

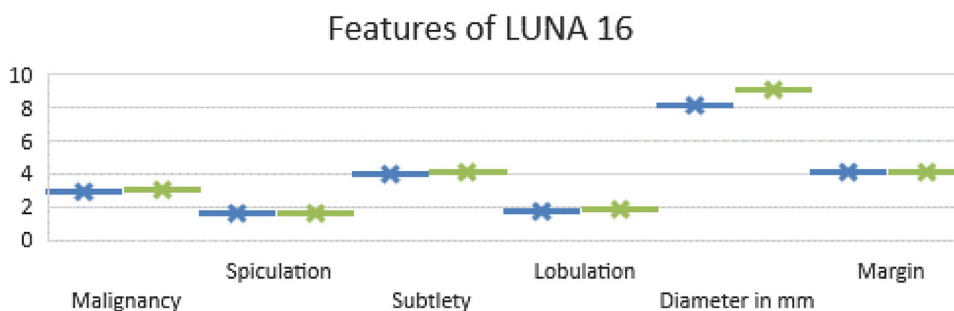
**Table 3** The features of the LUNA16 dataset in the standard deviation format

Characteristics	Training set	Testing set
Malignancy	$2.96 \pm 0.96$	$3.04 \pm 1.02$
Speculation	$1.61 \pm 0.80$	$1.66 \pm 0.88$
Subtlety	$3.92 \pm 0.84$	$4.08 \pm 0.79$
Lobulation	$1.74 \pm 0.74$	$1.83 \pm 0.81$
Diameter in mm	$8.14 \pm 4.59$	$9.08 \pm 5.25$
Margin	$4.04 \pm 0.84$	$4.07 \pm 0.78$

Table 3 shows the various feature extraction values from the LUNA16 database. Malignancy shows the range of presence of characteristics within the node. Speculation specifies the coordinates outline of the node. Subtlety is the region around the nodule. Lobulation is the shape and its characteristics of the nodule. The length of the nodule is calculated by the diameter and it is in mm. Margin indicates the area of the nodule region that is clear. Histograms of the LUNA16 dataset are shown in Fig. 7.

The benchmark dataset, the database resource initiative containing a CT scan with a slice thickness of 2.6 mm, was not included in the dataset. A total of 888 images were considered for the experimentation purpose. The images of LIDC/IDRI [47] were annotated by four experienced radiologists, and a

**Fig. 7** Histogram of the LUNA16 dataset and nodule sizes



two-phase annotation process was used for the process and it is a benchmark. The nodules of size above 3 mm were considered by all four radiologists. In total, 1186 annotations were present in the annotations file in the LUNA16 dataset and a property file that is enhanced which indicates the properties of the nodules. After post-dispensation, a total of 1167 CT scan metaphors consistent truth minced masks were portioned into two separate testing and training sets separately as 244 and 922 correspondingly. As represented in Table 4, the two sets are indistinguishable statistics distributions and their features. Table 4 shows the illustrations of various lung nodules on the LUNA16 dataset.

## 4.2 Estimation performance

The dice similarity index coefficient score is the main limitation performance matrix for the evaluation of the proposed DB-NET segmentation model. To calculate the outcome of the two segmentation, the most common performance metric was the dice similarity coefficient (DSC). And positive predictive value  $Pq$  and sensitivity were used as supplementary assessment parameters. The assessment performance system of measurement is articulated below:

$$DC = \frac{2 * V(Pq \cap Qr)}{V(Pq) + V(QR)}. \quad (6)$$

$$Sens = \frac{V(Pq \cap Qr)}{V(Pq)}. \quad (7)$$

$$PP = \frac{V(Pq \cap Qr)}{V(Qr)}. \quad (8)$$

Here “ $Pq$ ” is used to represent “ground truth label,” “ $Qr$ ” is for the results of segmentation of images, and “ $V$ ” is used for the voxels units measured in terms of volume size.


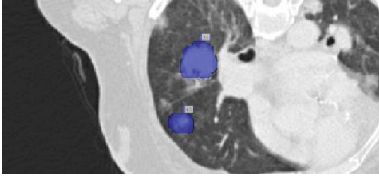
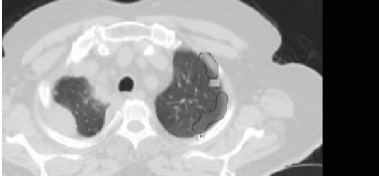
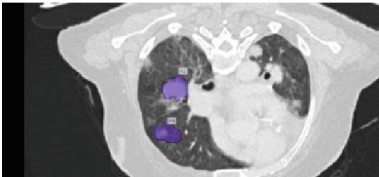
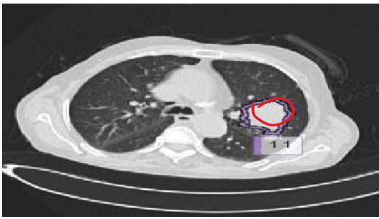
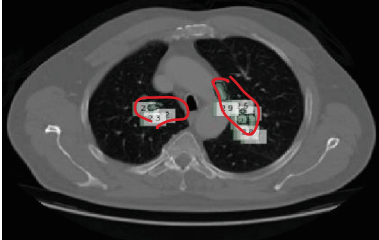
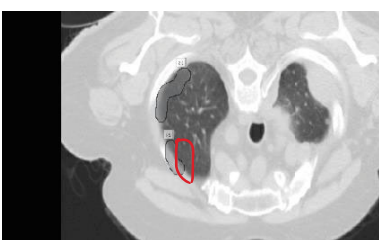
## 4.3 Execution details

In the simulations, Mish activation was utilized for efficient model training, and data augmentation was done on

the LUNA16 benchmark dataset to improve the proposed model’s performance and resilience. To avoid overfitting using the model, we implemented a new technique that involves ending strategy training early if the model’s performance does not increase. Model training will be stopped after every 20 epochs. Adam’s optimizer was utilized to get the most out of the system. All this research was done with the PyTorch 1.8 stable version of the Deep Learning Framework GPU version, Python 3.8 programming language for development, and a CUDA capable NVIDIA GPU for increased training and performance. Experiments were conducted on a Microsoft Azure infrastructure with four CPUs and a 1 TB SSD, and the training process took about nine hours to complete.

We mention techniques in the kernel to aid in a deeper understanding of the problem statement and data visualization. Matplotlib, NumPy, skimage, and pydicom are the libraries that will be used to interpret, process, and visualize data in the model. The images are ( $z, 512, 512$ ) pixels in size, with  $z$  representing the number of slices in the CT scan that varies depending on the scanner’s resolution. Because of the high computational constraints, such large images cannot be directly fed into convolution network architectures. We need to figure out which areas are more likely to develop cancer. We narrow down our search area by segmenting the lungs first and then eliminating the low-intensity areas. Because there is no homogeneity in the lung area, similar densities in the lung structures, and different scanners and scanning protocols, segmenting lung structures is a complicated subject. The segmented lungs can also be used to identify lung nodule candidates and regions of relevance that could aid in better CT scan classification. Since there are nodules attached to blood vessels or present at the lung region’s border, locating the lung nodule regions is a difficult task. Cutting 3D voxels around lung nodule candidates and moving them through a 3D CNN trained on the LUNA16 dataset can be used to further classify them. The position of the nodules in each CT scan is included in the LUNA16 dataset, which can be used to train the classifier.

**Table 4** Illustration of the various lung nodules present in the LUNA16 dataset

S. no	Nodule type	Nodule image
1	Small node	
2	GGO node	
3	Calcific node	
4	Cavitary	
5	Juxta-vascular	
6	Juxta-pleural	
7	Isolated	

## 5 Results and discussion

### 5.1 Ablation study

The ablation study experiment was based on the U-NET semantic segmentation architecture that had been planned. The ablation experiment checks whether each component of the DB-NET architecture is for the effective performance of the proposed algorithm. The experiment results of the ablation study are tabulated in Table 5.

#### 5.1.1 The outcome of mish activation function

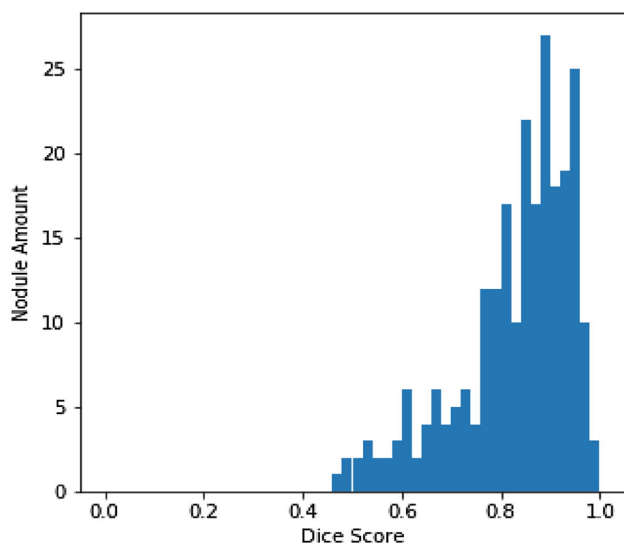
Mish activation functions were compared with the ReLU activation functions of the original U-NET architecture instead of the ReLU activation functions of the U-NET architecture. U-NET indicates the Mish activation function in conjunction with other U-NET Mish indications. The original dice similarity index score of the U-NET segmentation model is 77.84%. Subsequently, adding the Mish activation to the DB-NET architecture model achieved the dice similarity index score of 88.89%; the implementation of the solution was a success. Although the increase in performance gained when the Mish activation function was introduced to the proposed architecture was slightly mediocre, it can be seen that when these functions were employed, the performance increase was considerably better. Therefore, we have to consider the likelihood that the Mish model activation lags in the DB-NET experiment.

#### 5.1.2 Outcome of ReLU activation function

ReLU activation functions were implemented with the proposed DB-NET segmentation architecture, which performed slightly lesser to the Mish Activation. Thus, with the addition of the ReLU activation function to the DB-NET architecture, a dice coefficient of 88.89% was achieved. When we compared the ReLU with Mish activation functions, a difference of 4.38% variation can be observed. Thus, it can be observed that Mish outperformed the ReLU activation function.

#### 5.1.3 Outcome of BiFPN with ReLU activation function

From Table 2, by replacing the basic backbone of the U-NET semantic segmentation with the bi-directional feature network with ReLU activation function, the situation can be experimental that the architecture shows good development of 79.22%. Thus, we can observe that the Mish with a



**Fig. 8** Lung CT scan ratio of each image on the LUNA16 dataset

bi-direction feature network outperforms the remaining two methods. The major disadvantages of BiFPN with ReLU are computationally challenging and high overhead performance with marginally inferior dice coefficient percentage.

### 5.1.4 Deduction of the ablation study

In Table 2, with the reflection of the dice coefficient index of the DB-NET model (81.83%), it is manifest that the proposed DB-NET has shown the noteworthy development over U-NET + ReLU, U-NET + ReLU + BiFPN, and U-NET + Mish Activation. The sensitivity, positive predictive value, and dice coefficient of the proposed model were proved complete with the ablation study. Figure 8 shows the ‘histogram’ of the dice coefficient value index values. As shown in Fig. 8, the ‘histogram’ of the dice coefficient value index values and the whole quantity of nodes, each of which was counted and centered on every trial in the test set, was designed for easier assessment of the production of the DB-NET model on the test set.

## 5.2 Overall performance

In Fig. 9, we deduced the (a) clustered image after the image pre-processing. The detection results of lung nodules. (b) Results of lung nodule detection. (c) The overall effect of lung nodule segmentation. (d) The detection results of a lung nodule. (e) Local effect diagram of lung nodule segmentation. (f) Image of lung nodules whose segmentation effect is very accurate.

The majority of nodules, as shown in Fig. 9, have a DSC value of at least 0.8, which may be claimed with high confidence.

The dice index results were related to the novel performance of the U-NET architecture to validate the U-NET +

ReLU + BiFPN results. With a DSC of 77.84%, the U-NET model achieved outstanding results. But the proposed model can get even better results with a DSC of 82.82%, which is proving to be impressive in the segmentation challenge. Due to the DB-NET model’s decreased number of parameters when compared to the original U-NET design, the model has proved its capacity for competent feature abstraction and segmentation. Table 6 shows the results of the proposed algorithm.

In addition, it was also evaluated whether or not the results obtained in isolating difficult cases such as attached nodes (juxta-pleural and juxta-vascular) and nodes with a small size, such as smaller than 2 cm, could be helpful. Table 3 is the page with the standard deviation for DSC results. From the data presented in Table 5, it can be concluded that the DB-NET model’s abilities to correctly segment nodules of all different sizes are not reliant on the type of node, and it achieves remarkably well on nodes of small dimensions.

## 5.3 Visualization of results

When testing the outcomes, they found a relationship between the success of the proposed method and the effectiveness of other methods to depict that feature as shown in Table 7. Regarding the segmentation effectiveness of the five radiotherapists who worked on the LUNA16 trail dataset, the three radiologists have a radiologist segmentation effectiveness score of 81.26% and it can be seen that the DB-NET model outperforms human experts. A comparison of the proposed DB-NET model was also made with the U-NET and various additional convolution network models that have been developed recently, including the latest ResNet152V2.

The LUNA16 dataset contains stimulating study cases such as small nodules, cavitory nodes, juxta vascular, and juxta pleural nodes, and thus the performance of the DB-NET model is demonstrated in this case. As a result, it can be inferred that the anticipated prototypes for the DB-NET, the organization’s plans for it, appear well executed across various classifications of nodes, including nodes with diameters less than 6 mm.

## 5.4 Feature analysis for DB-NET architecture

This section details the filters, wraps, and embeds of these 34 features that were extracted after the semantic image segmentation. The LUNA16 dataset was chosen based on its features. The performance of the feature subsets was assessed using data from the LUNA16 testing and validation process. There was a total of 35 feature subsets evaluated. Both the MFCCs that were used and the results are included in the dataset. The entire feature set is indexed between ten and twenty times throughout the application (of 34 features). More information on the MFCC can be found in Table 8. There is a role

**Table 5** Ablation study on the LUNA16 testing set using the U-NET model

S. no	Method	Dice coefficient (%)	Sensitivity (%)	Positive predictive value (%)
1	U-NET	77.84 ± 21.79	78.98 ± 25.53	83.54 ± 22.55
2	U-NET + BFPN	81.22 ± 23.02	79.89 ± 25.85	84.89 ± 22.89
3	U-NET + ReLU	78.84 ± 12.52	84.45 ± 13.56	77.32 ± 14.45
4	U-NET + ReLU + BFPN	79.22 ± 12.36	91.69 ± 13.78	77.94 ± 14.35
5	DB-NET	88.89 ± 11.71	90.24 ± 13.15	77.92 ± 17.89

**Table 6** The segmentation results of the proposed architecture

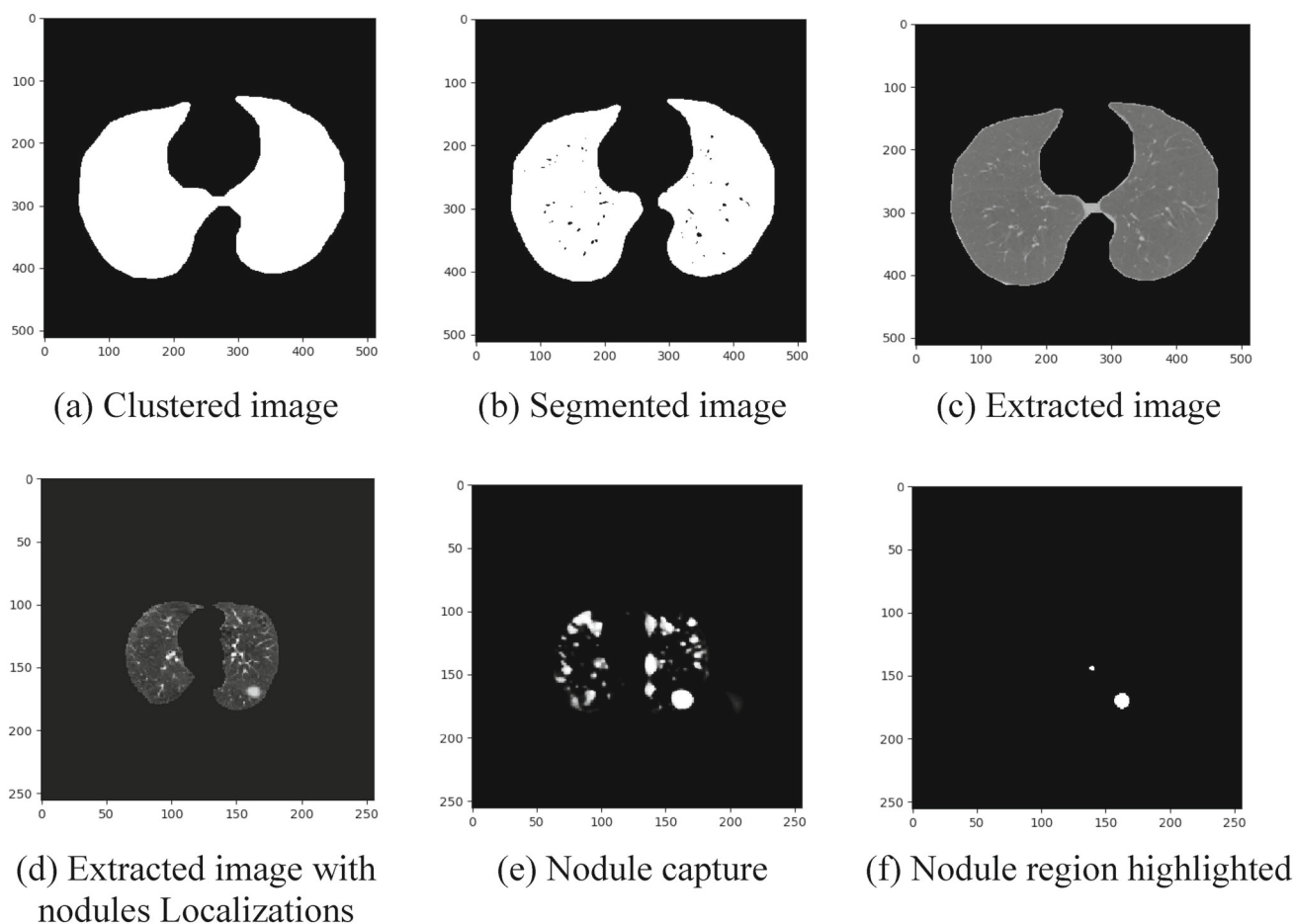
Performance	Benchmark LUNA16 dataset			
	When $N = 60$ images attached	When $N = 200$ images non-attached	Greater than 6 mm images	Less than 6 mm images
Dice coefficient	88.89 ± 11.71	90.24 ± 13.15	77.92 ± 17.89	75.63 ± 16.98

**Table 7** The measurable segmentation outcomes of the proposed model compared to other comparative models

Authors	Architectures	Dice coefficient (%)	Sensitivity (%)	Positive predictive value (%)
Zhitao Xiao et al. (2020)	3D-Res2U-NET	81.22 ± 22.02	79.89 ± 24.85	84.89 ± 22.89
Raghavendra Selvan et al. (2019)	U-NET-GNN	78.84 ± 12.52	84.45 ± 13.56	77.32 ± 14.45
Pius Kwao Gadosey et al. (2020)	Stripped-Down U-NET (SD-UNET)	79.22 ± 12.36	91.69 ± 13.78	77.94 ± 14.35
Sirojbek Safarov et al. (2021)	A-DenseUNet	80.23 ± 23.02	77.88 ± 24.85	79.89 ± 22.89
S Niranjn Kumar et al. (2021)	U-NET	77.84 ± 21.79	78.98 ± 24.53	82.54 ± 21.55
Kadia, Dhaval Dilip et al. (2021)	Advanced U-NET	79.22 ± 22.02	79.89 ± 24.85	81.89 ± 22.89
Dina M. Ibrahim et al. (2021)	ResNet152V2 + Gated Recurrent Unit (GRU)	78.22 ± 22.02	79.89 ± 24.85	82.89 ± 22.89
Proposed work	Proposed DB-NET Architecture	88.89 ± 11.71	90.24 ± 13.15	77.92 ± 17.89

**Table 8** Features selected by PySckit Library concerning performance with the proposed architecture

Model	Indices of selected features	Features	Dice coefficient	Mel frequency cepstral coefficient
3D-Res2UNET	[3,10,13,28,29,32]	[12,13,14]	81.22 ± 22.02	5
UNET-GNN	[1,3,6,9,17,19,30]	[12,13,14]	78.84 ± 12.52	7
Stripped Down UNET (SD-UNET)	[2,5,9,5,8,9,16,19]	[12,13,14]	79.22 ± 12.36	11
A Dense U-NET	[1,3,5,8,23,27,28]	[12,13,14]	80.23 ± 23.02	6
U-NET	[3,10,13,28,29,32]	[12,13,14]	77.84 ± 21.79	5
Advanced U-NET	[3,10,13,28,29,32]	[12,13,14]	79.22 ± 22.02	5
ResNet15V2_Gated Recurrent Unit (GRU)	[1,3,6,9,17,19,30]	[12,13,14]	78.22 ± 22.02	7
Proposed DB-NET Architecture	[3,10,13,28,29,32]	[12,13,14]	88.89 ± 11.71	6



**Fig. 9** The visual segmentation of the proposed algorithm on various heterogeneities of lung nodules

for MFCC in nearly every biomedical classification system [48–50]. It also includes a summary of the fundamental concepts of medical imaging sampling.

It was possible to implement a recursive function elimination algorithm using the RFE class of the PySckit image library. The estimator and the number of source functions that it can use are defined by two independent parameters each. The estimator is supervised, and the `coef_` attribute indicates how important a feature is to the estimator. The use of U-NET-GNN, Stripped-Down-U-NET, A Dense U-NET, and U-NET estimators are all examples of this. U-NET-GNN and 3D-res2U-NET are not suitable for use as estimators due to their significant characteristics.  $N$  is the number of features on which a user would like to stop and rest that is known as the "stopover parameter." We only considered the features that were the most effective to determine the best 12, 13, and 14 features for each model. The performance of the models is summarized in Table 7. The features of the RFE class are selected with RFE class support, which retrieves the feature indices that have been selected as the best. Table 7 shows the MFCCs selected for each model, as well as the indices of the 14 best features selected (via RFE) for each model.

## 6 Conclusions and future work

This work describes a simplified DB-NET architecture for lung nodule segmentation. The aim of the paper was to show how a weighted bidirectional feature network can be used to make a modified U-NET architecture that works well for lung nodule segmentation (DB-NET). The U-Net architecture is the backbone of the model, which collects and decodes feature maps. The Bi-FPN is a feature enricher that combines features from different scales. With a dice similarity coefficient of 88.89% for the LUNA16 dataset, the suggested method did a good job segmenting lung nodules after the results were looked at and shown. For example, the DB-NET model separates cavitory nodules, GGO nodules, small nodules, and juxta-pleural nodules well. Future work will focus on making a 3D capsule network based on DB-NET components for fully automated classification of lung cancer that is malignant.

**Funding** The author declares that they do not have any funding or grant for the manuscript.

## Declarations

**Conflict of Interest** The authors declare that they do not have any conflict of interests that influence the work reported in this paper.

**Ethical approval** No animals were involved in this study. All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

## References

- Center's for Disease Control and Prevention: Lung cancer statistics (2021). <https://www.cdc.gov/cancer/lung/statistics/>
- Mayo Clinic Staff: Tests and diagnosis (2021). <http://www.mayoclinic.org/diseases-conditions/lung-cancer/basics/tests-diagnosis/con-20025531>
- Long, F.: Microscopy cell nuclei segmentation with enhanced U-NET. *BMC Bioinform.* **21**, 8 (2020). <https://doi.org/10.1186/s12859-019-3332-1>
- Ayalew, Y.A., Fante, K.A., Mohammed, M.: Modified U-NET for liver cancer segmentation from computed tomography images with a new class balancing method. *BMC Biomed. Eng.* **3**, 4 (2021). <https://doi.org/10.1186/s42490-021-00050-y>
- Gianchandani, N., Jaiswal, A., Singh, D., Kumar, V., Kaur, M.: Rapid COVID-19 diagnosis using ensemble deep transfer learning models from chest radiographic images. *J. Ambient. Intell. Humaniz. Comput.* (2020). <https://doi.org/10.1007/s12652-020-02669-6>
- Satyanarayana, K.V., Rao, N.T., Bhattacharyya, D., Hu, Y.C.: Identifying the presence of bacteria on digital images by using asymmetric distribution with k-means clustering algorithm. *Multidimens. Syst. Signal Process.* (2021). <https://doi.org/10.1007/s11045-021-00800-0>
- Su, R., Zhang, D., Liu, J., Cheng, C.: MSU-net: multi-scale U-Net for 2D medical image segmentation. *Front. Genet.* **12**, 639930 (2021). <https://doi.org/10.3389/fgene.2021.639930>
- Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (2015)
- Rajagopalan, K., Babu, S.: The detection of lung cancer using massive artificial neural network based on soft tissue technique. *BMC Med. Inform. Decis. Mak.* **20**, 282 (2020). <https://doi.org/10.1186/s12911-020-01220-z>
- Joshua, E.S.N., Chakkravarthy, M., Bhattacharyya, D.: An extensive review on lung cancer detection using machine learning techniques: a systematic study. *Revued' Intelligence Artificielle* **34**(3), 351–359 (2020). <https://doi.org/10.18280/ria.340314>
- Mohammed, K.K., Hassanien, A.E., Afify, H.M.: A 3D image segmentation for lung cancer using V.Net architecture based deep convolutional networks. *J. Med. Eng. Technol.* **45**(5), 337–343 (2021). <https://doi.org/10.1080/03091902.2021.1905895>
- Baek, S., He, Y., Allen, B.G., et al.: Deep segmentation networks predict survival of non-small cell lung cancer. *Sci. Rep.* **9**, 17286 (2019). <https://doi.org/10.1038/s41598-019-53461-2>
- Sori, W.J., Feng, J., Godana, A.W., et al.: DFD-Net: lung cancer detection from denoised CT scan image using deep learning. *Front. Comp. Sci.* **15**, 152701 (2021). <https://doi.org/10.1007/s11704-020-9050-z>
- Saood, A., Hatem, I.: COVID-19 lung CT image segmentation using deep learning methods: U-NET versus SegNet. *BMC Med. Imaging* **21**, 19 (2021). <https://doi.org/10.1186/s12880-020-00529-5>
- Zhao, C., Han, J., Jia, Y., Gou, F.: Lung nodule detection via 3D U-Net and contextual convolutional neural network. *EasyChair Preprints* (2018). <https://doi.org/10.29007/bgkld>
- Hossain, S., Najeeb, S., Shahriyar, A., Abdullah, Z.R., Ariful Haque, M.: A pipeline for lung tumor detection and segmentation from CT scans using dilated convolutional neural networks. In: *2019 IEEE International Conference on Acoustics, Speech and Signal Processing* (2019). <https://doi.org/10.1109/icassp.2019.8683802>
- Zhou, Z., RahmanSiddiquee, M.M., Tajbakhsh, N., Liang, J.: UNet++: a nested U-NET architecture for medical image segmentation. In: Stoyanov, D., et al. (eds.) *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA 2018, ML-CDS 2018. Lecture Notes in Computer Science*, vol. 11045. Springer, Cham (2018). [https://doi.org/10.1007/978-3-030-00889-5\\_1](https://doi.org/10.1007/978-3-030-00889-5_1)
- Tong, G., Li, Y., Chen, H., Zhang, Q., Jiang, H.: Improved U-NET network for pulmonary nodules segmentation. *Optik* **174**, 460–469 (2018). <https://doi.org/10.1016/j.ijleo.2018.08.086>
- Chen, K., Xuan, Y., Lin, A., Guo, S.: Lung computed tomography image segmentation based on U-Net network fused with dilated convolution. *Comput. Methods Progr. Biomed.* **207**, 106170 (2021). <https://doi.org/10.1016/j.cmpb.2021.106170>
- Ruikar, D.D., Santosh, K., Hegadi, R.S., et al.: 5K+ CT Images on fractured limbs: a dataset for medical imaging research. *J. Med. Syst.* **45**(4), 51 (2021). <https://doi.org/10.1007/s10916-021-01724-9>
- Almeida, G., Tavares, J.M.R.: Deep learning in radiation oncology treatment planning for prostate cancer: a systematic review. *J. Med. Syst.* **44**(10), 179 (2020). <https://doi.org/10.1007/s10916-020-01641-3>
- Hwang, J.H., Lee, K.B., Choi, J.A., et al.: Quantitative analysis methods using histogram and entropy for detector performance evaluation according to the sensitivity change of the automatic exposure control in digital radiography. *J. Med. Syst.* **44**(10), 183 (2020). <https://doi.org/10.1007/s10916-020-01652-0>
- Agarwal, M., Saba, L., Gupta, S.K., et al.: A novel block imaging technique using nine artificial intelligence models for COVID-19 disease classification, characterization and severity measurement in lung computed tomography scans on an italian cohort. *J. Med. Syst.* **45**(3), 28 (2021). <https://doi.org/10.1007/s10916-021-01707-w>
- Yang, J., Zhu, J., Wang, H., Yang, X.: Dilated MultiResUNet: Dilated multiresidual blocks network based on U-Net for biomedical image segmentation. *Biomed. Signal Process. Control* **68**, 102643 (2021). <https://doi.org/10.1016/j.bspc.2021.102643>
- Chiu, T.W., Tsai, Y.L., Su, S.F.: Automatic detect lung node with deep learning in segmentation and imbalance data labeling. *Sci. Rep.* **11**, 11174 (2021). <https://doi.org/10.1038/s41598-021-90599-4>
- Amarasinghe, K.C., Lopes, J., Beraldo, J., Kiss, N., Bucknell, N., Everitt, S., Jackson, P., Litchfield, C., Denehy, L., Blyth, B.J., Siva, S., Michael, M., David, B., Jason, L., Nicholas, H.: A deep learning model to automate skeletal muscle area measurement on computed tomography images. *Front. Oncol.* **11**, 135 (2021). <https://doi.org/10.3389/fonc.2021.580806>
- Do, N.-T., Jung, S.-T., Yang, H.-J., Kim, S.-H.: Multi-level Seg-Net model with global and patch-based X-ray images for knee bone tumor detection. *Diagnostics* **11**(4), 691 (2021). <https://doi.org/10.3390/diagnostics11040691>. (MDPI AG)
- Lin, X., Jiao, H., Pang, Z., Chen, H., Wu, W., Wang, X., Xiong, L., Chen, B., Huang, Y., Li, S., Li, L.: Lung cancer and granuloma identification using a deep learning model to extract 3-Dimensional radionics features in CT imaging. *Clin. Lung Cancer* **22**(5), e756–e766 (2021). <https://doi.org/10.1016/j.clcc.2021.02.004>

29. Jalali, Y., Fateh, M., Rezvani, M., Abolghasemi, V., Anisi, M.H.: ResBCDU-NET: a deep learning framework for lung CT image segmentation. *Sensors* **21**(1), 268 (2021). <https://doi.org/10.3390/s21010268>. (MDPI AG)
30. Yoo, S.J., Yoon, S.H., Lee, J.H., Kim, K.H., Choi, H.I., Park, S.J., Goo, J.M.: Automated lung segmentation on chest computed tomography images with extensive lung parenchymal abnormalities using a deep neural network. *Korean J. Radiol.* **22**(3), 476–488 (2021). <https://doi.org/10.3348/kjr.2020.0318>
31. Rocha, J., Cunha, A., Mendonça, A.M.: Conventional filtering versus U-NET based models for pulmonary nodule segmentation in CT images. *J. Med. Syst.* **44**, 81 (2020). <https://doi.org/10.1007/s10916-020-1541-9>
32. Nemoto, T., Futakami, N., Yagi, M., Kumabe, A., Takeda, A., Kunieda, E., Shigematsu, N.: Efficacy evaluation of 2D, 3D U-Net semantic segmentation and atlas-based segmentation of normal lungs excluding the trachea and main bronchi. *J. Radiat. Res.* **61**(2), 257–264 (2020). <https://doi.org/10.1093/jrr/rzz086>
33. Bouget, D., Jørgensen, A., Kiss, G., et al.: Semantic segmentation and detection of mediastinal lymph nodes and anatomical structures in CT data for lung cancer staging. *Int. J. Comput. Assist. Radiol.* **14**, 977–986 (2019). <https://doi.org/10.1007/s11548-019-01948-8>
34. Park, J., Yun, J., Kim, N., et al.: Fully automated lung lobe segmentation in volumetric chest ct with 3D U-NET: validation with intra- and extra-datasets. *J. Digit Imaging* **33**, 221–230 (2020). <https://doi.org/10.1007/s10278-019-00223-1>
35. Shi, L., Ma, H., Zhang, J.: Automatic detection of pulmonary nodules in CT images based on 3D res-I network. *Vis. Comput.* **37**(6), 1343–1356 (2021). <https://doi.org/10.1007/s00371-020-01869-7>
36. Li, Y., Wang, Z., Yin, L., Zhu, Z., Qi, G., Liu, Y.: X-net: a dual encoding–decoding method in medical image segmentation. *Vis. Comput.* (2021). <https://doi.org/10.1007/s00371-021-02328-7>
37. Dai, W., Erdt, M., Sourin, A.: Detection and segmentation of image anomalies based on unsupervised defect repair. *Vis. Comput.* **37**(12), 3093–3102 (2021). <https://doi.org/10.1007/s00371-021-02257-5>
38. Jin, J., Zhu, H., Zhang, J., Ai, Y., Zhang, J., Teng, Y., Xie, C., Jin, X.: Multiple U-Net-based automatic segmentations and radiomics feature stability on ultrasound images for patients with ovarian cancer. *Front. Oncol.* **10**, 614201 (2021). <https://doi.org/10.3389/fonc.2020.614201>
39. Rahman, M.F., Tseng, T.-L.B., Pokojovy, M., Qian, W., Totada, B., Xu, H.: An automatic approach to lung region segmentation in chest X-ray images using adapted U-Net architecture. *Phys. Med. Imaging* (2021). <https://doi.org/10.1117/12.2581882>
40. Ratnaraj, R.R., Thamilarasi, V.: U-NET: convolution neural network for lung image segmentation and classification in chest X-ray images. *INFOCOMP J. Comput. Sci.* **20**(1), 101–108 (2021)
41. Ghazipour, A., Veasey, B., Seow, A., Amini, A.: 3D U-net for registration of lung nodules in longitudinal CT scans. In: *Medical Imaging 2021: Computer-Aided Diagnosis* (2021). doi:<https://doi.org/10.1117/12.2581755>
42. Ronneberger, O., Fischer, P., Brox, T.: U-NET: convolutional networks for biomedical image segmentation olaf ronneberger, philipp fischer. In: *Thomas Brox Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Springer, LNCS, vol. 9351, pp. 234–241 (2015). [arXiv:1505.04597](https://arxiv.org/abs/1505.04597)
43. Lin, L., Wu, J., Cheng, P., Wang, K., Tang, X.: BLU-GAN: bi-directional ConvLSTM U-NET with generative adversarial training for retinal vessel segmentation. In: Gao, W., et al. (eds.) *Intelligent Computing and Block Chain. FICC 2020. Communications in Computer and Information Science*, vol. 1385. Springer, Singapore (2021). [https://doi.org/10.1007/978-981-16-1160-5\\_1](https://doi.org/10.1007/978-981-16-1160-5_1)
44. Kabir, S., Sakib, S., Hossain, M. A., Islam, S., Hossain, M.I.: A convolutional neural network based model with improved activation function and optimizer for effective intrusion detection and classification. In: *2021 International Conference on Advance Computing and Innovative Technologies in Engineering* (2021). <https://doi.org/10.1109/icacite51222.2021.9404584>
45. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. *J. Big Data* **6**, 60 (2019). <https://doi.org/10.1186/s40537-019-0197-0>
46. Neal Joshua, E.S., Bhattacharyya, D., Chakkravarthy, M., Byun, Y.-C.: 3D CNN with visual insights for early detection of lung cancer using gradient-weighted class activation. *J. Healthc. Eng.* (2021). <https://doi.org/10.1155/2021/6695518>
47. Cui, L., Li, H., Hui, W., et al.: A deep learning-based framework for lung cancer survival analysis with biomarker interpretation. *BMC Bioinform.* **21**, 112 (2020). <https://doi.org/10.1186/s12859-020-3431-z>
48. Liu, X., He, J., Song, L., Liu, S., Srivastava, G.: Medical image classification based on an adaptive size deep learning model. *ACM Trans. Multimed. Comput. Commun. Appl.* **17**(3s), 102 (2021). <https://doi.org/10.1145/3465220>
49. Wang, Y., Feng, Z., Song, L., Liu, X., Liu, S.: Multiclassification of endoscopic colonoscopy images based on deep transfer learning. *Comput. Math. Methods Med.* **2021**, 2485934 (2021). <https://doi.org/10.1155/2021/2485934>
50. Liu, X., Song, L., Liu, S., Zhang, Y.: A review of deep-learning-based medical image segmentation methods. *Sustainability* **13**(3), 1224 (2021). <https://doi.org/10.3390/su13031224>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.



**Dr. Debnath Bhattacharyya** received Ph.D. (Tech., CSE) from the University of Calcutta, Kolkata, India. Currently, Dr. Bhattacharyya is associated with K L Deemed to be University, KLEF, Guntur, AP, India, as a Professor of Computer Science and Engineering. His research areas include Image Processing, Pattern Recognition, Bio-Informatics, Computational Biology, Evolutionary Computing, and Security. He published 200 + research papers in various reputed International Journals and Conferences. He published six textbooks for Computer Science as well. He is a member of IEEE, ACM, ACM SIGKDD, IAENG, and IACSIT.





**Dr. N. Thirupathi Rao** received Ph.D. from Andhra University, Visakhapatnam, India. Currently, Dr. N. Thirupathi Rao is associated with Vignan's Institute of Information Technology, Visakhapatnam-530049, India, as Associate Professor of Computer Science and Engineering of the Institute year 2015. His research areas include Image Processing, Pattern Recognition, Queuing Models, Stochastic Modeling, and Soft Computing. He published 85 + research papers in

various reputed International Journals and Conferences. He is a member of ACM, IE, CSI, and ISPS.



**Dr. Eali Stephen Neal Joshua** received Ph.D. from Lincoln University College, Kuala Lumpur, Malaysia. Currently, Dr. E. Stephen Neal Joshua is associated with Vignan's Institute of Information Technology(A), Visakhapatnam-530049, India, as Assistant Professor of Computer Science and Engineering of the Institute year 2017. His research areas include Deep Learning, Bio-Medical Image Analysis, Pattern Recognition, Bio-Informatics, Evolutionary

Computing. He published 10 + research papers in various reputed International Journals and Conferences. He is a member of ACM, IAENG, CSI, and ISPS.



**Yu-Chen Hu** received his Ph.D. Degree in computer science and information engineering from the Department of Computer Science and Information Engineering, National Chung Cheng University, Chiayi, Taiwan. Currently, Dr. Hu is a professor in the Department of Computer Science and Information Management, Providence University, Taichung City, Taiwan. He is a senior member of IEEE. He joins the editorial boards of *Advances in Multimedia*, *Electronics*, *IET Image*

*Processing*, *Mathematical Problems in Engineering*, etc. His research interests include digital forensics, information hiding, image and signal processing, data compression, information security, computer network, artificial intelligence, and information management.