Nodule Classification Using Custom Build 3D Convolution Neural Network Model

Sijo Thomas,* Dr. Sajimon Abraham,† Praveen Kumar,‡ Faizal Basheer,§ Jithinmary Raphel¶
Mahatma Gandhi University, Kerala, 686560, India

Abstract

The classification of lung nodules plays a pivotal role in the early detection and prognosis assessment of lung-related diseases. This paper introduces an innovative approach to nodule classification by leveraging the capabilities of a 3D Convolutional Neural Network (3D CNN) model. The primary objective of this method is to enhance the precision and efficiency of nodule classification, ultimately contributing to improved medical diagnosis and treatment planning.

Our research introduces customized 3D augmentation techniques and a distinct 3D CNN model with finely-tuned parameters. The main focus of the model is to determine the optimal CT patch size for nodule classification while minimizing the training requirements on the proposed custom model. Our research begins by creating patch slice counts ranging from 64 to 24 and 2D spaces of 96x96, ensuring that the nodule center is positioned at the center of each patch. To enhance performance and reduce data transfer bottlenecks between the GPU and HDD, each data chunk is cached in the local environment.

Key Words: nodule classification; augmentation; Lung Cancer;Computed Tomography; Convolutional Neural Network,

1 Introduction

Deep learning models require substantial amounts of data and a significant computational capacity. In the earlier stages, such computational resources were prohibitively expensive and not accessible to a wide range of researchers. However, the recent availability of high-performance computing environments, such as GPUs, has had a profound impact on the processing of models that harness the capabilities of deep learning [1]. This development has been particularly influential in fields like medical image processing and Computer-Aided Diagnoses, where deep learning models have seen extensive use [2].

Deep learning-based systems now play a crucial role in identifying specific regions of interest in medical images, such as CT scans and X-rays, which require in-depth examination by medical professionals. Additionally, they can generate automated reports that aid doctors in correlating their observations [3].

Lung cancer ranks as the most prevalent form of cancer among humans [4] [5]. Each year, lung cancer claims more lives than breast, prostate, and colon cancers combined [5]. Achieving effective treatment necessitates the early detection of cancer. During the initial stages of cancer development, cancer cells manifest as tiny nodules. Given the extensive volume of computed tomography data, medical professionals must dedicate substantial effort to extract and analyze various features, such as size, morphology, contours, interval growth between CT examinations, multiplicity, and location [6].

Among the various deep learning models available, Convolutional Neural Networks (CNNs) have achieved a remarkable breakthrough by reducing computational parameters and enhancing accuracy [7]. Two-dimensional CNNs have shown superior performance when applied to 2D data. However, the application of two-dimensional CNNs is limited to single-slice CT (Computed Tomography) data, potentially resulting in the loss of inter-slice information [8][9].

Considering the size and spatial distribution of malignant nodules, they often extend across multiple slices of CT scans. This necessitates the use of 3D CNN models, but it comes with the added requirement of substantial GPU resources to handle the increased computational load.

In this paper, we present a method for addressing classification challenges in 3D data without causing excessive strain on GPU resources. Our research involves conducting experiments with segments of CT scans containing multiple dimensions, using a custom-built CNN model developed from the ground up. The model experimentation begins with a chunk size of 96x96x64 (XxYxZ), and we explore different combinations of dimensions to determine the most effective chunk dimension for accurate nodule classification. Additionally, we introduce custom augmentation techniques, such as rotation, flip, and offset, to enhance the dataset without putting undue stress on the hardware.

*Research Scholar, School of Computer Sciences. Email: sijothakadiyel@gmail.com
†Research Supervisor, School of Computer Sciences. Email: sajimonabraham@mgu.ac.in
‡Research Scholar, School of Computer Sciences, Email: praveenplavila@yahoo.co.in
§Research Scholar, School of Computer Science, School of Computer Science. Email: faizalbs777@gmail.com
¶Senior Quality Analyst, McFadyen Digital. Email: jraphel@mcfadyen.com

The experiments are assessed within a local environment using a single GPU with 8GB of memory (RTX 3060 Ti) and an Intel i5 processor with 16 GB of RAM.

Our primary contributions to this study can be summarized as follows:

1. We have estimated an optimized chunk of CT scan data to maximize the capabilities of the 3D CNN model.
2. We have introduced a custom-built 3D CNN model tailored for efficient classification tasks.
3. We have introduced effective data augmentation techniques specifically designed for CT scan data, all while ensuring GPU resources are not overwhelmed.
4. We have successfully conducted precise nodule classification on the LUNA / LIDC-IDRI dataset, achieving benchmark results.

The remainder of this paper is structured as follows: Section 2 provides an overview of related research, Section 3 elaborates on the proposed method, Section 4 presents experimental results and analysis, and finally, Section 5 offers the conclusion of this paper.

2 Related Work

This study reviewed multiple categories of articles published on lung nodule classification. The major categories of articles are the deep learning approach based on CNN and the deep learning approach based on non-CNN models. In the traditional non-deep learning approach, researchers used strategies available in computer vision with derivatives of machine learning models. One of the leading developments was the use of Gabr Local Binary Pattern [10] - LBP with support Vector Machine [11] algorithm to predict the probability of nodule.

Hussein [13] worked on creating a 3D CNN model for the classification of nodules. Due to the 3D volumetric nature of CT scans, and the unavailability of the pre-trained model at that time, authors trained their models on sports datasets with more than 1 million videos. In a later stage, they fine-tuned the results on the LIDC dataset.

Author Nibali [14] used a transfer learning strategy for their Resnet18 model. The model was trained on the CIFAR 10 data-set for additional generalization of the pre-trained network. Based on the views of the nodule, for example, axial, coronal, and sagittal, the authors have introduced a three-layer network approach for each view. In the final stage, outputs from three networks are fed into a fully connected layer for final classification. The additional overhead of three network models's memory and training time was a major drawback observed there.

Author Liu[15] proposed a multi-view approach for the classification of nodules. By introducing three scales and four views, the authors are able to introduce a new data set containing 12 unique images of one slice of a CT scan. The authors trained the model on each of the images separately and combined the model for final training. This method also introduced additional overhead of memory and training time.

Author Wentao Zhu[9] DeepLung is a comprehensive system that comprises two core components: one for detecting nodules and the other for classifying these nodules as either benign or malignant. Given the three-dimensional nature of lung CT data and the efficiency of dual path networks (DPN), two distinct deep 3D DPNs have been developed to handle nodule detection and classification tasks. To be specific, a 3D Faster Regions with Convolutional Neural Net (R-CNN) has been tailored for nodule detection, utilizing 3D dual path blocks and adopting a U- net-like encoder-decoder structure to effectively capture nodule features. In the realm of nodule classification, the approach utilizes gradient boosting machine (GBM) in conjunction with 3D dual path network features.

Author Al-Shabi[16] proposal involves the utilization of Residual Blocks employing a 3 × 3 kernel size to extract local features and the incorporation of Non-Local Blocks for the extraction of global features. The Non-Local Block is particularly advantageous in extracting global features without the need for an excessively large number of parameters. The fundamental concept underpinning the Non-Local Block is the application of matrix multiplications among features within the same feature maps.

Table 1: Performance of leading classification methods and baseline models

Method	Accuracy	AUC
DeepLung	90.44	97.13
Local-Global	88.46	95.62
Resnet50	77.62	86.82
Resnet18	78.21	86.41
Densenet121	84.57	92.50
Multi-Crop	89.27	94.32

3 Proposed

3.1 Methods

Our custom CNN model was implemented utilizing the PyTorch deep-learning library[23]. The model execution took place on a local GPU environment, specifically the NVIDIA RTX 3060 Ti with an 8 GB capacity. The operating system used was Ubuntu OS 22.04, and the Python version employed was 3.9 via Anaconda.

In order to effectively handle the dataset, it was partitioned into a 70% training set and a 30% validation set. Given the considerable and fluctuating slice count within the CT scans, it was not feasible to accommodate full-size CT scans within our limited GPU environment. To address this constraint, we

structured our experiment to work with different slice counts, including 64, 48, 32, and 24 slices.

For each of these slice counts, we explored various 2D patch sizes, specifically 96x96, 64x64, 48x48, 32x32, and 24x24. Due to the constraints imposed by our hardware infrastructure, we maintained a fixed batch size of 16 across all patch sizes. These patches were derived using a mathematical approximation of the annotated nodule center, ensuring that it was positioned at the center of the patch in both 2D and 3D spaces.

Learning rate, hyper tuning parameter is tuned from $1x10^{-2}$ to $1x10^{-4}$[24]. The first 30 epochs are run with $1x10^{-2}$ and series of epochs 31 to 60 with $1x10^{-3}$ and the final epochs 61 to 75 with $1x10^{-4}$. The approach used for the learning rate selection was to introduce faster changes in the initial training phases and fine-tune in the last stage of training[25].

We followed an approach of Kaiming initialization of neural network weight to explicitly initialize the weight of kernels[28]. Since the CNN model relies on ReLU for activation, during the Kaiming initialization process, opted ReLU for non-linearity. Adam optimizer are used with default values for β_1 and β_2[26]. Checkpoints are saved on each model performance on each patch size with model state and neural weights.

3.2 Dataset Description

The dataset utilized in this study is the LIDC-IDRI [17] dataset, which has been made available by the National Cancer Institute, NIH. This dataset is renowned for being the most expansive and comprehensive public repository of lung nodules. It encompasses a compilation of 1,018 CT scans, derived from a total of 1,010 patients. Due to its substantial scale and public accessibility, the LIDC-IDRI dataset serves as a fitting resource for the development, comparison, and validation of diverse deep-learning methodologies. Moreover, it holds a significant presence in the existing literature, being frequently referenced [13] [15] in studies.

The image collection process encompassed contributions from four different institutions, each employing separate CT scanners. This approach introduced significant diversity in various image attributes. One notable area of diversity lies in image resolution, as evidenced by variations in pixel spacing across different CT scanners. Additionally, the thickness of CT scan slices ranged from 0.45 to 5.0 mm.

Leveraging a heterogeneous dataset like LIDC-IDRI for algorithm design presents the benefit of ensuring robustness to previously unencountered or broadly applicable data. This robustness arises because the algorithms are trained on a wide spectrum of cases, offering versatility in handling diverse scenarios.

Within the LIDC-IDRI dataset, the malignancy likelihood of each nodule underwent evaluation by a panel consisting of four highly experienced radiologists. Initially, all nodules present in each CT scan were annotated, and their boundaries were meticulously documented in individual XML files. It's worth noting that these annotations were restricted to nodules with diameters spanning from 3 to 30 mm.

Interestingly, despite the thorough evaluation by the four experienced radiologists, variations in the nodule annotations were observed, as noted in prior research. Specifically, only a subset of nodules received annotations by the majority of radiologists, which means they were annotated by at least three out of the four experts.

The radiologists also provided a malignancy rating for each nodule, employing a scale ranging from 1 to 5, where higher values indicated higher levels of malignancy (1 represented a benign nodule, and 5 indicated a highly malignant one). To consolidate the ratings provided by the radiologists, we computed the median malignancy score. Nodules receiving ratings below three were categorized as benign, while those with ratings exceeding 3 were classified as malignant. Following a common practice in similar studies, nodules with a rating of exactly 3 were omitted from the analysis due to their indeterminate malignancy status.

3.3 Preprocessing

The LIDC-IDRI dataset initially came in .mhd format along with corresponding raw data files. Our initial step involved converting these .mhd files into a tensor data type, while carefully considering variations in coordinate systems, vertical spacing, and differing slice counts.

In this research, we applied a clamping operation to the Hounsfield unit (HU) values of the scans, specifically restricting values below -1000 or above 400. This practice is well-established in the literature and is commonly used to exclude regions corresponding to air and bone tissues [25][22].

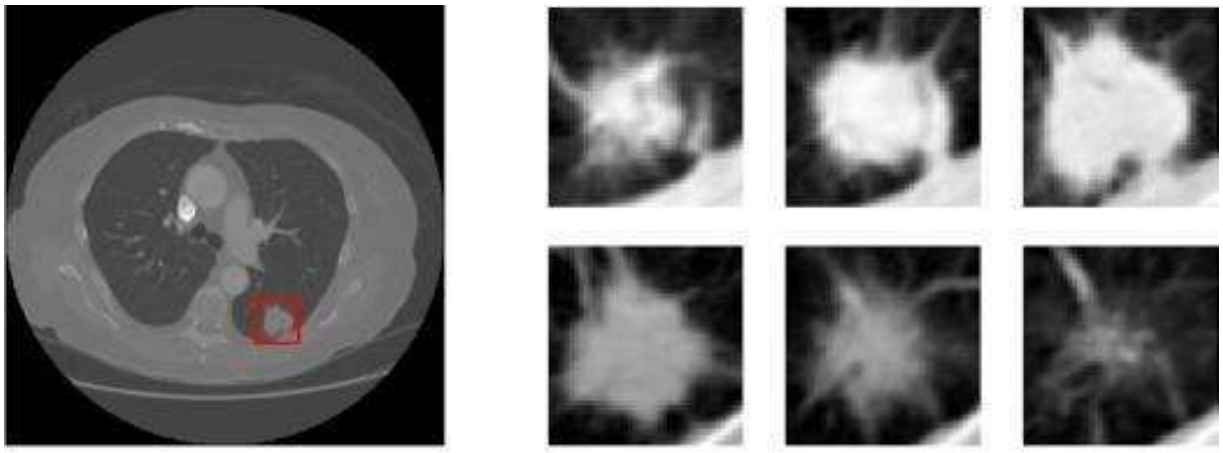
As part of our initial data preprocessing, we performed batch-level normalization of pixel intensities to ensure consistent mean and variance across the dataset. To precisely identify nodules, which are the areas of interest in CT scans, we utilized annotations to pinpoint the center of each nodule. We developed a versatile conversion function to prepare data chunks, ranging in size from 96x96x64 to 24x24x24, for input into our neural network.

To optimize performance and reduce data transfer bottlenecks between the GPU and HDD, each chunk of data was cached in the local environment.

3.4 Data Augmentation

The LIDC-IDRI dataset is imbalanced in terms of the nodule and non-nodule counts: 40,772 out of 62,492 patches are lung nodules, and 21,720 out of 62,492 patches are non-nodules [18]. Such scenarios make actual nodules have less impact on the training of networks. The resolution was to introduce and implement augmentation on chunks of CT scans. The following augmentation has been applied on the patches of CT.

Rotation: We carried out rotations at random angles of 0°, 90°, 180°, and 270°. The rotation logic was executed within a loop, repeating three times. During each iteration, an additional 90° angle was added based on a randomly generated number.



(a) A 2D lung CT slice. (b) Demonstration of a nodule at different Z-axis positions.

Figure 1: Represents sample CT image in 2D and 3D space

Augmentation of CT scans through flipping involves applying various transformations to the original scans to generate additional training data. Flipping typically includes horizontal or vertical flips. Horizontal Flip: This augmentation technique involves flipping the CT scan horizontally along its vertical axis. Vertical Flip: In contrast, a vertical flip flips the CT scan vertically along its horizontal axis[19].

Augmentation of CT scans through offset involves applying spatial translations or shifts to the original scans. This technique can create variations in the position of structures or objects within the CT images, which can help improve the robustness and generalization of machine learning models trained on these images[20]. The CT scan is shifted horizontally and/or vertically by a certain number of pixels. This simulates a change in the position of the patient or the scanner during image acquisition. Translations can be applied in both positive and negative directions along the x and y axes[19]

Augmentation techniques, which have a significant impact on the shape of nodules, are avoided [21].

3.5 Network Architecture

Figure 3 represents the overall architecture and methodology of research. The approach is modularized into 3 stages. In Stage 1, a 3D neural network model is developed using the PyTorch library 1.13. The model includes four blocks of 3D CNN where each set block has two 3D neural layer with ReLU activation function as represented in figure 2. Scale down has been applied using MaxPool 3D with kernel size and stride with 2 on the end of each block. Each layer of the 3D neural layer will have kernel size 3x3x3, stride 1, and padding 1.

The initial block comprises one input channel and yields eight output channels. Within this block, the first 3D CNN layer takes one input channel and produces eight output channels.

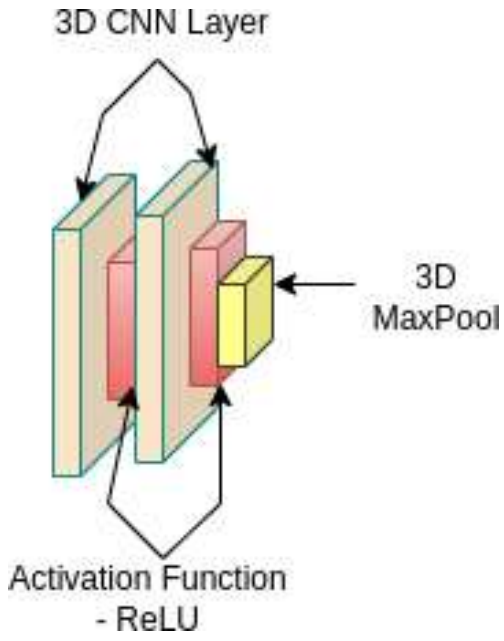


Figure 2: Represents unit block of our 3D CNN model

The output is then subjected to a ReLU activation function before being fed into the second 3D CNN layer. This second layer possesses eight input channels and generates eight output channels, once again going through a ReLU activation function. The outcome of this activation function is subsequently passed through a 3D Max pooling layer, where the pooling operation reduces the dimensionality by dividing the three-dimensional

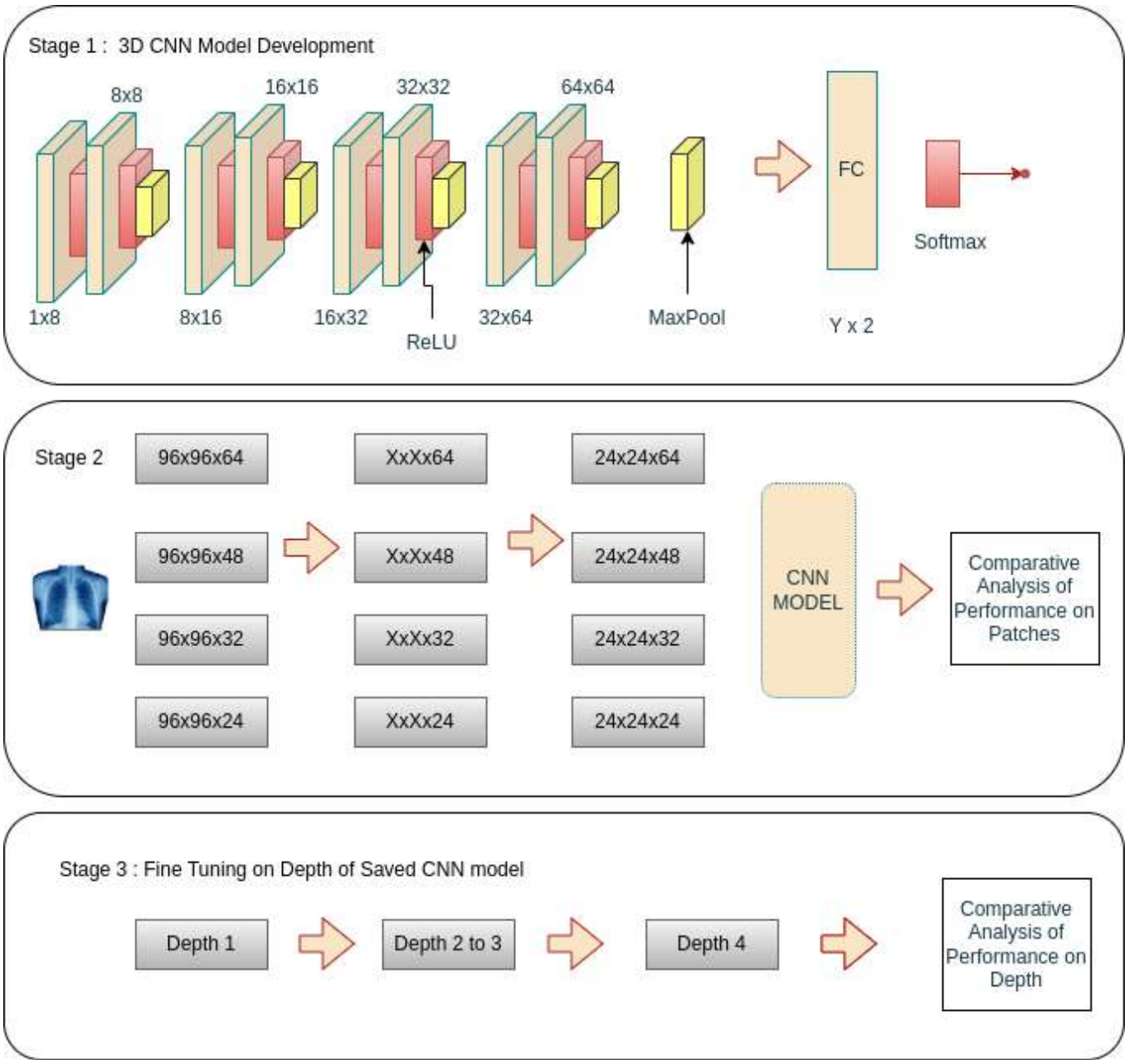


Figure 3: Represents the overall architecture of the study. In stage 1, the initial design of the 3D CNN model is built. Models are initialized with random parameters. In stage 2, the built model is evaluated on the different patch sizes and evaluate the model performance. The best model is saved for stage 3. In stage 3, the model is fine-tuned over different depths of CNN model

input into cuboidal regions and computing the maximum value in each region [source: network:maxpool].

The second block involves eight input channels and results in sixteen output channels. In this block, the first 3D CNN layer takes eight input channels and produces sixteen output channels, followed by a ReLU activation. The output is then fed into the second 3D CNN layer, which operates with sixteen input channels to yield sixteen output channels. Again, a ReLU activation is applied before the result is passed through a 3D Max pooling layer for downsampling.

Moving on to the third block, it encompasses sixteen input channels and delivers thirty-two output channels. The initial 3D CNN layer in this block takes sixteen input channels and generates thirty-two output channels, followed by ReLU activation. The subsequent 3D CNN layer within this block operates with thirty-two input channels to produce thirty-two output channels, once again undergoing ReLU activation. The result from this activation is then channeled through a 3D Max pooling layer for downsampling.

The fourth block has thirty-two input channels and yields sixty-four output channels. Here, the first 3D CNN layer in this block takes thirty-two input channels and generates sixty-four output channels, followed by ReLU activation. Subsequently, the second 3D CNN layer within this block operates with sixty-four input channels to yield sixty-four output channels, going through ReLU activation once more. The output from this activation function is passed through a 3D Max pooling layer for downsampling purposes.

The model's head consists of a single layer of fully connected linear neurons with two output units, which is subsequently followed by the application of a softmax activation function. In contrast, the model's tail includes batch-level normalization, where mean and standard deviation values are computed separately for each dimension across mini-batches.

During stage 2, the model undergoes evaluation on diverse patches extracted from chest scans, each with different dimensions. Promising patches that exhibit favorable performance will be chosen for further refinement during stage 3. In stage 3, the fine-tuning process takes place on the same dataset, with adjustments being made to various depths of the model.

4 Experimental Results and Analysis

Table 4 represents the performance of the classification model on a patch size of 96x9x96. Two-dimension size 96x96 with slice length of 64 to 24. Considering the model capacity and large patch size compared to nodule size, classification performance was not showing significant improvement compared to baseline models and other leading methods. Table 5 represents the performance of the classification model on a patch size of 96x96x64. Two-dimension size 64x64 with slice length of 64 to 24. Among the various patch sizes, 64x64x48 and 64x64x32 lead the classification performance.

Table 2: Model's detailed block architecture

Block	Layer	Input / Output	Kernel	Stride / Padding
Block 1	CNN	1 / 8	(3,3,3)	(1,1,1) (1,1,1)
Block 1	ReLU	-	-	-
Block 1	CNN	8 / 8	(3,3,3)	(1,1,1) (1,1,1)
Block 1	ReLU	-	-	-
Block 1	MaxPool3D	-	(2,2,2)	(2,2,2) (0,0,0)
Block 2	CNN	8 / 16	(3,3,3)	(1,1,1) (1,1,1)
Block 2	ReLU	-	-	-
Block 2	CNN	16 / 16	(3,3,3)	(1,1,1) (1,1,1)
Block 2	ReLU	-	-	-
Block 2	MaxPool3D	-	(2,2,2)	(2,2,2) (0,0,0)
Block 3	CNN	16 / 32	(3,3,3)	(1,1,1) (1,1,1)
Block 3	ReLU	-	-	-
Block 3	CNN	32 / 32	(3,3,3)	(1,1,1) (1,1,1)
Block 3	ReLU	-	-	-
Block 3	MaxPool3D	-	(2,2,2)	(2,2,2) (0,0,0)
Block 4	CNN	32 / 64	(3,3,3)	(1,1,1) (1,1,1)
Block 4	ReLU	-	-	-
Block 4	CNN	64 / 64	(3,3,3)	(1,1,1) (1,1,1)
Block 4	ReLU	-	-	-
Block 4	MaxPool3D	-	(2,2,2)	(2,2,2) (0,0,0)

Table 6 represents the performance of the classification model on a patch size of 48x48. Two-dimension size 48x48 with slice length of 64 to 24. Among the various patch sizes, 48x48x48 and 48x48x32 lead the classification performance.

Table 7 represents the performance of the classification model on a patch size of 32x32xZ Two-dimension size 32x32 with slice length of 64 to 24. Among the various patch sizes, 32x32x48 and 32x32x64 lead the classification performance, and patch size 32x32x48 has the highest accuracy across all the patches.

Table 3: model architecture for head and tail (32x32x48)

Block	Layer	Input / Output	Kernel	Stride / Padding
Tail	BatchNorm	-	-	- -
Head	Linear	1152 / 2	-	- -
Head	SoftMax	2/1	-	- -

Table 4: Performance of patch size 96x96 on CNN model

Patch Size	Accuracy	AUC
96x96x64	78.00	79.01
96x96x48	79.00	80.10
96x96x32	74.00	73.76
96x96x24	73.00	72.06

Table 8 represents the performance of the classification model on a patch size of 24x24xZ Two-dimension size 24x2 with slice length of 64 to 24. The performance of the classification model was below the other patches except the 96x96 patch.

The experimental results on the validation dataset show that patch sizes 32x32x48 and 48x48x32 have better performance yield compared to other patch sizes. These two patches are further taken for stage 3, the fine-tuning stage where depth-level fine-tuning of the CNN model is performed. Models are fine-tuned with depths 1 to 3 on the epoch of 50s. The model will be re-trained with all other layers of the model kept intact except layers up to specified depth. The table 9 and 10 shows the model performance of best classification patches on various depths

Table 9 represents the performance of the classification model on a patch size of 48x48x32 with fine-tuning at various depths. Among the various depth, depth 2 lead the classification performance.

Table 10 represents the performance of the classification model on a patch size of 32x32x48 with fine-tuning at various depths. Among the various depth, depth 3 lead the classification performance.

5 Discussion

The main objective of this work was to utilise the LIDC dataset for training and evaluating the acquired results. Additional validation can be attained through the implementation of extensive experiments that utilise diverse datasets. This approach serves to mitigate any biases and bolster the overall strength and reliability of the results. Further enhancements can be investigated by effectively

Table 5: Performance of patch size 64x64 on CNN model

Patch Size	Accuracy	AUC
64x64x64	81.04	81.00
64x64x48	84.09	84.00
64x64x32	84.69	83.34
64x64x24	70.03	73.10

Table 6: Comparative performance of the model on best slice-count

Patch Size	Accuracy	AUC
48x48x64	83.90	84.00
48x48x48	90.89	91.09
48x48x32	91.29	91.98
48x48x24	81.01	81.24

implementing augmentation techniques, activation functions, and incorporating them in conjunction with loss functions, among other considerations.

6 Conclusions and Future work

In our research paper, we introduced a specialized CNN model designed for the classification of nodules as either benign or malignant. The rationale behind developing this custom CNN model was to gain a deeper understanding of the hardware needs and achieve enhanced accuracy in performance. Given the substantial size of CT scans, there could be challenges in processing them in resource-limited GPU environments. To address this, we devised a solution: generating CT scan patches centered around the annotated center of the nodules. Our study revealed that the 32x32x48 patch size yielded superior performance results when used with our custom classification model.

The augmentation strategy was created with the aim of minimizing the machine's workload and expediting network training through faster networking. Additionally, the augmentation approach has been devised to facilitate parallel execution, relieving the CPU from excessive burdens.

Furthermore, to enhance the utilization of GPU bandwidth, we pre-cache CT patches prior to conducting experiments. This approach helps minimize data loading latency on the GPU.

Our observations indicate that both higher-dimensional and lower-dimensional patch sizes yield classification accuracy that falls below that of middle-sized patches. The reduced accuracy observed in lower-dimensional patches may be attributed to the limited information contained within them. On the other hand, the diminished accuracy in higher-dimensional patches

Table 7: Performance of patch size 32x32 on CNN model

Patch Size	Accuracy	AUC
32x32x64	88.12	87.27
32x32x48	92.10	93.23
32x32x32	87.09	88.14
32x32x24	88.98	87.12

Table 8: Performance of patch size 24x24 on CNN model

Patch Size	Accuracy	AUC
24x24x64	86.12	85.28
24x24x48	85.95	85.00
24x24x32	82.63	80.12
24x24x24	80.18	79.65

Table 9: Performance of fine-tuning on patch size 48x48x32

Patch Size	Accuracy	AUC
Depth 1	90.98	91.34
Depth 2	92.45	92.19
Depth 3	92.61	93.24
Depth 4	91.56	90.42

Table 10: Performance of fine-tuning on patch size 32x32x48

Patch Size	Accuracy	AUC
Depth 1	92.98	91.34
Depth 2	93.45	94.19
Depth 3	95.61	95.24
Depth 4	90.56	89.42

could be due to the model's smaller receptive field and/or the need for more training epochs to effectively process the richer information in the data. Conducting further research in this domain would necessitate additional computational resources and in-depth analysis.

An alternative approach to consider is combining the best-performing patches with a common fully connected linear layer within the CNN model, which may lead to improved results. This avenue presents an opportunity for future research exploration in this field.

Additionally, in the context of multi-modal feature extraction, it is possible to customise the model architecture to accommodate different inputs, such as MRI and X-rays. This customization allows for the inclusion of additional characteristics related to nodules, hence enhancing the accuracy of classification. One significant obstacle that we anticipate encountering is the need for a one-to-one correspondence between computed tomography (CT) scans and their corresponding magnetic resonance imaging (MRI) or X-ray images in order to validate findings.

Acknowledgments

We extend our heartfelt gratitude to Professor Dr. Sajimon Abraham, my esteemed research supervisor, whose patient guidance, unwavering encouragement, and invaluable critiques have been the cornerstone of this research endeavor. His dedication to our growth and development as researchers has been instrumental in shaping this work.

Furthermore, we wish to convey our sincere appreciation to the National Cancer Institute and the Foundation for the National Institutes of Health for their pivotal role in creating and maintaining the LIDC-IDRI database, which has served as an indispensable resource for our study. Their commitment to facilitating access to this invaluable database has significantly enriched the scope and depth of our research.

References

[1] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, 2012, pp. 1097–1105.

[2] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, J. Liang, Convolutional neural networks for medical image analysis: Fine tuning or full training?, IEEE Transactions on Medical Imaging 35 (5) (2016) 1299–1312.

[3] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, J. Dean, A guide to deep learning in healthcare, Nature medicine 25 (1) (2019) 24–29.

[4] W. Stewart, B. and Wild, C.P. (eds.), International Agency for Research on Cancer, "World Cancer Report 2014," 2014.

[5] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," CA. Cancer J. Clin., vol. 68, no. 1, pp. 7–30, Jan. 2018.

[6] A. McWilliams et al., "Probability of Cancer in Pulmonary Nodules Detected on First Screening CT," N. Engl. J. Med., vol. 369, no. 10, pp. 910–919, Sep. 2013.

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Adv. Neural Inf. Process. Syst., pp. 1–9, 2012.

[8] W. Shen, M. Zhou, F. Yang, D. Yu, D. Dong, C. Yang, Y. Zang, J. Tian, Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification, Pattern Recognition 61 (2017) 663–673.

[9] W. Zhu, C. Liu, W. Fan, X. Xie, Deeplung: Deep 3d dual path nets for automated pulmonary nodule detection and classification, in: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE, 2018, pp. 673–681.

[10] Z. Fan, S. Yang, C. Weidong, L. Min-Zhao, Z. Yun, H. Heng, S. Shimin, M. J. Fulham, D. D. Feng, Lung nodule classification with multilevel patch-based context analysis, IEEE Trans Biomed Eng 61 (4) (2014) 1155–1166

[11] S. M. B. Netto, A. C. Silva, R. A. Nunes, M. Gattass, Automatic segmentation of lung nodules with growing neural gas and support vector machine, Computers in Biology and Medicine 42 (11) (2012) 1110–1121

[12] W. Sun, B. Zheng, and W. Qian, "Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis," Comput. Biol. Med., vol. 89, no. January, pp. 530–539, 2017.

[13] S. Hussein, K. Cao, Q. Song, and U. Bagci, "Risk stratification of lung nodules using 3D CNN-based multi-task learning," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 10265 LNCS, pp. 249–260, 2017.

[14] A. Nibali, Z. He, and D. Wollersheim, "Pulmonary nodule classification with deep residual networks," Int. J. Comput. Assist. Radiol. Surg., vol. 12, no. 10, pp. 1799–1808, 2017.

[15] X. Liu, F. Hou, H. Qin, and A. Hao, "Multi-view multi-scale CNNs for lung nodule type classification from CT images," Pattern Recognit., vol. 77, pp. 262–275, 2018.

[16] Al-Shabi M, Lan BL, Chan WY, Ng KH, Tan M. Lung nodule classification using deep Local-Global networks. Int J Comput Assist Radiol Surg. 2019 Oct;14(10):1815-1819. doi: 10.1007/s11548-019-01981-7. Epub 2019 Apr 24. PMID: 31020576.

[17] S. G. Armato et al., "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, 2011.

[18] Li W., Cao P., Zhao D., Wang J. Pulmonary nodule classification with deep convolutional neural networks on computed tomography images. Computational and Mathematical Methods in Medicine. 2016;2016:7.

[19] Soni M, Singh AK, Babu KS, Kumar S, Kumar A, Singh S. Convolutional neural network based CT scan classification method for COVID-19 test validation. Smart Health (Amst). 2022 Sep;25:100296. doi: 10.1016/j.smhl.2022.100296. Epub 2022 Jun 11. PMID: 35722028; PMCID: PMC9188200.

[20] Khalifa, Nour Eldeen, Mohamed Loey, and Seyedali Mirjalili. "A comprehensive survey of recent trends in deep learning for digital images augmentation." Artificial Intelligence Review (2022): 1-27.

[21] Tran GS, Nghiem TP, Nguyen VT, Luong CM, Burie JC. Improving Accuracy of Lung Nodule Classification Using Deep Learning with Focal Loss. J Healthc Eng. 2019 Feb 4;2019:5156416. doi: 10.1155/2019/5156416. PMID: 30863524; PMCID: PMC6378763.

[22] X. Xu, C. Wang, J. Guo, Y. Gan, J. Wang, H. Bai, L. Zhang, W. Li, Z. Yi, MSCS-DeepLN: Evaluating lung nodule malignancy using multi-scale cost-sensitive neural networks, Med. Image Anal. 65 (2020) 101772. https://doi.org/10.1016/j.media.2020.101772.

[23] A. Paszke et al., "Automatic differentiation in PyTorch.".

[24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[25] M. Al-Shabi, H. K. Lee and M. Tan, "Gated-Dilated Networks for Lung Nodule Classification in CT Scans," in IEEE Access, vol. 7, pp. 178827-178838, 2019, doi: 10.1109/ACCESS.2019.2958663.

[26] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," Dec. 2014

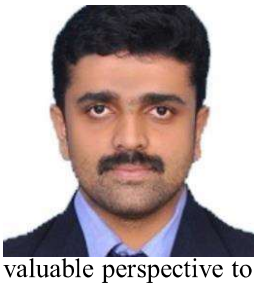
[27] Zhao, W., Jiang, W. and Qiu, X. Deep learning for COVID-19 detection based on CT images. Sci Rep 11, 14353 (2021). https://doi.org/10.1038/s41598-021-93832-2

[28] K. He, X. Zhang, S. Ren and J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015, pp. 1026-1034, doi: 10.1109/ICCV.2015.123.

Sijo Thomas serves as a research scholar affiliated with the School of Computer Sciences at Mahatma Gandhi University in Kerala, India. In addition to his academic pursuits, he assumes the role of a consultant software architect.His software solutions are utilised by prominent state universities, private universities, and autonomous colleges in India. Additionally he has academic experience of more than 5 years as faculty of science in colleges. Mr.Sijo Thomas holds a Master's degree in computer applications from Mahatma Gandhi University in Kerala, and he has also achieved a Master of Philosophy degree from Bharathidasan University in Tamil Nadu, India.

Dr.Sajimon Abraham earned his Ph.D. in Computer Science with a specialization in Spatio-Temporal Data Mining from Mahatma Gandhi University in Kerala, India. Starting in 2016, he has been serving as a research supervisor in the research department of the School of Computer Science. Additionally, Dr. Abraham holds the prestigious role of Honorary Director at the University Centre for International Co-operation at Mahatma Gandhi University, Kerala. Furthermore, he serves as a professor of computer and IT at the School of Management and Business Studies at Mahatma Gandhi University, Kerala. His distinguished academic career boasts an impressive total of 92 publications in the field of computer science and engineering.

Praveen Kumar holds the position of a research scholar at the School of Computer Sciences, Mahatma Gandhi University, located in Kerala, India. Additionally, he serves as an associate professor at a Government Aided College, boasting an impressive 24 years of teaching experience in undergraduate programs and 11 years in postgraduate programs. His academic interests primarily revolve around Spatio-Temporal Data Mining and the application of Artificial Intelligence in the realm of Human Rights. Sijo has made notable contributions to his field with three publications in international journals and one paper presented at an internal conference organized by the Indian Institute of Management (IIM).



Faizal Basheer is currently a research scholar affiliated with the School of Computer Sciences at Mahatma Gandhi University in Kerala, India. In addition to his academic pursuits, he actively serves as a consultant specializing in machine learning and artificial intelligence. Faizal brings a valuable perspective to his work with five years of experience as a faculty member in the field of computer science. His educational background includes a Master's degree in computer science and engineering, which he earned from Mahatma Gandhi University in India.

Jithinmary Raphel possessing a decade's worth of expertise in Information Technology (IT) as a quality engineer, is a committed professional deeply enthusiastic about the realms of Deep Learning and Artificial Intelligence (AI). Jithinmary's academic foundation was established at LBS College of Engineering in Kerala, India, where she earned a Bachelor of Technology (B.Tech) degree in Computer Science and Engineering.