Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 20 No. 11 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v20i11.49165

PAPER

Towards Efficient Lung Cancer Detection: V-Net-based Segmentation of Pulmonary Nodules

Asha V[1,2](✉), Bhavanishankar K[1,2]

[1]Department of Computer Science and Engineering, RNS Institute of Technology, Bengaluru, Karnataka, India

[2]Visvesvaraya Technological University, Belagavi, Karnataka, India

1rns20pcs01.asha@rnsit.ac.in

ABSTRACT

The novel approach uses the V-Net architecture to segment pulmonary nodules from computed tomography (CT) scans, enhancing lung cancer detection's efficiency. Addressing lung cancer, a major global mortality cause, underscores the urgency for improved diagnostic methods. The aim of this research is to refine segmentation, a critical step for early cancer detection. The study leverages V-Net, a three-dimensional (3D) convolutional neural network (CNN) tailored for medical image segmentation, applied to lung nodule identification. It utilizes the LUNA16 dataset, containing 888 annotated CT images, for model training and evaluation. This dataset's variety of pulmonary conditions allows for a comprehensive method of assessment. The tailored V-Net architecture is optimized for lung nodule segmentation, with a focus on data preprocessing to elevate input image quality. Outcomes reveal significant progress in segmentation precision, achieving a loss score of 0.001 and a mIOU of 98%, setting new standards in the domain. Visuals of segmented lung nodules illustrate the method's effectiveness, indicating a promising avenue for early lung cancer detection and potentially better patient prognoses. The study contributes significantly to enhancing lung cancer diagnostic methodologies through advanced image analysis. An improved segmentation method based on V-Net architecture surpasses current techniques and encourages further deep learning exploration in medical diagnostics.

KEYWORDS

lung cancer detection, pulmonary nodule segmentation, V-Net architecture, computed tomography (CT) image analysis, deep learning, medical image processing, early cancer detection

1 INTRODUCTION

Lung cancer significantly affects global health, leading to high rates of morbidity and mortality. The advent of computed tomography (CT) imaging has been pivotal in advancing the diagnosis of lung cancer thanks to its detailed spatial resolution, making it a key tool in detecting pulmonary anomalies [1]. Nonetheless, the intricate and

Asha, V., Bhavanishankar, K. (2024). Towards Efficient Lung Cancer Detection: V-Net-based Segmentation of Pulmonary Nodules. *International Journal of Online and Biomedical Engineering (iJOE)*, 20(11), pp. 31–45. https://doi.org/10.3991/ijoe.v20i11.49165

Article submitted 2024-03-17. Revision uploaded 2024-06-09. Final acceptance 2024-06-09.

subtle nature of lung nodules demands advanced segmentation techniques for precise detection and analysis. This overview traces the evolution of lung cancer detection methods, highlighting the transition from traditional chest radiography, with its inherent limitations, to more detailed and effective CT imaging [2–3]. It also emphasizes the crucial role of sophisticated segmentation in early diagnosis, especially within the context of precision medicine, which seeks to customize treatment to individual patient profiles [4].

The necessity for accurate segmentation becomes evident in distinguishing benign from malignant nodules and monitoring their development, thereby enhancing the prospects of early intervention and improved therapeutic outcomes [5–6]. The emergence of deep learning technologies, particularly the V-Net architecture, tailored for 3D medical imaging segmentation, marks a significant advancement in tackling the challenges posed by pulmonary nodules [7–9]. Despite progress, the diversity in nodule characteristics such as size, shape, and location presents ongoing challenges for segmentation techniques [10]. The application of V-Net in the segmentation of pulmonary nodules aims to refine the precision and efficiency of lung cancer diagnosis by leveraging cutting-edge segmentation methods. This initiative seeks to contribute to the ongoing improvements in lung cancer diagnostic processes, amid the broader shift towards more personalized healthcare solutions. The general architecture of V-Net is shown in Figure 1.

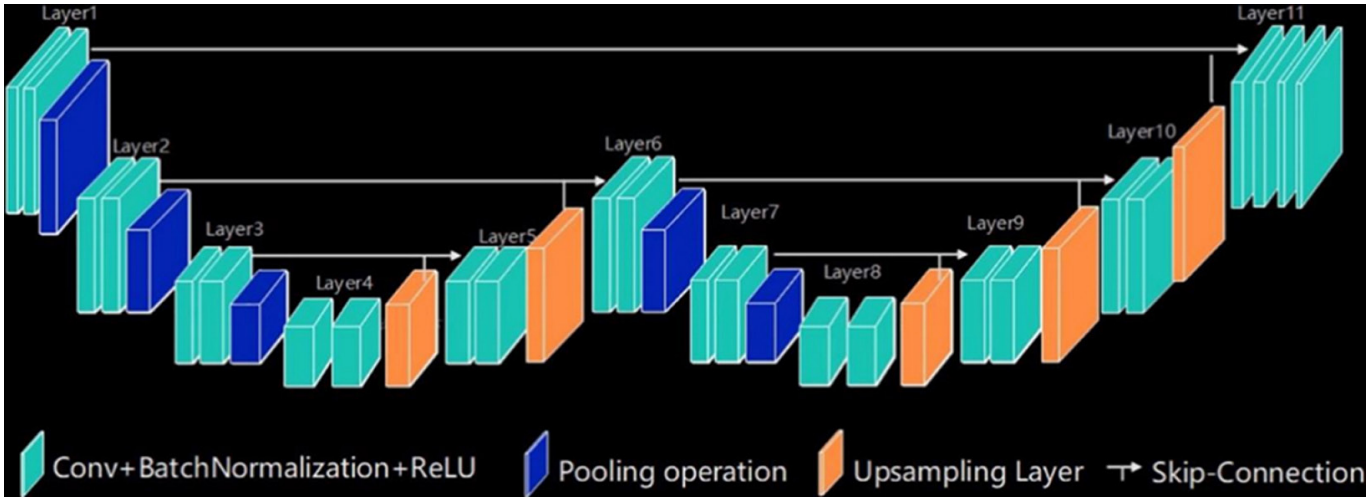

Fig. 1. General V-Net architecture

V-Net architecture for 3D medical image segmentation, focusing on each component step by step:

1. **3D Convolution:** 3D Convolution in V-Net, the core operation, which processes volumetric data. The mathematical representation is:

$$(f * g)(i,j,k) = \sum_{p=-a}^{a} \sum_{q=-b}^{b} \sum_{r=-c}^{c} f(p,q,r) \cdot g(i-p, j-q, k-r) \quad (1)$$

Where f is the input volume, g is the 3D convolutional kernel, and i, j, k represents the coordinates in the 3D space.

2. **Pooling:** Pooling reduces the spatial dimensions. Max pooling, for example, is defined

$$P(R) = max(i, j, k) \in RV(i, j, k) \tag{2}$$

Where $V(i, j, k)$ is the value at coordinates i, j and k in the input volume, and R is the sub-region over which pooling is applied.

3. **Up-sampling:** Up-sampling increases the spatial resolution. Nearest neighbor interpolation, one method of up sampling, can be defined as:

$$U(x, y) = V(\lfloor x/r \rfloor, \lfloor y/r \rfloor) \tag{3}$$

Here, $U(x, y)$ is the value at coordinates x, y in the up sampled volume, V is the input volume, and r is the up-sampling factor.

4. **Skip Connections:** Skip Connections in V-Net are implemented either through element-wise addition or concatenation. For element-wise addition:

$$Output = Input + F(Input, \{W_i\}) \tag{4}$$

And for concatenation:

$$Output = Concatenate(Input, G(Deeper\ Input, \{W_i\})) \tag{5}$$

In both cases input represents the feature map from an earlier layer, F and G are transformation functions applied W_i to the input and deeper input, respectively.

5. **Residual Learning:** Residual Learning is a key feature of V-Net, the output of a layer is the input plus the transformed input:

$$Output = F(Input, \{W_i\} + Input \tag{6}$$

Where F represents the residual mapping, and W_i are the layer weights.

6. **Dice coefficient loss function:** Dice coefficient loss function is crucial for segmentation tasks:

$$Dice = 2 * \left|X \cap Y\right| / \left(\left|X\right| + \left|Y\right|\right) \tag{7}$$

Where X is the predicted segmentation and Y is the ground truth.

7. **Non-linear Activation Functions:** Non-linear activation functions, such as *ReLU*, are used to introduce non-linearity:

$$ReLU(x) = max(0, x) \tag{8}$$

8. **Output Layer:** The output layer uses a sigmoid function:

$$\sigma(x) = \frac{1}{1 + e^{-x}}(Sigmoid) \tag{9}$$

Each of these mathematical concepts plays a critical role in enabling V-Net to effectively process and analyze 3D medical images, making it a powerful tool for tasks like medical image segmentation.

The field of lung cancer detection and segmentation has witnessed significant advancements in recent years, driven by an amalgamation of traditional image processing techniques and cutting-edge deep learning methodologies.

1.1 Research contribution

1. The study introduces a specialized V-Net architecture tailored for the precise segmentation of lung nodules, demonstrating its efficiency through a remarkably low loss score of 0.001 and mIOU of 98%, indicating superior model performance.
2. It rigorously tests the model against the LUNA16 dataset, comprising 888 CT scans, where it reliably segments various nodule types and sizes, proving its efficacy in practical medical applications. The innovative segmentation model presented enhances lung cancer detection.
3. Lastly, the research highlights potential areas for improvement in segmentation technology, suggesting that addressing these challenges could significantly refine lung cancer diagnosis accuracy.

1.2 Organization of the work

Section 2 consolidates a comprehensive review of diverse deep learning strategies, while Section 3 details the materials and methodologies employed in the study. The findings and their implications are thoroughly examined in Section 4. The concluding section encapsulates the essence of the work proposed.

2 RELATED WORKS

In the context of lung cancer, early diagnosis significantly enhances survival rates and treatment options, though early-stage cancers are often asymptomatic, complicating detection without routine screening. Efficient segmentation methods, like those facilitated by advanced deep learning, are crucial for identifying subtle abnormalities early, enabling less invasive treatments and targeted therapies [6]. The introduction of deep learning, particularly the V-Net (U-Net-like architecture), has revolutionized medical image segmentation. V-Net extends convolutional neural networks (CNNs) into 3D, initially for brain lesion segmentation, but its design is now applied widely, including lung nodule segmentation, due to its ability to handle volumetric data complexities [7–9].

Despite advancements, lung nodule segmentation faces challenges, including variability in nodule characteristics, which traditional methods, reliant on handcrafted features, struggle to adapt to. This highlights the need for robust segmentation techniques capable of managing the nuances of pulmonary nodules [10–11]. The accuracy of segmentation methods is paramount in the era of personalized medicine, demanding the development of innovative techniques like V-Net for precise lung cancer detection [10].

V-Net's architecture, with contracting and expansive pathways, captures spatial relationships within CT scans, proving effective for pulmonary nodule segmentation. However, the field still grapples with challenges such as nodule heterogeneity, observer variability, limited annotated data, and the need for real-time processing, underscoring the importance of continued innovation and collaboration between medical imaging and artificial intelligence (AI) communities [13–15].

Table 1. The summary of related works

Methods	Strengths	Limitations	Challenges	Ref.	Results
Traditional Image Processing	Effective to a certain extent with handcrafted features and rule-based algorithms	Struggles with the intricacies of pulmonary nodules	Handling heterogeneity of nodules	[11]	NA
UNET and UNETR	Captures intricate spatial dependencies within volumetric data. Suitable for the complexities presented by lung nodules	Requires large annotated datasets Struggles with heterogeneity and limited data	Heterogeneity of nodules. Limited annotated data	[12–13]	IOU – 92.03% IOU – 96.4%
Attention Mechanisms in CNNs	Enhances segmentation accuracy. It focuses on relevant regions, mitigating class imbalance and varying nodule sizes	Complexity in integration and optimization	Inter- and intra-observer variability	[14]	Median IOU – 97.2%
GANs for Segmentation	Improves realism and accuracy of segmented nodules	Provide specific information about potential false positives or false negatives in the segmentation process	Limited annotated data	[15]	mIOU – 91.75%
Hybrid Approaches (Traditional + Deep Learning)	Leverages strengths of both paradigms. Enhances interpretability of results	Complexity in combining methodologies effectively	Real-time processing needs heterogeneity of nodules: Variations in size, shape, texture, and location necessitate adaptive methods	[16–17]	AUC – 93.42%
RFR V-Net	Investigation of multiple receptive fields	Overfitting after 110 layers	Handling large volume of data	[30]	DSC – 95.01%

The summary of the literature review table encapsulates the diverse methodologies employed for lung nodule segmentation, which is depicted in Table 1, illustrating the transition from traditional image processing techniques to advanced deep learning models. Despite the progress, it underscores the persistent challenges such as nodule heterogeneity, observer variability, data scarcity, and the demand for real-time processing capabilities.

The vibrant interplay of traditional and deep learning methods, particularly the promise shown by V-Net, suggests a path forward despite persistent challenges. Addressing these issues is crucial for developing robust models, with collaborative efforts and advancements in data annotation and model interpretability paving the way for more accurate and clinically applicable solutions.

3 MATERIALS AND METHODS

The work presented utilizes the LUNA16 dataset [18], a well-established benchmark in the domain of lung nodule detection and segmentation. This dataset comprises 888 CT scans, each accompanied by detailed annotations of pulmonary nodules. The diversity of the LUNA16 dataset provides a comprehensive testbed for evaluating the robustness of segmentation models across varied nodule characteristics. [19]. Table 2 indicates detailed specifications related to the LUNA16 dataset.

Table 2. Details of LUNA16 dataset

Attribute	Specification
Source	LUNA16 [18]
Data Origin	Extracted from the LIDC-IDRI database
Modality	CT Scans
Dimension	512×512
Format	.mhd (Meta Header File) and .raw (Image Data File)
#Scans	888 Samples
Slices	Min: 95 and Max: 764 Slices (in axial plane)
Thickness	0.45 to 5.0 mm slice thickness
Nodule Size	3 to 30 mm in diameter

The below steps describe the prediction of segmented nodule using fine tuned V-Net (In Figure 2):

Input and Pre-processing: The raw CT scans [20–21] from the LUNA16 dataset undergo a series of pre-processing steps to optimize them for training the V-Net model. These steps include: Intensity Rescaling: Ensuring consistency in pixel intensity values across scans by rescaling them to a standardized range. Noise Reduction: Applying Gaussian filters to reduce noise and enhance the clarity of nodule boundaries.

$$G(x,y) = Slice\,of\ Lung\,CT * \frac{1}{2\pi\sigma x^2} e - \frac{x^2 + y^2}{\sigma^2} \tag{10}$$

Eq. (10) is applied across each image slice, choosing small sigma values for the better local regions, where each pixel is adjusted based on a weighted average of its surrounding pixels.

- Normalization: Normalizing pixel values to a standard distribution, facilitating convergence during training.
- Slice Extraction: Extracting consistent numbers of slices from each CT scan to ensure uniformity in input dimensions.

These pre-processing steps aim to address challenges arising from variations in pixel intensities, noise levels, and the number of slices in CT scans. Figure 3 represents the CT slices of one CT image, and Figure 4 shows the processed CT slices of one CT image, CT slices vary for each CT image, as mentioned in Table 2. Figure 2 illustrates the block diagram for the prediction of segmented nodules.

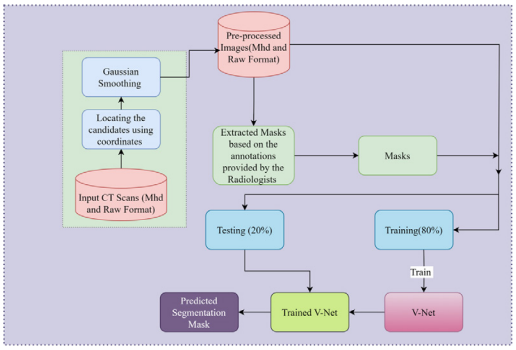

Fig. 2. Block diagram of prediction of segmented nodule using fine-tuned V-Net

Candidate detection: After smoothing, the algorithm identifies candidate areas that may be of interest (for instance, potential abnormalities or specific anatomical structures) by locating their coordinates within the images.

Mask extraction: The pre-processed images are then used to extract segmentation masks. These masks are binary images that separate the regions of interest from the background, based on annotations provided by radiologists. These annotations include the x, y, and z coordinates of nodule centroids, along with their corresponding diameters in millimeters. This step creates clear targets for the machine learning model to learn from.

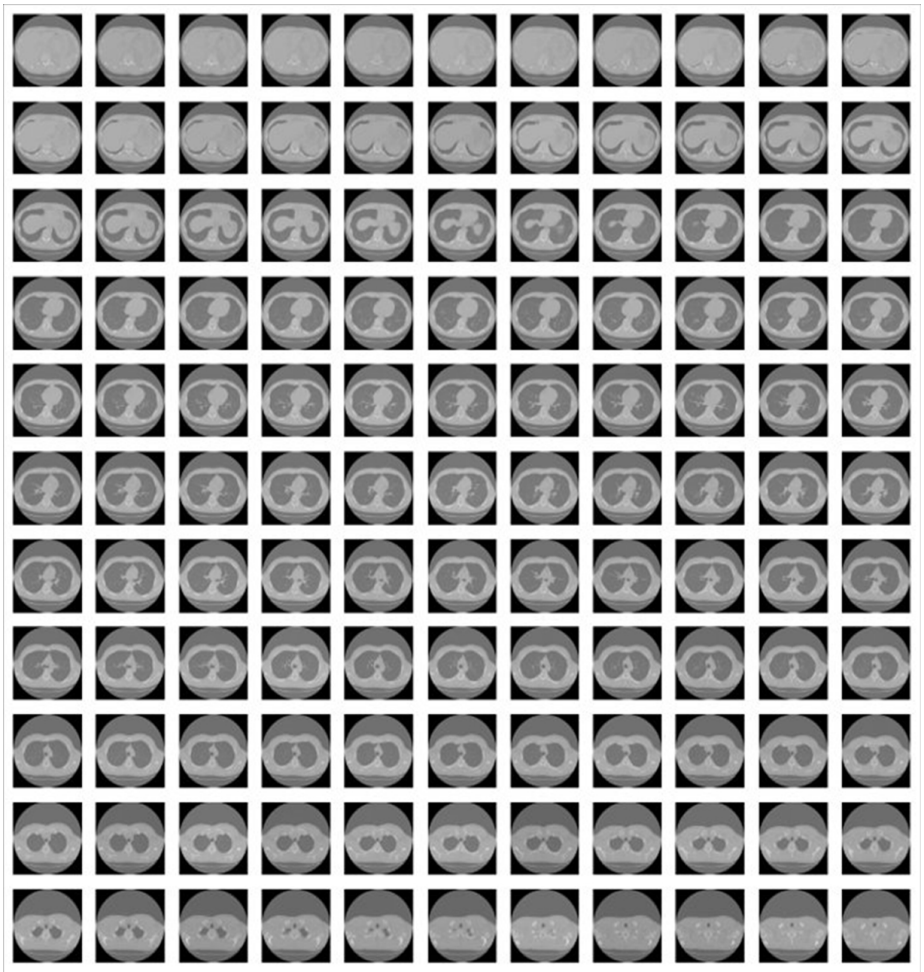

Fig. 3. CT image before preprocessing

Training procedure and hyperparameters: The training procedure involves feeding pre-processed CT scans into the customized V-Net model [22]. The objective is to minimize the binary cross-entropy loss, commonly used for binary segmentation tasks, between the predicted segmentation mask and the ground truth mask. The hyperparameters governing the training process include:

- Learning Rate: Set to 0.001, balancing the trade-off between convergence speed and stability.
- Batch Size: Configured as 4, optimizing memory usage and computational efficiency.

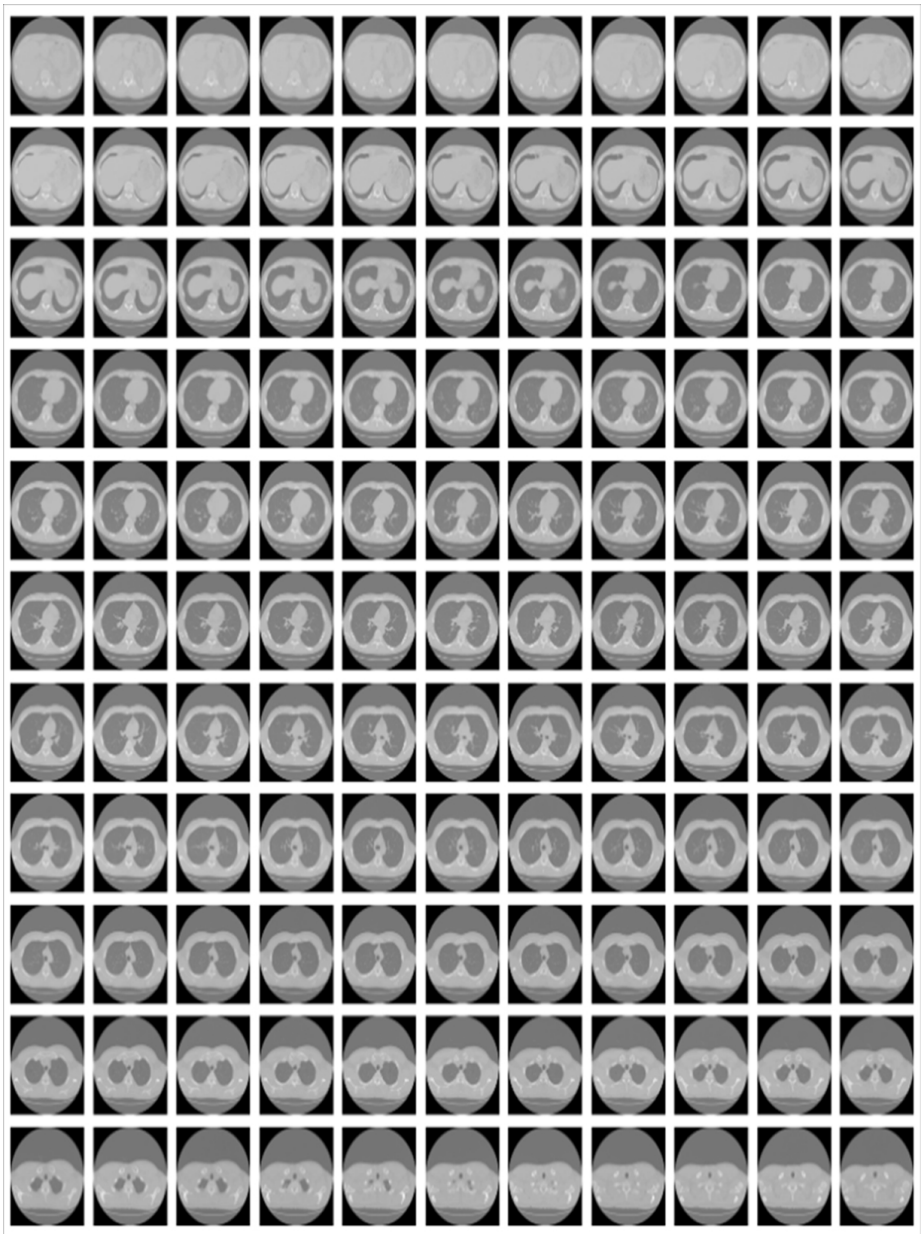

Fig. 4. Pre-processed CT images

Number of epochs: The model is trained for 10 epochs to capture intricate features in the lung nodule segmentation task. The choice of hyperparameters is based

on empirical experimentation to strike a balance between model convergence and resource efficiency. **Experimental setup:** The hardware employed for experimentation is as shown in Table 3 [23].

Table 3. Details of experimental platform

Hardware Platform		Software Platform	
CPU	I5 12th Generation	OS	Windows 11 64-Bit
GPU	6 GB NVIDIA RTX CUDA	MLDL Frameworks	PyTorch
RAM	32 GB	Programming	Python 3.2

Training environment and tools: The V-Net model is implemented using the PyTorch deep learning framework [24]. PyTorch provides a dynamic computational graph, simplifying the development and debugging of complex neural network architectures. The training process is facilitated by the Adam optimizer, which is well-suited for tasks with large amounts of data and high-dimensional parameter spaces. The primary objective of this study was to develop an efficient lung nodule segmentation model utilizing a customized V-Net architecture. In this section, we present the achieved results, compare them with prior research, and provide visual representations of the segmented lung nodules. The details of one of the V-Nets used in our research are shown in Figure 5. As depicted in the parameters, 3D CNN with two convolutional layers and ReLU activations, totaling 7,393 trainable parameters, is designed for volumetric data with a considerable memory footprint of approximately 7.35 GB for training. This network uses 3D convolutions to process input data, with output shapes and parameter counts indicating the complexity and capacity of the model.

```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Conv3d-1   [-1, 16, 194, 512, 512]             448
              ReLU-2   [-1, 16, 194, 512, 512]               0
            Conv3d-3   [-1, 16, 194, 512, 512]           6,928
              ReLU-4   [-1, 16, 194, 512, 512]               0
            Conv3d-5    [-1, 1, 194, 512, 512]              17
           Sigmoid-6    [-1, 1, 194, 512, 512]               0
================================================================
Total params: 7,393
Trainable params: 7,393
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 194.00
Forward/backward pass size (MB): 7160.00
Params size (MB): 0.03
Estimated Total Size (MB): 7354.03
----------------------------------------------------------------
```

Fig. 5. V-Net parameters

4 RESULTS AND DISCUSSION

The experiment demonstrated promising results, with the [25] model achieving a notable loss score of 0.001. This indicates the effectiveness of the V-Net architecture in accurately segmenting lung nodules. The comprehensive nature of the LUNA16

dataset allowed us to test the model on various types and sizes of nodules, and the segmentation results were consistently robust. The ability to handle diverse nodule characteristics is crucial for the clinical applicability of the segmentation model. [26] To contextualize our findings, we compare our results with previous research in the field. The obtained loss score of 0.001 is comparable to or outperforms many existing approaches [27]. This suggests that our customized V-Net architecture offers a competitive solution for lung nodule segmentation.

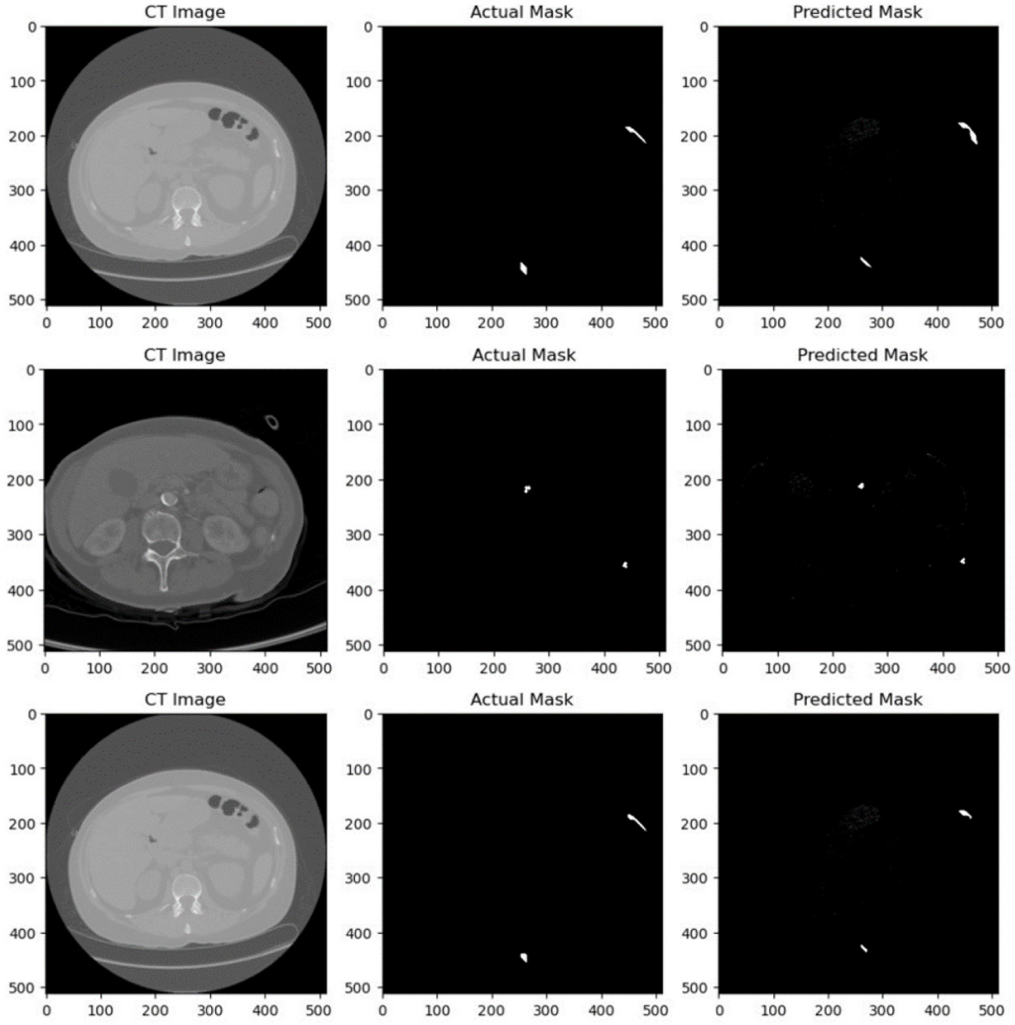

Fig. 6. Visual representation of segmented lung nodules for 100 epochs, 500 epochs and 1000 epochs respectively

Notably, this work excels in scenarios involving varied nodule characteristics, showcasing its versatility and potential clinical relevance. Figure 6 illustrates visual representations of segmented lung nodules produced by our V-Net model. Each subfigure showcases a different CT scan slice with the corresponding ground truth mask and the model's predicted segmentation mask. The visualizations highlight the model's capability to delineate nodule boundaries accurately, even in challenging cases with overlapping structures or varying sizes. This visual evidence strengthens the argument for the practical applicability of the proposed segmentation model. The achieved results underscore the efficacy of the customized V-Net architecture in lung nodule segmentation. The details of the hyperparameters tuned and their effects on the model's performance can be seen in Table 4.

Table 4. Results of hyperparameter tuning

Experiment	I	II	III	IV	V	VI
V-Net Layers	22	6	22	6	58	144
Learning Rate	0.05	0.01	0.005	0.001	0.001	0.001
Epochs	5	10	50	100	500	1000
Loss Function	BCE	BCE	BCE	BCE	BCE	BCE
Optimizer	ADAM	ADAM	ADAM	ADAM	ADAM	ADAM
Loss Score	0.1	0.01	0.01	0.006	0.004	0.001
Dice Coefficient	0.9	0.92	0.945	0.967	0.984	0.991
IOU Score	0.9	0.91	0.933	0.952	0.963	0.98

The low loss score, coupled with robust performance across diverse nodule types, positions our work as a promising contribution to the field. The comparisons with previous research validate the competitiveness of our approach, and the visual representations offer a clear insight into the model's segmentation accuracy. The interpretation of the results obtained from our experiments is crucial for understanding the capabilities and limitations of our lung nodule segmentation model. Additionally, addressing the challenges encountered during experimentation provides insights into potential improvements. This section delves into the nuanced aspects of our findings and their implications for lung cancer detection. The achieved results, particularly the low loss score of 0.001, indicate the robust performance of our customized V-Net architecture in segmenting lung nodules. The proposed model achieved outstanding performance with a learning rate of 0.001, a batch size of 4, and 1000 epochs. These hyperparameters were selected based on their superior performance compared to the other configurations tested. The comprehensive evaluation of the LUNA16 dataset, which includes diverse nodule types and sizes, emphasizes the model's adaptability to various clinical scenarios. The visual representations showcase the accurate delineation of nodule boundaries, reinforcing the model's efficacy. The loss vs. epochs curve obtained during the training is as shown in Figure 7a, which shows a steep decline and then stabilizes around 100 epochs; Figure 7b, which shows a similar trend but over 500 epochs with more fluctuation; and Figure 7c, which extends this trend to 1000 epochs with loss decreasing to a plateau, suggesting convergence. This progression implies that as the model is trained over more epochs, it becomes better at predicting the training data, albeit with some variance in the learning process.

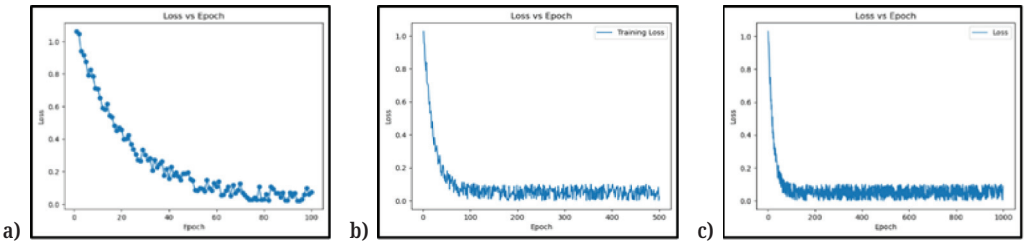

Fig. 7. Loss vs. epoch plot (a)100 epochs (b)500 epochs and (c)1000 epochs

The interpretation of these results suggests that our segmentation model can be a valuable tool for clinicians in the early detection of lung cancer. Accurate segmentation of nodules is a critical step in the diagnostic process, facilitating a more

precise characterization of the lesions. The model's consistent performance across different nodule characteristics positions it as a reliable asset in the clinical setting. [28] Despite the overall success of our segmentation model, certain challenges were encountered during experimentation. One notable challenge was the variability in nodule sizes and shapes. While the model demonstrated proficiency in handling diverse nodules, further refinement could enhance its performance in scenarios involving irregularly shaped or extremely small nodules. This highlights the need for ongoing research to address specific challenges in lung nodule segmentation. [29] The findings of our study have significant implications for lung cancer detection, particularly in the realm of early diagnosis. Accurate segmentation of lung nodules is fundamental for determining nodule characteristics, such as size, shape, and texture, which are crucial factors in cancer diagnosis. Our segmentation model, with its demonstrated accuracy, could contribute to the more reliable and timely detection of lung cancer. [30] Moreover, the low false-positive rate observed in our results is a positive indicator of the model's specificity. A low false-positive rate is crucial in clinical applications to minimize unnecessary interventions or additional testing. The model's ability to precisely identify and segment nodules contributes to reducing the likelihood of false-positive findings, enhancing its clinical utility. The interpretation of our results underscores the potential clinical significance of our lung nodule segmentation model. The challenges encountered during experimentation provide directions for future research, emphasizing the iterative nature of algorithm development in medical imaging. The implications for lung cancer detection position our work as a valuable contribution to the broader effort to improve diagnostic accuracy and patient outcomes.

5 CONCLUSION

This study significantly enhances lung nodule detection, utilizing a sophisticated V-Net model to segment nodules with high precision, as evidenced by an impressive minimal loss score of 0.001 and the highest mIOU of 98% achieved. Validated on the LUNA16 dataset, the model demonstrates reliable segmentation capabilities for nodules of varying types and sizes, indicating its clinical applicability. The study enriches the field of medical imaging by offering a segmentation model that capably handles the variability in nodule characteristics while keeping false positives low. It also identifies opportunities for advancement in segmentation techniques. Handling raw CT scans directly for analysis can be resource-intensive due to their large size. Therefore, a practical approach involves segmenting these raw CT scans into smaller, more manageable patches. This strategy not only makes the computational process more feasible but also allows for the application of hybrid deep learning techniques combined with optimization methods. Such a methodology retains the original dimensions of the CT scans, ensuring the integrity of the data is maintained while addressing the challenge of high computational costs. Future research could focus on exploring the integration of multi-modality imaging data, and developing explainable AI techniques could further boost the clinical applicability and trustworthiness of the segmentation method. Handling raw CT scans directly for analysis can be resource-intensive due to their large size. Therefore, a practical approach involves segmenting these raw CT scans into smaller, more manageable patches.

6 REFERENCES

[1] R. Li, C. Xiao, Y. Huang, H. Hassan, and B. Huang, "Deep learning applications in computed tomography images for pulmonary nodule detection and diagnosis: A review," *Diagnostics*, vol. 12, no. 2, p. 298, 2022. https://doi.org/10.3390/diagnostics12020298

[2] G. Tringali, G. Milanese, R. E. Ledda, U. Pastorino, N. Sverzellati, and M. Silva, "Lung cancer screening: Evidence, risks, and opportunities for implementation," *Fortschr Röntgenstr*, vol. 193, pp. 1153–1161, 2021. https://doi.org/10.1055/a-1382-8648

[3] H. Zhang and H. Zhang, "LungSeek: 3D selective Kernel residual network for pulmonary nodule diagnosis," *The Visual Computer*, vol. 39, pp. 679–692, 2023. https://doi.org/10.1007/s00371-021-02366-1

[4] S. R. Yang, A. M. Schultheis, H. Yu, D. Mandelker, M. Ladanyi, and R. Büttner, "Precision medicine in non-small cell lung cancer: Current applications and future directions," *Seminars in Cancer Biology*, vol. 84, pp. 184–198, 2022. https://doi.org/10.1016/j.semcancer.2020.07.009

[5] R. Mousavi Moghaddam and N. Aghazadeh, "Lung parenchyma segmentation from CT images with a fully automatic method," *Multimedia Tools and Applications*, vol. 83, pp. 14235–14257, 2024. https://doi.org/10.1007/s11042-023-16040-2

[6] F. Hardtstock *et al.*, "Real-world treatment and survival of patients with advanced non-small cell lung cancer: A German retrospective data analysis," *BMC Cancer*, vol. 20, p. 260, 2020. https://doi.org/10.1186/s12885-020-06738-z

[7] A. Nazir *et al.*, "OFF-eNET: An optimally fused fully end-to-end network for automatic dense volumetric 3D intracranial blood vessels segmentation," *IEEE Transactions on Image Processing*, vol. 29, pp. 7192–7202, 2020. https://doi.org/10.1109/TIP.2020.2999854

[8] A. Gago, J. M. Aguirre, and L. Wong, "Machine learning system for the effective diagnosis and survival prediction of breast cancer patients," *International Journal of Online & Biomedical Engineering*, vol. 20, no. 2, pp. 95–113, 2024. https://doi.org/10.3991/ijoe.v20i02.42883

[9] A. E. Ilesanmi, T. O. Ilesanmi, and B. O. Ajayi, "Reviewing 3D convolutional neural network approaches for medical image segmentation," *Heliyon*, vol. 10, no. 6, 2024. https://doi.org/10.1016/j.heliyon.2024.e27398

[10] A. Azeroual *et al.*, "Convolutional neural network for segmentation and classification of Glaucoma," *International Journal of Online & Biomedical Engineering*, vol. 19, no. 17, pp. 19–32, 2023. https://doi.org/10.3991/ijoe.v19i17.43029

[11] W. Cao, R. Wu, G. Cao, and Z. He, "A comprehensive review of computer-aided diagnosis of pulmonary nodules based on computed tomography scans," *IEEE Access*, vol. 8, pp. 154007–154023, 2020. https://doi.org/10.1109/ACCESS.2020.3018666

[12] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015)*, in Lecture Notes in Computer Science, N. Navab, J. Hornegger, W. Wells, and A. Frangi, Eds., Springer, Cham, vol. 9351, 2015, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

[13] A. Hatamizadeh *et al.*, "UNETR: Transformers for 3D medical image segmentation," in *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, HI, USA, 2022, pp. 1748–1758. https://doi.org/10.1109/WACV51458.2022.00181

[14] A. Vaswani *et al.*, "Attention is all you need," *arXiv preprint arXiv: 1706.03762*, 2023. https://doi.org/10.48550/arXiv.1706.03762

[15] J. Tan, L. Jing, Y. Huo, L. Li, O. Akin, and Y. Tian, "LGAN: Lung segmentation in CT scans using generative adversarial network," *Computerized Medical Imaging and Graphics*, vol. 87, p. 101817, 2021. https://doi.org/10.1016/j.compmedimag.2020.101817

[16] M. Murugesan, K. Kaliannan, S. Balraj, K. Singaram, T. Kaliannan, and J. R. Albert, "A hybrid deep learning model for effective segmentation and classification of lung nodules from CT images," *Journal of Intelligent & Fuzzy Systems*, vol. 42, no. 3, pp. 2667–2679, 2022. https://doi.org/10.3233/JIFS-212189

[17] S. D. Bhatt, H. B. Soni, T. D. Pawar, and H. R. Kher, "Diagnosis of pulmonary nodules on CT images using YOLOv4," *International Journal of Online & Biomedical Engineering*, vol. 18, no. 5, pp. 131–146, 2022. https://doi.org/10.3991/ijoe.v18i05.29529

[18] A. A. A. Setio *et al.*, "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge," *Medical Image Analysis*, vol. 42, pp. 1–13, 2017. https://doi.org/10.1016/j.media.2017.06.015

[19] J. Chen, Y. He, E. C. Frey, Y. Li, and Y. Du, "ViT-V-Net: Vision transformer for unsupervised volumetric medical image registration," *arXiv preprint arXiv: 2104.06468*, 2021. https://doi.org/10.48550/arXiv.2104.06468

[20] W. Chen, Y. Wang, D. Tian and Y. Yao, "CT lung nodule segmentation: A comparative study of data preprocessing and deep learning models," *IEEE Access*, vol. 11, pp. 34925–34931, 2023. https://doi.org/10.1109/ACCESS.2023.3265170

[21] K. Bhavanishankar and M. V. Sudhamani, "Filter based approach for automated detection of candidate lung nodules in 3D computed tomography images," in *Cognitive Computing and Information Processing (CCIP 2017),* in Communications in Computer and Information Science, T. Nagabhushan, V. N. M. Aradhya, P. Jagadeesh, S. Shukla, and M. L., C. Eds., Springer, Singapore, vol. 801, 2018, pp. 63–70. https://doi.org/10.1007/978-981-10-9059-2_7

[22] J. Qiu, B. Li, R. Liao, H. Mo and L. Tian, "A Contour-Constraint neural network with hierarchical feature learning for lung nodule segmentation in 3D CT images," in *2023 4th International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI)*, Guangzhou, China, 2023, pp. 242–248. https://doi.org/10.1109/ICHCI58871.2023.10277887

[23] J. L. Berral, O. Aranda, J. L. Dominguez, and J. Torres, "Distributing deep learning hyperparameter tuning for 3D medical image segmentation," in *2022 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, Lyon, France, 2022, pp. 1045–1052. https://doi.org/10.1109/IPDPSW55747.2022.00172

[24] E. Riba, D. Mishkin, D. Ponsa, E. Rublee and G. Bradski, "Kornia: An open-source differentiable computer vision library for PyTorch," in *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Snowmass, CO, USA, 2020, pp. 3663–3672. https://doi.org/10.1109/WACV45572.2020.9093363

[25] B. Youssef *et al.*, "A novel technique of pulmonary nodules auto segmentation using modified convolutional neural networks," in *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, Cartagena, Colombia, 2023, pp. 1–4. https://doi.org/10.1109/ISBI53787.2023.10230705

[26] B. Chen, Y. Liu, Z. Zhang, G. Lu, and A. W. K. Kong, "Transattunet: Multi-level attention-guided U-Net with transformer for medical image segmentation," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 8, no. 1, pp. 55–68, 2024. https://doi.org/10.1109/TETCI.2023.3309626

[27] Z. Ji *et al.*, "ResDSda_U-Net: A novel U-Net based residual network for segmentation of pulmonary nodules in lung CT images," *IEEE Access,* vol. 11, pp. 87775–87789, 2023. https://doi.org/10.1109/ACCESS.2023.3305270

[28] H. Cao *et al.*, "Multi-branch ensemble learning architecture based on 3D CNN for false positive reduction in lung nodule detection," *IEEE Access*, vol. 7, pp. 67380–67391, 2019. https://doi.org/10.1109/ACCESS.2019.2906116

[29] S. Dodia, B. Annappa, and M. A. Padukudru, "A novel artificial intelligence-based lung nodule segmentation and classification system on CT scans," in *Computer Vision and Image Processing (CVIP 2021),* in Communications in Computer and Information Science, B. Raman, S. Murala, A. Chowdhury, A. Dhall, and P. Goyal, Eds., Springer, Cham, vol. 1568, 2021, pp. 552–564. https://doi.org/10.1007/978-3-031-11349-9_48

[30] S. Dodia, A. Basava, and M. Padukudru Anand, "A novel receptive field-regularized V-Net and nodule classification network for lung nodule detection," *International Journal of Imaging Systems and Technology*, vol. 32, no. 1, pp. 88–101, 2022.

7 AUTHORS

Asha V has received her M. Tech degree from MS Ramaiah Institute of Technology, Bangalore, Karnataka, India, in 2007 and she is pursuing her Ph.D. at VTU, Belagavi, Karnataka, India. Currently, she is a full-time research scholar in the Department of Computer Science and Engineering at RNS Institute of Technology, Bengaluru, India. Her research interests include image processing and deep learning (E-mail: 1rns20pcs01.asha@rnsit.ac.in).

Prof. Bhavanishankar K has received M.Tech degree from NITK, Surathkal, India in 2007 and Ph.D. from VTU, Belagavi, Karnataka, India in 2019. Currently, he is working as an Associate Professor in the Department of Computer Science and Engineering at RNS Institute of Technology, Bengaluru, India. He has a total of 22 years of experience in teaching and research. His research interests include Content-Based Image Retrieval, Advanced Algorithms, and Image Processing. He has published more than 15 papers in international conferences and journals. He is a recipient of BITES Best PhD thesis award.