

Refined Deep Convolutional Neural Network for Lung Nodule Segmentation in Computed Tomography Histopathological Scans

Grace John M¹ and Dr. S Baskar²

¹Assistant Professor Vimal Jyothy Engineering College Kannur, India and PhD Scholar Karpagam Academy of Higher Education

²Assistant Professor Karpagam Academy of Higher Education

Abstract—One of the most deadly types of cancer is lung cancer. In the field of medicine, the creation of a dependable and automated method for separating lung nodules from computed tomography (CT) scans is extremely important. An Optimized Deep Convolutional Neural Network (ODCNet) framework is presented in this study with the goal of improving lung nodule segmentation accuracy. Our methodology makes use of multiple essential elements: (i) a simple thresholding algorithm to reduce noise and artifacts in skin lesion images; (ii) the use of a deep convolutional neural network (DCNN) to precisely segment each pixel within lung images; and (iii) the incorporation of the Adaptive Moment (Adam) optimizer to enhance the segmentation efficiency of the suggested algorithm. By carefully assessing the real-world dataset

Index Terms— Segmentation, Convolutional Neural network, Histopathology, Deep learning, Lung cancer.

I. INTRODUCTION

Globally, lung cancer is the deadliest cancer, with 18.4% of cancer-related deaths in 2018 coming from the disease; this is almost equal to the combined number of deaths from colon and breast cancer. Considerable progress is still required in screening and early detection, even though medications like tyrosine kinase inhibitors and immune checkpoint inhibitors have improved survival rates for some individuals. In radiation, a major treatment for lung cancer, automated detection and segmentation are essential because accurate tumor location and electron density calculations are necessary for efficient treatment planning. In order to minimize planning time and enable adaptive treatment modifications, precise, automated segmentation is crucial because mistakes in segmentation can result in large dosage calculation disparities.

Manual segmentation of lung nodules in CT scans is challenging and dependent on pathologist expertise, often resulting in inconsistent results due to the complex morphology of nodules. This complexity drives the need for automated systems to improve diagnostic accuracy and efficiency. Deep learning has revolutionized medical data analysis, with models like deep convolutional neural networks (DCNNs), radial basis function networks, and generative adversarial networks showing promise in lung nodule segmentation. These models utilize preprocessing techniques such as denoising, contrast enhancement, and spatial filtering to improve image quality. Deep learning also enhances feature extraction and classification of segmented images, though integrating these models into clinical workflows remains challenging.

Recent developments in deep learning offer improved segmentation and generalization performance, with DCNNs demonstrating superior results. This study presents a deep learning framework for segmenting lung nodules from CT scans, followed by a review of relevant deep learning-based classification techniques, a detailed description of the proposed algorithm, experimental setup, results, and conclusions.

II. RELATED WORKS

Lung cancer continues to be the most deadly type of cancer worldwide, accounting for 18.4% of all cancer-related deaths in 2018. This percentage is similar to the combined death toll from breast and colon cancers. There are still several obstacles in the way of lung cancer screening and early detection, even with advancements in treatment like tyrosine kinase inhibitors and immune checkpoint inhibitors boosting survival rates for some patient groups. In radiotherapy, which is a popular treatment for lung cancer, automated detection and segmentation play a crucial role since precise tumor localization and electron density calculations are necessary for treatment planning. Negative effects on treatment outcomes can result from significant differences in the therapeutic dose caused by segmentation errors. Manual lung nodule segmentation in ct images is time-consuming and depends on the pathologist's skill, which frequently yields unreliable results. Lung nodule shape is complicated, making manual diagnosis even more difficult. This has prompted research into automated solutions to increase the efficiency and accuracy of diagnosis. Especially when compared to manual methods, deep learning promises faster and more accurate identification. Deep convolutional neural networks (DCNN), radial basis function networks, generative adversarial networks, recurrent neural networks, multilayer perceptrons, and deep belief networks are among the deep learning models that have shown promise in lung nodule segmentation. In order to increase image quality and feature extraction, these models usually require preprocessing processes including denoising, image augmentation, spatial filtering, and color normalisation. Because segmentation separates important components including lymphocytes, stroma, and nuclei, which are necessary for detecting cancer cells, it is critical for the diagnosis of lung cancer. Under the direction of medical oncologists, efficient feature engineering improves the performance. Hence, this work proposes an enhanced deep convolutional neural network model for segmenting lung nodule from computer tomography.

III. PROPOSED ODCNET SEGMENTING ALGORITHM

A. Data Preprocessing

Certainly, the original dermoscopic photographs present high-resolution images that pose computational challenges for processing. Additionally, eliminating noise, artifacts, and air bubbles (resulting from gel/oil application, light reflection, etc.) presents a complex task for automatic cancer detection. This study employs a straightforward thresholding approach to address noise and artifact removal from the images. Within the Simple Thresholding Algorithm (STA), each pixel is assessed and classified as an artifact based on a constraint outlined in Equation (1).

$$\{(IM(x, y) > \psi_1) \text{ and } \{(IM(x, y) - IM_{avg}(x, y)) > \psi_2\} \} \quad (1)$$

where IM represents the image with pixel (x, y) . The term $IM_{avg}(x, y)$ represents the mean brightness of the adjacent picture element which is calculated by the local mean filter with sizes of 12×12 . The parameters ψ_1 and ψ_2 are predefined threshold values. In this work, we set $\psi_1 = 0.87$ and $\psi_2 = 0.096$ obtained from [32]. The actual intensity of a pixel in the image is defined as Equation (2).

$$A_p = A_p0 + \varepsilon \quad (2)$$

here A_p0 denotes actual pixel intensity and ε represents the noise in that pixel. We can achieve $A_p = A_p0$ when the mean value of artifacts and noise is zero

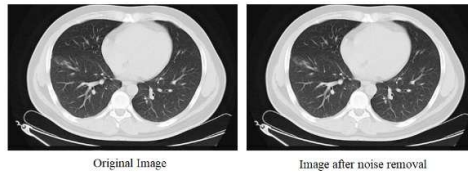


Figure 1. Removing noise and artifacts from the input image

Effective DCNN models require extensive datasets for training to yield accurate results. However, in the healthcare domain, the collection of large datasets is hindered by privacy concerns. Therefore, dataset augmentation techniques are employed to artificially increase the instances in the learning dataset by introducing slight modifications to existing images. Techniques such as oversampling or data warping are utilized to augment the number of images in the learning database or mitigate the risk of overfitting [33]. In this study, dataset augmentation involves altering image variables such as scaling, flipping, arbitrary cropping, rotation, and color-shifting. The parameters utilized in this research are summarized in Table 1.

TABLE I. PARAMETERS SELECTED FOR DATA AUGMENTATION

Parameter	Value	Action
Rotation_range	10	Rotate the input image
Horizontal_flip	True	Flips the picture horizontally
Channel_shift_range	10	Arbitrarily moves channel parameters to change the color
Shear_range	0.2	Stretch the image
Height_shift_range	0.2	Image is arbitrarily shifted in the vertical direction
Width_shift_range	0.2	The picture is arbitrarily moved horizontally
Zoom_range	0.2	Zoom out or in of the image from the midpoint
Fill_mode	closest	The value of the neighboring picture element is designated to fill the empty values

By performing data augmentation, we generate around 6000 images in each type to generate 38,600 pictures in the learning dataset.

B. Nodule segmentation using DCNN

Various architectures of deep learning models are utilized in medical data classification, each with unique variations, although their fundamental components remain consistent. Typically, deep learning classifiers consist of five fundamental layers: convolutional, activation, pooling, fully-connected (dense), and output layers. The convolutional layer aims to learn feature maps (channels) from the dataset by applying diverse kernels (filters or feature detectors), preserving their spatial relationships. Within this architecture, each pixel in a channel corresponds to a region of neighboring pixels in the preceding unit, termed the receptive field of the artificial perceptron (node or neuron). A new channel is generated by convolving the input with a trained filter and subsequently applying a pixel-wise nonlinear transfer operation to the convolved results. Every spatial position of the input employs the filter to produce a channel, aggregating comprehensive channel information from various filters utilized in the model. The feature value at position (x, y) in the channel of the l th layer is defined by Equation (1).

$$\eta_{p,q,n}^l = \omega_n^{lT} I_{p,q}^l + \beta_n^l \quad (1)$$

where ω_n^l is the weight vector and β_n^l is the bias term of the n^{th} kernel of the l^{th} layer, correspondingly. The term $I_{p,q}^l$ is the input patch focused at (p,q) of the l^{th} layer. $\eta_{p,q,n}^l$ represents the shared channel which is generated by the kernel ω_n^l . This weight disseminating technique can decrease the computational overhead and make the model simpler to train. The activation (transfer) operation adds nonlinearities to the classifier which is required for multi-level classifiers to identify nonlinear attributes. Consider $a(\cdot)$ is the nonlinear transfer operation. The activation value ($a_{p,q,n}^l$) of convoluted attribute $\eta_{p,q,n}^l$ can be computed by Equation (2).

$$a_{p,q,n}^l = a(\eta_{p,q,n}^l) \quad (2)$$

In general, hyperbolic tangent, sigmoid, and rectified linear unit (ReLU) [15, 16] are used as transfer functions. However, ReLU is one of the most widely used non-saturated transfer operations and it is defined as given in Equation (3).

$$a_{p,q,n}^l = \max(\eta_{p,q,n}^l) \quad (3)$$

The subsampling (pooling) operation is employed to decrease the size of the attribute space obtained from the convolutional function. This function targets to get shift-invariance by decreasing the resolution of the channels.

The pooling layer is generally located between two convolution layers. Every channel of a sub sampling layer is coupled to its respective channel of the previous convolution unit. The pooling operation ($pool(.)$) of a particular channel $a_{p,q,n}^l$ is defined by Equation (4).

$$\rho_{p,q,n}^l = pool(a_{r,s,n}^l), \text{ for all } (r,s) \in R_{pq} \quad (4)$$

where R_{pq} is a local region about the point (p,q) . Average or Max pooling are typical pooling operations used in the deep learning classifiers. We can excerpt high-level attributes gradually by performing many convolutional and subsampling operations. The filters in the first convolution block are intended to identify lower-level attributes like curves and edges, while the filters in the higher convolution blocks are trained to extract more abstract attributes. Following the integration of convolutional and subsampling blocks, deep learning networks typically incorporate one or more dense layers, aimed at capturing higher-level attributes. These layers establish connections between all nodes in the preceding unit and each node in the current layer to generate comprehensive semantic data. However, this operation may be replaced by a 1×1 convolutional operation in certain cases, rendering it unnecessary. The final component of deep learning classifiers is the output layer, where classification is often facilitated using the softmax operator. Optimization of variables pertaining to a specific process is achieved by minimizing an appropriate loss function corresponding to that process. Suppose we have k number of data, the anticipated input-output correlations is defined by $\{(I(k), J(k); k \in [1, 2, 3 \dots k])\}$, in which $I(k)$ is the k^{th} input data, $J(k)$ is its equivalent target class and $O(k)$ is the output of the classifier. The loss function of the classification model can be calculated using Equation (5).

$$L = \frac{1}{K} \sum_{k=1}^K l(\emptyset; I(k), O(k)) \quad (5)$$

Training deep learning classifiers involves optimizing system variables by minimizing the loss function, which serves as a global optimization problem. In this study, we propose a DCNN-based classifier for melanoma identification in dermoscopic images. Our framework, termed ODCNet, comprises a 7-layer DCNN architecture, as illustrated in Figure 4. This architecture consists of five convolutional and two fully connected blocks. MaxPooling is applied after each convolutional block to compute the maximum value for patches of a feature map, while batch normalization layers normalize the results from the convolutional blocks. The classifier utilizes the rectified linear unit (ReLU) as the activation function. Initially, the first layer employs 8 convolutional filters with a dimension of 3×3 , followed by batch processing to standardize the input and enhance training speed. Subsequently, ReLU is applied, followed by 2×2 MaxPooling, iterated four times. The number of filters increases to 16, 32, 64, and 128 in successive blocks. The output from the last convolutional layer is then compressed to create a fully connected layer with 256 neurons, succeeded by a fully connected block with 2 neurons. Finally, the SoftMax layer categorizes input images into defined tags. Data augmentation techniques are employed to address overfitting issues in the model, while the cross-entropy loss is calculated using Equation (6).

$$\chi = - \sum_{cl}^{\mathcal{C}} \mathcal{T}_i \log(\mathcal{P}_i) \quad (6)$$

In Equation (7), $\mathcal{T}_i$ and $\mathcal{P}_i$ are the ground truth and expected tags for every class (cl) in $\mathcal{C}$. This work considers two categories, (i.e., $\mathcal{C} = 2 \rightarrow$ melanoma and benign). Hence, the cross-entropy loss is calculated using Equation (7).

$$\chi = -t_1 \log(p_1) - (1 - t_1) \log(1 - p_1) \quad (7)$$

where t_1 is ground truth, p_1 is predicted results.

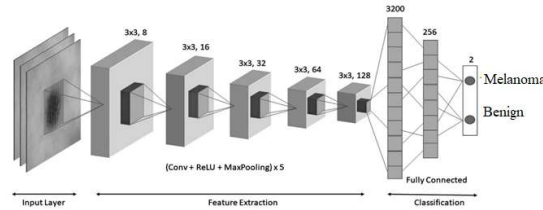


Figure 2 Architecture of Proposed DCNN

C. Optimization

Reducing classification errors is crucial in segmentation problems to enhance the efficiency of the classifier. In our approach, we utilize the Adam optimizer to fine-tune the hyper parameters of the model. This optimization technique calculates the adaptive learning rate for all variables involved in the learning process. Known for its simplicity and efficiency, Adam incorporates first-order gradients with minimal storage requirements, facilitating stochastic optimization. Particularly suited for machine learning tasks involving high-dimensional feature spaces and large datasets, Adam computes learning rates independently for different variables based on approximations that incorporate first- and second-order moments. The following equations are employed to optimize the model.

$$a_t = \gamma_1 a_{t-1} - (1 - \gamma_1) \rho_t$$

$$b_t = \gamma_2 b_{t-1} - (1 - \gamma_2) \rho_t^2$$

$$\Delta \omega_t = -\eta \frac{a_t}{\sqrt{b_t + \zeta}} \rho_t$$

$$\omega_{t+1} = \omega_t + \Delta \omega_t$$

For experimentation, we implement the Adam optimizer with $\rho_t = 0.0001$, decay = 0.0, $\gamma_1 = 0.9$, $\gamma_2 = 0.999$, amsgrad = false, and $\zeta = \text{zero}$.

IV. EXPERIMENTS AND RESULTS

To assess performance, we conducted empirical analysis on a system comprising a 3.6 GHz Intel Core i7-4790 CPU with 16GB memory, running the Windows 10 operating system. The effectiveness of our classification method is evaluated by comparing experimental results with six related models, namely U-Net [12], MultiResUNet [13], Deeply Supervised U-Net [19], and Deeply Supervised MultiResUNet [14].

A. Datasets Acquisition

Assessing the efficacy of any deep learning (DL) algorithm necessitates access to a vast database capable of yielding optimal solutions. In our study, we utilize a dataset comprising labeled lung CT scans sourced from the TCIA database, derived from the National Cancer Institute Cancer Analysis Consortium Lung Cohort [20]. These scans are accompanied by patient details encompassing genomic, proteomic, and clinical data. Stored in DICOM (Digital Imaging and Communications in Medicine) format, the scans feature definitive labels including imaging parameters, gender, birth date, study dates, among others. In total, our study incorporates 5043 scans obtained from various patients, comprising 48 series dedicated to lung tumor analysis

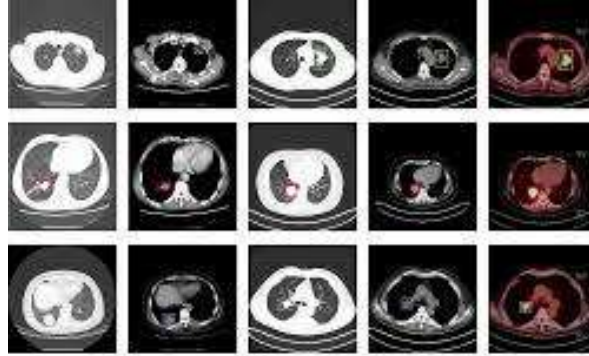


Figure 3 Sample Lung Cancer Images

B. Performance measure

We assess the accuracy of the proposed method quantitatively using performance metrics including the Dice similarity coefficient (DSC) and intersection over union (IoU). These metrics gauge the degree of overlap between the segmentation results and a manually annotated reference standard. The DSC is computed using the equation provided below.

$$\text{DSC} = \frac{2 \times |R \cap S|}{|R| + |S|}$$

where R is the manually annotated reference standard and S is the result of segmentation.

The Jaccard similarity score (JSS) serves as a performance metric for evaluating the effectiveness of data classification techniques. Specifically, the Intersection over Union (IoU) metric quantifies the similarity between predicted data and ground truth within a dataset. It is defined as follows:

$$JSS = \frac{|R \cap S|}{|R \cup S|}$$

IV. CONCLUSION

Lung cancer is the most detrimental type of malignancy. Developing an automated and credible system to segment lung nodule from a computed tomography (CT) scan is a very expedient tool in the healthcare industry. This study proposes an Optimized Deep Convolutional Neural Network (ODCNet) framework for segmenting lung nodule with enhanced accuracy. Our ODCNet exploits (i) a simple thresholding algorithm to evade noise and artifacts from the skin lesion images and a dataset augmentation method to upsurge the number of instances in the dataset artificially; (ii) the deep convolutional neural network (DCNN) to segment each pixel of the lung image; and (iii) Adaptive Momentum (Adam) optimizer to enhance the segmentation efficiency of the proposed segmentation algorithm. The effectiveness of the ODCNet model is carefully studied on the real-world dataset called The Cancer imaging Archive (TCIA) dataset and its effectiveness is compared with some modern segmentation algorithms with respect to evaluation metrics. The experimental results reveal that ODCNet significantly outdoes other prevailing segmentation algorithms with better volume error, dice similarity coefficient (DSC), Jaccard similarity score (JSS), spatial overlap, under and over segmentation rates.

REFERENCES

- [1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, Nov. 2018, doi: 10.3322/caac.21492.
- [2] N. Saito et al., "Computed tomography image-based bone marrow-sparing intensity-modulated radiotherapy for pelvic cancer," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 78, no. 4, pp. 957–964, Nov. 2010, doi: 10.1016/j.ijrobp.2009.11.012.
- [3] X. Zhen et al., "Deep convolutional neural network with transfer learning for rectum toxicity prediction in cervical cancer radiotherapy: a feasibility study," *Phys. Med. Biol.*, vol. 62, no. 21, p. 8246, Nov. 2017, doi: 10.1088/1361-6560/aa8e4c.
- [4] G. Litjens et al., "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.
- [5] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979, doi: 10.1109/TSMC.1979.4310076.
- [6] C. Shorten and T. M. K. Handerson, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.
- [7] T. Alshawi, M. Geoghegan, A. Heron, and N. Deighton, "The impact of color normalization in digital image analysis of routine HER2-stained breast cancer sections," *J. Pathol. Inform.*, vol. 9, no. 1, pp. 13–23, 2018, doi: 10.4103/jpi.jpi_45_18.
- [8] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. Upper Saddle River, NJ: Pearson, 2018.
- [9] M. M. Rahman, M. K. Hasan, and R. Islam, "A new approach for enhancing the performance of deep learning based lung nodule detection," *IEEE Access*, vol. 7, pp. 20183–20196, Jan. 2019, doi: 10.1109/ACCESS.2019.2896261.
- [10] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. MICCAI*, 2015, pp. 234–241.
- [11] B. Çiçek et al., "3D U-Net: Learning dense volumetric segmentation from sparse annotation," in *Proc. MICCAI*, 2016, pp. 424–432.
- [12] Z. Zhou et al., "UNet++: A nested U-Net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, 2018, pp. 3–11.
- [13] A. Shaffie, A. Soliman, X. A. Fu, et al., "A novel technology to integrate imaging and clinical markers for non-invasive diagnosis of lung cancer," **Sci. Rep.**, vol. 11, no. 1, p. 4597, 2021.
- [14] S. P. Primakov, A. Ibrahim, J. E. van Timmeren, G. Wu, S. A. Keek, M. Beuque, et al., "Automated detection and segmentation of non-small cell lung cancer computed tomography images," **Nat. Commun.**, vol. 13, no. 1, p. 3423, 2022. doi: 10.1038/s41467-022-30841-3.
- [15] G. C. Pereira, M. Traugher, R. F. Muzic Jr., "The Role of Imaging in Radiation Therapy Planning: Past, Present and Future," **NIH**, 2014, p. 231090. doi: 10.1155/2014/231090.