OPEN ACCESS
International Journal of Innovations in Science & Technology

Early Detection and Classification of Lung Cancer using Segment Anything Model 2 and Dense Net

Hafza Eman[1], Syed M. Adnan[1], Wakeel Ahmad[1], Ishtiaque Mahmood[2]
[1]University of Engineering and Technology, Taxila, Pakistan.
[2]Oman College of Management and Technology, Oman.
*** Correspondence:** 23-ms-csc-1@students.uettaxila.edu.pk

Citation | Eman. H, Adnan. S. M, Ahmad. W, Mahmood. I, "Early Detection and Classification of Lung Cancer using Segment Anything Model 2 and Dense Net", IJIST, Vol. 7 Issue. 1 pp 476-492, March 2025
DOI | https://doi.org/10.33411/ijist/202571476492
Received | Feb 12, 2025 **Revised |** Feb 28, 2025 **Accepted |** Mar 3, 2025 **Published |** Mar 5, 2025.

Lung cancer is one of the most perilous diseases worldwide with high incidence and low survival rates due to late diagnosis. Accurate detection and diagnosis of lung nodules is important for early-stage detection. Machine learning and deep learning techniques have greatly improved the precision of lung nodule segmentation and classification in Computed Tomography (CT) images. The study presents a novel approach to segmenting and classifying nodules by combining foundational models with deep learning architectures. We have used the Segment Anything Model (SAM2) to segment lung nodules and Dense Net to classify them as benign and malignant. SAM2 has been tested on the datasets using different prompts to achieve better results. Foundational Models and Deep Learning architecture's integration significantly improved Computer-Aided Detection (CADe) and Computer-Aided Diagnosis (CADx) in medical images. Experimental results proved the effectiveness of the proposed model for early-stage detection and classification of lung nodules from CT scans. SAM2 model achieves a Dice Similarity Coefficient (DSC) of 97.87% and an Intersection over Union (IoU) of 95.82% for segmentation, and the Dense Net model's classification accuracy is 97.34%. The experimental results demonstrate the performance of our model compared to existing techniques.

Keywords: SAM2, Transfer Learning, Vision Transformer Model, Bounding Box Prompts, Computed Tomography (CT) scans

Introduction:

Lung cancer is among the most life-threatening diseases globally, characterized by a high incidence and low survival rates due to late detection. It accounts for 18% of all cancer cases and is responsible for approximately 1.8 million deaths each year. According to the Global Cancer Observatory (GLOBOCAN), lung cancer holds the highest rank in both incidence and mortality rates worldwide [1]. The situation in Asia is even more severe, with a lung cancer incidence rate of 63.1%, a mortality rate of 62.9%, and a 5-year prevalence rate of 63.9% [1].

Medical imaging techniques such as Computed Tomography (CT) and X-rays have helped in lung nodule detection and classification [2]. Due to the extensive use of CT scans and the complexity of medical images, Computer-Aided Detection (CADe) and Computer-Aided Diagnosis (CADx) play a crucial role in supporting radiologists by enhancing the accuracy of lung nodule detection and diagnosis [3]. Traditional image processing techniques such as thresholding, region-based methods, and edge detection face limitations when handling the complexity of medical images [4].

Deep Learning plays an important role in lung cancer detection and classification for better treatment and follow-up procedures [5]. Deep Learning models learn features from medical images and precisely segment nodules from medical images such as CT scans and X-rays. The advancements in Deep Learning Models have significantly enhanced the performance of medical images [6]. Convolution Neural Networks (CNNs) can analyze complex medical image architectures but have limitations when dealing with 3D data. To overcome this issue, 3D CNN architectures [7] have been developed to process the length, height, width, and depth of 3D scans, enabling the extraction of meaningful features. However, these models demand extensive training to comprehend the intricate structures of medical images [8]. For instance, the development of DHEA-Net [9] consists of dual encoder-based architecture that uses CT scans and coronal views to improve segmentation accuracy. The refined UNET architecture was developed by [10] with similar architectural improvements [11]. Another systematic review and meta-analysis evaluated the diagnostic accuracy of deep learning models with lesion-wise sensitivity [12].

Foundational Models [13] have been trained on extensive data, models can generalize well on unseen tasks. Meta-AI segmentation models are well known for their zero-shot capabilities. Segmentation Anything Model (SAM) has been trained on multiple modalities and can segment unseen tasks in a zero-shot manner accurately [14]. A variant of SAM, called MedSAM [15] has demonstrated improved accuracy across various medical datasets. SAM2 offers superior segmentation capabilities compared to SAM and is also capable of segmenting videos [16]. SAM2 features a memory bank that tracks objects across frames, enabling continuous segmentation even when objects temporarily disappear and reappear on the screen. Lung nodule classification is also a challenging task for the diagnosis of lung cancer.

Computer Vision and Machine Learning models have great contributions to the diagnostics of many diseases. In a study, LCP- CNN [17] was trained on CT scans to generate malignancy scores and to classify nodules as benign and malignant with less false positive rate. This study [18] adopted a self-supervised learning technique by utilizing an adaptive slice selection model for pre-processing and then utilizing the self-supervised approach to learn features for classification. Progressive Growing Channel Attentive Non-Local (ProCAN) [19] network was developed for nodule classification. A channel-wise attention mechanism is applied, and the model is trained to improve learning capability. AlexNet architecture was introduced by [20] with modifications in layer ordering and hyperparameter adjustments to improve the performance of the model. Pre-processing steps such as zero centering, normalization, and segmentation were also implemented. Ren et al. [21] developed an ensemble framework by using LeNEt, GoogleNet, AlexNet, ResNet, VGG16, and DenseNet

for lung cancer classification. Dosovitskiy developed CNN and an attention-based model [22]. Deep learning approaches have shown significant results in medical image processing systems. However, accuracy and high false positive rates have been a challenge in handling medical images. To overcome the issue, we presented a fusion of SAM2 with the deep learning architecture DenseNet.

Objectives:

The objectives of this research are as follows:

1. Develop an automated framework for nodule detection and classification.
2. Improve segmentation and classification accuracy using Segment Anything Model 2 and DenseNet.
3. Validate the model on diverse datasets and compare the proposed model with existing techniques.
4. Investigate the impact of foundational models on medical images.

Material and Methods:

Dataset:

In this study, the LUNA16 [23] dataset was utilized, it consists of 888 volumetric CT scans which is a subset of the Lung Image Database Consortium (LIDC) [24] containing a total of 1018 scans. LIDC and LUNA16 are both used in lung cancer detection and diagnosis systems. The LUNA16 dataset stores scans in MetaImage (.mhd) format, divided into 10 subsets, with an annotation file that provides the real-world coordinates and diameters of nodules. Figure 1 shows sample images of the LUNA16 dataset. The LIDC dataset consists of DICOM files, which are preprocessed to segment a complete lung CT scan into individual lung lobes. Four radiologists annotated the nodules and calculated radiomic features for calculating malignancy scores which are helpful in the classification of nodules as benign and malignant. Figure 2 shows sample images from the LIDC dataset.

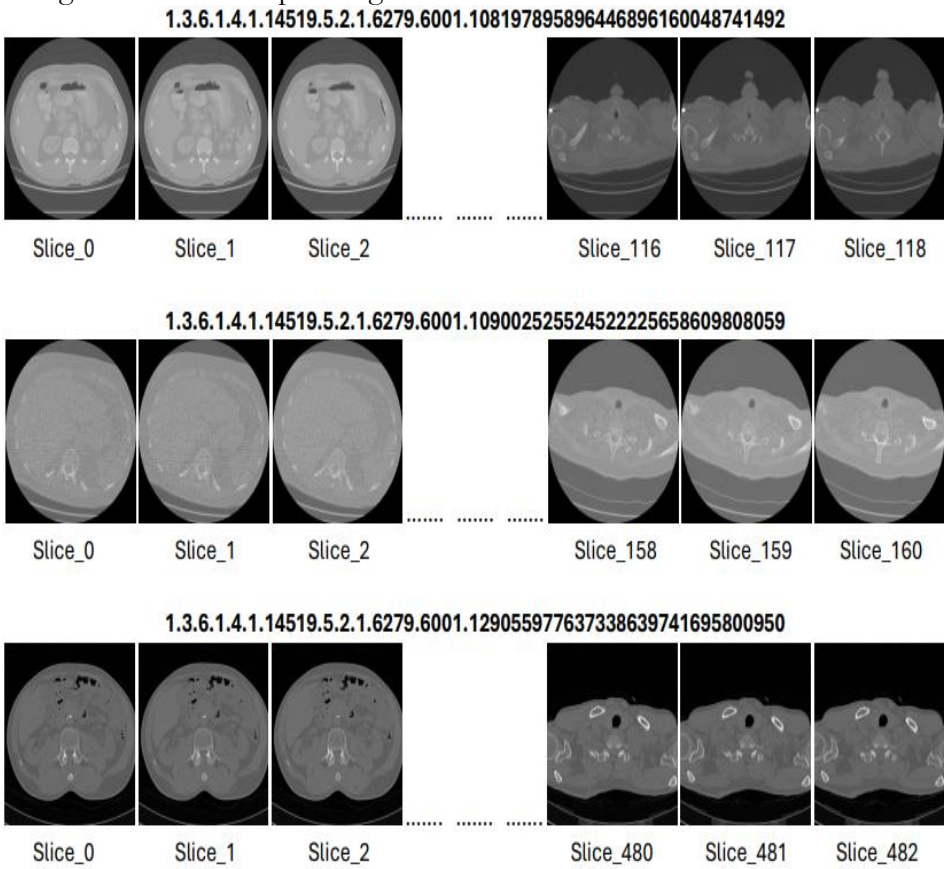

Figure 1: Sample images of LUNA16 dataset

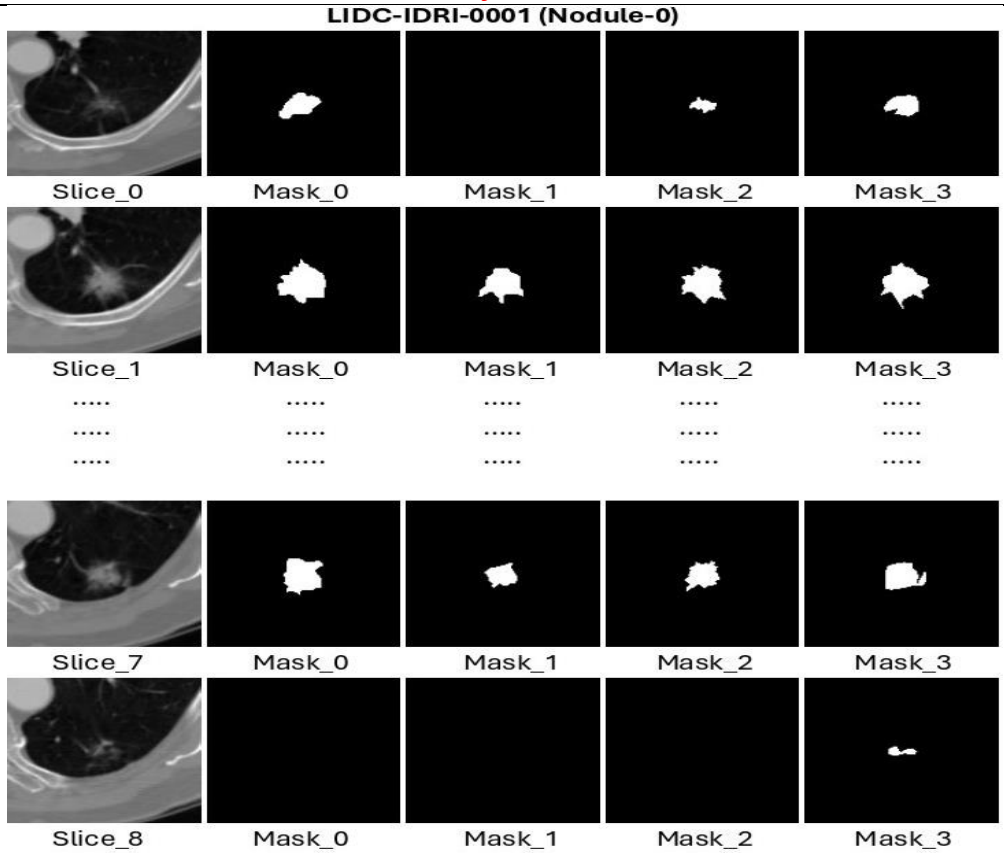

Figure 2 Sample images of the LIDC-IDRI dataset

Methodology:

We utilized SAM2 for the segmentation of lung nodules and DenseNet for lung nodule classification as benign and malignant. Figure 3 illustrates the proposed research methodology while Figure 4 presents the step-by-step flow diagram for this study.

Preprocessing:

CT scans have varying resolutions because of different scanners and protocols used. These varying resolutions affect the segmentation accuracy of the model. The two types of variabilities that arise in CT scans are Spatial Resolution Variability which is a difference in pixel spacing or slice thickness and Intensity Variability which is a difference in voxel intensity distribution due to the different protocols of each scanner. To overcome these issues, we have applied the following preprocessing techniques.

Spatial Normalization or Resampling:

Resampling converts all images to common voxel spacing so that all anatomical structures appear the same size across all images. The resampling factor can be calculated using Equations 1 and 2.

$$resampling\_factor\_x = \frac{o_x}{d_x} \qquad (1)$$

$$resampling\_factor\_y = \frac{o_y}{d_y} \qquad (2)$$

Where o_x and o_y are the original spacing and d_x and d_y are the desired spacing.
The new image size can be calculated using equations 3 and 4.

$$H' = H \times resampling\_factor\_y \qquad (3)$$

$$W' = W \times resampling\_factor\_x \qquad (4)$$

The new image size will be (H', W').

$$y = 255\frac{x - min_x}{max_x - min_x} \qquad (5)$$

Intensity Normalization:

CT scans have intensity variations due to the different scanners used which can downgrade the model's performance. To normalize intensity distributions for better contrast of a CT scan, equation 5 is used. Figure 5 shows the sample CT scan before and after applying normalization techniques.

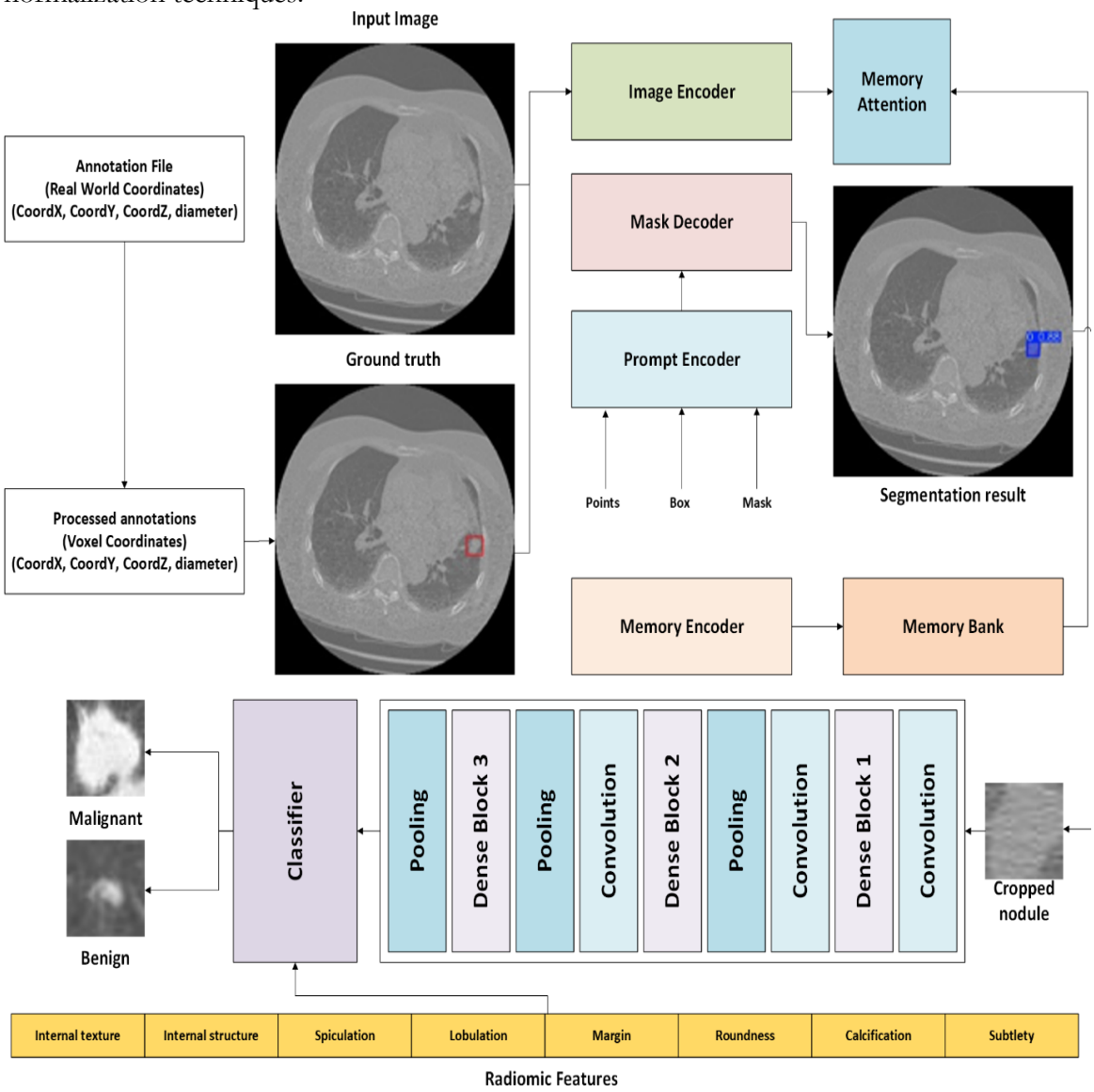

Figure 3 Proposed methodology

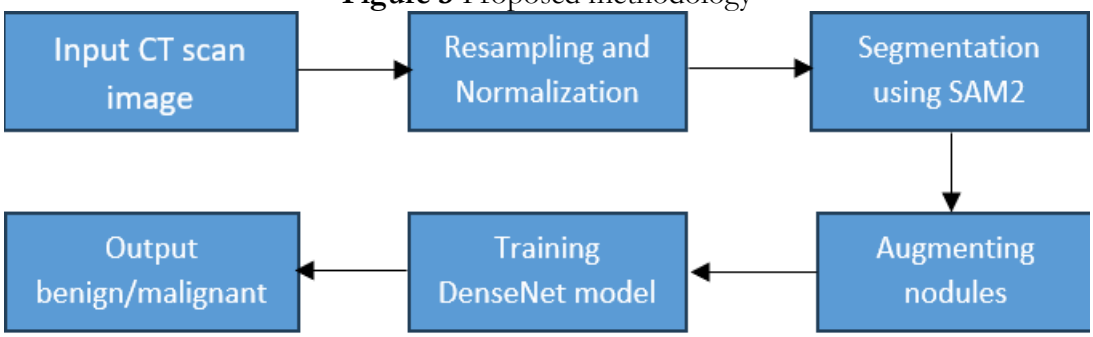

Figure 4. Flow Diagram for Lung Nodule Segmentation and Classification

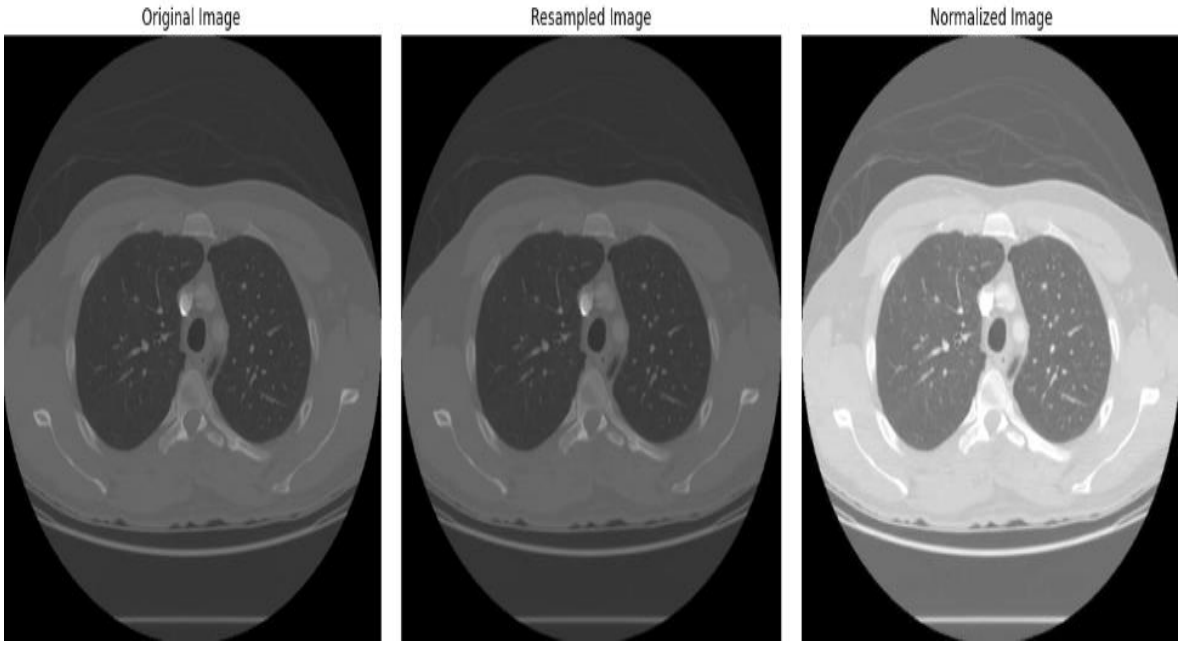

Figure 5 Sample image before and after preprocessing

SAM2 for Segmentation of Lung Nodules:

Segment Anything Model is a Vision Transformer (ViT) [22] based model trained on extensive datasets and has great generalization ability on unseen tasks. In this research, we have used SAM2 [16] which has better segmentation results than SAM. Before applying SAM2 to our datasets, the LUNA16 dataset must first undergo preprocessing. The LUNA16 dataset's annotation contains real-world coordinates that need to be converted to voxel coordinates. We utilized equation 6 to convert real-world coordinates to voxel coordinates. Figure 6 presents the code used to convert real-world coordinates into voxel coordinates.

$$VoxelCoord[i] = \frac{WorldCoord[i] - Origin[i]}{Spacing[i]}, for\ i = 0,1,2 \quad (6)$$

Where:

- VoxelCoord[i] is the voxel coordinates along with dimension.
- WorldCoord[i] is the world coordinates along the dimension, already available in the annotation file.
- Origin[i] is the origin of the CT scan available in image metadata.
- Spacing[i] is the spacing between voxels in a CT scan, also available in image metadata.

```
# world coordinates to voxel coordinates
voxel_coords = world_to_voxel([coordX, coordY, coordZ], origin, spacing)
voxel_coords = [int(round(c)) for c in voxel_coords]
print(f"World coordinates: ({coordX}, {coordY}, {coordZ})")
print(f"Voxel coordinates: {voxel_coords}")

# range affected by the nodule
z_nodule = voxel_coords[2]
z_radius = int(round(diameter_mm / 2 / spacing[2]))
z_start = max(0, z_nodule - z_radius)
z_end = min(array.shape[0], z_nodule + z_radius + 1)
print(f"Nodule affects slices: {z_start} to {z_end - 1}")

radius = int(round(diameter_mm / 2 / spacing[0]))
print(f"Nodule radius in pixels: {radius}")
```

Figure 6 Real-world coordinates to voxel coordinates

From a complete volumetric CT scan, we extracted slices containing nodules and fed them into SAM2 along with annotations for segmentation purposes. SAM2 is capable of segmenting objects in a zero-shot manner. Zero-shot segmentation capability of the SAM2 model plays an important role in lung nodule segmentation. Unlike traditional models, which require a large amount of manually annotated data and complex training, SAM2 eliminates the need for extensive annotation and training. Lung nodule variabilities such as size, shape, and textures make it even more challenging. SAM2 zero-shot ability allows segmentation without specific training which makes it better for clinical applications. SAM2 can generalize across diverse medical imaging conditions and nodule variations which allows it to capture details that traditional supervised models may miss. Its zero-shot ability also makes it more robust and transferable. It follows a promptable segmentation approach and can segment objects based on different types of prompts such as boxes, points, and masks. It is a transformer-based architecture and allows better generalization across various domains. Vision transformers can extract various spatial and contextual information at the encoder part. The prompts are converted to embeddings at a prompt encoder that helps the segmentation process. Lastly, the mask decoder generates a segmentation mask. The features from the image encoder and embeddings from the prompt encoder help in generating segmentation masks. These masks are refined through the attention mechanism of transformers.

Data Augmentation:

Medical image datasets are often imbalanced which can affect the model's performance and cause a bias towards the majority class. To overcome the issue, data augmentation techniques such as rotations and flipping have been applied to the dataset to make it balanced. Data augmentation techniques increase the diversity of a dataset. Different scanners have different orientations. Applying random rotations to the dataset will increase the model's rotational invariance. Slight patient movements can add variations to medical images. Applying small angle rotations will help in better generalization of a model. Horizontal and vertical flipping can be applied to handle lateral variations when the patient is lying on different sides and images are taken from different planes. Figure 7 shows some augmentation results.

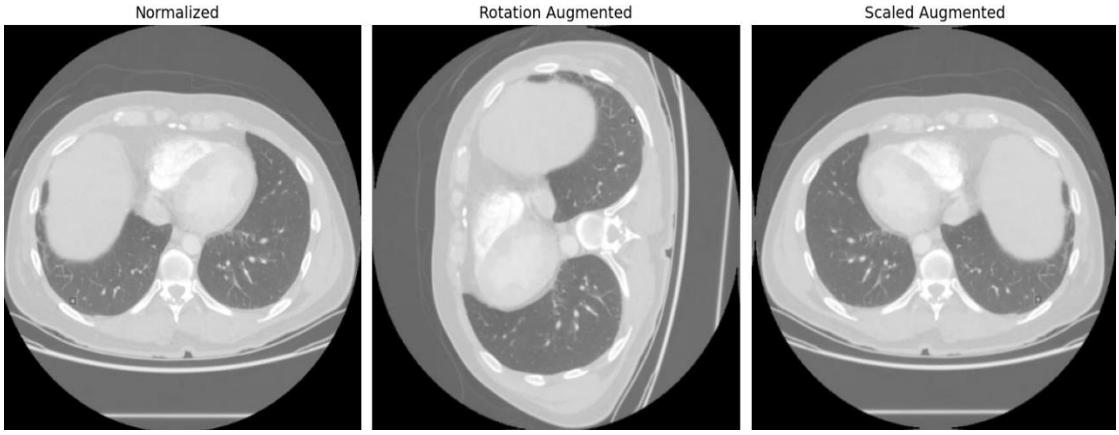

Figure 7. Sample augmented images

DenseNet Model for Classification:

In this study, we also utilized DenseNET-121 for nodule classification as benign and malignant. DenseNET-121 is known for its densely connected layers which improve gradient flow and encourage feature reuse. DenseNet reduces information loss and improves learning by establishing direct connections between each layer and its subsequent layers. This property of DenseNet makes it useful in medical imaging tasks where details are critical for classification, and we cannot afford to lose any details. DenseNET has high parameter efficiency and better feature extraction capabilities which make it capable of solving nodule

classification tasks. The DenseNET architecture consists of four dense blocks where features are concatenated. This makes the model learn intricate patterns of lung CT scans. Figure 8 shows the architecture of the DenseNet-121 model. The model is finetuned on a dataset with GPU available on Google Colab which helps in reducing training and inference time. DenseNET is designed for multi-classification. Since our research focuses on the binary classification of benign and malignant nodules, the final classification layer was modified to contain only two neurons. A learning rate of 0.001 was applied, with a learning rate scheduler implemented to gradually reduce the rate for improved adjustments during training stages. Adam optimizer and Stochastic Gradient Descent (SGD) were utilized, with a batch size of 32 to balance memory efficiency and training stability. Since this is a binary classification task, Cross Entropy loss was used for evaluation. The model was trained for 100 epochs. Table 1 shows hyperparameters used in model training.

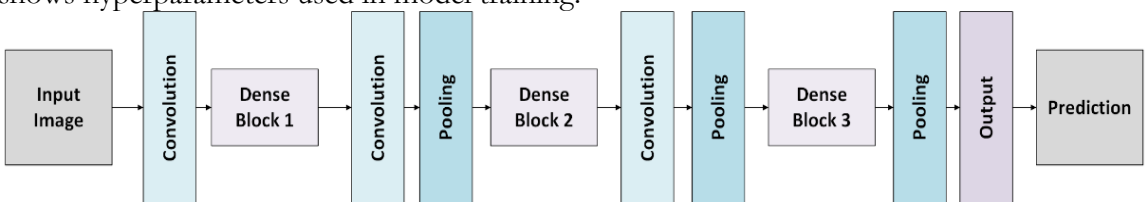

Figure 8 DenseNet-121 Architecture

Table 1. DenseNet-121 model training hyperparameters

Parameter	Value
Optimizer	Stochastic Gradient Descent (SGD)
Loss Function	Cross Entropy
Epochs	100
Batch size	32
Learning rate	0.001

Evaluation Parameters for Segmentation:

Dice Score: The Dice Score is also known as the Dice Similarity Coefficient (DSC). The Dice score is used to evaluate the similarity between predicted segmentation and ground truth. It is calculated by taking twice the TP divided by the total pixel count of the image. The Dice Score is calculated using Equation 7.

$$Dice\ Score = \frac{2 \times TP}{(2 \times TP) + FP + FN} \qquad (7)$$

Intersection over Union: IoU measures the overlapping region between predicted results and ground truths. IoU is calculated by taking the overlapping region (intersection) between actual and predicted output divided by the total region (union). The IoU is calculated using Equation 8.

$$IoU = \frac{TP}{TP + FP + FN} \qquad (8)$$

Loss Function for Classification:

Binary Cross Entropy Loss: To classify nodules as benign and malignant, we have used Binary Cross Entropy Loss (BCE) to compute loss. BCE is also known as Binary Log Loss. It is a loss function that measures the difference between predicted binary results and the actual binary result. BCE is calculated using equation 9.

$$BCE = -\frac{1}{N}\sum_{i=1}^{N}(ti.\log(pi) + (1 - ti)\log(1 - pi)) \qquad (9)$$

Where:

T_i is the true label for instance i.

p_i is the predicted label for instance i.

Result and Discussion:

The segmentation results of the SAM2 model were assessed using standard image segmentation evaluation metrics including Dice Similarity Coefficient (DSC), Intersection over Union (IoU), and Sensitivity. The model was assessed on two datasets LIDC-IDRI and LUNA16 using mask and box prompts to guide the segmentation process.

On the LUNA16 dataset, which consists of complete CT scans, SAM2 achieved an average performance of 95.72% Dice Score, 91.77% Intersection over Union (IoU), and 96.42% Sensitivity using the box prompt. The LUNA16 dataset consists of full CT scans where nodules are present within lung regions. Although SAM2 performed well, the presence of irrelevant structures slightly impacted the performance. Figure 9 shows the segmentation results on the LUNA16 dataset.

While, the LIDC-IDRI dataset, which does not have complete CT scans, shows a 97.87% Dice Score, 95.82% IoU, and 97.9% Sensitivity. Figure 10 shows segmentation results on the LIDC-IDRI dataset. Segmentation results on the LIDC-IDRI dataset were better than those on the LUNA16 dataset because the LIDC-IDRI dataset provided lung CT scans already cropped to lobes, defining the Region of Interest (ROI). This preprocessing removed unnecessary details, enabling the model to segment nodules more effectively.

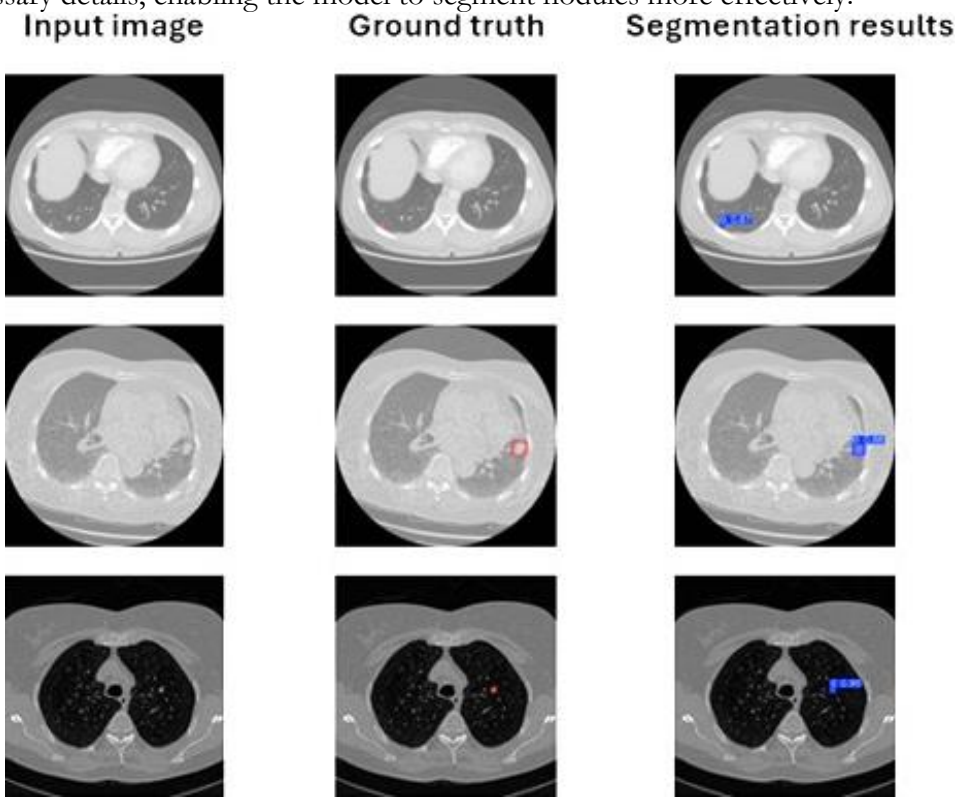

Figure 9. Segmentation results on the LUNA16 dataset

To further validate the performance of SAM2, its performance was compared with other state-of-the-art segmentation techniques. Table 2 shows the comparison of SAM2 with different segmentation models. SAM2 has fast inference time compared to other traditional segmentation networks which requires extensive task-specific fine-tuning. SAM2 zero-shot ability can segment nodules in a very short time making it suitable for clinical applications without the need for extensive training on institution-specific data. Prompt base segmentation allows flexible user inputs which make it adaptable across various clinical scenarios. Although SAM2 achieves better results, it requires GPU memory during inference which adds computational cost in resource-constrained systems.

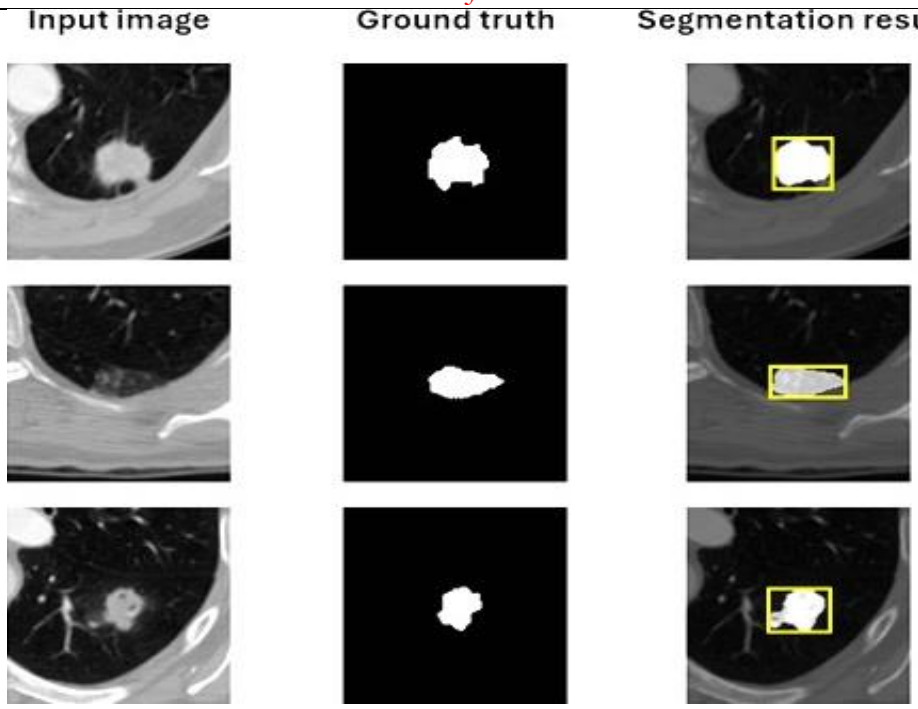

Figure 10. Segmentation results on the LIDC-IDRI dataset

Table 2. Comparison with different segmentation models

Techniques	Dataset	Dice Score%	IoU%
Dual Branch Residual Network [25]	LIDC-IDRI	82.74	---
RFRVNet [21]	LUNA16	95.01	83.00
EFCM [26]	LUNA16	97.10	91.96
SAM [27]	LUNA16	97.08	95.60
Ours	**LUNA16**	**95.72**	**91.77**
Ours	**LIDC-IDRI**	**97.87**	**95.82**

After segmentation, the detected lung nodules were classified into benign and malignant using the DenseNet model. Each radiologist has defined radiomic features for each nodule and based on these radiomic features, a malignancy score between 1-5 is calculated. The nodules having a malignancy score greater than 3 are considered malignant while those with a malignancy score less than 3 are benign. DenseNet model is trained using these malignancy scores for 100 epochs. The proposed model achieved an overall classification accuracy of 97.34% and a validation accuracy of 95.81%. The proposed classification model achieved better results compared to state-of-the-art methods. Table 3 represents the proposed model's performance analysis compared to state-of-the-art methods.

Table 3. Comparison with different classification models

Classification Method	Accuracy %	Sensitivity %	Specificity %
CLIP [28]	70.96	86.77	56.33
GCA+WIRN [29]	94.32	91.49	93.69
ResNet50 [30]	92.56	93.78	90.14
Collaborative Deep Learning [30]	93.24	91.92	91.37
Proposed Model	**97.34**	**98.52**	**96.15**

Discussion:

SAM2 has demonstrated superior performance compared to traditional deep learning-based segmentation models. Unlike traditional segmentation models that require extensive

training for downstream tasks, SAM2 foundational model architecture segment nodules without retraining. Traditional models must be trained on specific datasets which limits their generalization ability for other datasets. On the other hand, SAM2 can adapt to new datasets because it has been trained on diverse modalities during pretraining. SAM2 prompt-based segmentation ability provides flexibility in guiding the segmentation process. In contrast, traditional models rely on fixed pre-defined features or pixel-wise annotations which require radiologists' expertise and are time-consuming. The prompt-based approach allows the model to focus on the Region of Interest (ROI), minimizing segmentation errors. Unlike conventional CNN-based segmentation models that extract hierarchical features, SAM2 utilizes Vision Transformers (ViTs) architecture to capture long-range dependencies to recognize nodules' texture, shapes, and sizes which are important for accurate segmentation. SAM2 showed a slightly lower Dice Score for the LUNA16 dataset due to challenges in detecting nodules from nearby tissues or vessels with slightly similar textures. While SAM2 provides real-time segmentation, it requires high GPU memory which is a constraint for clinical deployment. Traditional deep learning models although slower are computationally more efficient. Our DenseNet model also achieved better classification accuracy outperforming other state-of-the-art methods. The sensitivity score demonstrates that nodules are correctly classified. While our approach significantly improved lung nodule segmentation and classification, there are several areas for future improvement. SAM2 works effectively on Box and masks prompt but this does not make it fully automated. Therefore, further work can be done by introducing text prompts for the SAM2 model to segment lung nodules in a fully automated manner. Also, the datasets are collected from specific demographic groups that do not represent a diverse global patient population. Also, LUNA16 has nodules with a diameter ≥3mm which makes the model less sensitive to small or less distinct nodules.

Conclusion:

The study presented anautomated framework for lung nodule detection and classification using the SAM2 with DenseNet-121 architecture. The combination of SAM2 with the deep learning model demonstrated better performance in nodule segmentation using box and mask prompts and classification tasks. Bounding box and Masks helped in efficient segmentation and results demonstrate the effectiveness of this approach. The research underscores the Segment Anything Model 2 capabilities to effectively segment nodules which can facilitate in Computer Aided Systems. While SAM2 has certain limitations, it can be a promising tool for computer-aided lung cancer detection. The proposed DenseNet classifier also outperforms existing techniques with higher diagnostics accuracy. Figure 11 shows the train and validation loss chart. Future work will focus on improving small nodule segmentation, multimodal integration, and computational efficiency for clinical deployment.

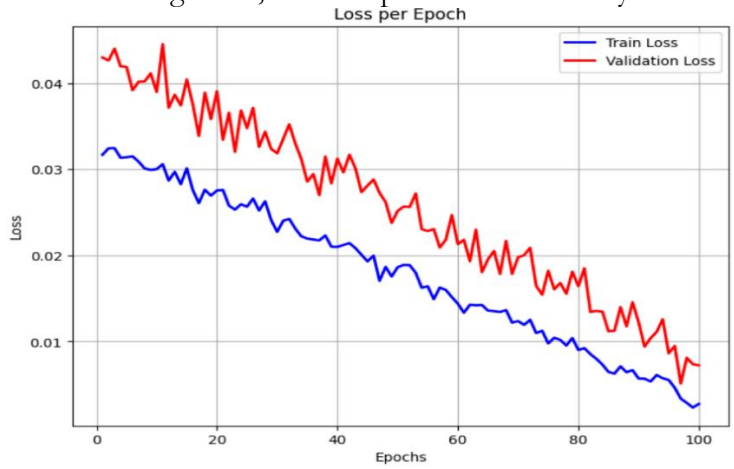

Figure 11 Train and validation loss chart

Acknowledgment:
The Department of Computer Science, University of Engineering and Technology, Taxila facilitated us with the Image Processing Lab and other resources for the experimentation of this research.

Author's Contribution:
Hafza Eman contributed to the methodology, manuscript writing, and interpretation of results.
Syed Muhammad Adnan proposed the topic and performed statistical analysis.
Wakeel Ahmad contributed to the statistical analysis and interpretation of results.
Ishtiaque Mahmood conducted the literature review and referencing.

Conflict of interest. No

Project details. No

References:

[1] B. F. Ferlay J, E.M., Lam F, Laversanne M, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, "Global Cancer Observatory: Cancer Today," *Int. Agency Res. Cancer*, 2024, [Online]. Available: https://gco.iarc.fr/today/en

[2] G. D. Rubin, "Lung nodule and cancer detection in computed tomography screening," *J. Thorac. Imaging*, vol. 30, no. 2, pp. 130–138, Mar. 2015, doi: 10.1097/RTI.0000000000000140.

[3] Ted Way, Heang-Ping Chan, Lubomir Hadjiiski , Berkman Sahiner, Aamer Chughtai, Thomas K. Song, Chad Poopat, Jadranka Stojanovska , Luba Frank, Anil Attili, Naama Bogot, and E. A. K. Philip N. Cascade, "Computer-Aided Diagnosis of Lung Nodules on CT Scans:: ROC Study of Its Effect on Radiologists' Performance," *Acad. Radiol.*, vol. 17, no. 3, pp. 323–332, 2010, doi: https://doi.org/10.1016/j.acra.2009.10.016.

[4] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017, doi: https://doi.org/10.1016/j.media.2017.07.005.

[5] Lulu Wang, "Deep Learning Techniques to Diagnose Lung Cancer," *Cancers (Basel)*, vol. 14, no. 22, p. 5569, 2022, doi: 10.3390/cancers14225569.

[6] Md. Eshmam Rayed, S.M. Sajibul Islam, Sadia Islam Niha and M. F. M. Jamin Rahman Jim, Md Mohsin Kabir, "Deep learning for medical image segmentation: State-of-the-art advancements and challenges," *Informatics Med. Unlocked*, vol. 47, p. 101504, 2024, doi: https://doi.org/10.1016/j.imu.2024.101504.

[7] B. O. A. Ademola E. Ilesanmi, Taiwo O. Ilesanmi, "Reviewing 3D convolutional neural network approaches for medical image segmentation," *Heliyon*, vol. 10, no. 6, p. e27398, 2024, [Online]. Available: https://www.cell.com/heliyon/fulltext/S2405-8440(24)03429-7?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS2405844024034297%3Fshowall%3Dtrue

[8] B. G. Satya P Singh, Lipo Wang, Sukrit Gupta, Haveesh Goli, Parasuraman Padmanabhan, "3D Deep Learning on Medical Images: A Review," *Sensors (Basel)*, vol. 20, no. 18, p. 5097, 2020, doi: 10.3390/s20185097.

[9] Y.-G. S. Muhammad Usman, "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, 2023, doi: https://doi.org/10.3390/s23041989.

[10] Z. Wang, J. Men, and F. Zhang, "Improved V-Net lung nodule segmentation method based on selective kernel," *Signal, Image Video Process.*, vol. 17, no. 5, pp. 1763–1774, Jul. 2023, doi: 10.1007/S11760-022-02387-W/METRICS.

[11] F. L. J. Hou, C. Yan, R. Li, Q. Huang, X. Fan, "Lung Nodule Segmentation Algorithm With SMR-UNet," *IEEE Access*, vol. 11, pp. 34319–3433, 2023, doi:

10.1109/ACCESS.2023.3264789.

[12] Y.-T. W. Ting-Wei Wang, Chih-Keng Wang, Jia-Sheng Hong, Heng-Sheng Chao, Yuh-Min Chen, "Deep Learning in Thoracic Oncology: Meta-Analytical Insights into Lung Nodule Early-Detection Technologies," *Cancers (Basel).*, vol. 17, no. 4, p. 621, 2025, doi: https://doi.org/10.3390/cancers17040621.

[13] J. Z. Xiaoguang Tu, Zhi He, Yi Huang, Zhi-Hao Zhang, Ming Yang, "An overview of large AI models and their applications," *Vis. Intell.*, vol. 2, no. 34, 2024, doi: https://doi.org/10.1007/s44267-024-00065-8.

[14] R. G. Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár, "Segment Anything," *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2023, doi: https://doi.org/10.48550/arXiv.2304.02643.

[15] Y. J. Junde Wu, Wei Ji, Yuanpei Liu, Huazhu Fu, Min Xu, Yanwu Xu, "Medical sam adapter: Adapting segment anything model for medical image segmentation," *arXiv Prepr. arXiv2304.12620*, vol. 1, 2023, doi: https://doi.org/10.48550/arXiv.2304.12620.

[16] C. F. Nikhila Ravi, Valentin Gabeur, Yuan-Ting Hu, Ronghang Hu, Chaitanya Ryali, Tengyu Ma, Haitham Khedr, Roman Rädle, Chloe Rolland, Laura Gustafson, Eric Mintun, Junting Pan, Kalyan Vasudev Alwala, Nicolas Carion, Chao-Yuan Wu, Ross Girshick, Piotr Dollár, "Sam 2: Segment anything in images and videos," *arXiv Prepr. arXiv2408.00714*, 2024, doi: https://doi.org/10.48550/arXiv.2408.00714.

[17] M. O. Marjolein A. Heuvelmans, Peter M.A. van Ooijen, Sarim Ather, Carlos Francisco Silva, Daiwei Han, Claus Peter Heussel, William Hickes, Hans-Ulrich Kauczor, Petr Novotny, Heiko Peschl, Mieneke Rook, Roman Rubtsov, Oyunbileg von Stackelberg, Maria T. Tsakok, "Lung cancer prediction by Deep Learning to identify benign lung nodules," *Lung Cancer*, vol. 154, pp. 1–4, 2021, doi: https://doi.org/10.1016/j.lungcan.2021.01.027.

[18] H. Huang, R. Wu, Y. Li, and C. Peng, "Self-Supervised Transfer Learning Based on Domain Adaptation for Benign-Malignant Lung Nodule Classification on Thoracic CT," *IEEE J. Biomed. Heal. Informatics*, vol. 26, no. 8, pp. 3860–3871, Aug. 2022, doi: 10.1109/JBHI.2022.3171851.

[19] K. S. Mundher Al-Shabi and M. Tan, "ProCAN: Progressive growing channel attentive non-local network for lung nodule classification," *Pattern Recognit.*, vol. 122, p. 108309, 2022, doi: https://doi.org/10.1016/j.patcog.2021.108309.

[20] S. A. Mahmood and H. A. Ahmed, "An improved CNN-based architecture for automatic lung nodule classification," *Med. Biol. Eng. Comput.*, vol. 60, no. 7, pp. 1977–1986, Jul. 2022, doi: 10.1007/S11517-022-02578-0/METRICS.

[21] S. Dodia, A. Basava, and M. Padukudru Anand, "A novel receptive field-regularized V-net and nodule classification network for lung nodule detection," *Int. J. Imaging Syst. Technol.*, vol. 32, no. 1, pp. 88–101, Jan. 2022, doi: 10.1002/IMA.22636.

[22] A. Dosovitskiy *et al.*, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," *ICLR 2021 - 9th Int. Conf. Learn. Represent.*, Oct. 2020, Accessed: May 16, 2024. [Online]. Available: https://arxiv.org/abs/2010.11929v2

[23] Arnaud Arindra Adiyoso Setio , Alberto Traverso , Thomas de Bel, Moira S.N. Berens, Cas van den Bogaard, Cerello Piergiorgio, Hao Chen and R. van der G. Qi Dou, Maria Evelina Fantacci, Bram Geurts, "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge," *Med. Image Anal.*, vol. 42, pp. 1–13, 2017, doi: https://doi.org/10.1016/j.media.2017.06.015.

[24] S. Armato *et al.*, "The Lung Image Database Consortium (LIDC) and Image Database

Resource Initiative (IDRI): A Completed Public Database of CT Scans for Lung Nodule Analysis," *Med. Phys.*, vol. 37, no. 6Part6, pp. 3416–3417, Jun. 2010, doi: 10.1118/1.3469350.

[25] Haichao Cao, Hong Liu, C.-C. H. Enmin Song, X. X. Guangzhi Ma, Renchao Jin, and J. Lu, "Dual-branch residual network for lung nodule segmentation," *Appl. Soft Comput.*, vol. 86, p. 105934, 2020, doi: https://doi.org/10.1016/j.asoc.2019.105934.

[26] S. Dodia, B. Annappa, and M. A. Padukudru, "A Novel Artificial Intelligence-Based Lung Nodule Segmentation and Classification System on CT Scans," *Commun. Comput. Inf. Sci.*, vol. 1568 CCIS, pp. 552–564, 2022, doi: 10.1007/978-3-031-11349-9_48.

[27] B. K. Asha V, "Advanced Lung Nodule Segmentation and Classification for Early Detection of Lung Cancer using SAM and Transfer Learning," *arXiv:2501.00586v1*, 2024, doi: https://doi.org/10.48550/arXiv.2501.00586.

[28] F. Shaukat, S. M. Anwar, A. Parida, V. K. Lam, M. G. Linguraru, and M. Shah, "Lung-CADex: Fully Automatic Zero-Shot Detection and Classification of Lung Nodules in Thoracic CT Images," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* , vol. 15241 LNCS, pp. 73–82, 2025, doi: 10.1007/978-3-031-73284-3_8.

[29] H. Y. Wenju Wang, Shuya Yin, Fang Ye, Yinan Chen , Lin Zhu, "3D global coordinate attention wide inverted ResNet network for pulmonary nodules classification," *BMC Pulm Med*, vol. 24, no. 1, p. 465, 2024, doi: 10.1186/s12890-024-03272-7.

[30] A. H. Alshamrani K, "Classification of Chest CT Lung Nodules Using Collaborative Deep Learning Model," *J. Multidiscip. Healthc.*, p. 1459—1472, 2024, doi: https://doi.org/10.2147/JMDH.S456167.