Original Article

Deep Learning-Based Segmentation Of Lung Images For Accurate Diagnosis

Smd Shafi[1], Sathiya Kumar C [2]*

[1,2] School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India;
Corresponding Author: Sathiya Kumar C
Email: sathiyakumar.c@vit.ac.in
DOI: 10.47750/pnr.2022.13.S05.455

Abstract

As lung cancer continues to be the foremost cause of death worldwide, it is crucial to explore various diagnostic approaches such as computed tomography, magnetic resonance imaging, and radiography. We aim to develop an advanced image processing method that utilizes segmentation algorithms to distinguish between CT scan images of lung cancer. To assess the technique's effectiveness, we will compare three segmentation methods to the "ground truth" established by an oncologist, utilizing accuracy, precision, recall, and F-score tests. Our methodology involves information gathering, image segmentation, and area enlargement as the primary image processing techniques. Ultimately, the goal is to provide a dependable and efficient diagnostic tool for lung cancer through research. Various image segmentation techniques, such as k-means clustering, Otsu's thresholding, and watershed segmentation, were successfully utilized to separate lung images. The region growth method was then applied to measure the lung area accurately. A performance study was conducted to evaluate the efficiency of the segmentation algorithm. Processing medical MRI or CT scans present unique challenges in computer vision, as it requires careful consideration of spatial information to ensure accurate alignment and orientation of volumes. To streamline medical image computing pipelines and enable more deep learning research, the TorchIO platform was developed.

Keywords: Medical image computing, Deep learning, MRI, CT, k-means clustering.

1. INTRODUCTION

As a result of advancements in medical imaging technology, clinical photographs today comprise an ever-increasing number of data. Clinical pictures may quantitatively convey the relevant information for organ disorders. However, quantification in some clinics is often done manually. The combination of computer-aided diagnostic (CAD) and automated detection techniques is becoming an active study area since it can reduce the required effort and speed up human work. A piecewise-defined curve is referred to as a contour. A contour is made up of an ordered collection of data points that are linked by segments. The resulting shape has the potential to be readily stated and used to assist with many issues. Problem with the image's segmentation to begin with. The process of segmenting pictures is one of the most difficult challenges encountered in the field of image processing and analysis.

Since the segmentation of target picture objects often occurs before any other stages of image analysis, any faults in accurately recognizing object borders significantly influence all operations after them. This article will focus on chest radiographs, which the general public often calls chest X-ray photos. Many studies have attempted to solve the problem of accurately segmenting the lung component in chest X-ray photos; moreover, the findings of these investigations could have been more satisfactory.

The problem has been studied because it is essential. This is true in many circumstances, particularly with automated lung area extraction. This is especially true for the segmentation of the lungs, which may undergo significant alterations over time and is influenced by a variety of diseases and the substantial aging-related changes that occur over time. As extensive population screening was moved into big data, it made the problem of precise and automated segmentation much more challenging. Neural networks that use convolutional and deep learning

methods. Recently, there has been a discernible rise in the number of people interested in the Deep Learning strategy. This strategy is often considered a subset of several machine learning methodologies. It is typical to conceive of it in this manner since it employs algorithms that use numerous processing layers to express high-level abstractions in data, and this way of thinking about it may be considered conventional [1].

Convolutional neural networks are often employed with some processing and abstraction layers (CNN) to explain the corresponding computational architecture better. The fact that CNN has achieved the first superhuman visual pattern recognition results in a few domains and has won numerous official international pattern recognition competitions since 2009 explains why deep neural networks, in general, and CNN, have attracted so much attention. CNN achieved these results in a small number of domains. Since 2009, CNN has won numerous official international pattern recognition competitions. Since 2009, CNN has been victorious in many prestigious international pattern recognition competitions. Neural networks include a convolutional layer in addition to an encoder-decoder layer.

The early approaches for segmenting pictures that used techniques from deep learning were analogous to those previously examined in the context of prior research into image processing and pattern recognition. They intended to modify the architectures of the deep learning programs such that specific classes may be directly assigned to certain pixel areas or tiny image patches. The purpose behind the creation of Senet was to partition pictures at the pixel level semantically. The core trainable segmentation engine comprises three layers: the encoder network, the auxiliary decoder network, and the pixel-wise classification layer. To carry out pixel-wise classification, the task assigned to the decoder network is to transform feature maps generated by encoders with a low resolution into input-resolution feature maps with a complete solution. The Encoder Decoding Convolutional Neural Network (ED-CNN) created as a consequence of this investigation might be considered later in the generalizing neural networks. These include the need for information that defines the physical features of the voxel, the absence of labels for large datasets, and the significant costs associated with data processing. A process referred to as data augmentation is used to modify the amount of information included in the training datasets. While doing training, using image subvolumes or patches is more successful than using traditional methods since it requires fewer processing resources [2].

Deep learning strategies and ED-CNN neural networks were used in this investigation to assess participants' abilities to recognize the lung component in chest X-ray images. The results of this study were positive. This study was conducted as part of a more extensive screening program for people throughout the community to check for lung and heart problems. Consequently, computational services for a worldwide gateway on lung TB hosted by Amazon and X-Ray image databases that include up to millions of items were built. From an application point of view, this study was carried out in the context of mass screening of the population for lung and heart diseases. Because it is not feasible nor practical to immediately assess the efficacy of ED-CNN networks on the complete image database, the following three steps have been included in this investigation:

For training and testing purposes, several hundred chest shots manually segmented were taken from exploratory research that only used a limited collection of images—forming opinions on the ED-CNN networks' practicability and prospective value.

Modifications were made to ED-CNN networks, and extensive testing was conducted using separate training and test sets that each included thousands of instances. Testing lung segmentation in three-dimensional computed tomography (CT) images using a design based on an adaptation network.

After doing so, we are moving the generated software solutions to a powerful workstation equipped with modern GPUs and integrating them into the desired environment.

Consequently, this investigation primarily emphasizes the preliminary, exploratory stage of an extensive collection of future research and developments on lung segmentation in radiological pictures using innovative neural network methodologies.

2. LITERATURE

This research separates the lungs in X-ray chest photographs (also known as chest radiographs), and the results are provided in this journal. Deep Learning methods and encoder-decoder convolutional neural networks were used to accomplish this task (EDCNN). The Nvidia Titan X GPU, which has 3072 CUDA Cores and 12 GB of GDDR5 RAM, was used to perform computational tests. The accuracy of the segmentation that was created was compared to the accuracy of the segmentation that was done manually using Dice's score. The findings indicated that the accuracy was 0.962 on average, with the lowest value of 0.926 and the best value of 0.974, respectively, with a standard deviation of 0.008. The study was conducted to construct computational services for a global gateway on pulmonary tuberculosis and the large-scale population screening for lung and heart problems. Because of the outcomes of this study, based on analysis, it appears that ED-CNN networks can serve as a powerful tool for the task at hand—Automated segmentation of large-scale lung projects [3].

In contrast to the RGB images often used in computer vision, processing MRI or CT scans for medical reasons presents several challenges that must be overcome. The requirement for information to describe the physical features of voxels, a lack of labels for large datasets, and high processing costs are some of the issues that need to be addressed. The training datasets' size is artificially increased using a data augmentation process. While training, it is more efficient to use image subvolumes or patches since it uses fewer processing resources. Careful consideration must be given to spatial information to ensure that volumes are accurately aligned and orientated. Methods: We provide TorchIO, an open-source Python toolbox, to ease the rapid loading, pre-processing, augmentation, and patch-based sampling of medical images for deep learning. During the training process for neural networks, TorchIO uses the Porch design language and includes standard medical image processing libraries. This allows for the efficient management of images [4].

Creating, duplicating, tracing, and extending TorchIO transformations is straightforward. Since most transformations may be inverted, the library is excellent for enhancing test times and estimating aleatoric uncertainty regarding segmentation. We offer a broad selection of pre-processing, augmentation, and simulation processes and MRI-specific artifacts. The command pip install torchy from the Python Package Index is necessary to establish the paestablishPyPI). The command-line interface of this software allows users to modify image files without the need for Python to be installed. We also offer a graphical user interface inside a TorchIO extension in 3D Slicer so that users may see the outcomes of transformations once they have been applied. TorchIO was developed to aid academics in standardizing medical image processing pipelines and to reduce the time for deep learning research. It allows for experiments to be repeated and is version-controlled, which ensures that the software can be correctly recognized, which in turn supports open-science solid practices. Due to the modular nature of its architecture, the library may be used in conjunction with a variety of deep-learning frameworks for medical picture analysis [5].

Over the last two years, COVID-19 has been responsible for the loss of a shockingly high number of lives. The currently available imaging-based diagnostic approaches for identifying COVID-19 in multiclass pneumonia-type chest X-rays could be more useful in clinical practice. This is primarily because of the significant error rates associated with these techniques. Suppose we can attain a segmentation-based classification error rate of less than 5%, which is a usual benchmark for 510(K) regulatory reasons. In that case, our diagnostic system may be adjusted for use in clinical settings. This is what our theory predicts will be necessary to make this transition. This study proposes sixteen distinct deep learning-based segmentation-based classification strategies as an automated, fast, and accurate means of identifying COVID-19. The optimum network combination was determined by using the eight classification models VGG16, VGG19, Exception, Inception V3, Densenet201, NAS Net Mobile, Resnet50, and Mobile Net in conjunction with the two deep learning-based segmentation networks UNet and UNet+ [6].

Using Dice, Jaccard, area-under-the-curve (AUC), receiver operating characteristics (ROC), and Grad-CAM to validate it, the cross-entropy loss function was utilized to evaluate the system’s performance. Re- salts: The best segmentation model, UNet, has accuracy, loss, Dice, and Jaccard values that are respectively 96.35%, 0.15%, 94.88%, and 90.38%, and AUC values that are correspondingly 0.99 (p-value 0.0001) each. UNet+Xception was the segmentation-based classification model that performed the best, with accuracy, precision, recall, F1-score, and AUC values of 97.45%, 97.46%, 97.45%, and 97.43%, respectively. These numbers were derived from the model's performance (p-value 0.0001). Our approach performed far better than

the others regarding segmentation-based classification models. The UNet+Xception system demonstrated an average improvement of 8.27% over the remaining set of tests. If the assumption of a 5% mistake rate is correct, then segmentation-based categorization is a practical method that can potentially be used in clinical practice [7].

A computer-aided methodology detects and segments the cancer-affected area from the lung MR image. The two techniques used for this task are k-means clustering and fuzzy c-mean. Calculations are utilized to establish the tumor stage and the lung area it impacts. A technique known as two-level classification is used to recognize and classify strokes automatically. This approach uses characteristics from the wavelet domain in conjunction with the intensity enhancement method. They extracted and segmented using the K-means clustering technic with the Perona-Malik anisotropic diffusion model from the damaged photographs. Image segmentation is accomplished by applying the Optimal Fusion and K-means methods based on the pixels in a picture [8].

The value that will be selected as the threshold will be. When the value of a pixel in an image drops below the threshold, that pixel is replaced with either a white or a black pixel. In MR images, the lung's damaged portion is segmented using the thresholding method. A further use of thresholding is transforming grayscale pictures into binary images. It is not recommended to use thresholding since this method might sometimes cause one to ignore the lung-affected cells. In cardiac image analysis and general medical picture segmentation, CNN-based techniques are gradually replacing more conventional methods thanks to the advent of deep learning in the spring of 2014. For instance, it was possible to distinguish between the LA and the pulmonary veins in MR images using a multi-planar deep learning approach. A recurrent, fully convolutional neural network has been proposed to differentiate between the MRI and the LV [9].

Like this, the LV on the MRI was segmented by using a deep learning approach in combination with a deformable-model method. A process that combines localization and segmentation, called hybrid localization and segmentation, was applied inside a deep learning framework to segregate RV data. To this day, most deep learning approaches have only successfully split one or two cardiac components and are restricted to a single modality. When applied to natural pictures, the performance of several deep learning-based algorithms, including SegNet and DeepLab, is much superior to that of previous approaches. SegNet has described an encoder-decoder design. Under this architecture, the decoder component is responsible for unpooling and resizing, and it uses the data obtained from the max-pooling layer (in the encoder part). DeepLab's atrous convolution kernel can extract more valuable properties from images when used with those obtained from a conventional convolution kernel. Subsequently, based on the encoder-decoder design, U-Net was built by adding skip connections from the encoder to the decoder part of the architecture to ease the flow of information. This was done to build on the foundation laid by the encoder-decoder design. The DenseNet was improved by adding global connections between the encoder and the decoder and local skip connections inside each block of the encoder and decoder parts. PAN was designed to distinguish the pancreas from CT scans by learning high-level consistencies in images and pixel-level information. This was accomplished using deep learning [10].

3. CT IMAGING

3.1 Physics of CT scans

In computed tomography, often known as CT, X-ray beams are utilized to construct three-dimensional images of the human body's pixel intensities. When a cathode is heated, rays of high-energy electrons are emitted. These electrons subsequently release the point that they have stored as X-ray radiation. X-rays penetrate the human body's tissues before arriving at a detector located on the other side of the body. For instance, bones are a solid material that absorbs more radiation than soft tissues because of their greater thickness (i.e., fat). When X-rays reach the detector and are not absorbed by the body, they seem to us as a dark color, like the appearance of a black film (for example, in the air area within the lungs). On the other hand, the color white is the one that is utilized to symbolize dense tissues.

Figure 1: CT imaging can distinguish density differences and create a 3D body image

Segment the main vessels and compute the vessels over lung area ratio.

First, we do element-wise multiplication using CT imaging and the lung mask to isolate only the lungs. After that, the zeros from the element-wise multiplication are changed to a value of -1000 (AIR in HU), and at that point, the only intensities preserved as vessels are those that ter -500.

```
def create_vessel_mask(lung_mask, ct_numpy, denoise=False):
vessels = lung_mask * ct_numpy  # isolate lung area
vessels[vessels == 0] = -1000
vessels[vessels >= -500] = 1
vessels[vessels < -500] = 0
show_slice(vessels)
If denoise:
return denoise_vessels(lungs_contour, vessels)
show_slice(vessels)
return vessels
```

3.1.1 Ct Intensities and Hounsfield Units

To evaluate the X-ray absorption, the Hounsfield scale is used. We made the intensity of the water equal to zero and the intensity of the air similar to one thousand. It is essential to have an understanding that the scale used in the House field is absolute, in contrast to the relative scale used in the MRI, which ranges from 0 to 255. The image depicts many essential tissues and the intensity levels corresponding to those tissues. We must be aware that the images include noise.

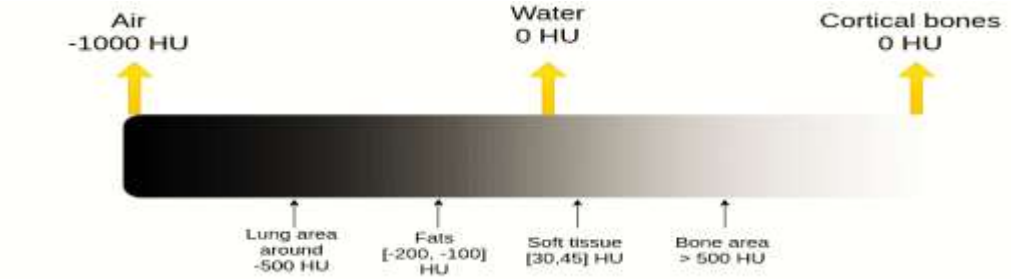

Figure 2: The numbers will be slightly off in the images.

Bones are incredibly intense. To get a more extensive maximum range, the image is often cropped. The maximum value may be 1000 since this is the most helpful number. The problem is that visualization tools only support numbers between 0 and 255. For medical diagnosis, it is not a good idea to visualize the whole Hounsfield scale, which ranges from minus 1000 to plus 1000. Instead, we focus on the underlying tissues to restrict this spectrum's extent.

3.1.2 CT data visualization: level and window

To reduce the House field range, the standard procedure for a medical image is to choose a level of central intensity, also known as the level, and a window, as shown in the illustration:

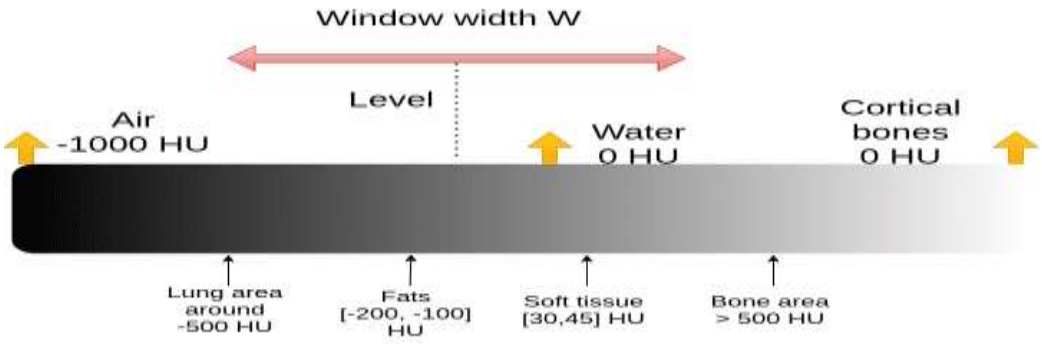

Figure 3: It is a very revolting convention for those who work with computers.

The lowest and highest values in the range are all that we require:

```
import matplotlib.pyplot as p  lt

import numpy as np

def show_slice_window(slice, level, window):

  Function to display an image slice

  Input is a numpy 2D array

  max = level + window/2

  min = level - window/2

  slice = slice.clip(min,max)

  plt.figure()

  plt.imshow(slice.T, cmap="gray", origin="lower")

  plt.savefig('L'+str(level)+'W'+str(window))
```

Consider the following picture if you are still not convinced that the same CT picture has more information than a regular image channel:

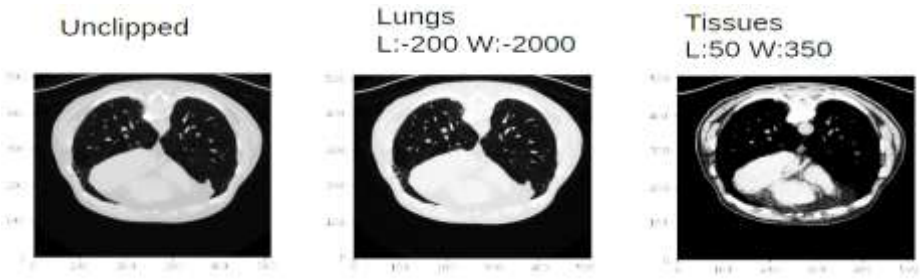

Figure 4: CT image

For reference, here is a list of visualization ranges:

Table 1: List of Visualization Ranges

Region/Tissue	Window	Level
Brain	81	39
Lungs	1400	-700
Liver	140	40
Soft Tissues	260	40
Bone	1900	500

3.1.3 Contour Detection on Private Datasets with Varying Learning Rates

To gain the related training length, the Dice coefficient, and E at various learning rates, the number of neurons in the hidden layer has been preconfigured to be set to 10, and the number of epochs has been preconfigured. The training periods that matched had varying learning rates ranging from 0.2 to 0.7. Since Lung A is worth less than Lung B and Lung C, the total time required for its training is cut down. When the learning rate increases, the overall picture of the training period shows it is becoming longer and longer until it reaches 0.5. After then, the duration of the training time is essentially unaltered. When the learning rate is increased from 0.4 to 0.5 in Lung B and Lung C with the correctly initialized learning rate, there is a little increase in the required time for training, but the BNNM converges quite rapidly.

The Dice coefficient, in addition to E, is at various learning rates. The dice coefficient is a representation of the overlap that exists between the detection contours obtained by the recommended algorithm and the contours that were drawn by hand by radiologists. The letter E represents the overall error associated with the suggested approach. The Dice coefficient and E move in opposing directions while the learning rate remains constant. USING DIFFERENT LEARNING RATES, the BNNM can rapidly calculate the appropriate Dice coefficient for Lung A. Even if there are just a few starting points, the maximum Dice coefficient can reach a value of 0.95 when the learning rate is 0.4. The rapid shift in the learning rate contributes to the instability of the iterative process in Lung B; when the learning rate is set to 0.3, the Dice coefficient curve suddenly begins to descend. When the curve angles are placed at the concave vertice, we can see in A that the optimal learning rate is 0.4 and that the E is reduced to 1.47 102. The approach will eventually converge, at which point the E will become almost constant. The comparison of the contours created by the proposed CPL-BNNM algorithm with the shapes made by hand at different learning rates [11].

3.2 CLUSTERING METHOD IN K

Image segmentation is a critical method that is used in medical imaging. Image segmentation makes it easy to identify the affected region and contributes to the treatment planning process when used in medical photographs. Several approaches, including edge detection, dual clustering, histogram-based algorithms, thresholding, and many more, may be used to split the region of interest (RoI) into individual segments. The recommended study applied the K-means clustering method for the segmentation. The strategy of learning without supervision is the one that is the least difficult. The value of K, which denotes the cluster, is selected at random. The value of k is

determined by a process of trial and error up to the point when optimum clustering is attained. The k value changes whether a CT image or an MR image is used, depending on which one is being seen.

3.2.1 Algorithm

1. Set the value of the number of clusters to k.

2. k clusters' centers may be chosen at random.

3. Determine the cluster's midpoint (center).

4. Every pixel's distance from the center of each cluster is determined.

5. The cluster with the shortest determined distance receives the pixel.

6. Or else, shift the pixel to the next cluster.

7. Calculate the center again.

8. Once the cluster's center does not change, repeat the whole procedure.

This is a flowchart for the K-means clustering algorithm. It explains how to calculate the distance between each pixel and each cluster and how to add the pixel to the group to which the distance is the least. Also, it demonstrates how to decide the number of sets to use. Go on in this manner until the cluster's center no longer moves.

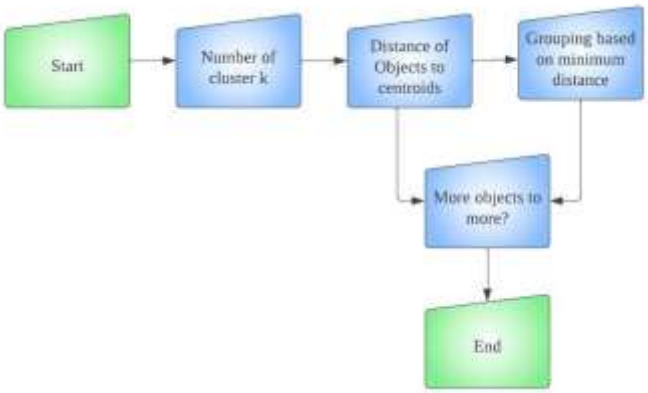

Figure 5: Flowchart of k-means clustering algorithm

4. METHODOLOGY

The four primary components of this project are the gathering of data, the masking of areas of interest, the segmentation of pictures, and the expansion of regions. The first step of this study is data collection, and the Advanced Medical and Dental Institute (AMDI) database at University Sains Malaysia was utilized as the source for the sample. The following sections describe the approaches used to characterize the morphology of the lung tumor. These methods include the area of interest masking, image segmentation, and region growth. Depicts the method flowchart for image segmentation and area enlargement. This may be found in the following: The technique for segmentation consisted of two parts: the first was a semi-automated segmentation, and the second was a manual segmentation. Manual segmentation is a time-honored technique that radiologists often use to distinguish between different types of lung tumors. The particulars of the approach are broken down into their parts in the following subsections.

4.1 Data Collection

A total of ten sets of images were collected: eight patients or persons with atypical lung disorders and two sets of pictures of people with normal lungs. Five example photographs were submitted by each respondent in total. The use of a Siemens SOMATOMS Definition Flash CT scanner obtained the picture. Before loading into the MATLAB application, the image was first organized in a file for later use. The present study analyzed 50 lung imaging samples using three different algorithms.

4.2 Region of Interest Masking

After selecting the value for the threshold, a masking technique was carried out on the threshold picture. Finally, the binary image was created by finding the pixel. Any numeric number that is less than the cutoff value has the potential to be split into two parts, with the presumption being that one of those parts is black and the other is white. Utilizing this method makes it much simpler to get the desired lung shape.

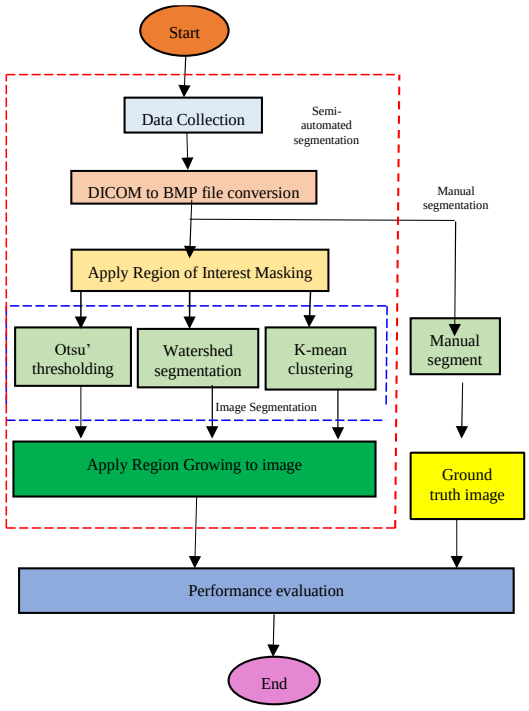

Figure 6: Flowchart of image segmentation and region-growing process.

4.3 Image segmentation

K-means clustering, Otsu's thresholding, and watershed were proposed as the three methods of segmentation that should be used in this investigation: Watershed, which aimed to highlight the foreground item or background at a particular image point. The original color photograph was converted into a grayscale form first. The segmentation function was generated using the magnitude of the gradient as its input. Most of the time, the low angle of an item could be discovered inside the object itself, whereas the high rise of its border could be found around the thing. After then, the object in the front was marked. Connecting the individual blobs that make up the object's pixels is necessary to determine the foreground marker. Each item's black patches and stem marks were removed using morphological techniques, creating flat maxima for each object. The skeleton was computed with the help of the foreground's influence, which produced a black background pixel. This was caused by the background markers being situated too close to the edge of the segmented item [12].

4.4 Cumulative Frequency Curves for Dice and Jaccard

The Jaccard index, or intersection over union (IoU), and the Dice coefficient, the F1-score, are the most important metrics when evaluating segmentation. The Dice coefficient is computed by taking the total number of pixels

present in both images and dividing that figure by the area of overlap that exists between the AI (predicted mask) and the GT (ground truth mask). The Jaccard index is calculated by taking the area of union between AI and GT and dividing it by the area where AI and GT overlap. The Dice and the Jaccard both have a high positive association with one another and are highly similar. The Dice and Jaccard cumulative frequency curves for the UNet [13] and UNet++ [14] models are shown in Figure 7.

Table 2: Cumulative Frequency Curves for Dice and Jaccard

Model	Test Accuracy (%)	Test Loss (%)	DDice (%)	Jaccard (%)	Area Error(mm2)	AUC (p-Value)
UNet	97.35	0.17	94.9	91.38	1.45	0.99 ($p < 0.001$)
UNet++	96.5	0.18	93.35	89.07	2.65	0.98 ($p < 0.001$)

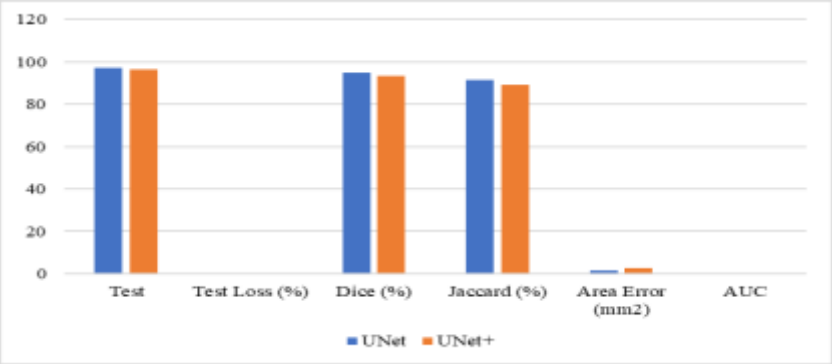

Figure 7: Cumulative Frequency Curves for Dice and Jaccard

In 80% of the scans, the Dice and Jaccard values for the UNet model were more than 0.96 and more significant than 0.93. On the other hand, in 80% of the scans for the UNet+ model, the Dice and Jaccard values were higher than 0.95 and more significant than 0.91. Consequently, the UNet model fared better regarding Dice and Jaccard than the UNet+ model.

4.5 Classification Results

The next step in achieving our mission was to successfully classify the segmented chest X-ray images into five distinct categories using the most suitable classification model we could devise. This would allow us to get the best possible results. To classify segmented lung images into the following five categories: VP, BP, TB, and normal, we used eight distinct incredibly successful deep neural networks. These networks were as follows: VGG16 and VGG19[15], Xception [16], InceptionV3 [17], Densenet201 [18], NASNetMobile [19], Resnet50 [20], and MobileNet [21]—a comparison of the eight different CNNs' individual performance characteristics. The Xception model performed the best overall, with an accuracy of 97.45% and weighted averages of 97.46% for Precision, Recall, and F1 score. Compared to other classifiers, the depth model known as Xception contains 81 different layers, making it the one with the most layers overall. Separate convolution layers are another characteristic of Xception that sets it apart from typical convolutional layers in terms of the amount of money spent on processing and the amount of memory required. The fact that Xception has these properties and superior accuracy makes it the model that best matches our CXR datasets. MobileNet's performance was ranked as the second most effective overall, with an accuracy score of 93.66%, a precision score of 93.87%, a recall score of 93.66%, and an F1 score of 93.60%—table 3 displays, according to the best-performing Xception model, the

various performance measures of each class. The class had the best Precision score, 98.88%, while the Bacterial Pneumonia class had the highest Recall score, 100%, and the Normal class had the highest F1 score, 98.55%.

Table 3: The weighted average of performance metrics by eight deep learning models for five-class classification of segmented chest X-ray images, VP, BP, TB, and normal.

DL Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
VGG16	88.24	88.88	88.26	89.82
VGG19	87.65	87.09	89.64	87.25
Xception	98.45	98.46	97.49	99.43
InceptionV3	90.98	90.9	90.88	91.47
Densenet201	83.07	82.32	82.05	80.15
NASNetMobile	92.97	95.01	93.97	95.78
Resnet50	90.05	91.01	90.11	89.11
MobileNet	94.66	93.89	95.66	95.11

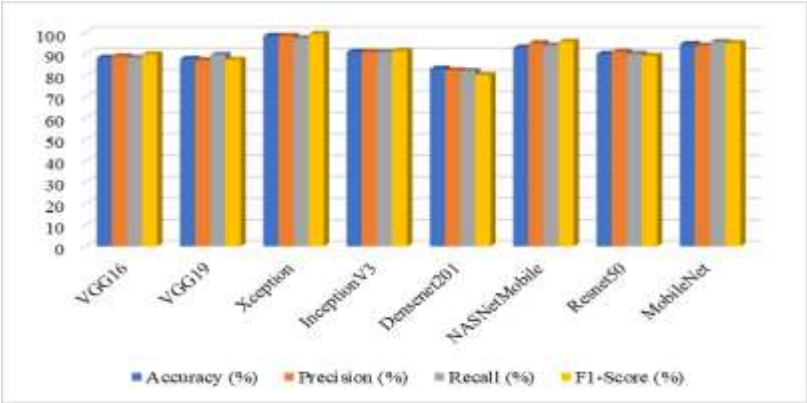

Figure 8: The weighted average of performance metrics by eight deep learning models for five-class classification of segmented chest X-ray images, VP, BP, TB, and normal.

4.6 UNet-200 epochs

In medical imaging, improving segmentation outcomes is one of the goals of the UNet-200 architecture, which is constructed on a fully convolutional network. The form is in the shape of a U. On the UNet, there are two routes: one for encoding and another for decoding. Surprisingly, these two routes are almost identical to one another. They will take on the structure of the output that we gave as the input when the UNet processes them. The three fundamental systems that comprise UNet are scaled-down, bottleneck, and scale-up. In autoencoders, the decoder derives the output from the compressed or encoded representation, formed when the neural network encoder compresses the input into a latent space model. This representation is made possible by the compression of the information by the encoder. On the other hand, in contrast to other encoder-decoder systems, the two parts are not split apart in this case. When moving fine-grained data from low-level analysis routes to high-level levels of a synthesis path, skip connections are often used to transport the data. Several pieces of information are necessary to generate precise reconstructions at a finer granular scale. The framework of UNet itself. After the implementation of two convolutional layers with kernel sizes of (3 3) in this design comes a contraction step performed by a MaxPool layer with a length of (2 2) for a total of 32 iterations, which results in an image of the form (32 32). Following that, they started upsampling the image to concatenate the transposed layer with the

required feature map and then transpose the layer. In addition to this, we based the output on the kernel size of the convolutional layer, which was set at 1. As a result of the sigmoid activation function found in the network's last layer, the network training approach uses the cross-entropy cost function [22].

Table 4: Results of the validation and training of the UNet-200 Model

Model	Loss	Dice coef	Specificity	Mean_iou	Sensitivity	Recall	Precision
UNet train	0.2526	0.6885	0.9923	0.3837	0.9668	0.995	0.9513
Unet val	0.3527	0.6765	0.9898	0.3635	0.972	0.981	0.944
Difference	0.0018	0.003	0.0025	0.0005	0.005	0.0065	0.0073

4.7 Method Based on Bronchial Trees and The Eleudican Distance

The suggested segment approximation technique is now used in a clinical investigation to investigate the potential benefits of computer-aided segment determination and visualization in planning segment resections. Using distance information to evaluate anatomical substructures was used in the past and is being used today to identify liver segments. The Euclidean distance to the portal vein splits the liver in this situation. The planning of liver interventions such as surgery for liver metastases and living donor liver transplantation has already been achieved effectively at several institutions using this strategy for the segmental subdivision of the liver. Some of these institutions include: Under the framework of our technique, the pulmonary artery acts as an approximation of a section. There is also the possibility of using the distance to the bronchial tree. We predict that the PA-based strategy will work better than the BT-based method because the pulmonary artery is more visible on CT scans and can be traced into peripheral locations. The PA-based strategy will be more effective in regions far from the lung hilum [23].

To put this notion to the test, we analyzed 11 CT scans and utilized a BT-based technique to locate the regions of the lungs being examined. In the case of the BT-based segment approximation, we also assessed the inaccuracy between the predicted segment boundaries and the intersegmental veins. The bronchial tree was segmented to the extent that it could be seen on CT in a manner comparable to the division of the pulmonary artery. After completing an automated bronchial tree segmentation, the missing visible bronchi branches were inserted via line tracing. It is essential to partition the lung lobes into segments, which must then be subdivided into the pulmonary parts. We have devised an automated method to separate the lobes and components of the lungs. The sources provide detailed explanations of the topics. Calculating lung and lobe segmentations during the preprocessing phase of the segment approximation approach for a single lobe, identifying the segmental branches of the pulmonary artery, and assigning each location in the considered lobe to its closest segmental branch using Euclidean distance information are the three steps that make up the segment approximation approach [24].

The culmination of this process results in the segmentation of the lobe. The following paragraphs explain the procedure for segmentally subdividing a single lobe into several lobes. As reliable automated pulmonary artery segmentations that can achieve the required level of accuracy are not currently on the market, pulmonary artery segmentation is obtained by using line tracing, which highly trained radiological assistants perform. This enables the most accurate and comprehensive results possible. Beginning in the hilum regions, the courses of the pulmonary artery branches are tracked and documented as they go outward toward the periphery and upward into the finest vessel branches—all visible pulmonary artery branches. The segmentation of the pulmonary artery is accomplished with the help of an internal software program. The Euclidean distance is used to convert the pulmonary artery that is selected for use in the lobe region. In line with the results of the distance transformation, the segmental branch of the pulmonary artery located closest to each point on the lobe is determined to be the appropriate one. After the end is assigned to the proper segment, the segmental subdivision of the lobe is said to be finished [25].

4.8 ADAPTIVE AND ADAM OPTIMIZER

One of the most used optimization approaches, stochastic gradient descent (SGD), is an algorithm that is both straightforward and efficient in various contexts. One disadvantage of the gradient is that it scales uniformly, or in the same way, in all directions (for each parameter). Another challenge is finding an appropriate number representing the learning pace (LR). Since it is a fixed parameter in SGD-based approaches and directly impacts the precision and velocity of neural network convergence, it is essential to ensure that LR is adjusted appropriately [26].

By allowing the LR to be changed at any point throughout the training process, several algorithms called "adaptive optimizers" have been created to solve this problem. Every iteration of LR is given a unique look by tailoring the calculations to the history of gradient shifts. Techniques such as ADAM, ADAGrad, and RMSProp are examples of this method. Adaptive optimizers are commonly employed in deep learning applications because of the shorter time needed for training sessions that they typically give. The same may be said for optimizers based on momentum [27].

Momentum optimizers were developed to speed up the convergence process by considering the changes during the previous iteration and applying a multiplier called "momentum." To evaluate how well our recommended optimizer stacks up against existing optimizers, we estimated the Dice Index (DI) and Cross-Entropy (CE) losses using the validation set. In proportion to the number of iterations, the CE and DI curves for the U-Net architecture are given for several different optimizers. From what can be gleaned from these graphs, the DI in U-Net with Adam optimizer experiences a sharp and precipitous rise at the outset before leveling to an almost constant level. Despite this, our recommended optimizer, CLMR (C lr=20, C mr=20), ultimately outperforms the ADAM curve, even though it does not initially learn as rapidly as the ADAM optimizer (this phenomenon is more apparent in CE curves) [28].

Table 5: Using online assessment and TaDI with the test data set

		Adam	Nesterov	CLR	CLMR
Enc_dec	RV	0.3275	0.131	0.3834	0.4332
UNet		0.8577	0.5968	0.8616	0.8823
DenseNet_1		0.8805	0.5933	0.896	0.8955
DenseNet_2		0.878	0.7235	0.8911	0.905
Enc_dec	Myo	0.1475	0.149	0.1693	0.1687
UNet		0.8629	0.6488	0.8586	0.863
DenseNet_1		0.8789	0.7172	0.8833	0.8963
DenseNet_2		0.8799	0.7197	0.8904	0.8999
Enc_dec	LV	0.4953	0.3261	0.497	0.5415
UNet		0.4049	0.7672	0.8936	0.936
Densenet_1		0.9377	0.8464	0.935	0.9395
DenseNet_2		0.9195	0.8447	0.9376	0.9479
Enc_dec	Ave.	0.323	0.1688	0.3496	0.3813
UNet		0.8815	0.671	0.8815	0.8938
DenseNet_1		0.8985	0.7522	0.7523	0.9104

DenseNet_2		0.8922	0.7623	0.9066	0.9177

Similarly, it took much work to initially determine an appropriate value for the momentum rate (MR). Despite this, several adaptive optimizers, such as the Adaptive Momentum Optimizer (ADAM), were ultimately created to change the MR and the LR dynamically. Adaptive optimizers such as ADAGrad and ADAM have been at the forefront of deep learning for several years, thanks to their rapid convergence. There is a prevalent misconception that adaptive optimizers will always converge to the same minimal sites as classic SGD algorithms. This is only sometimes the case. In other words, since more studies are being conducted and published recently, they are expected to have poorer out-of-sample behavior and generalization ability than non-adaptive optimizers. To achieve better generalization capabilities, the researchers went back to the SGD approaches that were used initially, but they also made new attempts to solve the issue of convergence speed. For instance, the YellowFin optimizer demonstrated that LR and MR were superior to ADAM in performance when hand-tuned sufficiently. While a proof-of-concept study supported the counter-intuitive idea of non-adaptive systems, it takes time and effort to hand-tune such rates in practical applications. An additional method was implementing a cyclic learning rate (CLR), which varied the learning rate by a cycle (such as a triangle or Gaussian) and provided a viable alternative to the requirement of human labor-intensive hand-tuning. This method was known as the "cyclic learning rate." The CLR has one and only one downside: the possibility that a predetermined momentum rate may limit the search state and preclude finding an optimum solution [29].

In addition, Adam and CLMR's training programs for various architectural designs' DI and CE curves. DenseNet 2 outperforms other topologies despite having fewer parameters than those. This is because of its superior performance. These curves bring attention to many significant facets of adopting a variety of architectures: To begin, the performance of the CLMR optimizer is superior to that of the ADAM optimizer across all architectures. This demonstrates the efficiency of the cyclic optimizer that was presented. DenseNet topologies are singular than U-Net architectures and Enc Dec architectures in terms of performance; however, U-Net and Enc Dec contain nearly ten times as many parameters as DenseNet systems. This demonstrates that overly parameterized topologies have the potential to saturate the training quickly, but the problem described above can be avoided by using local skip connections.

5. CONCLUSION

Comparative research for image segmentation for lung cancer diagnosis employing Otsu's thresholding, watershed segmentation, and k-mean clustering has been successfully finished. Even though the outcomes of these three ways of segmentation were substantially the same, the watershed segmentation did better than the other two. The investigation results showed that the thresholding proposed by Otsu was quite similar to the actual results. Consequently, it has the potential to be improved to become more reliable and effective. Future research should focus on extracting features to classify tumors and diagnose lung cancer. The most important advantage of the approach employed in this research is that the Deep Learning technique is universal enough to be applied to a wide range of medical image segmentation issues with very little need for adaptation. In addition, the method can be adapted to segment three-dimensional computed tomography (CT) images and solve other problems associated with medical image analysis. One of these problems is identifying "atypical" image areas, which are frequently associated with tumors and other abnormalities.

REFERENCES

[1] Prus, V. Kovalev, P. Vankevich. A method for lung segmentation in massive X-ray screening of the population. International Journal of Computer Assisted Radiology and Surgery, 3(1), 2009, p. 367- 368.

[2] S. Jaeger et al. Automatic tuberculosis screening using chest radiographs. IEEE Transactions on Medical Imaging, 33 (2), 2014, p. 233-245.

[3] V.A. Kovalev, A.A. Kalinovsky. Big Medical Data: Image mining, retrieval, and analytics, Proceedings of the International Conference on Big Data and Predictive Analytics, Belarus State University of Informatics and Radioelectronics, ISBN 978-985-543- 146-7, Minsk, Belarus, June 2015, pp. 33-46.

[4] Chowdhury, M.E.H.; Rahman, T.; Khandakar, A.; Mazhar, R.; Kadir, M.A.; Bin Mahbub, Z.; Islam, K.R.; Khan, M.S.; Iqbal, A.; Al Emadi, N.; et al. Can AI Help in Screening Viral and COVID-19 Pneumonia? IEEE Access 2020, 8, 132665–132676.

[5] Rahman, T.; Khandakar, A.; Qiblawey, Y.; Tahir, A.; Kiranyaz, S.; Kashem, S.B.A.; Islam, M.T.; Al Maadeed, S.; Zughaier, S.M.; Khan, M.S.; et al. Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. Comput. Biol. Med. 2021, 132, 104319.

[6] Kermany, D.S.; Zhang, K.; Goldbaum, M. Labeled optical coherence tomography (oct) and chest X-ray images for classification. Mendeley Data 2018, 2.
[7] Kermany, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.S.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell 2018, 172, 1122–1131.e9.
[8] Agarwal, M.; Agarwal, S.; Saba, L.; Chabert, G.L.; Gupta, S.; Carriero, A.; Pasche, A.; Danna, P.; Mehmedovic, A.; Faa, G.; et al. Eight pruning deep learning models for low storage and high-speed COVID-19 computed tomography lung segmentation and heatmap-based lesion localization: A multicenter study using COVLIAS 2.0. Comput. Biol. Med. 2022, 146, 105571.
[9] Saba, L.; Banchhor, S.K.; Londhe, N.D.; Araki, T.; Laird, J.R.; Gupta, A.; Nicolaides, A.; Suri, J.S. Web-based accurate measurements of carotid lumen diameter and stenosis severity: An ultrasound-based clinical tool for stroke risk assessment during multicenter clinical trials. Comput. Biol. Med. 2017, 91, 306–317.
[10] A. Ravishankar and A. Raj, "A Survey on Noise Reduction Techniques in Medical Images," Int. Conf. Electron. Commun. Aerosp. Technol., pp. 385–389, 2017.
[11] "Analysis of Filtering and Novel Technique for Noise Removal in MRI and CT Images," pp. 1–3, 2017.
[12] M. Vas and A. Dessai, "Lung cancer detection system using lung CT image processing," 2017 Int. Conf. Comput. Commun. Control Autom. ICCUBEA 2017, pp. 1–5, 2018.
[13] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015.
[14] Zhou, Zongwei, et al. "Unet++: A nested u-net architecture for medical image segmentation." Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4. Springer International Publishing, 2018.
[15] Simonyan, Karen, and Andrew Zisserman. "Intense convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).
[16] Chollet, François. "Xception: Deep learning with depthwise separable convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
[17] Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
[18] Huang, Gao, et al. "Densely connected convolutional networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
[19] Zoph, Barret, et al. "Learning transferable architectures for scalable image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
[20] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
[21] Howard, Andrew G., et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." arXiv preprint arXiv:1704.04861 (2017).
[22] I. F. Nizami, S. Ul Hasan, and I. T. Javed, "A wavelet frames + K-means based automatic method for lung area segmentation in multiple slices of CT scan," 17th IEEE Int. Multi Top. Conf. Collab. Sustain. Dev. Technol. IEEE INMIC 2014 - Proc., pp. 245–248, 2015.
[23] A. Buslaev, V.I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, A.A. Kalinin, Albumentations: Fast and Flexible Image Augmentations, Information 11 (2) (2020) 125, doi:10.3390/info11020125. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute
[24] M. Larobina, L. Murino, Medical image file formats, J. Digit. Imaging 27 (2) (2014) 200–206, doi:10.1007/s10278-013-9657-9.
[25] M.J. Cardoso, M. Modat, R. Wolz, A. Melbourne, D. Cash, D. Rueckert, S. Ourselin, Geodesic information flows spatially-variant graphs and their application to segmentation and fusion, IEEE Trans. Med. Imaging 34 (9) (2015) 1976–1988, doi:10.1109/TMI.2015.2418298.
[26] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J.V. Miller, S. Pieper, R. Kikinis, 3D slicer as an image computing platform for the quantitative imaging network, Magn. Reson. Imaging 30 (9) (2012) 1323–1341, doi: 10.1016/j.mri.2012.05.001.
[27] B.C. Lowekamp, D.T. Chen, L. Ibez, D. Blezek, The design of SimpleITK, Front. Neuroinformatics 7 (2013) 45, doi:10.3389/fninf.2013.00045.
[28] K. Weiss, T.M. Khoshgoftaar, D. Wang, A survey of transfer learning, J. Big Data 3 (1) (2016) 9, doi:10.1186/s40537-016-0043-6.
[29] V. Cheplygina, Cats or CAT scans: transfer learning from natural or medical image source data sets? Curr. Opin. Biomed. Eng. 9 (2019) 21–27, doi:10.1016/ j. cobme.2018.12.005.