IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# PULMONARY CT SCAN IMAGE SEGMENTATION

Boddu Srivatsa[1], Gousiya Begum [2]

[1] B.Tech Student, Department of CSE, MGIT, Hyderabad, bsrivatsa_cse180566@mgit.ac.in

[2] Assistant Professor, Department of CSE, MGIT, Hyderabad, gousiyabegum_cse@mgit.ac.in

***Abstract*--** For any medical diagnosis, to predict the lung diseases and for medication purposes which are needed to cure or to prevent the lung diseases, medical scans are used in the diagnosis stage. In order to do this, it is a prerequisite step to provide an accurate lung CT image analysis such as lung cancer detection.

Traditional method follows with human interaction where the doctor is supposed to examine the lungs region from the CT scan images of the lung. This may lead to several errors in predicting the disease or in medication that is to be given to the patient. This can be overcome by using the computer technology image segmentation where lung CT scan images can be segmented based on the requirement.

In this project, lung CT image segmentation is proposed by U-Net architecture rather than segmenting with the traditional algorithms which are already implemented.

***Keywords*—** Image Segmentation; U-Net Architecture; Deep Learning

## I. INTRODUCTION

The network assigns a label (or class) to each input image in an image classification task. But what if you wanted to know the shape of that object, which pixel belongs to which object, and so on? Segmentation is the term for this process. Medical imaging, self-driving cars, and satellite imaging are just a few of the applications for image segmentation.

Lung's segmentation is the process of focusing upon similar components and eliminating contrasting elements of the lung region. Lung Segmentation can be performed using machine learning, deep learning techniques or conventional methods. Conventional methods concern only about the effectiveness of algorithms as machine learning requires human efforts for feature extraction. Deep learning takes machine learning to the next step as features in the data are extracted by automation. One of the models of deep learning is the Convolutional Neural Networks (CNNs). These Convolutional Neural Networks are designed likewise in a manner firstly to reduce the size of the image in consecutive steps. This might generate losses as important information which is required might be lost. To overcome this issue, an alternate model of deep learning has been introduced which is the "U-Net Model". It was primarily developed for image segmentation on biomedical images. UNet architecture is different from the CNN architecture as UNet performs resizing of images as shrinking the image for top-bottom path and expanding for bottom-up path. This process helps in finding the features predominately and thus classify the images accurately.

Image segmentation of lungs using the appropriate Deep Learning models helps to segment the lung region which can be used in predicting any lung disease incurred and the occurrence of lung abnormalities. Segmentation is concentrated only on the pulmonary region. The lung segmentation plays a crucial role in medical applications as improper segmentation can lead to incorrect medical predictions.

The proposed work is as follows: Related work on lung segmentation has been discussed in the UNet model.
Some previously used techniques to segment images the pulmonary CT scan images are the KNN (K-nearest neighbor algorithm), where the segmentation process is not implemented using the neural networks.

## II. RELATED WORK

In 2021, "Lung CT image segmentation using Reinforcement Learning" was published by Parnia Gheysari, Mansoor Fateh and Mohen Rezvani. The techniques used are Q-learning algorithm and edge detection. The advantages are: Q-learning method produces the image without noise. The disadvantage is DQN requires an adequate amount of time to train a reinforcement learning agent and also the rewards for the action and environment must be defined effectively [1].

"Automatic clustering method to segment COVID-19 CT images" is developed by Mohamed Abd Elaziz, Mohammed A. A. AI-qaness, Songfeung Lu, Rehab Ali Ibrahim, Ahmed A. Ewees. It was published in 2021.They have used density peaks clustering (DPC) using generalized extreme value (GEV) distribution techniques. The drawback is it is difficult to determine the optimal number of clustering centers automatically without visualization [2].

"MSU-Net: Multi-Scale U-Net for 2D Medical Image Segmentation" is developed by Run Su, Deyun Zhang, Jinhuai Liu and Chuandong Cheng. It was published in 2021. They have used U-Net, multi-scale U-Net (MSU-Net). The drawback is that the feature maps which are extracted from the convolution kernel for different receptive fields are of the same scale [3].

"Eff-UNet: A Novel Architecture for Semantic Segmentation in Unstructured Environment" is developed by Bhakti Baheti, Shubham Innani, Suhas Gajre, Sanjay Talbar. They have used Encoders and Decoders. Advanced Driver Assistance Systems (ADAS). The drawback is that the learning process may be slowed down in the middle layers of deeper models, so there is some risk that the network is learning to ignore the layers where abstract features are represented [4].

## III. SYSTEM DESIGN

Scene interpretation, medical image analysis, robotic perception, video surveillance, augmented reality, and picture compression are all applications of image segmentation in image processing and computer vision. Various picture segmentation algorithms have been developed in the literature. We look at the similarities, strengths, and problems of different deep learning models, as well as the most commonly used datasets, describe results, and identify possible future research avenues in this field.

In our system, initially the image dataset's features are extracted by using a suitable feature extraction where U-Net is used, then all these features are stored in the image database. In a similar fashion the query image given by the user also undergoes feature extraction and its features are also saved.

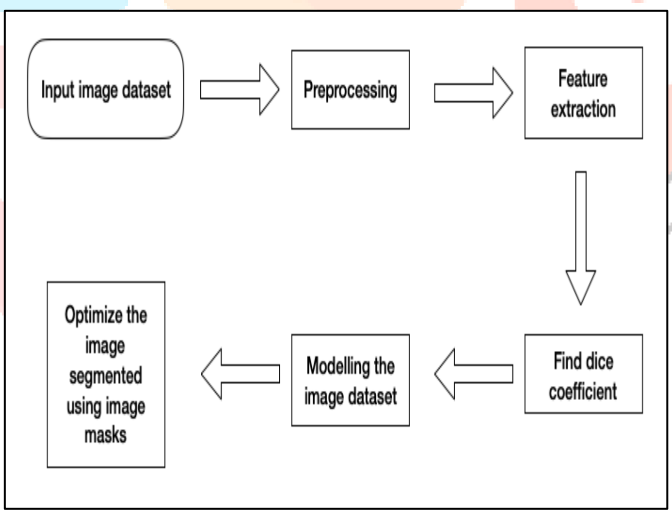


**Figure 1** System Architecture for Pulmonary CT Scan Image Segmentation

In Figure 1, initially the image dataset's features are extracted by using a suitable feature extraction where U-Net is used, then all these features are stored in the image database. In a similar fashion the query image given by the user also undergoes feature extraction and its features are also saved.

### 3.1 FEATURE EXTRACTION

The process of transforming raw data into numerical features that can be processed while preserving the information in the original data set is known as feature extraction. It produces better results than applying machine learning to raw data directly. Convolutional neural networks extract the relevant features of the input images and another neural network classifies the image features. The extracted features are processed as signals and are utilized by the neural networks for the process of classification.

Feature extraction can be done manually or automatically, as follows:

1. Identifying and describing the features that are relevant for a given problem, as well as implementing a method to extract those features, are all part of manual feature extraction. In many cases, knowing the background or domain can assist in making informed decisions about which features might be useful. Engineers and scientists have developed feature extraction methods for images, signals, and text after decades of research. The mean of a signal's window is an example of a simple feature.
2. Automated feature extraction extracts feature from signals or images without the need for human intervention using specialized algorithms or deep networks. When you need to move quickly from raw data to developing machine learning algorithms, this technique can be very useful. Automated feature extraction is an example of wavelet scattering.

Feature extraction has been largely replaced by the first layers of deep networks with the rise of deep learning – but mostly for image data. Feature extraction remains the first challenge in signal and time-series applications, requiring significant expertise before building effective predictive models.

3. After the images in the dataset are imported, the images are converted to a grayscale standard form of type float-32. The images are converted to gray scale with height and width constraints. Feature extraction is done on pixels using the cv2 INTER_LANCZOS4 and INTER_NEAREST. INTER_LANCZOS4 takes all the current value and neighbor pixel value, INTER_NEAREST takes all the neighboring pixels of the lung CT scan image. INTER_NEAREST uses nearest neighbor interpolation. INTER_LANCZOS4 is a sinusoidal method.

### 3.2 CONVOLUTIONAL LAYERS

Around the 1980s, CNNs were first developed and used. At the time, the most a CNN could do was recognize handwritten digits. It was primarily used in the postal sector to read zip codes, pin codes, and other similar information. The most important thing to remember about any deep learning model is that it requires a lot of data and a lot of computing resources to train. This was a major disadvantage for CNNs at the time, so they were limited to the postal sector and never made it into the world of machine learning.

Convolutional networks have shown to be the most exact and accurate machine learning approaches in difficult applications. The name "convolutional neural network" implies that the network works with convolution as a mathematical operation. The mathematical operation used in this implementation is mathematically linear. Convolutional networks use mathematically-linear convolution operations in at least one of their layers. Convolutional neural networks are neural networks that retrieve activation values from a volume for another volume layer using a convolution approach. Convolutional networks are a series of layers in which each layer converts one activation volume to a different activation volume with distinct functions. Convolutional layers are divided into three categories, each of which contains three layers: the convolution layer (conv layer), the fully connected layer, and the pooling layer. The dot product of the filter's output volume and the input volume aids in the creation of activation mapping in the forward pass using a convolutional layer. Later, activation of values is performed using the ReLU function to reduce negative values. To reduce the size, the pooling method is used. This process is iterated recursively, with the number of repetitions determined arbitrarily.

Input Layers: This is the layer where we give our model input. The total number of features in our data is equal to the number of neurons in this layer (number of pixels in the case of an image).

Hidden Layer: The hidden layer receives the input from the input layer. Depending on our model and data size, there could be a lot of hidden layers. The number of neurons in each hidden layer can vary, but they are usually more than the number of features. The output from each layer is computed by matrix multiplication of the previous layer's output with that layer's learnable weights, followed by the addition of learnable biases, and finally the activation function, which makes the network nonlinear.

The output of the hidden layer is then fed into a logistic function like sigmoid or SoftMax, which converts each class's output into a probability score for each class.

### 3.3 U-NET ARCHITECTURES (BASE U-NET)

U-net is a type of deep neural network that can be broadly thought of as an encoder and decoder for semantic images. U-net was originally invented and first used for biomedical image segmentation. Unlike classification where the end result of the deep network is the only important thing, semantic segmentation requires discrimination at pixel level as well as a mechanism to project the discriminative features learnt at different stages of the encoder onto the pixel space.

The encoder is the first half of the architecture, where convolution blocks are used to encode the input image into feature representations at many levels, followed by maxpool downsampling. Upsampling and concatenation precede conventional convolution operations in the decoder. The goal is to generate a dense classification by projecting the encoder's discriminative features (lower resolution) onto the pixel space (higher resolution).

There are two elements to the U-net network: The first is the standard CNN structure. The contracting path is made up of two 3x3 convolutions followed by a ReLU activation unit and a max-pooling layer in each block. This pattern is repeated multiple times. The actual context of U-net is found in the expansive route, in which each stage uses 2x2 up-convolution to resample the feature map. The feature map from the contracting path's corresponding layer is then cropped and concatenated into the upsampled feature map. Then there are two 3x3 convolutions in a row, followed by ReLU activation. To decrease the feature map to the appropriate number of channels and produce the segmented image, an additional 1x1 convolution is done at the end. Cropping is strongly recommended in order to reduce edge pixel features, which contain little contextual information. This creates a u-shaped network that transfers contextual information along it, allowing the model to separate items in a given area using context from a larger overlapping area.[5].

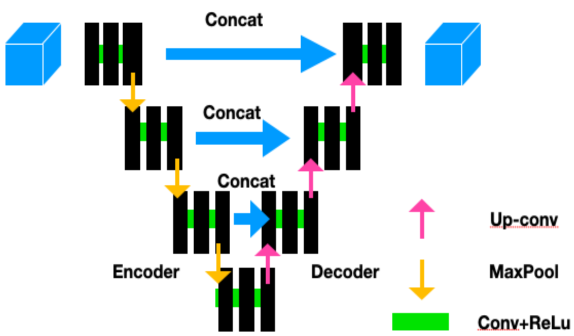


**Figure 2** Architecture of CNN

Figure 2 represents the architecture of Convolutional Neural Networks. The number of repetitions and steps are mentioned in the figure 2.

### 3.4 OPTIMIZATION

As the name implies, optimization is a method of solving a problem by fine-tuning a set of parameters to achieve an optimal solution toward a defined goal. Finding the maximum or minimum value of a function, for example, may be of interest. The function could be linked to a real-world issue, and the possibilities are endless. For example, the function could represent a company's profit, which should be maximized, or it could represent expenses, which should be minimized. The goal of using optimization techniques is to fine-tune certain parameters that affect profit and expenses, allowing for function maximization or minimization. "Optimization" refers to a technique for determining a set of input values for a function that yields a maximum or minimum function evaluation. Optimizers are algorithms or approaches that adjust the characteristics of a neural network, such as weights and learning rate, to reduce losses. These are in charge of minimizing losses and obtaining the most accurate outcomes possible. With the use of 2d masks pictures in the dataset, the images that are iteratively learned using the convolutional layer are further improved using Adam's optimiser.

- **Adaptive Gradient Algorithm:** The Adaptive Gradient algorithm, also known as AdaGrad, is a variant of the gradient descent optimization algorithm. It is intended to speed up the optimization process by reducing the number of function evaluations needed to reach the optimal solution. John Duchi and colleagues described the algorithm in their 2011 paper "Adaptive Subgradient Methods for Online Learning and Stochastic Optimization.". The gradient descent algorithm has the drawback of having the same step size (learning rate) for each variable or dimension in the search space. It's possible that a step size tailored to each variable will result in better performance, allowing for larger movements in dimensions with a consistently steep gradient and smaller movements in dimensions with less steep gradients. AdaGrad was created to test the idea of tailoring the step size for each dimension in the search space automatically. This is accomplished by first calculating a step size for a given dimension, and then using that step size to move that dimension using the partial derivative. After that, the procedure is repeated for each dimension in the search space.
- **Root Mean Square Propagation:** Root Mean Square Propagation (RMSProp) keeps per-parameter learning rates that are adjusted based on the average of recent gradient magnitudes for the weight. When used on online and non-stationary issues, this method works best. Propagation of the Root Mean Square RMS Prop is a technique for dampening y-axis motion and speeding up gradient descent, similar to Momentum. Let's call the Y-axis the bias b and the X-axis the weight W for a better understanding. Because we square the derivatives of both the w and b parameters, it's called Root Mean Square.

- RMSProp is intended to speed up the optimization process, for example, by reducing the number of function evaluations needed to reach the optimal solution, or to improve the capability of the optimization algorithm, for example, by producing a better final result.

It's related to Adaptive Gradient, or AdaGrad, a variant of gradient descent that's designed to investigate the idea of automatically tailoring the step size (learning rate) for each parameter in the search space. This is accomplished by first calculating a step size for a given dimension, and then using that step size to move that dimension using the partial derivative. After that, the procedure is repeated for each dimension in the search space. Adam combines the advantages of both the AdaGrad optimizer and the RMSProp optimizer.

The technique creates an exponential moving average of the gradient and the squared gradient, with the decay rates of these moving averages controlled by the constants beta1 and beta2.

- The moving averages' starting values, as well as beta1 and beta2 values, all tend to 1.0, resulting in a bias of moment estimates towards zero. Instead of calculating bias-corrected estimates, this bias is eliminated by first calculating biased estimates [6].

## IV. RESULTS & DISCUSSION

### 4.1 RESULTS

The pulmonary CT images are segmented using U-Net Architecture and are optimized using the Adam's Optimizer. In implementing the U-Net Architecture the image is best segmented if it gains maximum dice coefficient value.

The measure of similarities between two sets of values is called Dice coefficient (DC). It is given by the following equation

$$\text{Dice Coefficient} = 2*|AB|/( |A| + |B)|$$

where A is ground-truth and B is predicted values, |AB| is the intersection of elements and |A| and |B| are the cardinalities. When both the sets are identical the Dice coefficient value is 1.0 and 0.0 otherwise.

Loss function is used to know how well the network is performing the given task. The measure of the performance of the classifier is defined by Cross-entropy (CE). Cross entropy is also called log-loss.

The calculated dice coefficient was found to be 0.8823832.

### 4.2. OUTPUT SCREENS

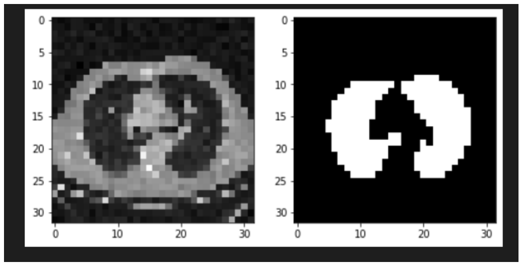

**Figure 3** Screenshot of Lung CT scan images after feature extraction.

Figure 3, shows the lung CT scan images after feature extraction. The above image is the output before calculating the dice coefficient.

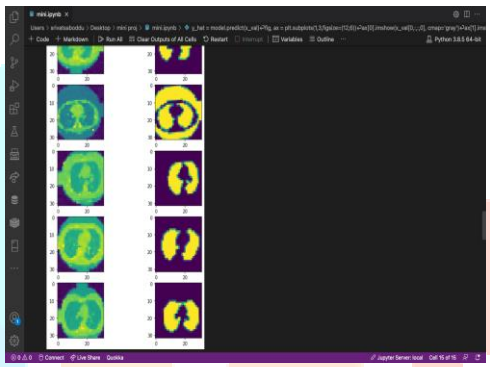

**Figure 4** Output of images after calculating the dice coefficient and using neural networks.

As shown in Figure 4, the output of the images that are generated according to masks of the CT scan image and with the CT scan image using the dice coefficient and neural networks.

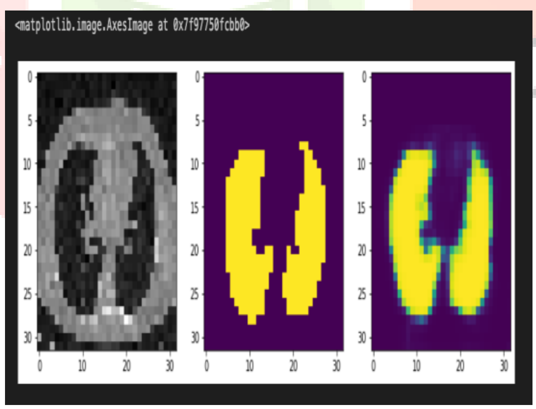


**Figure 5** Screenshot of page after clicking the button

Figure 5, shows the final output of the images that are generated after optimizing the trained segmented images of lung CT scan images using their masks.

## V. CONCLUSION

This project has proposed image segmentation can be done on CT scan images of lung with neural networks by using U-Net architecture rather than the traditional algorithms. This has been possible by computing the features present in the image by the utilization of state-of-the-art deep learning architecture. For experimental purposes, a dataset is used which contains 366 various lung CT scan images by different sizes of 256 x 384 or 384 x 256 pixels.

This project helps the users and medical staff a lot by saving their valuable time in the medical diagnosis stage and helps in predicting the appropriate medication which is to be used to prevent or cure lung diseases. This helps us to get a rough overview of the lung region.

## VI. REFERENCES

[1] https://www.worldscientific.com/doi/10.1142/S0218213021500056
[2] https://chadrick-kwag.net/cv2-resize-interpolation-methods/
[3] NAHIAN SIDDIQUE, (Graduate Student Member, IEEE), SIDIKE PAHEDING, COLIN P. ELKIN, (Member, IEEE), AND VIJAY DEVABHAKTUNI, (Senior Member, IEEE) Department of Electrical and Computer Engineering, Purdue University Northwest, Hammond, IN 46323, USA 2Department of Applied Computing, Michigan Technological University, Houghton, MI 49931, USA
[4] https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/
[5] https://towardsdatascience.com/unet-line-by-line-explanation-9b191c76baf5
[6] https://towardsdatascience.com/optimizers-for-training-neural-network-59450d71caf6