

Received 12 August 2023, accepted 2 September 2023, date of publication 8 September 2023,
date of current version 18 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3313174

RESEARCH ARTICLE

Integrated Deep Learning and Stochastic Models for Accurate Segmentation of Lung Nodules From Computed Tomography Images: A Novel Framework

BASSANT E. YOUSSEF^{1,2}, AHMED ALKSAS^{1,3}, AHMED SHALABY⁴, ALI H. MAHMOUD¹,
ERIC VAN BOGAERT⁵, NORAH SALEH ALGHAMDI⁶, ALYSSA NEUBACHER¹,
SOHAIL CONTRACTOR⁵, MOHAMMED GHAZAL⁷, (Senior Member, IEEE),
ADEL S. ELMAGHRABY³, AND AYMAN EL-BAZ¹, (Fellow, IEEE)

¹Department of Bioengineering, University of Louisville, Louisville, KY 40292, USA

²Department of Engineering Mathematics and Physics, Alexandria University, Alexandria 21544, Egypt

³Computer Science and Engineering Department, University of Louisville, Louisville, KY 40292, USA

⁴Lyda Hill Department of Bioinformatics, UT Southwestern Medical Center, Dallas, TX 75390, USA

⁵Department of Radiology, University of Louisville, Louisville, KY 40292, USA

⁶Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia

⁷Electrical, Computer, and Biomedical Engineering Department, Abu Dhabi University, Abu Dhabi, United Arab Emirates

Corresponding author: Ayman El-Baz (aselba01@louisville.edu)

This work was supported by Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, under Project PNURSP2023R40. The work of Mohammed Ghazal was supported by the Abu Dhabi University's Office of Research and Sponsored Programs under Grant #19300792.

This work involved human subjects or animals in its research. The authors confirm that all human/animal subject research procedures and protocols are exempt from review board approval.

ABSTRACT This paper introduces an innovative model for precise extraction of lung nodules from 3D computed tomography (CT) scans. Our approach comprises two essential preprocessing stages aimed at refining search accuracy and nodule segmentation. Initially, we leverage a two-level joint Markov-Gibbs random field (MGRF) model to delineate the lung region, effectively distinguishing lung wall nodules from the chest region with shared visual characteristics. Subsequently, employing a deep learning U-net technique, we pinpoint the region of interest (ROI) housing the lung nodule, minimizing the inclusion of surrounding lung tissues. Further enhancement comes from a 3D U-net, trained with a novel loss function to mitigate under- or over-segmentation issues. The resulting segmentation robustly outlines lung nodules in terms of morphology and volume metrics, validated by Dice coefficient (DCE), absolute volume difference (AVD), 95th-percentile Hausdorff distance (HD), sensitivity, and specificity metrics. To assess our approach, we conducted comprehensive experiments. Our evaluation encompasses *in vivo* data from 50 patients and employs 679 subjects from the publicly available dataset of the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI). The LIDC-IDRI dataset, a seminal resource for computer-aided diagnosis (CAD) in lung nodules, offers annotations enabling tasks like detection, segmentation, classification, and quantification. Our experiments showcase our model's superiority over existing deep learning methods, particularly evident in metrics such as the 95th-percentile HD and DCE. While limited demographic information constrains a comprehensive analysis, our approach's robust performance underlines its potential integration into nodule assessment AI systems.

The associate editor coordinating the review of this manuscript and approving it for publication was Hengyong Yu¹

INDEX TERMS Computed tomography (CT), joint stochastic model, lung nodules, Markov-Gibbs random field, segmentation, 3D-UNet.

I. INTRODUCTION

Lung cancer is responsible for 25% of all cancer cases in the United States. In 2023, an estimated 238,340 people in the United States will be diagnosed with lung cancer, and 127,070 lung cancer patients will die [1]. Developing an artificial-based system to analyze detected lung nodules based on objective imaging markers using the current image analysis framework plays an immense role in the early diagnosis of lung nodules, especially small lung nodules that are far away from large airways and rarely cause symptoms [2]. To extract these imaging, or radiomic, markers such as shape complexity, nodule size, and appearance inhomogeneities, the first step is to extract/segment the nodule from surrounding tissues, which will be this paper's main focus.

A. RELATED WORK

Several strategies have emerged to tackle the intricate task of segmenting lung nodules from the intricate structures present in computed tomography (CT) images. These strategies can be broadly classified into two categories: deformable model-based methods and deep learning-based approaches. In the following section, we delve into a comprehensive overview of these two categories, highlighting their distinctive features and contributions.

1) DEFORMABLE MODEL-BASED LUNG NODULE SEGMENTATION

Deformable models, renowned for their ability to capture complex shapes and boundaries, have found significant application in image processing and computer vision, including the segmentation of anatomical structures like lung nodules in CT images [3]. These models excel in scenarios where objects possess irregular boundaries or intricate contours, such as lung nodules [3].

Farag et al. [4] introduced a method employing active shape models (ASM) and template matching to detect lung abnormalities. Incorporating Markov-Gibbs random field (MGRF) models, [5] enhanced segmentation accuracy and robustness by capturing nuanced appearance variations.

Itai et al. [6] presented a multi-step methodology that involves pre-processing, lung segmentation, and nodule extraction using deformable models. Gonçalves et al. [7] explored Hessian-based techniques, employing a process that encompasses pre-processing, computation of the Hessian matrix, and rule-based elimination of false positives.

Shakir et al. [8] proposed a 3D segmentation technique combining hybrid level sets, leveraging intensity, texture, and edge features to improve accuracy. In contrast, Shakibapour et al. [9] adopted an unsupervised metaheuristic search approach for segmentation, integrating shape, intensity, and texture features through optimization techniques.

The approach by Roy et al. [10] stands out by merging a level-set model, effectively combining intensity-based and boundary-based attributes to accurately segment nodules within lung regions. Rakesh and Mahesh [11] presented a holistic approach involving thresholding, morphological operations, and region growing, allowing for multiple stages of nodule segmentation.

Savic et al. [12] adopted a fast-marching level set method, refining lung nodule segmentation through an initial threshold and subsequent fast-marching technique application.

2) DEEP LEARNING-BASED TECHNIQUES

Deep learning has emerged as a significant advancement in various fields, revolutionizing the way complex patterns are learned and interpreted from data. Within the realm of medical imaging, particularly in the analysis of CT scans of lungs, there is a prominent trend where deep learning models are being incorporated for nodule segmentation models [13], [14], [15], [16]. These models, including convolutional neural networks (CNNs), harness large labeled datasets to discern patterns that differentiate nodules from surrounding tissues. After training, these models autonomously segment nodules in new images. Several network architectures, such as U-Net, Mask R-CNN, and Dual-Branch Residual Networks (DBRN), have been proposed for this purpose.

Cao et al. [17] introduced the DBRN, harnessing two parallel branches for feature extraction and fusion, which enhances the representation of nodules. The work of Roy et al. [18] bridged deep learning and shape-driven level sets, using a CNN for nodule characteristic recognition and initializing the level set method.

Huang et al. [19] presented a two-stage deep CNN-based method for efficient nodule detection and segmentation. Aresta et al. [20] proposed iW-Net, combining a U-Net-based segmentation network with an interaction network to allow users to interactively refine segmentation results.

Singadkar et al. [21] designed a deep deconvolutional residual network that employs a two-pronged approach for feature capture and strategic connections, improving nodule segmentation. Dong et al. [22] introduced a novel multi-view secondary input residual CNN that capitalizes on diverse input angles for enhanced 3D lung tumor segmentation.

Xiao et al. [23] enhanced the conventional Res2Net with 3D U-Net, creating the 3D-Res2UNet model and incorporating multiple techniques for improved segmentation precision. Shi et al. [24] compared multiscale residual U-Net with fuzzy C-means clustering, demonstrating the superiority of the former.

Khan et al. [25] proposed a novel framework merging segmentation and classification through a VGG-SegNet architecture, extracting deep attributes for improved nodule detection. Yu et al. [26] introduced a comprehensive algorithm utilizing

3D-Res U-Net and 3D-ResNet50 for segmentation and recognition tasks.

Kido et al. [27] employed the N3D-FCN model for candidate extraction and refinement, achieving accurate lung nodule segmentation. Lung PAYNet, a pyramidal deep learning architecture with attention mechanisms, was presented by Bruntha et al. [28] for lung nodule segmentation.

Bhattacharjee et al. [29] introduced ResiUNet, combining U-Net and ResNet152 to enhance segmentation accuracy. Usman and Shin [30] proposed DEHA-Net, incorporating hard attention networks and adaptive ROI mechanisms for improved segmentation precision.

Halder and Dey [31] offered a framework integrating pre-processing, segmentation, and classification modules, using an innovative combination of techniques to address challenges at each stage.

Comparing these diverse techniques, it becomes evident that deformable model-based approaches excel in managing intricate contours, which are prevalent in lung nodules. On the other hand, deep learning-based methods offer the advantage of data-driven pattern recognition. Deformable model-based approaches often require manual feature engineering, while deep learning approaches automatically learn features from data. Notably, certain approaches, such as Roy et al.'s fusion of CNNs and shape-driven level sets [18], demonstrate the potential for synergizing deep learning and traditional techniques to achieve improved results.

B. CHALLENGES AND CONTRIBUTIONS

Despite their substantial progress, both deformable model-based and deep learning-based strategies encounter specific challenges. Deformable model-based methods may encounter difficulties in accurately segmenting certain nodule types, such as cavity and lung wall nodules, along with limitations in handling concave boundaries. Conversely, deep learning-based approaches are hampered by imbalanced data, where nodule size is considerably smaller than surrounding tissues, leading to reduced accuracy. To surmount these constraints, we present a novel U-Net-based approach capable of addressing the spectrum of nodule segmentation challenges. Additionally, we introduce a unique loss function that rectifies imbalanced data, contributing to heightened segmentation precision.

The key contributions of our work can be distilled into the following categories:

- 1) **Two-Level Joint Markov-Gibbs Random Field (MGRF) Model:** Our methodology revolves around a pioneering two-level MGRF model, adept at distinguishing lung wall nodules and enhancing segmentation accuracy. This model addresses the nuances posed by nodule resemblance, culminating in improved detection accuracy.
- 2) **Precision in Region of Interest (ROI) Extraction:** In conjunction with the aforementioned model, we introduce an innovative strategy employing a Fully Convolutional Network (FCN) for precise ROI extraction

around detected lung nodules. This precise extraction minimizes the inclusion of extraneous tissues, augmenting the overall accuracy of segmentation.

- 3) **3D U-Net Architecture and Novel Loss Function:** Recognizing the intricacies of 3D nodule segmentation, our custom 3D U-Net architecture is fortified with a novel loss function. This fusion of specificity and sensitivity metrics promotes a balanced outcome, mitigating both over- and under-segmentation and contributing to precise segmentation results.
- 4) **Equitable Data Handling with Customized loss Function:** In addition to algorithmic innovations, our methodology addresses the challenge of imbalanced data through a tailored loss function. This function ensures the representation of nodules of varying sizes, fostering an inclusive and unbiased segmentation outcome.

Throughout the forthcoming sections, our paper delves into the technical details of our methodology, unraveling the foundational steps that underpin accurate lung nodule extraction from CT images. By synergistically harnessing the strengths of deformable models and deep learning, our approach redefines lung nodule segmentation, promising robustness, accuracy, and significant advancements in medical diagnostics.

II. METHODS

Figure 1 illustrates the fundamental steps involved in the segmentation of detected lung nodules from CT images. These steps comprise: (i) separation of lung wall nodules from the chest region using a two-level joint MGRF model (ii) employing Fully Convolutional Network, a deep learning network, to extract a region of interest (ROI) centered at the detected lung nodules' core, (iii) segmenting the detected lung nodules from the CT images' extracted ROI using 3D U-Net with a novel loss function accounting for both specificity and sensitivity performance metrics to prevent over or under-segmentation of the detected lung nodules, and (iv) finally applying connectivity analysis to extract the largest connected component that represents the segmented detected lung nodules. Below, we will discuss in-details each of the aforementioned steps.

A. AUTOMATIC EXTRACTION OF LUNG REGIONS

In order to enhance the accuracy of segmenting lung nodules, particularly those located within the lung wall, the initial step involves the separation of the lung region from the chest. To accomplish this task, we have devised an approach that performs lung region segmentation from CT images. Our approach effectively captures the spatial correlation among lung voxels and the intensity distribution of voxels within lung tissues. For this purpose, we employ a two-level joint MGRF Model. The MGRF model represents a joint probabilistic distribution of the initial images and the targeted

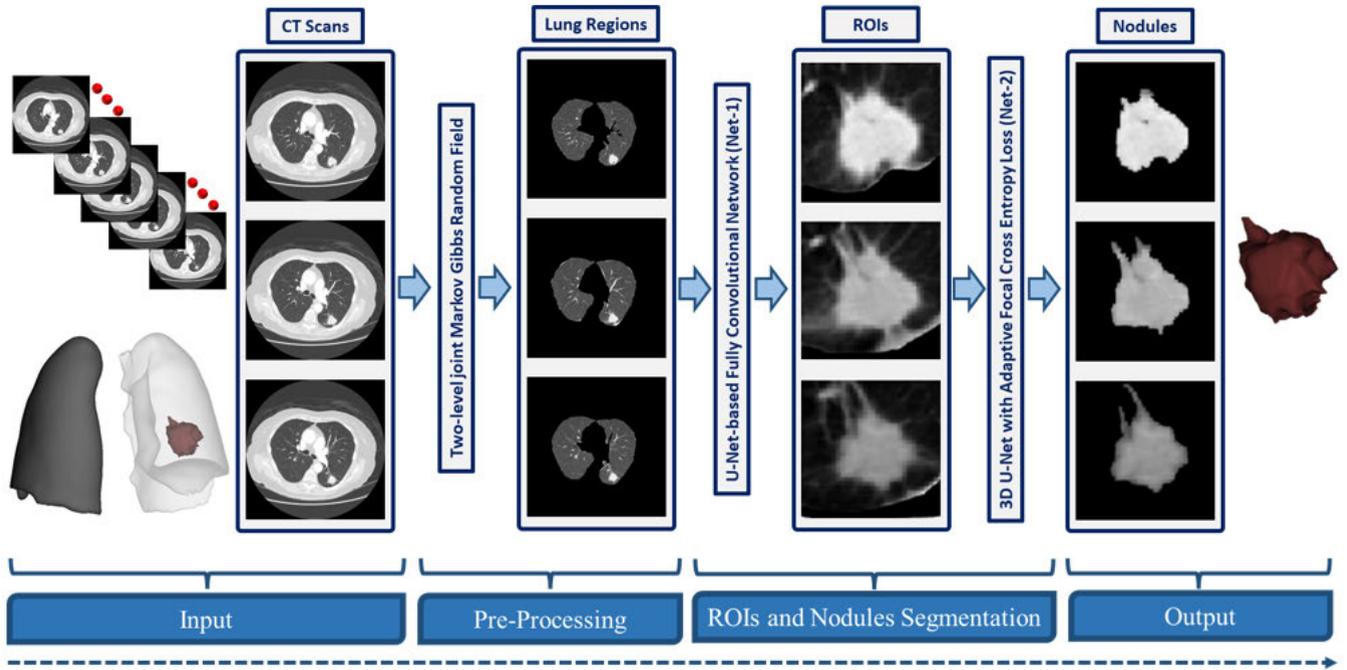


FIGURE 1. Proposed 3D-U-Net Framework for Pulmonary Nodule Segmentation: Beginning with Dataset Acquisition, followed by Lung Region Separation, ROI Extraction, Nodule Segmentation, and culminating in Connected Component Extraction.

region maps, as shown below:

$$P(\mathbf{g}, \mathbf{m}) = P(\mathbf{g}|\mathbf{m})P(\mathbf{m}) \tag{1}$$

where: a conditional distribution of images given a map, denoted by $P(\mathbf{g}|\mathbf{m})$, and an unconditional distribution of maps, denoted by $P(\mathbf{m})$. To estimate the maximum a posteriori probability (MAP) of the map, given an image \mathbf{g} , we maximize the log-likelihood function $L(\mathbf{g}, \mathbf{m})$ and obtain $\mathbf{m}^* = \arg \max_{\mathbf{m}} L(\mathbf{g}, \mathbf{m})$.

$$L(\mathbf{g}, \mathbf{m}) = \log P(\mathbf{g}|\mathbf{m}) + \log P(\mathbf{m}) \tag{2}$$

To estimate the map image (\mathbf{m}), this segmentation approach used a combination of a discrete Gaussian (LCDG) as a model for the CT data (\mathbf{g}) and an MGRF as a model for the map image (\mathbf{m}). For further mathematical details, please refer to [32]. Figure 2 demonstrates the lung segmentation in three directions - axial, sagittal, and coronal - using the two-level joint MGRF model.

B. AUTOMATIC ROI EXTRACTION

To ensure minimal overlap with other structures within segmented lung regions, we automated the selection of a region of interest (ROI) centered around the detected lung nodule’s centroid. This approach removes potential subjectivity associated with manual decisions. Our use of a fixed-size bounding box for region extraction not only eliminates the need for resource-intensive algorithms, ensuring a uniform computational load, but also enhances processing speed by enabling predictable ROI extraction, efficient parallel processing, and leveraging a U-Net-based network for rapid

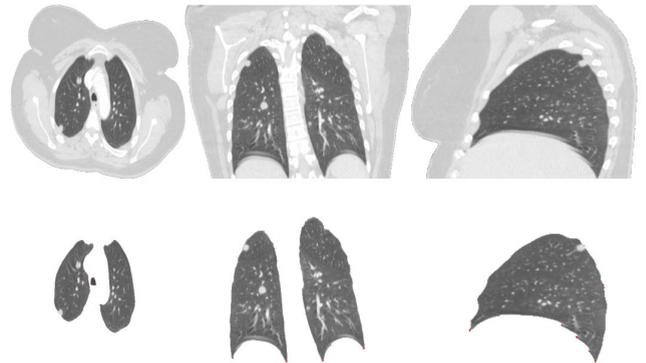


FIGURE 2. Visualization of 3D lung segmentation results projected onto 2D axial (left), coronal (center), and sagittal (right) planes. The upper row presents 2D profiles of the original CT images, while the lower row showcases our segmentation results.

nodule detection and segmentation. Our workflow employs a bounding box of dimensions $40 \times 40 \times 40$ voxels for ROI extraction, with the lung nodule centroid detected using a U-Net-based Fully Convolutional Network (Net-1), as illustrated in Figure 3.

Net-1 is trained to segment lung nodules from CT images within the extracted lung region. The ROI center is then aligned with the 3D segmented lung nodule centroid. However, this step is solely utilized to estimate the centroid for the cross-section with the maximum axial representation of the detected lung nodule. Once the centroid is estimated, a bounding box of size $40 \times 40 \times 40$ is extracted around it. It’s important to note that although training the Net-1

network requires time and effort to tune its hyper-parameters, the proposed ROI extraction method significantly outpaces alternative approaches involving the Hough Transform.

C. SEGMENTATION OF LUNG NODULES

The 3D U-Net architecture has been extensively studied in the literature, and its ability to handle 3D data, preserve spatial information, and generate accurate segmentations makes it a powerful tool for medical image analysis. This paper presents a framework that proposes the utilization of the 3D U-Net architecture for the segmentation of detected lung nodules within the extracted ROIs (see Figure 1). The 3D U-Net architecture is designed with an encoder-decoder network structure, incorporating skip connections to allow integrating both high- and low-level features. The encoder component, responsible for downsampling the input image, consists of convolutional and pooling layers. Conversely, the decoder component gradually increases the spatial resolution of the output using upsampling and convolutional layers. The skip connections establish connections between corresponding layers in the encoder and decoder, allowing for the integration of local and global contexts. To generate a probability map for the segmented lung nodules, a softmax activation function is employed in the network's final layer.

To extend the capabilities of the previous U-Net architecture, a modified version called

3D-U-Net was introduced [33]. Unlike the fully 2D architecture of the original U-Net, the 3D U-Net analyzes the entire image and produce full-resolution segmentations via incorporating a sequence of a contraction encoder and an expansion decoder. The input to the 3D U-Net is a 3D volume, employing 3D convolution, 3D maxpooling, and 3D upconvolution layers. In contrast, the 2D U-Net exclusively operates in the 2D domain.

The addition of max pooling layers in CNNs after individual convolutional layers aims to reduce the dimensionality of the image by decreasing the number of pixels in the output. This brings multiple benefits, including alleviating the computational load by reducing the resolution of the output from the convolutional layer. Simultaneously, the use of max pooling enables the network to examine larger regions of the image concurrently, while also reducing the risk of overfitting.

On the other hand, the proposed 3D U-Net is a deep neural network that produces compact volumetric segmentations by training with a small number of annotated 2D slices. The encoder pass shrinks the image at each step by doubling the number of channels and halving the spatial image size. The decoder pass in the proposed approach enlarges the spatial dimensions of the image while gradually reducing the number of feature channels, ultimately leading to the labeling layer. Through the utilization of U-Net's network architecture and data augmentation techniques, the learning models achieve impressive levels of abstraction, even with limited annotated examples. The architecture consists of an

analysis pathway and a synthesis pathway, both comprising four resolution stages. Within the analysis pathway, each layer incorporates two convolutions with dimensions of $3 \times 3 \times 3$, followed by the rectified linear activation function (ReLU). ReLU serves as a piecewise linear function that directly outputs the input if it is positive, and otherwise outputs zero. The popularity of ReLU stems from its ease of training and its ability to enhance overall performance, making it a widely used default activation function in various neural networks. Subsequently, a max-pooling operation with dimensions of $2 \times 2 \times 2$ and strides of two in each dimension is applied. The synthesis pathway begins each layer with an upconvolution operation of $2 \times 2 \times 2$ and strides of two in each dimension. This is followed by two convolutions with dimensions of $3 \times 3 \times 3$, each accompanied by a ReLU activation. Additionally, shortcut connections are established from layers in the analysis pathway with matching resolutions, contributing crucial high-resolution features to the synthesis pathway. Finally, a $1 \times 1 \times 1$ convolution is employed in the last layer to reduce the number of output channels to match the desired number of labels, which in this instance is three. The overall architecture consists of 19,069,955 parameters. To further improve the performance of our model, we fine-tune the 3D U-Net architecture using a novel adaptive loss function.

D. LOSS FUNCTION

During training, the U-Net model updates its parameters by backpropagating the gradient of the loss function with respect to the model parameters. The goal is to minimize the cross-entropy loss function, which leads to better accuracy in predicting the segmentation masks. The Cross Entropy loss function (L_{CE}) is widely used to train models. However, when training a model for segmentation using binary L_{CE} , it is crucial to ensure that the model accurately predicts each class. Given that t represents the ground truth label, with 1 denoting object (the nodule) and 0 denoting background, and p is the probability that a pixel belongs to the object class $t = 1$, we can define the Cross Entropy loss value using the following equation:

$$L_{CE} = - \underbrace{t \log(p)}_{\text{Nodules}} - \underbrace{(1-t) \log(1-p)}_{\text{Background}} \quad (3)$$

Cross-entropy loss treats all classes equally and assigns equal importance to each class, regardless of how well it is classified. Unfortunately, datasets with extreme imbalance between the segmentation classes can cause the model to underestimate the less represented class, resulting in poor performance.

To address this problem, Focal Loss [34] allows for more flexible predictions by providing the model with a margin of error. The Focal Cross Entropy Loss (L_{FCE}) enhances the ability to differentiate between simple and complex samples and addresses the class imbalance encountered by single-stage detectors. The focal loss uses the factor γ that

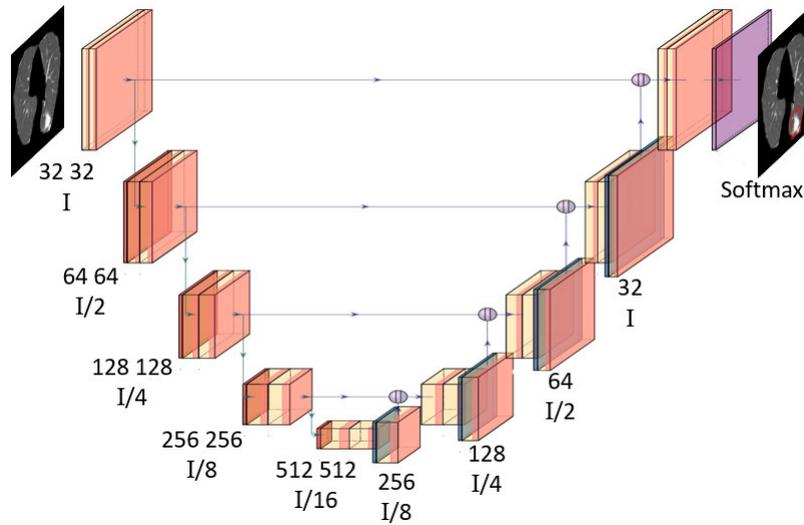


FIGURE 3. A visual representation for the architecture of the incorporated Fully Convolutional Network for the ROI localization.

modulates the impact of the standard L_{CE} . This modification allows the model to focus more on challenging examples and down-weight the influence of simpler ones, thus enhancing its ability to deal with class imbalance. The L_{FCE} function is shown in Equation 4.

$$L_{FCE} = \underbrace{-t(1-p)^\gamma \log(p)}_{\text{Nodules}} - \underbrace{(1-t)p^\gamma \log(1-p)}_{\text{Background}} \quad (4)$$

However, despite its effectiveness in capturing complex cases, the Focal Loss fails to fully address the trade-off between sensitivity and specificity inherent in cross entropy-based loss functions.

To address this issue, we propose a novel Adaptive Focal Cross Entropy Loss (L_{AFCE}), which merges the overall performance of each subject within the training patch rather than solely focusing on pixel losses. In other words, we aim to have a loss function that can consider not only pixel-level losses but also subject-level performance within the training patch.

The L_{AFCE} loss function, as defined in Equation 5, leverages adaptive factors β_t that correspond to the reciprocal within-class accuracy (sensitivity for $t = 1$ and specificity for $t = 0$) at the current training epoch. Additionally, balancing factors w_t are introduced to further refine the loss function’s behavior. These factors can be determined based on the frequency of the ground truth class t in the training data or, as in this study, tuned as hyperparameters through grid search methodology along with cross-validation.

$$L_{AFCE} = w_t \beta_t L_{FCE}, \quad (5)$$

The proposed L_{AFCE} loss function effectively balances sensitivity and specificity, which is particularly crucial in medical image segmentation tasks. By incorporating the overall subject-level performance within the training patch, L_{AFCE} takes into account class imbalances at both the pixel and subject levels. This approach allows for fine-grained control over

the contribution of each class to the overall loss, benefiting the model’s ability to accurately segment objects in medical images.

We applied the L_{AFCE} loss function to optimize the 3D-U-Net architecture (Net-2) for lung nodule segmentation on a dataset of lung CT scans. The architecture of the proposed 3D-U-Net model is depicted in Figure 4. Input CT scan regions of interest (ROIs) are pre-processed to match the network’s expected input size. Training data undergo augmentation through random rotations and flipping. The encoder pathway of Net-2 consists of pairs of $3 \times 3 \times 3$ convolutions followed by ReLU activations and $2 \times 2 \times 2$ max pooling layers with stride 2. Conversely, the synthesis pathway employs $2 \times 2 \times 2$ upsampling convolutional layers followed by ReLU activations except for the final, softmax layer.

E. NETWORK TRAINING AND HYPER-PARAMETERS ESTIMATION

Before commencing network training, network hyperparameters need to be defined, which include variables such as the number of kernels that determine the network parameters that govern network training (i.e., architecture, learning rate, optimizer, and epochs) and the loss function hyperparameters (i.e., γ , w_1 , and w_0) [35], [36]. In order to enhance the performance and generalizability of the network on new unseen data, a grid search methodology was implemented to determine the optimal values for various hyperparameters [37]. The search space for the initial number of kernels was defined as 8, 16, 32, while the learning rate of the Adam optimizer [38] was explored within the range of 0.01, 0.001, 0.0001. The learning momentum was set to a fixed value of 0.9, and the batch size was investigated across the values of 8, 16, 32. Additionally, the number of epochs

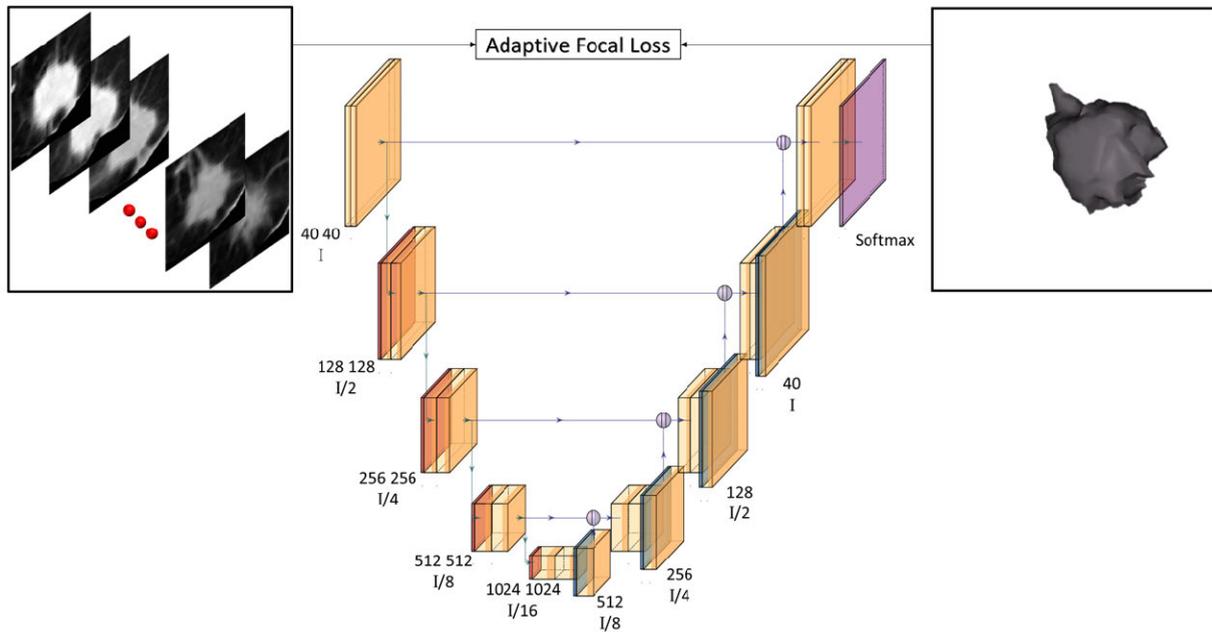


FIGURE 4. A visual representation for the architecture of the incorporated 3D U-Net with the proposed novel adaptive Focal Cross Entropy loss (L_{AFCE}).

ranged from 100 to 300, with a step size of 50, allowing for a comprehensive exploration of the parameter space. This meticulous grid search approach aimed to identify the optimal configurations for these hyperparameters, ultimately maximizing the network's performance and its ability to generalize to new data. This process entails training the model with all possible combinations of hyperparameter values, evaluating the model's performance, and selecting the combination of hyper-parameters that would yield the best segmentation performance.

F. IMPLEMENTATION CHALLENGES

The implementation of our methodology unveiled a spectrum of challenges, bridging the conceptual elegance of our approach with the intricate realities of application.

Navigating through the initial stages, we confronted the complexity of data preprocessing. This foundational step, essential for refining inputs and enhancing segmentation precision, demanded the creation of robust preprocessing pipelines to ensure a consistent and accurate data stream for subsequent phases. The integration of the two-level joint MGRF model and deep learning techniques introduced intricacies, requiring careful engineering to harmonize these components seamlessly. Parameter tuning for the MGRF model emerged as a nuanced challenge. Striking the right balance between sensitivity and specificity to distinguish lung wall nodules from the chest region demanded iterative experimentation and validation. The intricacies of lung nodule characteristics within the dataset underscored the need for a well-calibrated model that could adapt to variations effectively.

Parallely, training deep learning models, notably the custom 3D U-Net architecture, showcased its computational

appetite. Handling the volumetric 3D data, coupled with the intricacies of the architecture, necessitated judicious allocation of computational resources and meticulous fine-tuning. Ensuring that the novel loss function converged effectively further added to the training complexities. Additionally, addressing data imbalances through a tailored loss function unveiled its own set of considerations. Beyond the technical realm, ensuring the adaptability of our approach across diverse datasets and real-world scenarios required a comprehensive validation strategy.

Finally, integrating our segmentation framework into clinical workflows should face challenges transcending technical boundaries. The final clinical application should be decided by medical professionals, which highlights the need for expert validation before clinical deployment. Hence, collaborative efforts with medical professionals and technical teams will be pivotal to ensure that our solution not only aligned with medical standards but also offers a seamless user experience in actual medical settings.

III. EVALUATION METRICS

To evaluate the performance of the proposed segmentation framework, the Dice similarity coefficient (DSC), the 95th-percentile modified Hausdorff distance (HD), and the absolute volume difference (AVD) were used [39]. A more detailed explanation of these three metrics will be provided in the following sections (refer to Fig. 5).

A. DICE SIMILARITY COEFFICIENT

The level of agreement between the segmented item and the ground truth object is measured by the Dice similarity coefficient (DSC). The DSC takes into account both False Negative (FN) and False Positive (FP) segmentation errors,

making it a useful metric for evaluating the segmentation of small objects, which is the case in the proposed segmentation approach. The formula for calculating the DSC is provided in.

$$DSC = \frac{2TP}{2TP + FP + FN} \quad (6)$$

where TP, FP, and FN represent the true positive, false positive, and false negative values, respectively. A higher DSC value indicates superior segmentation performance, implying that the proposed approach accurately delineates the desired object with minimal error. This is reflected in lower values of False Positive and False Negative, affirming the efficacy of the segmentation method.

B. SENSITIVITY

The evaluation of segmentation algorithms often relies on the sensitivity metric, which is widely utilized to assess their performance. In the context of segmentation, sensitivity assesses the algorithm's ability to detect true positive pixels, indicating its accuracy in outlining the target object. The sensitivity metric is calculated as:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (7)$$

C. SPECIFICITY

In addition to sensitivity, specificity is another important metric for evaluating the performance of segmentation algorithms. Specificity measures the algorithm's accuracy in identifying true negative pixels, reflecting its capacity to exclude non-target pixels. The Specificity metric is calculated as:

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (8)$$

D. HAUSDORFF DISTANCE

The Hausdorff distance (HD) is another metric to evaluate segmentation performance. To evaluate how well the algorithm maintains the morphology of the object's edges, HD evaluates the distance between the borders of the ground truth (A_1) object and the segmented object (A_2). HD specifically calculates the greatest separation between the furthest point on one object and the nearest point on the other. A higher HD value indicates a greater distance between the boundaries of the ground truth object and the segmented object, indicating poorer segmentation accuracy. The HD metric is particularly useful in cases where segmentation errors are concentrated along the object boundaries. HD is calculated using the following formula:

$$HD(A_1, A_2) = \max(h(A_1, A_2), h(A_2, A_1)) \quad (9)$$

where A_1 is the ground truth object, A_2 is the segmented object, and $h(A_1, A_2)$ and $h(A_2, A_1)$ are the directed euclidean distance from A_1 to A_2 and from A_2 to A_1 , respectively. The 95th-percentile bidirectional Hausdorff Distance (HD) is employed as a metric in this paper to quantify the accuracy of segmentation.

E. ABSOLUTE VOLUME DIFFERENCE

The Absolute Volume Difference (AVD) calculates the absolute difference between the volume of the segmented region and the volume of the ground truth. This difference is then divided by the volume of the ground truth and multiplied by 100 to express the result as a percentage. The AVD provides an assessment of how closely the segmented volume aligns with the ground truth, enabling a quantitative evaluation of the accuracy of the segmentation algorithm. Since the segmented region's volume is more closely matched to the actual data, a lower AVD indicates better segmentation accuracy.

IV. EXPERIMENTAL RESULTS

We conducted our assessment using lung nodules sourced from the benchmark Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) dataset [40], a publicly available resource for research purposes. The LIDC-IDRI database comprises a curated compilation of CT scans, produced by a diverse array of scanner manufacturers including GE, Philips, Siemens, and Toshiba. These scans were acquired under varying x-ray tube current settings, ranging from 40 to 627 mA, and tube voltage values of 120, 130, 135, and 140 kV. The dataset also encompasses a range of slice thicknesses, spanning from 0.6 to 5.0 mm, with corresponding reconstruction intervals varying from 0.45 to 5.0 mm. These CT scans have been meticulously annotated by four radiologists from seven research organizations and eight companies. For the annotation process, each nodule underwent a comprehensive assessment process, which involved collaborative discussions among the radiologists to determine the final grading, segmentation boundaries, and clinical features. Notably, the nodules were graded using a discrete scale ranging from 0 to 5, signifying different levels of severity. A grade of 0 indicates the least severe malignancy status.

To construct our dataset, we worked closely with our medical collaborators and focused on selecting nodules with a minimum diameter of 3 mm. We placed emphasis on nodules that demonstrated a satisfactory level of agreement among the radiologists. It is important to emphasize that the final assessment of nodules relies on subjective grading rather than biopsy confirmation. To establish the ultimate grade for each nodule, we computed the average grading assigned by the four radiologists. Nodules with an average grade of less than 3.5 were classified as benign, while those with an average grade of 3.5 or higher were classified as malignant. This meticulous criterion led to the curation of a well-balanced dataset comprising 679 nodules, with 364 (53.6%) classified as benign and 315 (46.4%) classified as malignant.

For accurate segmentation, the final boundaries of the nodules were determined by identifying the common areas among the four segmentation boundary annotations provided by the radiologists. These refined segmentation boundaries represent the ground truth segmentation which are used for a comprehensive assessment of our proposed lung nodule segmentation approach. Finally, to ensure a robust evaluation,

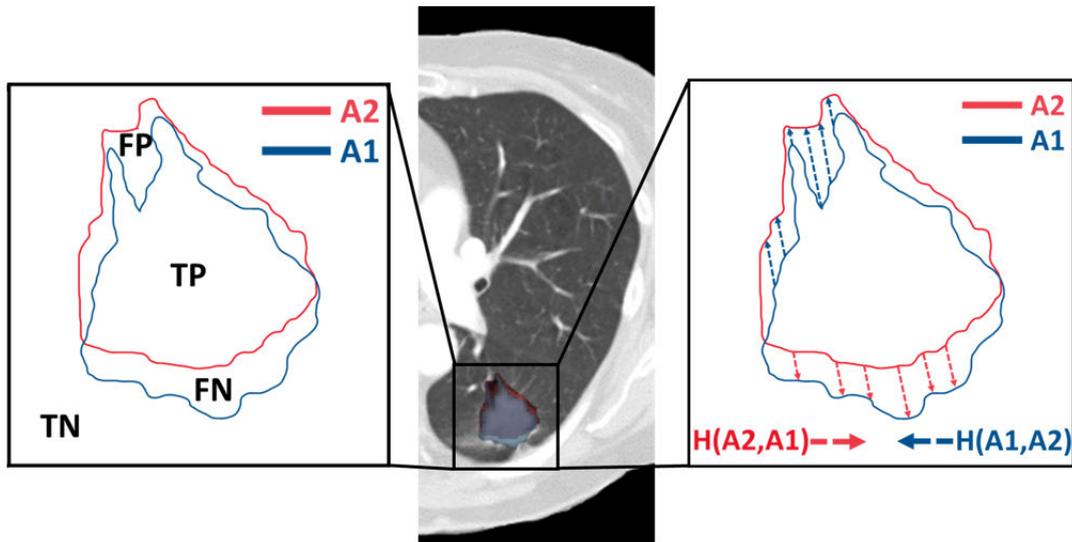


FIGURE 5. Schematic illustration of evaluation metrics calculations. On the right, the use of the 95th-percentile bidirectional Hausdorff Distance is depicted, with A_1 representing the ground truth object and A_2 the segmented object. $H(A_1, A_2)$ and $H(A_2, A_1)$ denote the directed Euclidean distances between these objects. On the left, accuracy metrics including true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) are detailed.

TABLE 1. The incorporated dataset categorized (Five categories) based on the median diameter of the whole nodule and the malignancy status.

	Median Diameter (mm)					Malignancy Status	
	<6	6 - 8	8 - 10	10 - 30	>30	Benign	Malignant
Subjects	70	96	89	275	149	364	315
Percentage (%)	10.3	14.1	13.1	40.5	22.0	53.6	46.4

we partitioned the dataset into two subsets: 452 subjects were allocated for training and validation, while the remaining 227 subjects were reserved for testing. Table 1 summarizes and groups the dataset based on the median diameter and the malignancy status.

Figures 6, 7, and 8 provides visual examples of the segmentation results obtained using the proposed approach with the L_{AFCE} loss function on three subjects with different nodules types (i.e., solid, cavity, and lung wall). The segmentation results show that our approach achieved accurate and consistent segmentation of the nodules, which are marked with red boundaries compared to the blue-marked ground truth. The segmentation results show that our approach achieved good performance in accurately identifying the boundaries of nodules with different types, even in cases where the nodules were irregularly shaped or had low contrast with the surrounding tissues.

The approach was quantitatively evaluated in terms of the aforementioned evaluation metrics, i.e., accuracy, Dice score, specificity, sensitivity, HD95, HD, and AVD.

We first performed segmentation using Net-2 along with the proposed loss function (L_{AFCE}). Then, for comparison, we experimented with two standard loss functions (i.e., L_{CE} and L_{FCE}) to obtain the optimization of results. To further validate the generalization abilities of the proposed model with

the novel loss function (L_{AFCE}), it was externally tested on a locally acquired dataset (with obtained Informed Consent) of 50 CT scans. These results are tabulated at Table 2.

Our model trained with the L_{CE} loss function achieved an accuracy of $95.34\% \pm 5.57\%$ on the LIDC/IDRI database. The Dice score was $62.57\% \pm 24.71\%$, indicating the degree of overlap between the predicted segmentation and the ground truth. The specificity, measuring the ability to correctly identify the background, reached $99.59\% \pm 0.76\%$, while the sensitivity, capturing the capability to detect lung nodules, was $53.20\% \pm 25.00\%$. The HD and 95th-HD, representing the maximum and 95th percentile of the Hausdorff distance, were $6.06 \text{ mm} \pm 4.46 \text{ mm}$ and $5.39 \text{ mm} \pm 4.25 \text{ mm}$, respectively. The AVD, indicating the average volume difference between the predicted and ground truth segmentations, was $21.36\% \pm 29.48\%$.

Using the L_{FCE} loss function, the model achieved improved performance across multiple metrics. The accuracy increased to $95.58\% \pm 3.95\%$, while the Dice score significantly improved to $81.71\% \pm 20.92\%$. The specificity remained high at $97.92\% \pm 1.37\%$, and the sensitivity improved to $77.58\% \pm 23.69\%$. The HD and 95th-HD were $6.16 \text{ mm} \pm 4.67 \text{ mm}$ and $3.12 \text{ mm} \pm 2.46 \text{ mm}$, respectively. The AVD decreased to 12.99

Our proposed L_{AFCE} loss function further enhanced the performance of the model. The accuracy reached $95.81\% \pm 2.78\%$, while the Dice score significantly improved to $93.64\% \pm 5.20\%$. The specificity was $95.99\% \pm 2.90\%$, and the sensitivity achieved remarkable accuracy of $93.30\% \pm 0.72\%$. The HD and 95th-HD were reduced to $4.41 \text{ mm} \pm 2.05 \text{ mm}$ and $1.22 \text{ mm} \pm 0.58 \text{ mm}$, respectively. The AVD was substantially reduced to $5.47\% \pm 3.03\%$. To further validate the generalizability of our model, we conducted testing

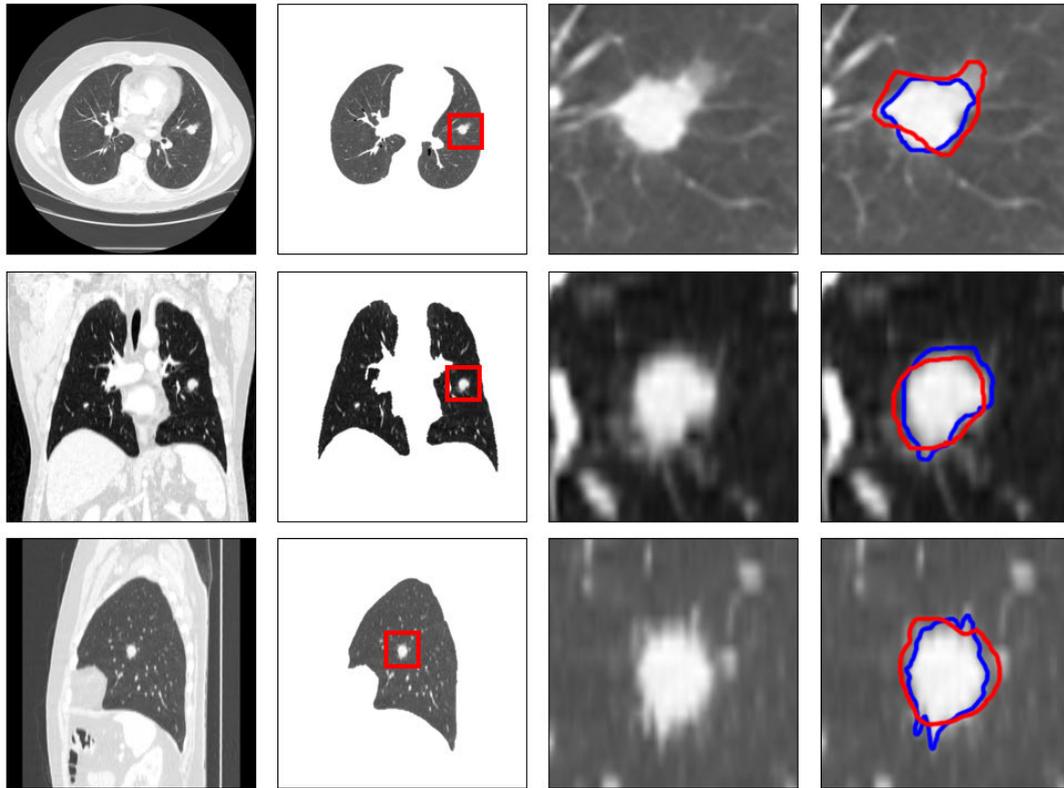


FIGURE 6. Lung nodules segmentation results for a solid nodule subject: original image cross-sections (first column), extracted lungs (second column), ROIs (third column), and final segmentation (fourth column) with color-coded edges (blue for ground truth, red for output). Visualized in 2D axial (first row), coronal (second row), and sagittal (third row) views.

on the independent dataset using the L_{AFCE} loss function. The results remained consistent, with an accuracy of $95.79\% \pm 3.11\%$ and a Dice score of $94.09\% \pm 04.16\%$. The specificity, sensitivity, HD, 95th-HD, and AVD were $95.80\% \pm 3.28\%$, $94.18\% \pm 05.59\%$, $5.52 \text{ mm} \pm 4.22 \text{ mm}$, $1.43 \text{ mm} \pm 1.14 \text{ mm}$, and $6.46\% \pm 04.74\%$, respectively.

The results demonstrate the effectiveness of our proposed approach in accurately segmenting lung nodules. The L_{AFCE} loss function significantly improved the accuracy, Dice score, and other performance metrics, indicating its superiority in accurately delineating lung nodules.

V. DISCUSSION

In this study, we present a novel approach that combines advanced deep learning techniques with connectivity analysis to achieve precise and reliable segmentation of lung nodules in CT images. Our method addresses the challenges posed by complex nodule shapes and concave boundaries, providing accurate results.

Our proposed approach leverages the power of deep learning algorithms, specifically a Fully Convolutional Network (FCN), which allows for efficient and effective extraction of relevant features. By focusing on the ROI centered on the identified lung nodules, our approach captures the intricate details necessary for accurate segmentation.

Segmentation of lung nodules is inherently challenging due to their variability in size, shape, and texture. Our study

employed a state-of-the-art 3D U-Net architecture, known for its ability to capture spatial dependencies in medical images. Furthermore, we introduced a novel adaptive loss function, tailored specifically for lung nodule segmentation, to optimize the training process and enhance the accuracy of our model. This loss function takes into account both the specificity and sensitivity performance metrics, allowing for better separation between nodular and non-nodular regions.

The experimental results obtained from the LIDC/IDRI database and the locally acquired dataset validate the effectiveness of our proposed approach. The segmentation results showcased in Figures 6, 7, and 8 demonstrate the capability of our model to accurately delineate lung nodules, as indicated by the blue contours representing the ground truth annotations and the red contours representing the predicted segmentations.

Quantitative evaluation of our approach revealed promising results. The L_{AFCE} loss function outperformed the other tested loss functions, achieving an accuracy of $95.81\% \pm 2.78\%$ on the LIDC/IDRI database and a Dice score of $93.64\% \pm 5.20\%$. These results highlight the ability of our model to accurately capture the nodular regions while maintaining a high specificity of $95.99\% \pm 2.90\%$. The sensitivity of $93.30\% \pm 07.72\%$ further demonstrates the model's proficiency in detecting lung nodules. The performance metrics of HD, 95th-HD, and AVD also provide valuable insights into the accuracy and robustness of our approach of accurately

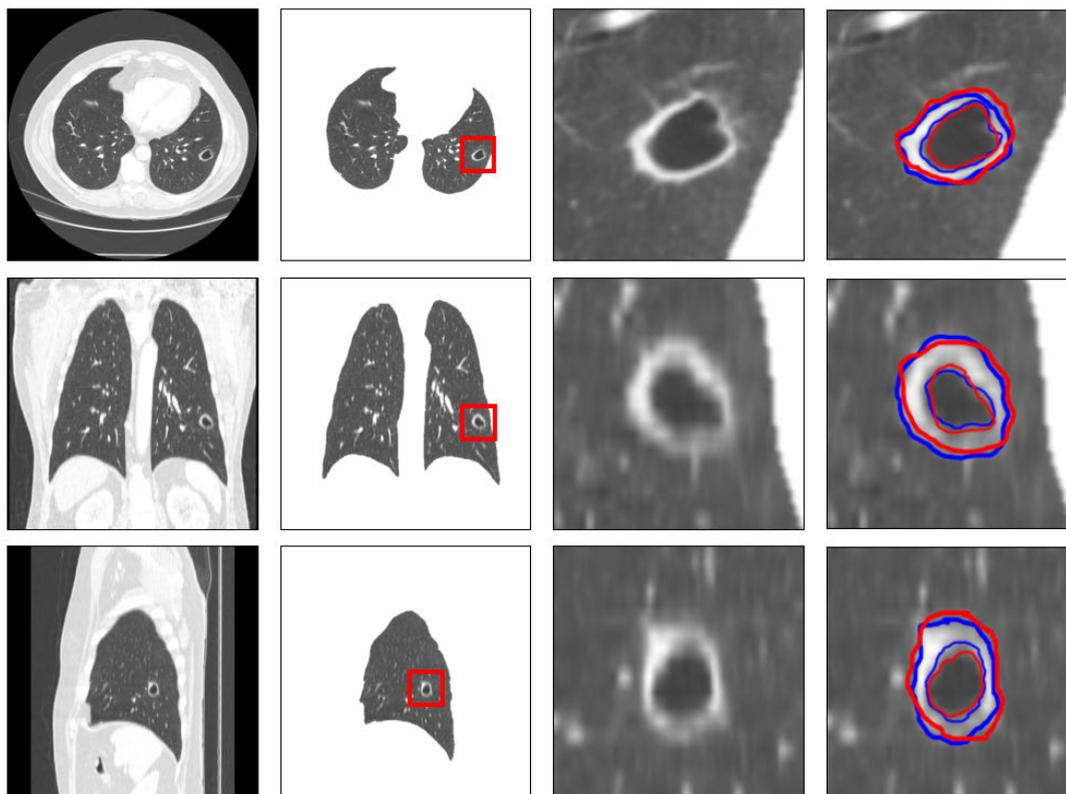


FIGURE 7. Lung nodules segmentation results for a cavity nodule subject: original image cross-sections (first column), extracted lungs (second column), ROIs (third column), and final segmentation (fourth column) with color-coded edges (blue for ground truth, red for output). Visualized in 2D axial (first row), coronal (second row), and sagittal (third row) views.

TABLE 2. Quantitative evaluation results on the LIDC/IDRI database and the locally-acquired dataset in terms of accuracy, Dice score, specificity, sensitivity, HD95, HD, and AVD. The evaluation involves two variants: L_{AFCE} and L_{AFCE} Test, both employing the proposed Adaptive Focal Cross Entropy loss function. In contrast, L_{CE} and L_{FCE} employ traditional loss functions. (Cross Entropy and Focal Cross Entropy, respectively.)

Model	Evaluation metric (Mean±SD)						
	Accuracy	Dice	Specificity	Sensitivity	HD	95 th -HD	AVD
L_{CE}	95.34 ± 5.57	62.57 ± 24.71	99.59 ± 0.76	53.20 ± 25.00	6.06 ± 4.46	5.39 ± 4.25	21.36 ± 29.48
L_{FCE}	95.58 ± 3.95	81.71 ± 20.92	97.92 ± 1.37	77.58 ± 23.69	6.16 ± 4.67	3.12 ± 2.46	12.99 ± 22.33
L_{AFCE}	95.81 ± 2.78	93.64 ± 05.20	95.99 ± 2.90	93.30 ± 07.72	4.41 ± 2.05	1.22 ± 0.58	05.47 ± 03.03
L_{AFCE} Test	95.79 ± 3.11	94.09 ± 04.16	95.80 ± 3.28	94.18 ± 05.59	5.52 ± 4.22	1.43 ± 1.14	06.46 ± 04.74

delineating the nodules from both volumetric and morphological aspects. The reduced HD and 95th-HD values of 4.41 mm ± 2.05 mm and 1.22 mm ± 0.58 mm, respectively, indicate that our model can closely approximate the boundaries of the lung nodules. The low AVD of 5.47% ± 3.03% signifies the minimal volume difference between the predicted and ground truth segmentations. Importantly, our approach exhibits strong generalizability as demonstrated by the consistent results on an independent dataset using the L_{AFCE} loss function. The accuracy, Dice score, specificity, and sensitivity metrics remained consistently high, validating the reliability of our model across different datasets.

One of the noteworthy aspects of our approach is its ability to handle various types of lung nodules encountered in clinical practice, including solid nodules, cavity nodules, and lung wall nodules. These types of nodules pose challenges due to their irregular shapes and low contrast compared to the surrounding tissues. However, our deep learning framework combined with connectivity analysis proved highly

effective in achieving consistent and accurate segmentation, even in these complex scenarios. This capability sets our approach apart from traditional deformable models, which often struggle to accurately segment diverse nodule types. Additionally, deep learning-based approaches have limitations, and imbalanced data, where the size of lung nodules is considerably smaller compared to the surrounding tissues, can pose challenges. To address this limitation, we introduced the Adaptive Focal Cross-Entropy Loss (L_{AFCE}), a novel loss function tailored to improve segmentation accuracy. By incorporating both specificity and sensitivity metrics, this loss function effectively mitigates problems related to over- or under-segmentation, resulting in improved segmentation accuracy.

One of the challenges in lung nodule segmentation lies in accurately separating nodules from vascular connections. In some cases, our approach may exhibit inclusion of vascular connections in the segmented nodule, leading to false positives. It is essential to acknowledge the potential

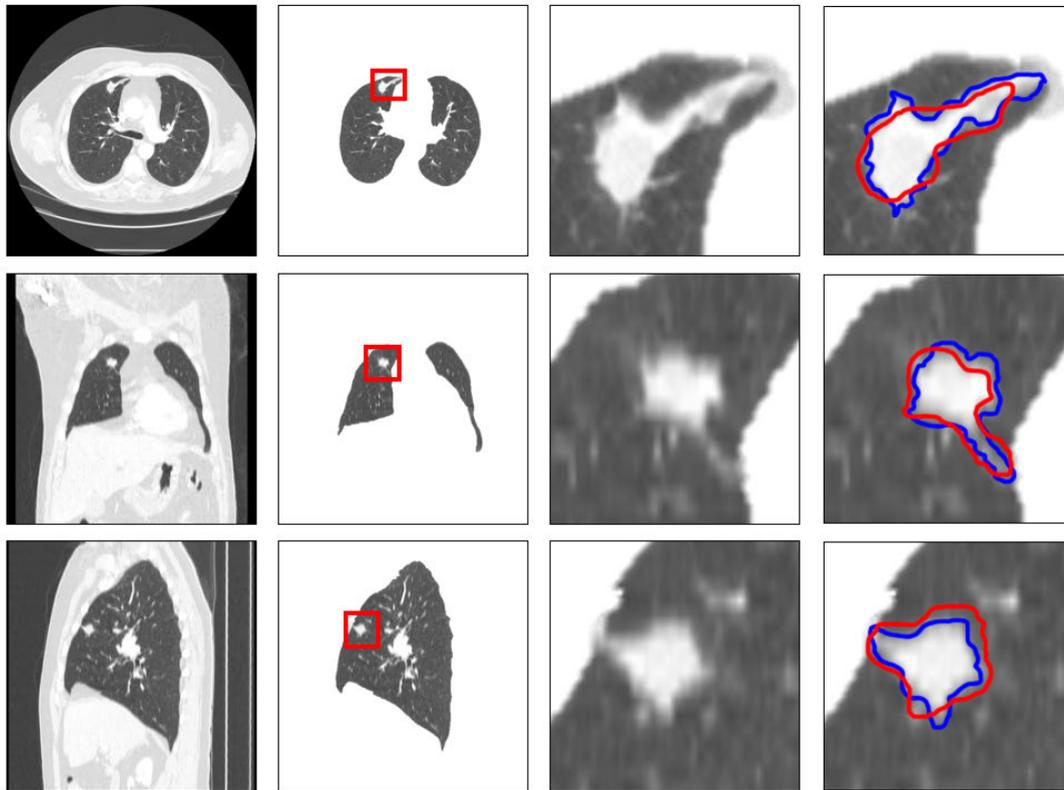


FIGURE 8. Lung nodules segmentation results for a wall nodule subject: original image cross-sections (first column), extracted lungs (second column), ROIs (third column), and final segmentation (fourth column) with color-coded edges (blue for ground truth, red for output). Visualized in 2D axial (first row), coronal (second row), and sagittal (third row) views.

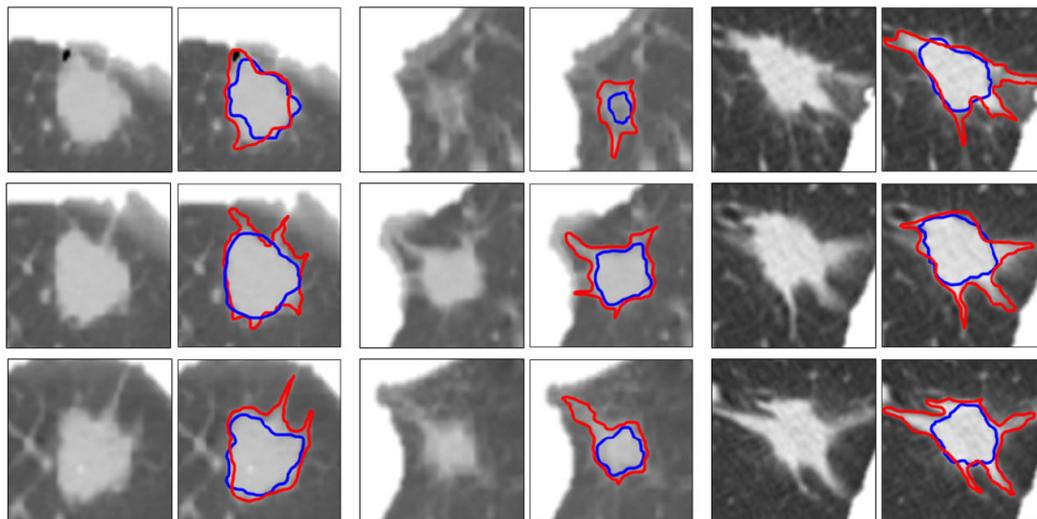


FIGURE 9. Mis-segmentation of vascular connections as lung nodules: Segmentation results for nine 2D samples from three different nodules. The first, third, and fifth columns display the ROIs of the three subjects. The second, fourth, and sixth columns show the final segmentation with color-coded edges, where blue represents the ground truth and red represents the output.

consequences of considering vascular connections as nodules, these false positives can result in unnecessary interventions or diagnostic confusion, potentially impacting patient management and treatment decisions. Therefore, it is crucial to continue exploring solutions to improve the accuracy and reliability of lung nodule segmentation, specifically addressing the issue of delineating nodules from vascular

structures. To overcome this challenge, further research is needed to develop advanced algorithms capable of accurately separating nodules from vascular structures. Algorithms that leverage not only the intensity information but also consider the spatial relationships and contextual information can be explored to improve the discrimination between these structures. For instance, incorporating vessels' analysis and

tracking, or other advanced image processing techniques can enhance the ability to distinguish nodules from vascular connections based on their distinct characteristics.

Figure 9 present an illustrative example where the vascular connection is incorrectly included as part of the segmented nodule.

By addressing this challenge, we can improve the reliability and clinical applicability of lung nodule segmentation algorithms. Advancements in this area will have a significant impact on early detection, diagnosis, and treatment planning for lung cancer patients.

Additionally, the field of transfer learning holds promise for improving the performance of deep learning models in medical imaging. The pretraining of models on large-scale datasets, such as ImageNet, followed by fine-tuning on specific medical datasets, has shown potential in improving both accuracy and efficiency. Exploring transfer learning techniques could enhance the generalizability of our model and facilitate its adoption in clinical practice.

VI. CONCLUSION

In this study, we proposed a novel approach for automatic segmentation of lung nodules using CT images. Our approach combines a two-level joint MGRF, deep learning, and connectivity analysis to accurately segment nodules from surrounding tissues, even when nodules have complex shapes or concave boundaries. We developed a new adaptive loss function, L_{AFCE} , which helps prevent over/under-segmentation of the detected lung nodules. The model exhibits high accuracy, reliability, and generalizability in segmenting lung nodules, as demonstrated by the superior results on the LIDC/IDRI database and our locally acquired dataset.

This accurate segmentation should aid clinicians to make informed decisions regarding treatment strategies, leading to improved patient management and survival rates. Our approach holds promise for enhancing patient outcomes and contributing to ongoing efforts in refining medical imaging techniques for lung cancer management.

ACKNOWLEDGMENT

(Bassant Youssef and Ahmed Alksas are co-first authors.)

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- [1] American Cancer Society. (2022). *Cancer Facts and Figures 2022*. Accessed: Aug. 11, 2023. [Online]. Available: <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2022/2022-cancer-facts-and-figures.pdf>
- [2] Cleveland Clinic. (2022). *Pulmonary Nodules*. Accessed: Aug. 11, 2023. [Online]. Available: <https://my.clevelandclinic.org/health/diseases/14799-pulmonary-nodules>
- [3] A. McWilliams, M. C. Tammemagi, J. R. Mayo, H. Roberts, G. Liu, K. Soghrati, K. Yasufuku, S. Martel, F. Laberge, and M. Gingras, *Level set Method in Medical Imaging Segmentation*, no. 10, 1st ed. Boca Raton, FL, USA: CRC Press, 2013, pp. 910–919.
- [4] A. A. Farag, A. El-Baz, G. Gimel'farb, and R. Falk, "Detection and recognition of lung abnormalities using deformable templates," in *Proc. 17th Int. Conf. Pattern Recognit.*, 2004, pp. 1–9.
- [5] A. A. Farag, A. El-Baz, G. Gimel'farb, R. Falk, M. A. El-Ghar, T. Eldiasty, and S. Elshazly, "Appearance models for robust segmentation of pulmonary nodules in 3D LDCT chest images," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI (Lecture Notes in Computer Science)*. Berlin, Germany: Springer, 2006, pp. 662–670.
- [6] Y. Itai, H. Kim, S. Ishikawa, S. Katsuragawa, T. Ishida, K. Nakamura, and A. Yamamoto, "Automatic segmentation of lung areas based on SNAKES and extraction of abnormal areas," in *Proc. 17th IEEE Int. Conf. Tools Artif. Intell. (ICTAI)*, 2005, pp. 1–9.
- [7] L. Gonçalves, J. Novo, and A. Campilho, "Hessian based approaches for 3D lung nodule segmentation," *Expert Syst. Appl.*, vol. 61, pp. 1–15, Nov. 2016.
- [8] H. Shakir, T. M. Rasool Khan, and H. Rasheed, "3-D segmentation of lung nodules using hybrid level sets," *Comput. Biol. Med.*, vol. 96, pp. 214–226, May 2018.
- [9] E. Shakibapour, A. Cunha, G. Aresta, A. M. Mendonça, and A. Campilho, "An unsupervised metaheuristic search approach for segmentation and volume measurement of pulmonary nodules in lung CT scans," *Exp. Syst. Appl.*, vol. 119, pp. 415–428, Apr. 2019.
- [10] R. Roy, P. Banerjee, and A. S. Chowdhury, "A level set based unified framework for pulmonary nodule segmentation," *IEEE Signal Process. Lett.*, vol. 27, pp. 1465–1469, 2020.
- [11] S. Rakesh and S. Mahesh, "Nodule segmentation of lung CT image for medical applications," in *Proc. 1st Int. Conf. Adv. Inf., Comput. Trends Data Eng. Global Transitions*, vol. 2, no. 1, 2021, pp. 80–83.
- [12] M. Savic, Y. Ma, G. Ramponi, W. Du, and Y. Peng, "Lung nodule segmentation with a region-based fast marching method," *Sensors*, vol. 21, no. 5, p. 1908, Mar. 2021.
- [13] X. Deng, E. Liu, S. Li, Y. Duan, and M. Xu, "Interpretable multimodal image registration network based on disentangled convolutional sparse coding," *IEEE Trans. Image Process.*, vol. 32, pp. 1078–1091, 2023.
- [14] S. Xiong, B. Li, and S. Zhu, "DCGNN: A single-stage 3D object detection network based on density clustering and graph neural network," *Complex Intell. Syst.*, vol. 9, pp. 3399–3408, 2023, doi: [10.1007/s40747-022-00926-z](https://doi.org/10.1007/s40747-022-00926-z).
- [15] Y. Zhuang, S. Chen, N. Jiang, and H. Hu, "An effective WSENET-based similarity retrieval method of large lung CT image databases," *KSII Trans. Internet Inf. Syst.*, vol. 16, no. 7, pp. 2359–2376, 2022.
- [16] S. Lu, B. Yang, Y. Xiao, S. Liu, M. Liu, L. Yin, and W. Zheng, "Iterative reconstruction of low-dose CT based on differential sparse," *Biomed. Signal Process. Control*, vol. 79, Jan. 2023, Art. no. 104204.
- [17] H. Cao, H. Liu, E. Song, C.-C. Hung, G. Ma, X. Xu, R. Jin, and J. Lu, "Dual-branch residual network for lung nodule segmentation," *Appl. Soft Comput.*, vol. 86, Jan. 2020, Art. no. 105934.
- [18] R. Roy, T. Chakraborti, and A. S. Chowdhury, "A deep learning-shape driven level set synergism for pulmonary nodule segmentation," *Pattern Recognit. Lett.*, vol. 123, pp. 31–38, May 2019.
- [19] X. Huang, W. Sun, T.-L. Tseng, C. Li, and W. Qian, "Fast and fully-automated detection and segmentation of pulmonary nodules in thoracic CT scans using deep convolutional neural networks," *Computerized Med. Imag. Graph.*, vol. 74, pp. 25–36, Jun. 2019.
- [20] G. Aresta, C. Jacobs, T. Araújo, A. Cunha, I. Ramos, B. van Ginneken, and A. Campilho, "IW-net: An automatic and minimalistic interactive lung nodule segmentation deep network," *Sci. Rep.*, vol. 9, no. 1, p. 11591, Aug. 2019.
- [21] G. Singadkar, A. Mahajan, M. Thakur, and S. Talbar, "Deep deconvolutional residual network based automatic lung nodule segmentation," *J. Digit. Imag.*, vol. 33, no. 3, pp. 678–684, Jun. 2020.
- [22] X. Dong, S. Xu, Y. Liu, A. Wang, M. I. Saripan, L. Li, X. Zhang, and L. Lu, "Multi-view secondary input collaborative deep learning for lung nodule 3D segmentation," *Cancer Imag.*, vol. 20, no. 1, p. 53, Aug. 2020.
- [23] Z. Xiao, B. Liu, L. Geng, F. Zhang, and Y. Liu, "Segmentation of lung nodules using improved 3D-U-Net neural network," *Symmetry*, vol. 12, no. 11, p. 1787, Oct. 2020.
- [24] J. Shi, Y. Ye, D. Zhu, L. Su, Y. Huang, and J. Huang, "Comparative analysis of pulmonary nodules segmentation using multiscale residual U-net and fuzzy C-means clustering," *Comput. Methods Programs Biomed.*, vol. 209, Sep. 2021, Art. no. 106332.

- [25] M. A. Khan, V. Rajinikanth, S. C. Satapathy, D. Taniar, J. R. Mohanty, U. Tariq, and R. Damaševičius, "VGG19 network assisted joint segmentation and classification of lung nodules in CT images," *Diagnostics*, vol. 11, no. 12, p. 2208, Nov. 2021.
- [26] H. Yu, J. Li, L. Zhang, Y. Cao, X. Yu, and J. Sun, "Design of lung nodules segmentation and recognition algorithm based on deep learning," *BMC Bioinf.*, vol. 22, no. S5, p. 314, Nov. 2021.
- [27] S. Kido, S. Kidera, Y. Hirano, S. Mabu, T. Kamiya, N. Tanaka, Y. Suzuki, M. Yanagawa, and N. Tomiyama, "Segmentation of lung nodules on CT images using a nested three-dimensional fully connected convolutional network," *Frontiers Artif. Intell.*, vol. 5, Feb. 2022, Art. no. 782225.
- [28] P. M. Bruntha, S. I. A. Pandian, K. M. Sagayam, S. Bandopadhyay, M. Pomplun, and H. Dang, "Lung_PAYNet: A pyramidal attention based deep learning network for lung nodule segmentation," *Sci. Rep.*, vol. 12, no. 1, p. 20330, Nov. 2022.
- [29] A. Bhattacharjee, R. Murugan, T. Goel, and S. Mirjalili, "Pulmonary nodule segmentation framework based on fine-tuned and pretrained deep neural network using CT images," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 7, no. 4, pp. 394–409, Apr. 2023.
- [30] M. Usman and Y.-G. Shin, "DEHA-net: A dual-encoder-based hard attention network with an adaptive ROI mechanism for lung nodule segmentation," *Sensors*, vol. 23, no. 4, p. 1989, Feb. 2023.
- [31] A. Halder and D. Dey, "Atrous convolution aided integrated framework for lung nodule segmentation and classification," *Biomed. Signal Process. Control*, vol. 82, Apr. 2023, Art. no. 104527.
- [32] S. Z. Li, *Markov Random Field Modeling in Computer Vision*. Cham, Switzerland: Springer, 2012.
- [33] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning dense volumetric segmentation from sparse annotation," in *Proc. Int. Conf. Med. Image Comput. Computer-Assisted Intervent.* Cham, Switzerland: Springer, 2016, pp. 424–432.
- [34] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 2, pp. 318–327, Feb. 2020.
- [35] S. F. Qadri, L. Shen, M. Ahmad, S. Qadri, S. S. Zareen, and S. Khan, "OP-convNet: A patch classification-based framework for CT vertebrae segmentation," *IEEE Access*, vol. 9, pp. 158227–158240, 2021.
- [36] M. Ahmad, Y. Ding, S. F. Qadri, and J. Yang, "Convolutional-neural-network-based feature extraction for liver segmentation from ct images," in *Proc. SPIE*, vol. 11179, 2019, pp. 829–835.
- [37] H. Alibrahim and S. A. Ludwig, "Hyperparameter optimization: Comparing genetic algorithm against grid search and Bayesian optimization," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jun. 2021, pp. 1551–1559.
- [38] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
- [39] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool," *BMC Med. Imag.*, vol. 15, no. 1, pp. 1–28, Dec. 2015.
- [40] S. G. Armato, "The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans," *Med. Phys.*, vol. 38, no. 2, pp. 915–931, Jan. 2011.



BASSANT E. YOUSSEF received the Ph.D. degree in mathematical modeling for social networks from the Department of Electrical and Computer Engineering Doctor of Engineering, Virginia Polytechnic Institute and State University (Virginia Tech), USA. She was with the Medical Institute of Research, Alexandria University, Egypt, where she is currently an Assistant Professor with the Faculty of Engineering, Department of Mathematical Engineering and Physics. She is a Visiting Research Scholar with the Bioengineering Department, UofL, Louisville, KY, USA. Her research interests include networks, deep learning, medical image applications and mathematical modeling, and stochastic processes.



AHMED ALKSAS received the B.Sc. degree in computer engineering and control systems from Mansoura University, Egypt, in 2013. He is currently pursuing the Ph.D. degree with the Computer Engineering and Computer Science Department. In 2019, he joined the University of Louisville. He has been actively working on medical imaging analysis, specifically, design of computer-aided diagnostic (CAD) systems for detection/classification of Hepatocellular carcinoma, Gliomas, and lung cancer. In addition, he is working on developing new techniques for automatic segmentation, early diagnosis, detection, and response prediction of different types of cancers, such as renal cancer, liver cancer, spine cancer, breast cancer, prostate cancer, and Willm's tumors. He is also working on the design of AI-based image processing techniques for extraction of vessels from OCTA scans and oil spills detection from satellite images. He has authored multiple technical peer-reviewed publications and also received many prestigious awards. In 2017, he was awarded the M.Sc. degree.



AHMED SHALABY received the B.S. and M.S. degrees in electrical engineering from Alexandria University, Egypt, in 2003 and 2009, respectively, and the Ph.D. degree in electrical engineering from the University of Louisville, Louisville, USA, in 2014. From 2016 to 2022, he was a Research Scientist with the Bioengineering Department, University of Louisville. He is currently a Senior Research Scientist with the University of Texas Southwestern Medical Center, Dallas, TX, USA. In BioImaging Laboratory, he had been working on three main projects: 1) Development of new machine learning approaches in order to extract novel features from functional MRI scans to build a personalized autism diagnosis system. 2) Development of new machine learning approaches in order to extract novel features from diffusion weighted-MRI scans to automatically distinguish malignant and benign prostate cancer. All experiments are conducted on DW-MR images. 3) Development of new machine learning approaches in order to extract novel features from OCT scans to automatically distinguish between normal and diseased retinas. All experiments are conducted on OCT images.



ALI H. MAHMOUD received the B.S. and M.S. degrees in electrical engineering from Alexandria University, Egypt, in 2005 and 2009 respectively, and the Ph.D. degree in electrical engineering from the University of Louisville, Louisville, USA, in 2014. He is currently a Researcher with the Bioengineering Department, University of Louisville. His research interests include computer vision and image processing, object detection and tracking, medical imaging, and facial biometrics.



ERIC VAN BOGAERT received the degree from the University of Louisville School of Medicine and the University of Louisville Diagnostic Radiology Residency Program. He completed a fellowship in abdominal imaging with the Mallinckrodt Institute of Radiology, Washington University in St. Louis. He is currently a Clinical Assistant Professor with the Department of Radiology, University of Louisville, where he practices as a Trauma and an Emergency Imaging specialist. His research interests include medical image analysis, medical informatics, and deep learning.

NORAH SALEH ALGHAMDI received the B.Sc. degree from Taif University, Taif, Saudi Arabia, and the M.Sc. and Ph.D. degrees from the Department of Computer Science, La Trobe University, Melbourne, Australia. She has been acting as the Vice-Dean of the Quality Assurance, since 2019. She is currently an Associate Professor with the Department of Computer Science, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University (PNU), Riyadh, Saudi Arabia. She is also the Director of the Businesses and Projects Management in her college. Her research interests include data mining, machine learning, text analytics, image classification, bioengineering, and deep learning. She has authored or coauthored many articles published in a well-known journals in the research field. She has participated in organizing the international conference on computing (ICC 2019). She is a member of the reviewer committee of several journals, such as IEEE, MDPI, Emerald, and Elsevier.



ALYSSA NEUBACHER is currently pursuing the Bioengineering degree with the J.B. School of Engineering, University of Louisville. She is also studying with the Dr. Hermann Friebes' Laboratory for the mathematical modeling of disease progression. Her experience includes researching with the BioImaging Laboratory for computer assisted diagnostic systems. Her research interests include finding innovative ways to diagnose and treat human diseases.



SOHAIL CONTRACTOR received the MBBS degree from the Grant Medical College, India, in 1992, where he gained a solid foundation in the field of medicine. He went on to pursue further training, completing an internship in internal medicine with Coney Island Hospital, New York, in 1996. He then, completed his residency in diagnostic radiology at the University of Medicine and Dentistry in New Jersey, in 2001, further honing his skills in the field. Driven by his passion for interventional radiology, he pursued a fellowship at the prestigious New York University Medical Center, in 2002, specializing in this advanced area of medical practice. With his exceptional educational background and extensive training, he is certified in diagnostic radiology and vascular interventional radiology. He is currently a Professor and the Chair of the Department of Radiology-University of Louisville, where he imparts his knowledge and expertise to future generations of medical professionals. He is a highly accomplished medical professional with expertise in diagnostic imaging and radiology, oncology, and cancer care. He is dedicated to providing comprehensive care to patients, with a focus on diagnostic imaging and radiology, oncology, and gastrointestinal cancer.



MOHAMMED GHAZAL (Senior Member, IEEE) received the B.Sc. degree in computer engineering from the American University of Sharjah, the M.Sc. and Ph.D. degrees in electrical and computer engineering (ECE) from Concordia University, Canada, in 2006 and 2010, respectively, and the President's Cup and the Ministry of Education's Shield for Creative Thinking from the American University of Sharjah. He is currently a Professor and the Chairperson of the Department of Electrical, Computer, and Biomedical Engineering, College of Engineering, Abu Dhabi University. He is the author of more than 200 publications in journals and international conferences, including Scientific Reports, Elsevier's *Renewable and Sustainable Energy Reviews*, IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, and IEEE TRANSACTIONS ON CONSUMER ELECTRONICS. His current research interests include AI in bioimaging, genomics, and smart systems. He is a member of ACM, BMES, and an ABET program evaluator for the Engineering Accreditation Commission. He received several awards, including Dhs1,000,000 from the United Arab Emirates Prime Minister's Office for an innovative mobile app during the 2015 Government Summit, the Distinguished Faculty Award from Abu Dhabi University, in 2014 and 2016, and the Abu Dhabi University Research Fellow, in 2018. He was awarded an Alexander Graham Bell Fellowship from NSERC and QFRNT from the Government of Quebec. He was inducted into the American University of Sharjah Hall of Fame by Sheikh Sultan Al Qasimi, in 2018.



ADEL S. ELMAGHRABY received the B.Sc. degree in computer science from Alexandria University, Egypt, in 1973, and the M.S. and Ph.D. degrees in computer science from the University of Wisconsin–Madison, in 1978 and 1982, respectively. He is currently the Speed School Director of the Industrial Research and Innovation, a Winnia Professor of CSE, and the Former Chairperson of the Computer Engineering and Computer Science Department, University of Louisville. He also held appointments with the Carnegie Mellon's Software Engineering Institute and the University of Wisconsin–Madison, and has advised more than 60 master's graduates and 24 doctoral graduates. His research and publications span intelligent systems, neural networks, cyber-security, and visualization and simulation. The IEEE-Computer Society has recognized his work with multiple awards, including a Golden Core membership.



AYMAN EL-BAZ (Fellow, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering, in 1997 and 2001, respectively, and the Ph.D. degree in electrical engineering from the University of Louisville, in 2006. He is currently a Professor with University Scholar and the Chair of the Department of Bioengineering, University of Louisville. He is also a Distinguished Professor with the University of Louisville and Alamein International University (UofL-AIU), New Alamein, Egypt. He has two decades of hands-on experience in the fields of bio-imaging modeling and non-invasive computer-assisted diagnosis systems. He has authored or coauthored more than 700 technical articles (168 journals, 44 books, 85 book chapters, 255 refereed-conference papers, 196 abstracts, and 36 U.S. patents). He was named as a fellow of Coulter, AIMBE, and NAI for his contributions to the field of artificial intelligence in medicine and biomedical translational research.

...