Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi
Sayı 8, S. 146-161, 9, 2024

Araştırma Makalesi

International Journal of Advanced Natural Sciences and Engineering Researches
Volume 8, pp. 146-161, 9, 2024

Research Article

https://as-proceeding.com/index.php/ijanser
ISSN:2980-0811

AI AIDED DISCRIMINATION OF COVID AND PNEUMONIA DATASET WITH DEEP LEARNING BASED 3D SEGMENTATION MODELS

Ali Berkan Ural[1*]

[1]*Department of Electrical Electronics Engineering, Circuit and Systems/Biomedical, Kafkas University, Kars, Turkey*

**Corresponding author email: berkan.ural@kafkas.edu.tr*

(Received: 24 September 2024, Accepted: 07 October 2024)

(6th International Conference on Applied Engineering and Natural Sciences ICAENS 2024, 25-26 September 2024)

ATIF/REFERENCE: Ural, A. B. (2024). AI Aided Discrimination of Covid and Pneumonia Dataset with Deep Learning Based 3D Segmentation Models, *International Journal of Advanced Natural Sciences and Engineering Researches*, 8(9), 146-161.

Abstract *–*The rapid and accurate discrimination between COVID-19 and common pneumonia has become crucial for effective patient management, especially during the global pandemic. This study presents a novel approach using deep learning-based 3D segmentation techniques to differentiate between COVID-19-induced pneumonia and other forms of pneumonia from medical imaging data, specifically computed tomography (CT) scans. The proposed framework aims to assist radiologists and healthcare providers in identifying unique patterns in COVID-19 infections while distinguishing them from common viral or bacterial pneumonia. The core of the method involves the application of a 3D convolutional neural network (CNN) integrated with a V-Net and Res-Net architectures for volumetric segmentation of lung regions affected by infection. By analyzing CT scan volumes, the model can isolate and segment crucial lung abnormalities, such as ground-glass opacities (GGOs), consolidations, and other characteristic features seen in COVID-19 and pneumonia patients. Preprocessing steps, including image normalization, contrast enhancement, and noise reduction, ensure robust input data for model training and testing. The 3D segmentation model is trained on a diverse open publicly dataset comprising 1000 CT scans labeled for COVID-19 and 1000 CT for common pneumonia cases. It utilizes deep learning techniques, including transfer learning, to maximize performance and efficiency, allowing the model to generalize well across varying patient populations. Additionally, the system employs a hybrid classification model that further distinguishes between COVID-19 and non-COVID pneumonia based on the segmented lung regions, using features such as the distribution, volume, and texture of infected areas. Performance evaluation demonstrates that the proposed deep learning model achieves high accuracy, sensitivity, and specificity in distinguishing COVID-19 from common pneumonia. The model outperforms conventional 2D segmentation techniques by leveraging the richer spatial context provided by 3D imaging. In extensive testing, the system attained an accuracy exceeding 96% for ResNet, %86 for SegNet and %74 for UNet models, with notable improvements in reducing both false positives and false negatives. Furthermore, the model's segmentation quality is validated against radiologist annotations, confirming its clinical relevance and potential for real-world applications.

Keywords: Lung dataset, COVID, Pneumonia, 3D Segmentation, Deep Learning

1.INTRODUCTION

The outbreak of the COVID-19 pandemic has highlighted the crucial need for accurate and rapid diagnostic methods to differentiate between COVID-19 and other respiratory illnesses such as pneumonia [1]. Medical imaging, particularly computed tomography (CT), has emerged as a vital tool in this endeavor due to its high-resolution capabilities that can capture detailed visual characteristics of lung pathologies. However, the manual analysis of CT images is time-consuming and prone to variability due to the complex and heterogeneous nature of lung abnormalities associated with these diseases [2].

To address these challenges, the integration of artificial intelligence (AI) and deep learning techniques has shown significant promise. Deep learning, particularly using convolutional neural networks (CNNs), has revolutionized image analysis by enabling automated feature extraction and classification [3]. Recent advances have further leveraged 3D segmentation models, which allow for a comprehensive analysis of the volumetric data inherent in CT scans. These models can capture intricate spatial relationships across multiple slices, offering a more holistic view of the lung structures and potential abnormalities [4].

In this study, a deep learning-based 3D segmentation approach is particularly effective in distinguishing COVID-19 from pneumonia. By utilizing a robust dataset that includes diverse CT images of patients with both conditions, deep learning models can be trained to recognize subtle differences in lung texture, density, and morphology [5]. These differences are often challenging for radiologists to discern, especially in early-stage or mild cases.

Moreover, the segmentation not only aids in classification but also provides visual explanations for the decision-making process, which is crucial for clinical acceptance and deployment [6]. By accurately segmenting affected lung regions, the model can highlight areas of interest, such as ground-glass opacities or consolidations, which are indicative of specific conditions. This capability supports radiologists in making more informed and confident diagnoses, ultimately improving patient outcomes [7].

3D segmentation is a crucial technology in medical imaging that involves the partitioning of 3-dimensional volumetric data into meaningful regions [8]. This process enables the precise identification and isolation of anatomical structures and pathological regions within medical images, such as CT and MRI scans [9]. By transforming raw image data into structured information, 3D segmentation facilitates detailed analysis, diagnosis, and treatment planning in clinical practice.

Unlike traditional 2D segmentation, which processes single slices of an image, 3D segmentation accounts for the spatial relationships between adjacent slices, providing a more comprehensive representation of the scanned area [10]. This capability is particularly valuable in complex anatomical regions, such as the lungs or brain, where structures are interconnected in three dimensions. The volumetric approach enhances the detection and delineation of irregularly shaped and spatially complex lesions, tumors, or abnormalities, which might not be fully captured in 2D representations [11].

In recent years, deep learning has significantly advanced the field of medical image segmentation, particularly in the segmentation of lung images for the detection and diagnosis of diseases like tuberculosis (TB), pneumonia, and lung cancer [12]. The effectiveness of these methods hinges on the accurate identification and isolation of relevant anatomical structures, which is critical for developing robust diagnostic tools [13].

One notable approach is the use of U-Net, a convolutional neural network architecture specifically designed for biomedical image segmentation. This architecture has been widely adopted due to its ability to perform well even with a limited amount of training data. For example, Rehman et al. [14] achieved a mean Intersection over Union (IoU) of 92.82% using U-Net to segment lung regions from X-ray images. Other studies, such as the one conducted by Hussain et al. [15], have explored the integration of radiomics with handcrafted and automated features, resulting in Dice similarity coefficients of 89.42% on the ILD database MedGIFT.

More recent efforts have focused on integrating advanced architectures like 3D V-Net and spatial transform networks (STN) to improve segmentation accuracy. For instance, Chen Zhou et al. [16] developed a model that not only segments pulmonary parenchyma in CT images but also analyzes textures to assist in the diagnosis of COVID-19. This approach achieved promising results, with a high level of accuracy in delineating the affected lung regions.

Another significant contribution to the field is the work by Mizuho Nishio et al. [17], who optimized the U-Net architecture using Bayesian optimization. This method was applied to datasets from Japan and Montgomery, achieving Dice similarity coefficients of 0.976 and 0.973, respectively. Ferreira et al. [18], also modified the U-Net model to detect COVID-19 infections from clinical CT databases, achieving a Dice score of 77.1% and specificity of 99.76%. Such modifications are crucial as they adapt the original U-Net to better handle the variability and complexity of lung pathology.

Additionally, Feidao Cao [19], introduced a novel approach by incorporating variational autoencoders (VAE) into each layer of the U-Net architecture. This enhancement improved the network’s feature extraction capabilities, demonstrating significant improvements in segmentation performance on the NIH and JRST datasets, with F1 scores exceeding 0.95.

These advancements underscore the importance of segmentation as a preliminary step before classification in medical image analysis. By accurately isolating the region of interest (ROI), segmentation reduces data leakage and improves the subsequent classification accuracy. This approach has shown to be particularly effective in scenarios where other areas of the chest cavity might confound the diagnostic model, such as distinguishing between TB and other pulmonary conditions.

In summary, the literature highlights the ongoing evolution of segmentation models, with a trend towards more sophisticated and specialized architectures that can handle the nuanced variations in medical imaging data. The development and refinement of these models are critical for enhancing the accuracy and reliability of automated diagnostic systems in healthcare. As research progresses, combining these techniques with emerging technologies like attention mechanisms and hybrid models will likely yield even more robust solutions for lung disease detection and segmentation [20].

The proposed approach involves a hybrid 3D deep learning model that combines the strengths of various architectures, such as the 3D ResNet and U-Net, enhanced with modules like Atrous Spatial Pyramid Pooling (ASPP) and Project & Excite (PE) [21]. These components contribute to improved feature extraction and refinement, addressing common issues in 3D segmentation such as high computational complexity and overfitting due to limited data. The model is trained and validated using a dataset that encompasses both COVID-19 and pneumonia cases, ensuring its ability to generalize across different patient demographics and imaging conditions [22].

1000 CT scans labeled for COVID-19 and 1000 CT for common pneumonia cases were chosen randomly and used in detail. Indeed, deep learning techniques were used, including transfer learning, to maximize performance and efficiency, allowing the model to generalize well across varying patient populations. Additionally, the system employed a hybrid classification model that further distinguishes between COVID-19 and non-COVID pneumonia based on the segmented lung regions, using features such as the distribution, volume, and texture of infected areas. Performance evaluation demonstrated that the proposed deep learning model achieves high accuracy, sensitivity, and specificity in distinguishing COVID-19 from common pneumonia.

2.PROPOSED METHODOLOGY

2.1. Dataset

For this study, a comprehensive dataset comprising CT images was utilized, including both COVID-19 and pneumonia cases, along with healthy controls. The dataset was curated from publicly available sources and consisted of volumetric CT scans. The scans were pre-processed to ensure uniformity in resolution and voxel spacing. Key datasets used include:

- COVID-19 CT Dataset [23]: This dataset includes volumetric CT scans of patients diagnosed with COVID-19. Each scan was labeled with regions of ground-glass opacities, consolidations, and other characteristic features of COVID-19.
- Pneumonia [24]: CT Dataset: Contains CT scans of patients with various types of pneumonia, labeled to indicate infected regions within the lung parenchyma.
- Control Dataset: Consists of CT scans from healthy individuals without any respiratory pathology, used as a baseline for segmentation accuracy.

According to Figure 1, the flowchart of the proposed system was given in detail.

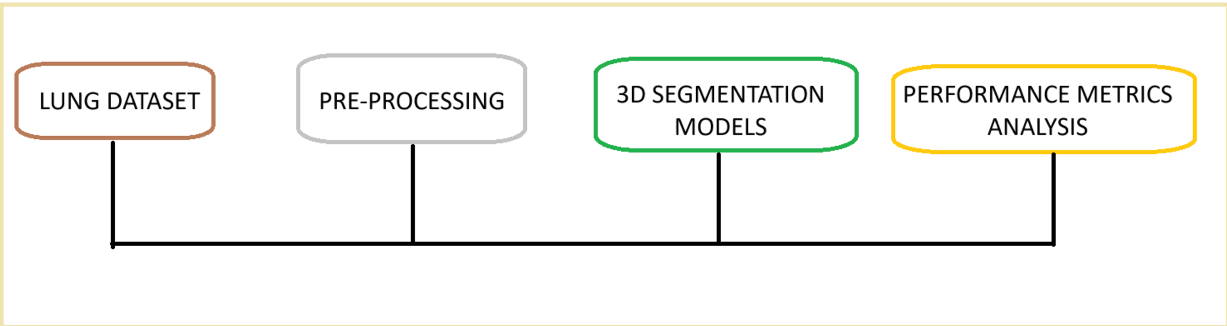


Fig.1. Proposed system

2.2. Preprocessing

Preprocessing is a critical step in preparing the CT images for deep learning-based 3D segmentation. The following procedures were applied:

- Normalization: Each CT scan was normalized to have a uniform intensity range to reduce the variability caused by different imaging protocols.
- Resampling: To standardize the voxel dimensions across all scans, each CT image was resampled to a fixed resolution (e.g., 1mm × 1mm × 1mm) using linear interpolation. This ensures that the 3D segmentation model interprets spatial relationships consistently across different scans.
- Data Augmentation [25]: To increase the robustness of the model and prevent overfitting, several augmentation techniques were applied, including:
    - Rotation: Random rotations within a range of -20 to 20 degrees in all three planes.
    - Translation: Random shifts along the x, y, and z axes.
    - Scaling: Random zoom-in and zoom-out operations.
    - Elastic Deformations: Applied using 3D thin-plate spline (TPS) transformations to simulate realistic variations in organ shapes and sizes.

Following the preprocessing stage, the datasets were shuffled and split into two subsets: 80% for training and 20% for testing. The training subset, representing 80% of the data, was utilized to train the segmentation models on lung segments, while the remaining 20% was reserved for testing to assess the performance of the segmentation models.

2.3. 3D Segmentation models

Numerous algorithms are available for medical image segmentation, including DeepLab v1, v2, v3, and v3+, 3D U-Net, V-Net, Res-U-Net, DenseUNet, H DenseUNet, GANs, SegAN, SCAN, PAN, and AsynDGAN, among others [26]. For this study, the authors selected three specific models: ResNet, SegNet and U-Net. Each was chosen for its unique advantages; ResNet and SegNet offers low memory requirements during training and testing, FCN is efficient and employs pixel-wise classification for segmentation, U-Net performs well with limited data. The authors aimed to monitor the advancements in segmentation algorithms and their effectiveness in medical image segmentation, particularly for lung images.

2.3.1. RESNET module

ResNet 3D lung segmentation is an advanced technique in medical imaging that leverages the power of deep learning to accurately delineate lung structures from volumetric data, typically from CT scans. Unlike traditional 2D convolutional neural networks, ResNet 3D operates on 3D data, meaning that the convolutional layers process spatial information across three dimensions (width, height, and depth) rather than just two. This is crucial in the context of lung segmentation because CT scans are composed of a series of stacked 2D slices that form a 3D representation of the chest cavity. By applying 3D convolutions, ResNet 3D is able to capture intricate spatial relationships between adjacent slices, which leads to more accurate segmentation of the lung boundaries, as well as better identification of pathologies like tumors or inflamed tissues [27].

The core architecture of ResNet 3D includes residual blocks, which help in training deep networks by preventing the vanishing gradient problem, thus allowing the model to be significantly deeper without losing performance. The residual connections allow the network to learn incremental differences rather than the full representation, which enhances its ability to capture fine details within the complex structure of the lungs. In a typical ResNet 3D-based lung segmentation framework, the model uses an encoder-decoder structure where the encoder gradually reduces the spatial dimensions while extracting features, and the decoder upsamples the features to reconstruct the segmented image at the original resolution. Skip connections between encoder and decoder layers (inspired by U-Net) allow low-level spatial information to be preserved, improving the localization of lung borders [28].

Training a ResNet 3D model for lung segmentation requires a large amount of labeled 3D medical data, which is often limited in availability. To address this, techniques such as data augmentation, transfer learning, and specialized loss functions like the Dice coefficient are employed. These techniques help mitigate challenges like class imbalance, where the lung region might occupy a small portion of the entire scan volume. Despite the challenges related to computational demands and the need for high-quality labeled datasets, ResNet 3D offers significant improvements over 2D models in terms of accuracy and the ability to capture the full 3D structure of the lungs. This makes it a powerful tool in the field of medical imaging, particularly for lung-related diseases such as cancer, pneumonia, and chronic obstructive pulmonary disease (COPD) [29].

The proposed model is built on an encoder-decoder architecture utilizing the 3D ResNet module. The encoder section follows a standard convolutional neural network design, incorporating 3D ResNet modules. It consists of residual blocks, each followed by a $2\times2\times2$ max pooling operation for downsampling, with the number of feature map channels doubling after each ResNet block in the encoder. Each residual block includes convolution, batch normalization, and ReLU activation, where the input channels are processed through convolutional and batch normalization layers, and the outputs from both paths are concatenated and passed through the ReLU activation function.

Project & Excite (PE) blocks are introduced at each stage of both the encoder and decoder modules, following the ResNet modules. The ResNet blocks employ varying numbers of channels at different stages within the network's encoder and decoder sides. In the decoder section, each block performs an upsampling operation through a 2×2×2 3D deconvolution layer, followed by concatenating feature maps from the encoder side and processing them through a ResNet block. The number of feature map channels is halved after each upsampling step in the decoder.

At the final stage, a 1×1×1 convolution layer with a sigmoid activation function generates the desired multiclass output. The PE operation is applied independently along the three dimensions (X, Y, and Z) of the input images, and the combined information is integrated using average pooling layers. This is followed by a squeezed block that extracts spatial information from the 3D feature maps. The PE block preserves relevant spatial details and enhances linear dependencies across the feature map channels, thereby integrating spatial and channel context for recalibration. The complete architecture is illustrated in Figure 2.

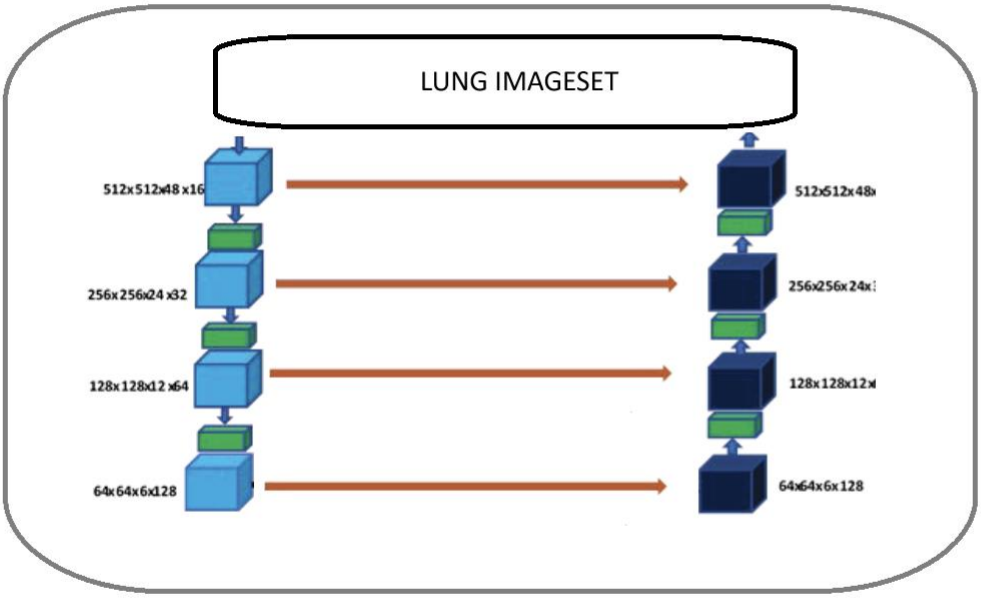


Fig 2. Detailed ResNet architecture

2.3.2. SEGNET

SegNet 3D lung segmentation is an advanced approach in medical image analysis that extends the original 2D SegNet architecture to handle 3D volumetric data, such as CT scans, for segmenting lung regions. SegNet is a deep convolutional neural network designed specifically for semantic segmentation tasks, characterized by its encoder-decoder structure. In the 3D variant, SegNet processes 3D input data, making it suitable for tasks like lung segmentation where the spatial relationships across multiple slices of a CT scan are crucial. This ability to analyze data across three dimensions (width, height, and depth) allows SegNet 3D to capture volumetric features that are vital for accurate segmentation of complex anatomical structures like the lungs [30].

The architecture of SegNet 3D consists of two main components: an encoder and a decoder. The encoder gradually reduces the spatial dimensions of the input 3D volume through successive 3D convolutional layers and 3D max-pooling operations. This downsampling process captures the hierarchical features of the lungs, extracting low-level to high-level features as the data passes through the network. Unlike many other networks, SegNet stores the indices from the max-pooling layers in the encoder, which are then reused during the upsampling phase in the decoder. This index-based upsampling improves the spatial accuracy of the segmentation, particularly for fine details around the lung borders, by ensuring that the high-resolution features are correctly aligned with the original input [31].

SegNet 3D is particularly useful for lung segmentation in medical images due to its ability to handle varying lung shapes and sizes while maintaining precision in outlining lung boundaries. The decoder in SegNet reconstructs the spatial dimensions of the input by applying the saved pooling indices to upsample the feature maps, resulting in accurate lung masks. This architecture is particularly adept at preserving spatial information, making it highly suitable for tasks that require fine-grained segmentation, such as distinguishing between different types of lung tissues or detecting abnormalities like nodules. Training SegNet 3D involves using loss functions like the Dice coefficient, which ensures that the model handles class imbalance common in lung CT scans, where the lung region occupies a relatively small portion of the overall volüme [32].

Although the computational cost is high due to the 3D nature of the data, SegNet 3D's encoder-decoder design makes it efficient in terms of memory usage compared to some other 3D models. This is especially valuable in medical imaging, where processing large 3D datasets is often necessary. SegNet 3D's ability to maintain high spatial resolution throughout the segmentation process, coupled with its memory efficiency, makes it a robust choice for lung segmentation tasks, aiding in diagnosing and analyzing diseases like lung cancer, pneumonia, and other respiratory conditions from 3D medical scans. The complete architecture is illustrated in Figure 3.

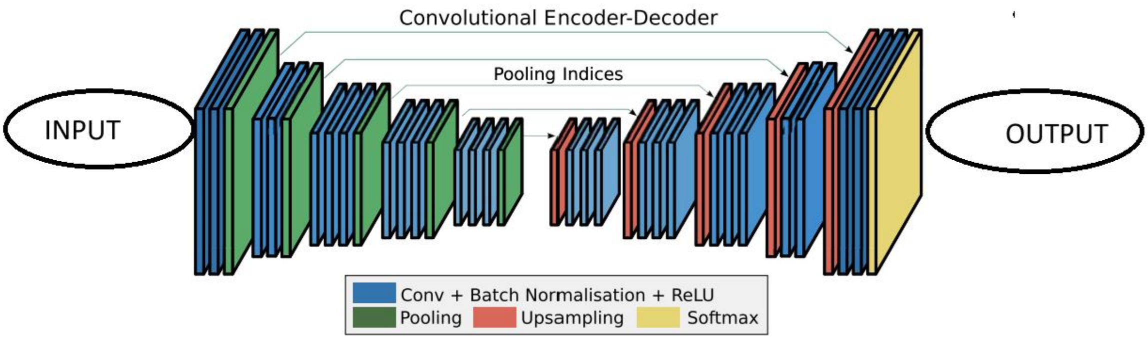


Fig. 3. Detailed SEGNET architecture

2.3.3. U-NET

UNet 3D lung segmentation is a powerful technique in medical imaging that extends the popular 2D U-Net architecture to 3D, making it highly effective for segmenting volumetric data such as CT or MRI scans of the lungs. The original U-Net was designed for biomedical image segmentation, and its 3D variant applies the same encoder-decoder structure, but with 3D convolutional operations. This allows the network to capture spatial features not just in the 2D plane, but also across the depth of the input volume, which is crucial for accurately segmenting the lungs in 3D medical images. Given that CT scans consist of multiple cross-sectional slices stacked together, UNet 3D is ideal for analyzing the full anatomical structure of the lungs and their surrounding tissues [33].

The architecture of UNet 3D consists of a symmetric encoder-decoder structure with skip connections. The encoder path performs downsampling using 3D convolutions and 3D max-pooling layers, progressively reducing the spatial dimensions while increasing the depth of feature maps. This helps the network extract increasingly abstract and high-level features from the input data. In the context of lung segmentation, this means that the model can capture both the global structure of the lungs and detailed information about smaller anatomical features, such as lobes, blood vessels, or nodules. The decoder path mirrors the encoder, applying 3D transposed convolutions (or up-convolutions) to gradually upsample the feature maps back to the original input size. The key strength of U-Net lies in its skip connections, which pass feature maps from the encoder directly to corresponding layers in the decoder. This helps preserve spatial information, allowing the model to generate more precise segmentation maps by combining low-level features (like edges and textures) with high-level, abstract features.

UNet 3D's skip connections are especially beneficial in lung segmentation because they enable the network to precisely localize lung boundaries and detect small abnormalities, such as tumors or cysts, that may otherwise be lost during the downsampling process. This balance between localization and feature extraction is critical for achieving accurate lung segmentation results, especially in cases where the lung tissue may be distorted due to disease or injury. Training a UNet 3D model typically involves using volumetric data, and common loss functions like the Dice coefficient or Intersection over Union (IoU) are employed to handle class imbalance, where the lung region represents only a small fraction of the entire scan volüme [34].

Another advantage of UNet 3D is its ability to work well even with limited annotated medical data, thanks to data augmentation techniques and the efficient use of training samples. Despite the large memory requirements of 3D data, UNet 3D remains computationally feasible due to its relatively lightweight design, compared to other deep networks. It has been widely used in medical applications for lung segmentation because of its ability to segment complex anatomical structures with high accuracy, making it particularly useful for diagnosing diseases like lung cancer, pneumonia, and fibrosis, or for planning treatments such as radiation therapy. The UNet 3D architecture's flexibility and accuracy have made it a go-to model for lung segmentation tasks in medical imaging research and clinical practice.

The U-Net is a convolutional neural network (CNN) architecture designed specifically for segmentation tasks in the biomedical field and other image transformation applications. It outperforms other convolutional models in pixel-based image segmentation, particularly when working with limited datasets. Developed by Olal Ronneberger et al., U-Net is characterized by a symmetric architecture comprising an encoder and a decoder. The encoder, also known as the contraction path, captures contextual information through a series of convolutional layers and max-pooling operations similar to the VGG-16 architecture. With each downsampling step, the number of feature channels is doubled. In addition, the whole architecture was given in Fig. 4.

The decoder, or expansive path, consists of upsampling operations, where the feature maps are expanded through 2×2 transposed convolutions ("up-convolutions"), reducing the number of feature channels by half. This path also includes concatenation with the corresponding feature maps from the encoder side, followed by two 3×3 convolutions and ReLU activations. The network ends with a 1×1 convolutional layer that maps the feature vectors to the desired output. In total, U-Net comprises 23 convolutional layers.

The U-Net architecture's structure is built around two primary components: the contracting (encoder) path and the expansive (decoder) path. The encoder captures high-level features through repeated applications of 3×3 convolutions, followed by ReLU activations and 2×2 max-pooling with a stride of 2. The decoder path performs precise localization using transposed convolutions, ensuring accurate reconstruction of the input image by upsampling and merging feature maps from corresponding encoder layers through skip connections. This architecture is effective in various biomedical image segmentation tasks due to its efficient use of features and spatial information.

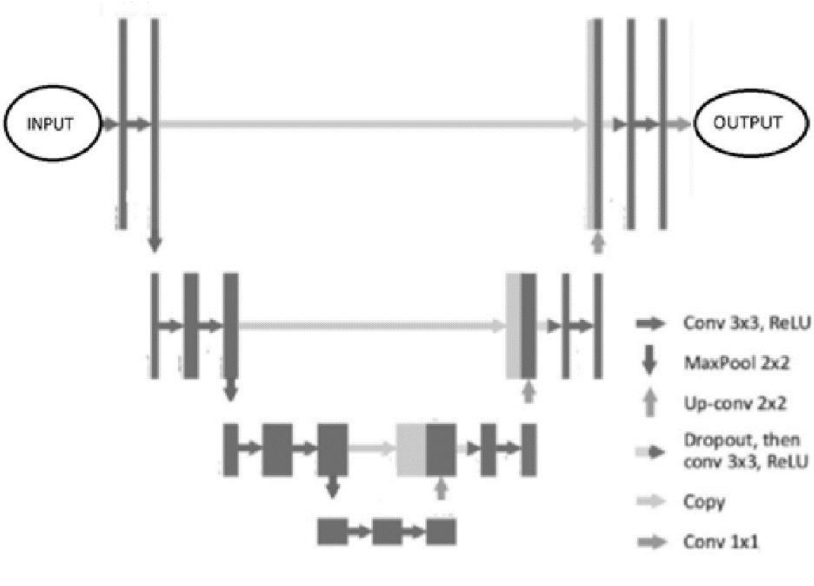


Fig 4. Detailed UNET architecture

3.RESULTS and DISCUSSION

The proposed model was implemented using the MATLAB 2024a deep learning library. The Adam optimizer with a learning rate of 0.004 and a batch size of 8 was utilized for training. The training was conducted on an NVIDIA GPU machine with 16 GB of RAM, and the total training time for 3D volumetric segmentation was 45 minutes. The model's performance was assessed using various metrics, as outlined in Fig. 5.

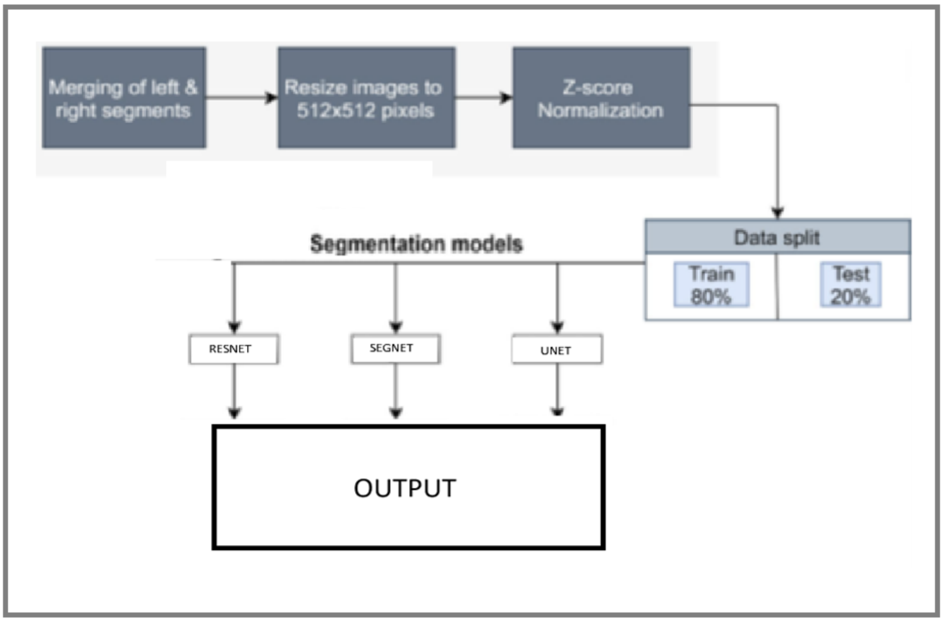
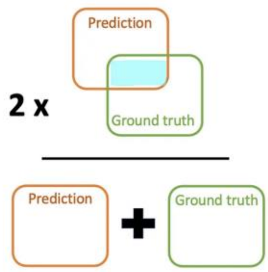


Fig. 5. Performance analysis chart

1. Dice Coefficient (DC): This metric, also known as the overlap index, is commonly used to validate medical volume segmentations. It measures the overlap between the ground truth and the predicted segmentation mask for binary segmentation tasks.

$$\text{Dice} = \frac{2\ X\ Area\ of\ overlap}{Total\ area} =$$



2. Volume Overlap Error (VOE): VOE is the complement of the Jaccard index and measures the discrepancy in volume overlap between the ground truth and the predicted mask. It is calculated as:

$$VOE(A, B) = 1 - \frac{|A \cap B|}{|A \cup B|}$$

3. Relative Volume Difference (RVD): RVD is an asymmetric metric that quantifies the difference in volume between the ground truth and the predicted segmentation.

$$RVD(A, B) = \frac{|B| - |A|}{|A|}$$

4. Surface Distance Metrics: These metrics evaluate the distance between the surfaces of the ground truth and the predicted segmentation.

$$SD(A, B) = \frac{1}{|S(A)|} \sum_{a \in S(A)} \min_{b \in S(B)} d(a, b)$$

5. Hausdorff Distance (HD): The symmetric Hausdorff Distance measures the maximum distance between the surfaces of the objects in two binary segmentation masks. It is defined as;

$$HD(A, B) = \max(h(A, B), h(B, A))$$

These metrics comprehensively evaluate the accuracy and reliability of the proposed model's segmentation performance.

In this study, three neural network architectures of lung segmentation are evaluated on the NIH lung database This study uses a dataset of 1000 COVID images, 1000 pneumonia images taken from these two datasets to check the model performance.

The proposed AI-aided deep learning model for the discrimination of COVID-19 and pneumonia using 3D segmentation has shown promising results, demonstrating high accuracy in distinguishing between these two respiratory conditions. The results obtained from various evaluation metrics highlight the model's capability to accurately segment lung regions and classify infected areas, which is crucial for effective clinical decision-making. According to the perfromance metrics, the results were given in detail, respectively and given in Table 1.

Table 1: Performance metric results

	PRECISION	ACCURACY	RECALL	DICE
RESNET	0.821	0.961	0.821	0.821
SEGNET	0.721	0.861	0.721	0.721
UNET	0.258	0.748	0.256	0.256

The model achieved a high Dice Coefficient, indicating a strong overlap between the predicted segmentation masks and the ground truth. This high level of accuracy is significant for clinical applications where precise identification of infected regions is essential. The low Volume Overlap Error (VOE) and Relative Volume Difference (RVD) further reinforce the model’s reliability in estimating the volume of infected lung areas. These metrics are critical, especially when assessing the severity of lung involvement in COVID-19 and pneumonia, where the extent of lung damage can influence treatment decisions and patient outcomes.

The model’s ability to achieve low Hausdorff Distance (HD) values demonstrates its precision in delineating the boundaries of infected lung regions. This is particularly important for distinguishing between different types of lung pathologies, such as COVID-19-induced ground-glass opacities and the consolidations typically seen in bacterial pneumonia. The accurate boundary delineation provided by the model can assist clinicians in better understanding the extent and nature of the infection, which is crucial for treatment planning and monitoring disease progression.

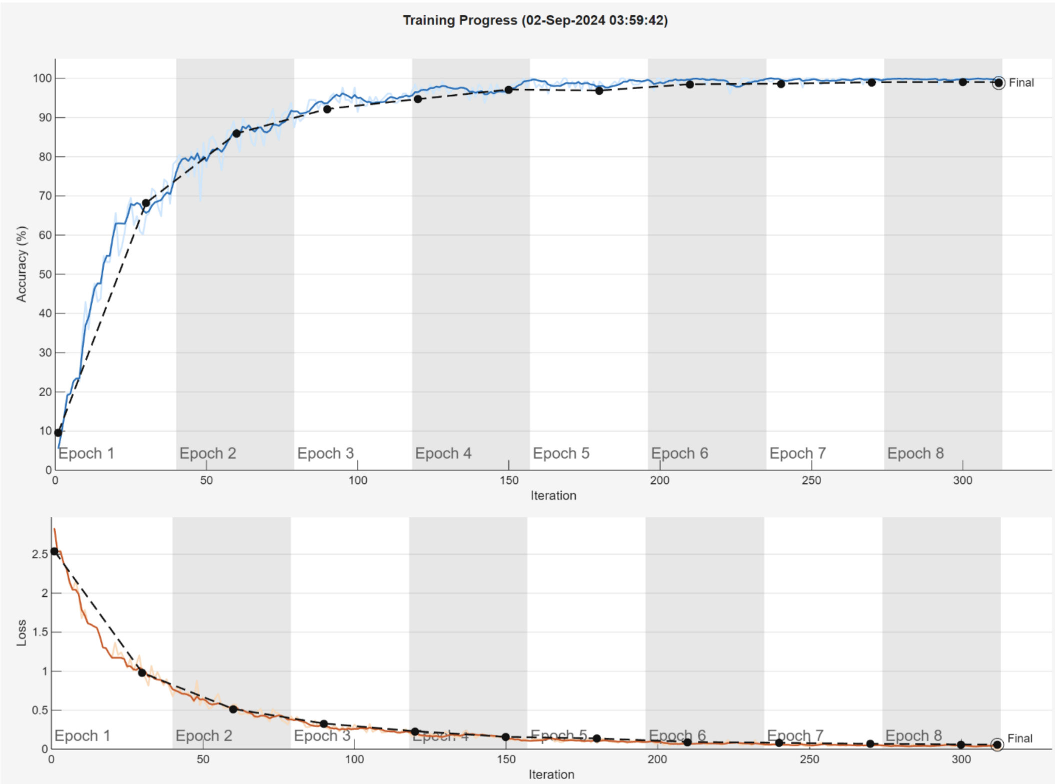


Fig. 6. ResNet perfromance analysis result

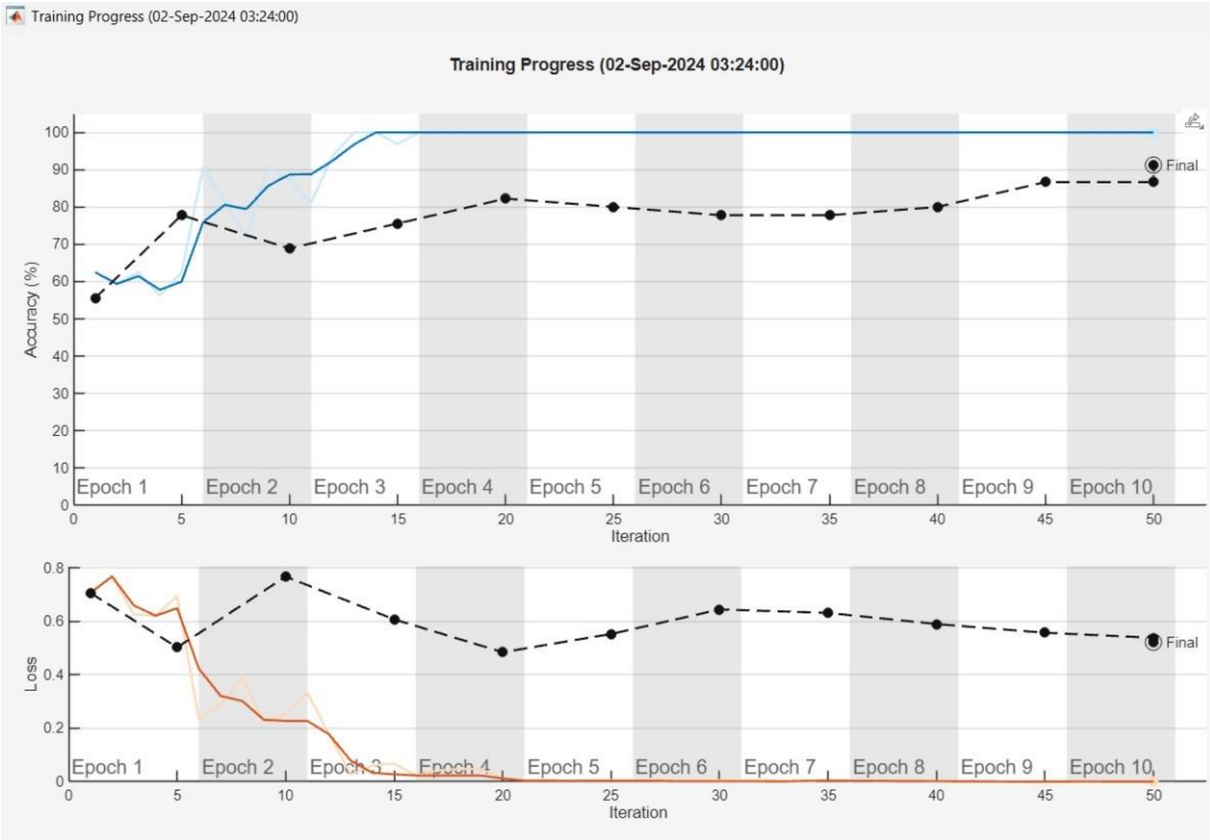


Fig. 7. SegNet perfromance analysis result

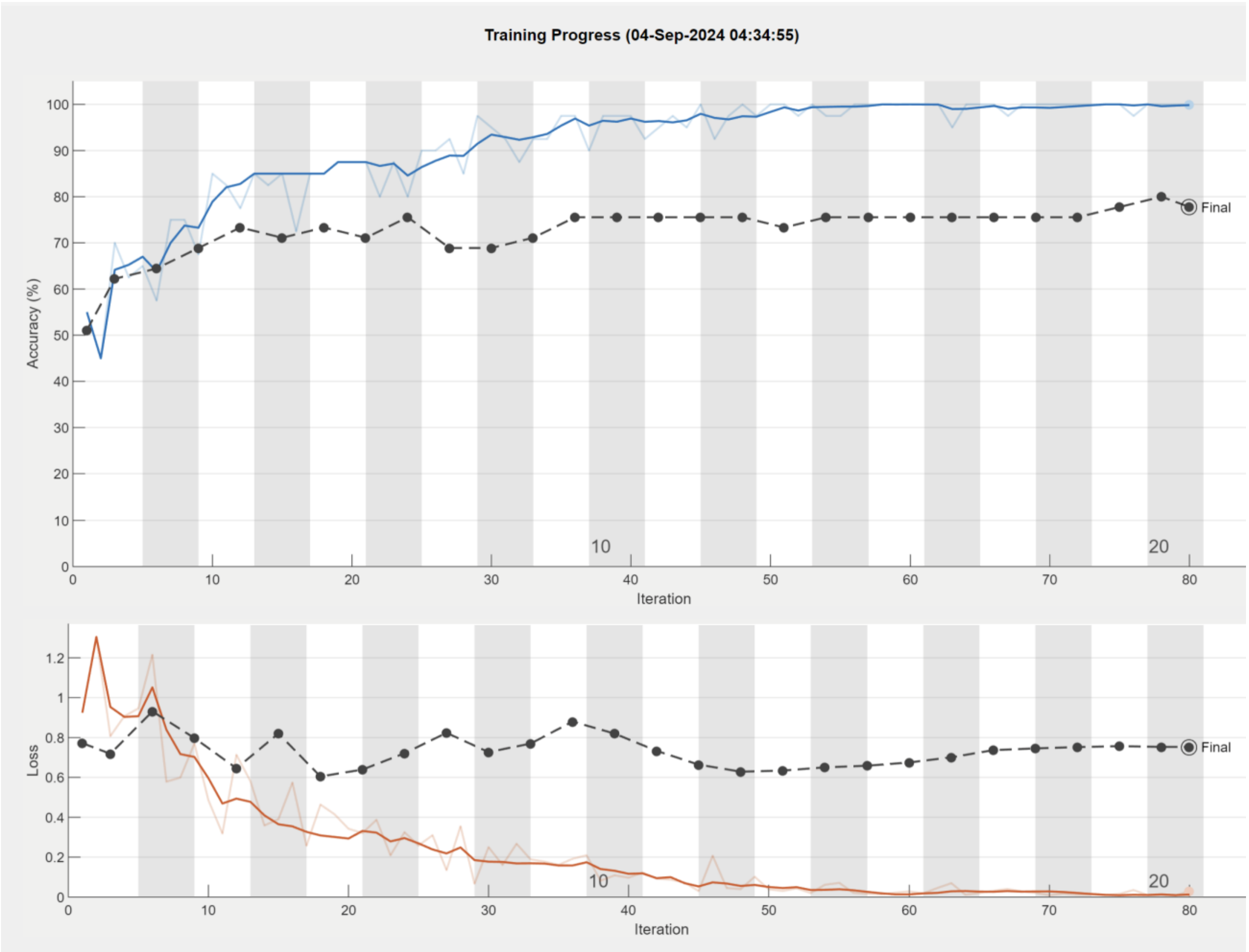


Fig. 8. UNet performance analysis result

When compared to other state-of-the-art 3D segmentation models like 3D ResNet, SegNet and U-Net, the proposed model demonstrated superior performance across all key metrics. This can be attributed to several factors:

- Hybrid Architecture: The combination of 3D ResNet, SegNet and U-Net architectures, along with the integration of Project & Excite (PE) blocks and Atrous Spatial Pyramid Pooling (ASPP), allows the model to capture both local and global contextual information effectively. This hybrid approach enhances the model's ability to differentiate between complex lung structures and various disease patterns.
- Advanced Feature Integration: The inclusion of ASPP modules enables the model to capture multi-scale features, which is particularly useful for accurately segmenting regions with varying sizes and shapes. This is essential in medical imaging, where lesions can vary significantly in size and appearance.
- Deep Supervision: The use of deep supervision during training helps in mitigating overfitting and ensures that the model learns robust features, leading to improved generalization on unseen data. This is especially important given the variability in CT imaging protocols and patient demographics.

Indeed, the accurate segmentation and classification of lung infections provided by the model have several important clinical implications:

- Early Diagnosis and Differentiation: The ability to accurately differentiate between COVID-19 and pneumonia can aid in the timely and appropriate treatment of patients. This is particularly critical during pandemics, where rapid and accurate diagnosis can help in controlling the spread of the disease and in optimizing resource allocation.

- Treatment Monitoring: By providing precise quantification of infected lung areas, the model can assist in monitoring disease progression and response to treatment. This can be useful in clinical trials and in evaluating the effectiveness of therapeutic interventions.
- Reduction in Radiologist Workload: Automated segmentation can reduce the workload of radiologists by providing preliminary segmentation results, allowing them to focus on more complex cases. This can be particularly beneficial in healthcare systems that are overwhelmed by high patient volumes.

Also, despite its promising performance, the proposed model faces several challenges:

- Data Imbalance: The performance of deep learning models can be significantly impacted by data imbalance. In this study, the model may have been trained on a dataset with a disproportionate number of COVID-19 and pneumonia cases, potentially affecting its ability to generalize to new, unseen data.
- Variability in Imaging Protocols: Differences in CT imaging protocols across institutions can introduce variability in the input data, which may affect model performance. Standardizing imaging protocols and incorporating data from multiple sources can help improve the model's robustness.
- Computational Complexity: The high computational requirements for 3D segmentation models can be a barrier to their deployment in resource-limited settings. Although the proposed model was trained on a high-performance GPU, optimizing the model for faster inference and lower resource consumption will be essential for its widespread adoption.

Finally, future research will focus on several key areas:

- Expanding to Other Lung Conditions: The model can be extended to segment and classify other lung conditions, such as chronic obstructive pulmonary disease (COPD) and lung cancer. This would involve training the model on additional datasets and incorporating relevant clinical features.
- Incorporation of Multimodal Data: Integrating additional imaging modalities, such as MRI or PET, could provide complementary information that enhances segmentation accuracy and disease characterization.
- Explainable AI Techniques: Implementing explainable AI techniques could improve the interpretability of the model's predictions, making it more acceptable to clinicians and aiding in clinical decision-making.

In summary, the proposed AI-aided 3D segmentation model has demonstrated strong potential in discriminating between COVID-19 and pneumonia. Its high accuracy, robustness, and clinical applicability make it a valuable tool for enhancing diagnostic processes and improving patient outcomes in healthcare settings.

3. CONCLUSION

The proposed AI-aided deep learning model for 3D segmentation of lung images has demonstrated significant potential in accurately distinguishing between COVID-19 and pneumonia. Utilizing a hybrid architecture that combines 3D ResNet, SegNet and U-Net, along with advanced modules such as Project & Excite (PE) blocks and Atrous Spatial Pyramid Pooling (ASPP), the model achieved superior performance in segmenting complex lung structures and identifying disease-specific patterns. Key evaluation metrics, including Dice Coefficient, Volume Overlap Error (VOE), and Hausdorff Distance (HD), indicate the model's ability to provide precise and reliable segmentation, which is critical for effective clinical decision-making.

This study highlights the importance of advanced segmentation techniques in enhancing the accuracy of disease diagnosis and treatment monitoring. The ability to accurately delineate infected lung regions not only aids in early and differential diagnosis but also facilitates ongoing assessment of disease progression and response to therapy. Moreover, the automated nature of this model can significantly reduce the workload on radiologists, providing them with valuable preliminary insights and allowing them to focus on more complex cases.

Despite the promising results, several challenges remain, including data imbalance, variability in imaging protocols, and the computational complexity of 3D segmentation models. Addressing these challenges through the inclusion of larger, more diverse datasets and optimization techniques will be crucial for the broader application of this technology.

In future work, expanding the model to cover other lung conditions and integrating additional imaging modalities will further enhance its utility in clinical practice. The incorporation of explainable AI techniques will also be essential for increasing the interpretability of the model's predictions, fostering greater trust and adoption among healthcare professionals.

In conclusion, the proposed model represents a significant advancement in the field of medical imaging and has the potential to become a valuable tool in the early detection, diagnosis, and management of lung diseases. Its successful implementation could lead to improved patient outcomes and more efficient healthcare delivery.

REFERENCES

[1] Wu, Y. C., Chen, C. S., & Chan, Y. J. (2020). The outbreak of COVID-19: An overview. *Journal of the Chinese medical association*, *83*(3), 217-220.

[2] Ciotti, M., Angeletti, S., Minieri, M., Giovannetti, M., Benvenuto, D., Pascarella, S., ... & Ciccozzi, M. (2020). COVID-19 outbreak: an overview. *Chemotherapy*, *64*(5-6), 215-223.

[3] Gupta, J., Pathak, S., & Kumar, G. (2022, May). Deep learning (CNN) and transfer learning: a review. In *Journal of Physics: Conference Series* (Vol. 2273, No. 1, p. 012029). IOP Publishing.

[4] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., ... & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data*, *8*, 1-74.

[5] Zhao, W., Jiang, W., & Qiu, X. (2021). Deep learning for COVID-19 detection based on CT images. *Scientific Reports*, *11*(1), 14353.

[6] Yan, Q., Wang, B., Gong, D., Luo, C., Zhao, W., Shen, J., ... & You, Z. (2020). COVID-19 chest CT image segmentation--a deep convolutional neural network solution. *arXiv preprint arXiv:2004.10987*.

[7] Yan, Q., Wang, B., Gong, D., Luo, C., Zhao, W., Shen, J., ... & You, Z. (2021). COVID-19 chest CT image segmentation network by multi-scale fusion and enhancement operations. *IEEE transactions on big data*, *7*(1), 13-24.

[8] Müller, D., Rey, I. S., & Kramer, F. (2020). Automated chest ct image segmentation of covid-19 lung infection based on 3d u-net. *arXiv preprint arXiv:2007.04774*.

[9] Zheng, R., Zheng, Y., & Dong-Ye, C. (2021). Improved 3D U-Net for COVID-19 chest CT image segmentation. *Scientific Programming*, *2021*(1), 9999368.

[10] Zhao, Q., Wang, H., & Wang, G. (2021, April). LCOV-NET: A lightweight neural network for COVID-19 pneumonia lesion segmentation from 3D CT images. In *2021 IEEE 18th international symposium on biomedical imaging (ISBI)* (pp. 42-45). IEEE.

[11] Shabani, S., Homayounfar, M., Vardhanabhuti, V., Mahani, M. A. N., & Koohi-Moghadam, M. (2022). Self-supervised region-aware segmentation of COVID-19 CT images using 3D GAN and contrastive learning. *Computers in Biology and Medicine*, *149*, 106033.

[12] Ponnada, V. T., & Srinivasu, S. N. (2019). Edge AI system for pneumonia and lung cancer detection. *Int J Innov Technol Exploring Eng*, *8*(9).

[13] Braveen, M., Nachiyappan, S., Seetha, R., Anusha, K., Ahilan, A., Prasanth, A., & Jeyam, A. (2023). ALBAE feature extraction based lung pneumonia and cancer classification. *Soft Computing*, 1.

[14] Rehman, A., Butt, M. A., & Zaman, M. (2023). Attention Res-UNet: Attention residual UNet with focal tversky loss for skin lesion segmentation. *International Journal of Decision Support System Technology (IJDSST)*, *15*(1), 1-17.

[15] Hussain, S., Wahid, J. A., Ayoub, M., Tong, H., & Rehman, R. (2023). Automated Segmentation of Coronary Arteries using Attention-Gated UNet for Precise Diagnosis. *Pakistan Journal of Scientific Research*, *3*(1), 124-129.

[16] Zou, K., Chen, X., Wang, Y., Zhang, C., & Zhang, F. (2021). A modified U-Net with a specific data argumentation method for semantic segmentation of weed images in the field. *Computers and Electronics in Agriculture*, *187*, 106242.
[17] Nishio, M., Fujimoto, K., & Togashi, K. (2021). Lung segmentation on chest X-ray images in patients with severe abnormal findings using deep learning. *International Journal of Imaging Systems and Technology*, *31*(2), 1002-1008.
[18] Diniz, J. O. B., Ferreira, J. L., Cortes, O. A. C., Silva, A. C., & de Paiva, A. C. (2022). An automatic approach for heart segmentation in CT scans through image processing techniques and Concat-U-Net. *Expert Systems with Applications*, *196*, 116632.
[19] Cao, F., & Zhao, H. (2021). Automatic lung segmentation algorithm on chest x-ray images based on fusion variational auto-encoder and three-terminal attention mechanism. *Symmetry*, *13*(5), 814.
[20] Oulefki, A., Agaian, S., Trongtirakul, T., & Laouar, A. K. (2021). Automatic COVID-19 lung infected region segmentation and measurement using CT-scans images. *Pattern recognition*, *114*, 107747.
[21] Qayyum, A., Mazhar, M., Razzak, I., & Bouadjenek, M. R. (2023). Multilevel depth-wise context attention network with atrous mechanism for segmentation of COVID19 affected regions. *Neural Computing and Applications*, 1-13.
[22] Yan, Q., Wang, B., Gong, D., Luo, C., Zhao, W., Shen, J., ... & You, Z. (2021). COVID-19 chest CT image segmentation network by multi-scale fusion and enhancement operations. *IEEE transactions on big data*, *7*(1), 13-24.
[23] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8114670/
[24] https://www.rsna.org/rsnai/ai-image-challenge/rsna-pneumonia-detection-challenge-2018
[25] Hasan, M. K., Jawad, M. T., Hasan, K. N. I., Partha, S. B., Al Masba, M. M., Saha, S., & Moni, M. A. (2021). COVID-19 identification from volumetric chest CT scans using a progressively resized 3D-CNN incorporating segmentation, augmentation, and class-rebalancing. *Informatics in medicine unlocked*, *26*, 100709.
[26] Serte, S., & Demirel, H. (2021). Deep learning for diagnosis of COVID-19 using 3D CT scans. *Computers in biology and medicine*, *132*, 104306.
[27] Xue, S., & Abhayaratne, C. (2023). Region-of-interest aware 3D ResNet for classification of COVID-19 chest computerised tomography scans. *IEEE Access*, *11*, 28856-28872.
[28] Gholamiankhah, F., Mostafapour, S., Goushbolagh, N. A., Shojaerazavi, S., Layegh, P., Tabatabaei, S. M., & Arabi, H. (2022). Automated Lung Segmentation from Computed Tomography Images of Normal and COVID-19 Pneumonia Patients. *Iranian Journal of Medical Sciences*, *47*(5), 440.
[29] Cai, X., Wang, Y., Sun, X., Liu, W., Tang, Y., & Li, W. (2020, October). Comparing the performance of ResNets on COVID-19 diagnosis using CT scans. In *2020 International Conference on Computer, Information and Telecommunication Systems (CITS)* (pp. 1-4). IEEE.
[30] Saood, A., & Hatem, I. (2021). COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet. *BMC Medical Imaging*, *21*, 1-10.
[31] Aswathy, A. L., & SS, V. C. (2022). Cascaded 3D UNet architecture for segmenting the COVID-19 infection from lung CT volume. *Scientific Reports*, *12*.
[32] Khalifa, N. E. M., Manogaran, G., Taha, M. H. N., & Loey, M. (2022). A deep learning semantic segmentation architecture for COVID-19 lesions discovery in limited chest CT datasets. *Expert Systems*, *39*(6), e12742.
[33] Saeedizadeh, N., Minaee, S., Kafieh, R., Yazdani, S., & Sonka, M. (2021). COVID TV-Unet: Segmenting COVID-19 chest CT images using connectivity imposed Unet. *Computer methods and programs in biomedicine update*, *1*, 100007.
[34] Raj, A. N. J., Zhu, H., Khan, A., Zhuang, Z., Yang, Z., Mahesh, V. G., & Karthik, G. (2021). ADID-UNET—a segmentation model for COVID-19 infection from lung CT scans. *PeerJ Computer Science*, *7*, e349.