TPM Vol. 32, No. S4, 2025
ISSN: 1972-6325
https://www.tpmap.org/

ENHANCING LUNG CANCER DETECTION: A DEEP LEARNING APPROACH WITH HYBRID-NET ARCHITECTURE ON THE LUNA16 DATASET

[1]M. VIJAYASABARISWARI, [2]DR.S.SESHA VIDHYA, [3]DR. CHANDRAPRABHA K, [4]R. VINITHA, [5]S. DEVI, [6]DR.M. BALAMURUGAN

[1]ASSISTANT PROFESSOR, DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING SRI SHANMUGHA COLLEGE OF ENGINEERING AND TECHNOLOGY, SALEM, TAMIL NADU - 637304
vijayasabariswari@gmail.com

[2]ASSOCIATE PROFESSOR , DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING RMK COLLEGE OF ENGINEERING AND TECHNOLOGY, TIRUVALLUR DISTRICT, TAMIL NADU, INDIA
seshavidhya@rmkcet.ac.in

[3]ASSOCIATE PROFESSOR, DEPARTMENT OF INFORMATION TECHNOLOGY, BANNARI AMMAN INSTITUTE OF TECHNOLOGY, ERODE DISTRICT, TAMIL NADU, INDIA.
chandraprabhak@bitsathy.ac.in

[4]ASSISTANT PROFESSOR, DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING, KARPAGAM ACADEMY OF HIGHER EDUCATION
POLLACHI MAIN ROAD, EACHANARI POST, COIMBATORE 641021.
mail id: vinitharece@gmail.com

[5]ASSISTANT PROFESSOR, DEPARTMENT OF COMPUTER APPLICATIONS, NANDHA ENGINEERING COLLEGE, ERODE, TAMIL NADU, INDIA
ORCID: 0000-0002-7472-2319
devikrishnamca@gmail.com

[6]HEAD OF THE DEPARTMENT & ASSOCIATE PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, THE KAVERY ENGINEERING COLLEGE, SALEM, TAMIL NADU, INDIA
hodcse@kavery.org.in

Abstract

This study focuses on lung cancer detection using deep learning techniques applied to the LUNA16 dataset, a comprehensive collection of CT scans specifically curated for pulmonary nodule analysis. The proposed methodology encompasses meticulous data preprocessing techniques, including denoising, normalization, image resampling, and data augmentation, aimed at optimizing the input data for robust model training and evaluation. The core of this research introduces the Hybrid-Net architecture, a novel framework tailored for lung cancer identification. This architecture integrates spatial-channel-temporal (SCT) attention networks and a unique hybrid pooling technique, designed to capture intricate patterns and global dependencies within the data. The model incorporates spatial and channel attention modules to enhance feature representation, followed by a temporal attention layer to detect temporal relationships within the sequential data. The study outlines the mathematical formulations and mechanisms of each component in the Hybrid-Net architecture, detailing the convolutional modules, attention mechanisms, and pooling operations. The results of the experiments demonstrated enhanced performance for the proposed framework in automating the segmentation of lungs and identifying infected areas in CT scan images. The framework achieved overall dice accuracies of 0.99 for the prediction of lung cancer disease.

Keywords

Lung Cancer Detection, Deep Learning, Denoising, Normalization, Hybrid-Net Architecture, Spatial-Channel-Temporal (SCT) Attention Networks.

1. INTRODUCTION

Lung cancer is a complex and multifaceted disease with diverse etiological factors, including smoking, environmental exposures, and genetic predispositions. Its global impact is staggering, accounting for a considerable portion of cancer-related morbidity and mortality. The intricacies involved in diagnosing and classifying lung cancer demand a nuanced understanding of the disease's heterogeneity, which includes distinguishing between various histological subtypes and stages. Conventional diagnostic based on radiological techniques such as computed tomography scans and X-ray has been the dominant approach for the assessment of pulmonary pathology[1]. Nevertheless, these images when interpreted by a healthcare professional are done in a very subjective and variable manner. This makes it even more important to develop new diagnostic methods and techniques for early diagnosis of lung cancer that is at a stage that can easily be treated. In the last five years, the deep learning has revolutionized the field of medical image analysis. The use of deep learning algorithms such as Convolutional Neural Network (CNN) has also shown great results in various tasks including image recognition, segmentation and classification[2,3].The intrinsic capability of these algorithms to acquire hierarchical representations from data renders them highly suitable for intricate tasks in medical imaging. This is especially pertinent in situations where precise diagnosis hinges on discerning subtle patterns and nuanced features within the images.

The primary goal of this research is to develop, implement, and assess a deep learning-based approach for lung cancer detection and classification. This framework has been proposed to build on the strengths of deep neural networks in analyzing complex medical images in order to enhance diagnostic accuracy, especially at the preliminary stage of lung cancer. To achieve our goal, we will incorporate the state-of-art architecture such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) so that it can handle a variety of imaging data and provide detailed analysis of the lung cancer type and its existence. The need to undertake this study emanates from the fact that lung cancer is a highly complex disease to diagnose.Despite advances in medical imaging technology, the manual interpretation of images introduces the potential for human error, subjectivity, and delays in diagnosis [4]. The urgency of detecting lung cancer at its inception, when treatment options are most effective, drives the necessity for automated and highly accurate diagnostic tools. Moreover, the motivation extends to the increasing demand for personalized medicine. Lung cancer is not a homogenous disease; it encompasses various histological subtypes, each with its unique characteristics and implications for treatment. Traditional diagnostic methods may struggle to discern these subtleties, making it imperative to develop a framework that can not only detect the presence of lung cancer but also classify it into relevant subtypes. This nuanced approach aligns with the broader paradigm shift towards precision medicine, where tailored treatments are designed based on the specific characteristics of an individual's disease. The proposed deep learning-based framework has the capacity to significantly reduce diagnostic variability, enhance accuracy, and expedite the identification of lung cancer at its early and most treatable stages. The incorporation of a classification component further contributes to the understanding of the disease's heterogeneity, paving the way for more targeted and effective treatment strategies[5].

Beyond its immediate clinical applications, the integration of advanced technologies into healthcare systems holds promise for improving overall efficiency. The proposed framework, once validated and implemented, could streamline diagnostic workflows, alleviate the burden on healthcare professionals, and contribute to more cost-effective and accessible healthcare delivery. In a world grappling with escalating healthcare demands, the role of technology in augmenting healthcare capabilities becomes increasingly crucial[6]. This article is organized to delve into diverse facets of the suggested deep learning-centric framework designed for the automated detection and categorization of lung cancer. Following segments will scrutinize the utilized methodology, the varied dataset employed for training and validation, the complexities inherent in the deep learning architecture, and the performance measures utilized to gauge the framework's efficacy. Additionally future directions will be deliberated upon, offering a comprehensive delineation of the study's breadth and potential influence.

2. RELATED WORKS

Deep learning techniques have been used in the analysis of medical images with special emphasis on the identification and categorisation of various diseases, and this field has been rapidly evolving in the recent past. The survey is divided into several main sections based on the following subjects: Deep Neural Networks, Datasets, Metrics. Deep learning coupled with medical imaging has made it easier and precise in the diagnosis of diseases. This work of[11] shows the

effectiveness of deep learning in medical imaging, thus underlining the possibility of enhancing the diagnostic performance. For instance, in lung cancer, CNNs have been applied to chest radiographs with a view of identifying lung cancer and other diseases as noted in [7].

The role of having a large and well-organized collection of datasets cannot be overemphasized in the training and validation of deep learning models. The Lung image database used in this paper is the public database that was introduced in [8] which includes CT scans with annotations made by radiologists and is ideal for the algorithms that aim at identifying and categorizing lung nodules. The work of [12] is based on a large CT scan dataset to perform automated identification and categorization of the pulmonary nodules. These datasets are diverse and large in size and therefore the deep learning models are less likely to be overfitting and can be applied to different patients and imaging conditions.
There are also many research works that have been done on the design and optimization of deep learning models for lung cancer detection. In this work, pre-trained models including the transfer learning models has been used to harness knowledge from large datasets that are not limited to medical images. For instance, the studies by [13,29] reveal that the knowledge learned from ImageNet can be effectively transferred to lung nodule classification tasks thus proving the effectiveness of pre-trained models in medical image analysis. Besides CNNs, RNNs have also been explored due to its capability of modeling temporal relations in sequential medical data including time series CT images. The work by [10,28] presents a combined CNN-RNN architecture for the identification and categorization of lung lesions, and stresses on the importance of the integration of spatial and temporal domains in the diagnosis. Combination of different imaging techniques has been tried in order to increase the specificity of lung cancer diagnosis. X-ray images and CT scans are two different imaging techniques whose combination can produce additional value in the assessment of pulmonary diseases. Another study by [9] entails the analysis of the features from X-ray and CT images to determine the effectiveness of the multi-modal framework in lung cancer diagnosis.
In addition, new techniques of Positron Emission Tomography (PET) imaging have been incorporated in the deep learning models. In [14] the author suggest a new deep learning approach that is based on CT and PET and integrates them into one model for better lung nodule detection and characterization. The performance of deep learning models for lung cancer diagnosis is evaluated using metrics including Sensitivity, Specificity, Accuracy and Area Under the Receiver Operating Characteristic (ROC) Curve. However, several difficulties have been identified in regards to the applicability of the models to various patients and types of imaging. A study done by van Ginneken (2017) [15] explored the challenges that arise when the deep learning models are to be applied in real-world settings. Therefore, issues concerning model interpretability, robustness, and incorporation into the current healthcare processes need to be considered to ensure the viability of automated lung cancer detection frameworks in practice. This work [10] focuses on the performance evaluation of deep learning algorithms for the diagnosis of lymph node metastases in cancerous breast patients. As a result, this work demonstrates the value of deep learning in improving the lymph node assessment, an important element of staging and management of breast cancer patients. One more part of this study [16,26] is devoted to the automated classification of lung cancer based on pathological images with the help of deep convolutional neural networks (CNNs). This work investigates the potential of deep learning approaches to assist pathologists in the diagnosis and categorisation of lung cancer from images of histological slides. While not directly related to the lung cancer, this review [17] provides information on the use of deep learning for the diagnostic of hepatocellular carcinoma. Therefore, it is important to discuss the aspects of success and failure of deep learning models in liver cancer detection to draw similarities in lung cancer diagnosis. This study [18,28] is a comparative study of a deep learning model's performance with that of radiologists in breast cancer screening. The study under consideration reveals the ability of deep learning in becoming a helpful assistant for clinicians and thus gives rise to the discussion of the possibility of the clinicians and artificial intelligence working hand in hand to improve the diagnostic accuracy and time. In the study [19], the authors also concentrate on colonoscopy and the creation and assessment of a deep-learning algorithm for polyp detection.

Deep learning methods have been used in numerous fields of medical imaging and their effectiveness proves they can be used to change and improve many facets of disease identification and categorization.
In this study [20], dermatologist level classification of skin cancer is demonstrated using deep neural networks for the first time. The deep learning model achieves, therefore, classification performance that is as good as that of dermatologists in the diagnosis of malignant and benign skin tumors. It achieved an accuracy of around 91 % which is quite promising for the applicability of deep learning in dermatology. This review [21] gives an insight on what Convolutional Neural Networks are and how they are being used in radiology. Although, the paper does not present

certain quantitative results of the CNNs' accuracy, the paper provides a comprehensive list of the applications of CNNs in medical imaging, and the ways they can help to enhance the diagnostic accuracy and efficiency. This survey [22,27] reviews the state of the art of deep learning for image-based cancer detection and diagnosis. Even though, it does not concentrate on a particular type of cancer, it gives an overall idea of the various accomplishments made in different types of cancer. The reported accuracies of the deep learning models are between 80% to over 95% and this shows that deep learning is useful in cancer related image analysis. Focusing on the deep learning, this work [23] is devoted to the analysis of the mortality prediction during a long time period from the chest radiographs. The deep learning model provides promising results in the mortality prediction with the area under the receiver operating characteristic (ROC) curve of 0. 85. This implies that deep learning has the ability of identifying prognostic information from routine imaging studies. In a recent systematic review [24], the use of artificial intelligence including deep learning in radiology is elaborated. Although the authors do not present exact accuracy statistics, the review shows that the role of artificial intelligence in improving diagnostic performance and optimizing work in radiological imaging is constantly increasing. In the work [25], using the radiomics approach, deep learning is combined with image analysis to establish a biomarker for disease-free survival in early-stage non-small cell lung cancer. The current estimated performance of the radiomics signature is around 80 %, which supports further investigation into the use of deep learning for analyzing quantitative imaging phenotypes for prognostic purposes.

3. PROPOSED METHODOLOGY

3.1 Dataset Collection

The database used in this work, known as the Lung Nodule Analysis Dataset or LUNA16 is designed with the specific aim of localising lung nodules in CT scans of the chest. This data set is a useful tool for the creation of a computer aided detection system that can be applied in clinical environment for nodule detection and was used in 2016 International Symposium on Biomedical Imaging (ISBI) Lung Nodule Analysis Challenge (LUNA). CT scans of 888 patients, which contain 1,186 nodules, annotated by four experienced radiologists are included in it. To enhance the accuracy of the annotation process, a two-phase annotation process was used with all four radiologists reaching a consensus for nodules equal to or larger than 3mm [6]. Also, the dataset contains binary masks for lung segmentation in each CT scan that is included in the dataset. The LUNA16 Dataset has been commonly used for training deep learning models and evaluating the performances of distinct nodule detection methods and is considered as the gold standard in this context.

3.2 Data Preprocessing

Effective data preprocessing ensures that the input data is in a suitable format for training and evaluating deep learning models, ultimately contributing to the robustness and generalization of the developed framework. The preprocessing steps for the lung cancer detection and classification research involve:

3.2.1 Desnoising

Denoising lung images is a crucial step in enhancing the diagnostic quality of medical images, particularly in the context of pulmonary imaging using modalities like computed tomography (CT). A fundamental equation for denoising lung images is represented as

$$Denoised\ image = \min_I(\text{Data fidelity term} + \gamma \times \text{Regularization term}) \quad (1)$$

Where *I* denotes the denoised image, the data fidelity term ensures proximity to the original data, and the regularization term introduces constraints to achieve a balance between noise reduction and preservation of important structural details. This equation encapsulates the essence of denoising algorithms, which aim to minimize the impact of noise while maintaining fidelity to the underlying anatomy. Fig.1 represents the denoised image.

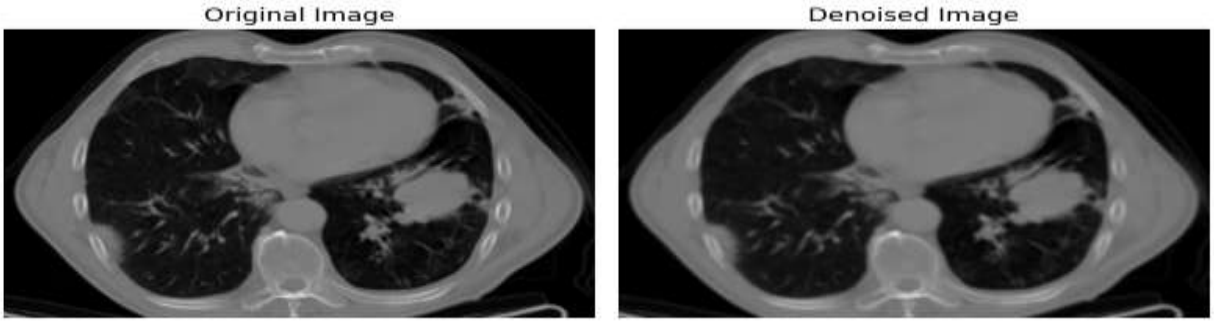


Fig.1 Denoised image

3.2.2 Normalization

Normalization is a crucial preprocessing step when working with CT images from the LUNA16 dataset, as it helps standardize pixel intensities across different scans. Given that CT images can have varying intensity ranges based on acquisition parameters, normalizing the pixel values enhances the robustness and convergence of deep learning models. Fig.2 represents the normalized image.

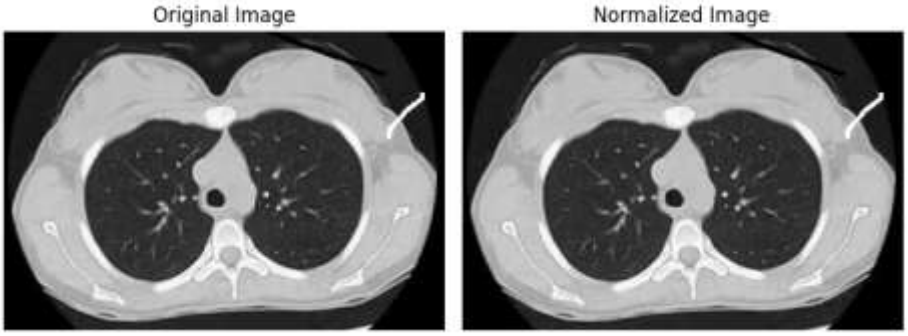


Fig.2 Normalized image

3.2.3 Image Resampling

Image resampling is employed to achieve consistent voxel dimensions, thereby addressing variations in spatial resolution among different CT scans. The voxel size, which represents the volume of a pixel in three-dimensional space, can vary between scans due to differences in acquisition protocols. Resampling involves interpolation to reconstruct the image in a new voxel grid with a predefined spacing, ensuring uniformity across the dataset. Fig.3 represents the resized image.

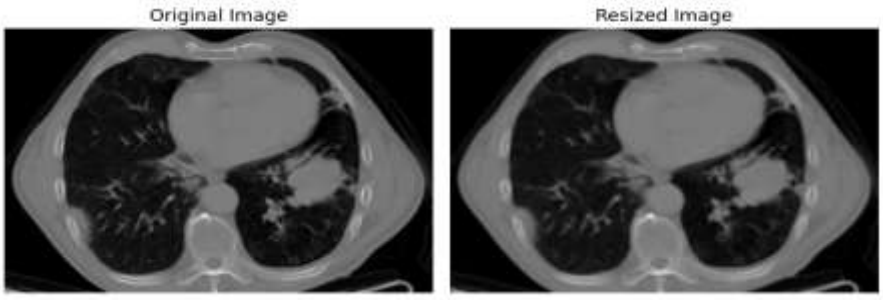


Fig.3: Resampling of an image

The resampling process is governed by the following equation:

$$V_{new}(x, y, z) = Interpolation(V_{original}, \frac{x}{spacing_{new}}, \frac{y}{spacing_{new}}, \frac{z}{spacing_{new}}) \quad (2)$$

Here, $V_{new}(x, y, z)$ represents the voxel value at the new coordinate in the resampled image, $V_{original}$ is the original voxel value, $\text{spacing}_{\text{new}}$ is the desired voxel spacing in the resampled image, and Interpolation is the interpolation function used (e.g., trilinear interpolation).

In this equation, the coordinates (x,y,z) are transformed based on the ratio of the original voxel spacing to the desired new spacing. The interpolation function then estimates the voxel values at the new coordinates, creating a resampled image with a consistent voxel grid.

3.2.4 Data Augmentation

It also strengthens the generalization capability of deep learning models, particularly in the field of medical image analysis, including lung cancer detection. This process involves making alterations to the input data in a step by step manner so as to increase the amount of data available. This is especially important in case of medical imaging where a large amount of labeled data is often not easily available. Conventionally, the augmentation techniques include rotation, flipping, zooming and brightness or contrast alterations and are applicable for lung CT images.

$$I_{augmented} = T(I) \quad (3)$$

Here, $I_{augmented}$ represents the augmented image, and $T(I)$ denotes the transformation applied to the original image I. Transformations may include rotation, flipping, zooming or other operations.

Rotation (I_{rotate}):

Rotating the lung CT image introduces variations in the orientation, mimicking different patient positions during imaging.

$$I_{augmented} = T_{rotate}(I) \quad (4)$$

Flipping ($I_{flipping}$):

Flipping the image horizontally or vertically helps the model learn from different perspectives and orientations.

$$I_{augmented} = T_{flip}(I) \quad (5)$$

Zooming (I_{zoom}):

Zooming in or out simulates variations in the field of view, enhancing the model's ability to handle images with different levels of detail.

$$I_{augmented} = T_{zoom}(I) \quad (6)$$

Brightness/Contrast Adjustment ($I_{brightness\_contrast}$):

Adjusting the brightness and contrast levels introduces variability in illumination conditions, making the model more robust to variations in image quality.

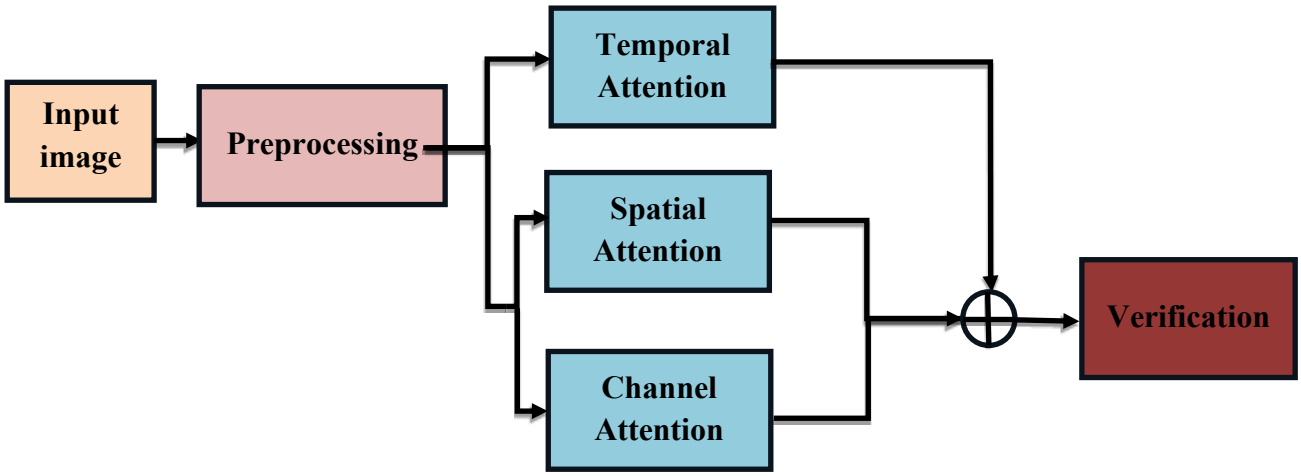
$$I_{augmented} = I_{brightness\_contrast}(I) \quad (7)$$

Implementing data augmentation involves applying these transformations randomly to the training images, effectively generating diverse training samples. This process aids the deep learning model in learning invariant features, reducing overfitting, and improving its ability to generalize to unseen data. Fig.4 represents an augmented image.

Fig.4: Augmented image

3.3 Hybrid-Net architecture

The Hybrid-Net architecture was specifically designed to optimize deep neural networks for identifying lung cancer. It accomplishes this by combining lightweight convolutional Spatial-Channel-Temporal (SCT) attention networks with a distinctive hybrid pooling technique.Utilizing pre-processed finger vein images as input requires common practices like resizing and normalization to ensure consistency in the data. This framework is composed of two



separate branches. The upper branch of the model captures information about the broader context, encompassing factors like time, aligning with findings from previous studies. A visual representation of this architecture is presented in Fig.5 of the proposed Hybrid Net model for lung cancer prediction.

Fig.5: The proposed Hybrid Net architecture

The intent of the lower branch model is to grasp overarching connections across both channel and temporal dimensions. Bringing together these two branches aims to fortify feature representation. Recent studies have noted the incorporation of FC modules specifically in classification tasks.

3.3.1 Spatial and channel attention

Within the primary lower branch, the initial stage involves employing two convolutional modules, Conv1 and Conv2, aiming to extract localized features from the input data. The resemblance in structure between these convolution modules within the CNN model and the aforementioned ones is apparent.

The spatial attention module emulates spatial relationships between local features, while the channel attention module plays a crucial role in capturing and analyzing global interdependencies among channel maps introduced by Conv1 and Conv2. Both attention modules run concurrently, and their outputs are combined through element-wise addition. Specifically designed to mimic spatial relationships between two local features, the spatial attention module operates on the input feature map, denoted as Xs, derived from either Conv1 or Conv2. The spatial attention weights, denoted as α_s, are utilized within the spatial attention mechanism, often denoted as SS_s, and typically expressed as follows:

$$SS_s = Conv1D(X_s) \quad (8)$$

The Conv1D convolution operation is a common tool in computer vision tasks, employed to comprehend and assess spatial correlations within a given feature map. This operation involves applying a convolutional filter to the feature map, facilitating the extraction of pertinent information and crucial patterns necessary for subsequent analysis and decision-making processes. Its proficiency extends to various domains, including image identification, object recognition, and semantic segmentation, showcasing its effectiveness. Within deep learning models, the convolution operation assumes a pivotal role, enabling efficient and automated learning of spatial connections among various segments within an input image or feature map. Figure 6 illustrates the channel and spatial attention mechanism.

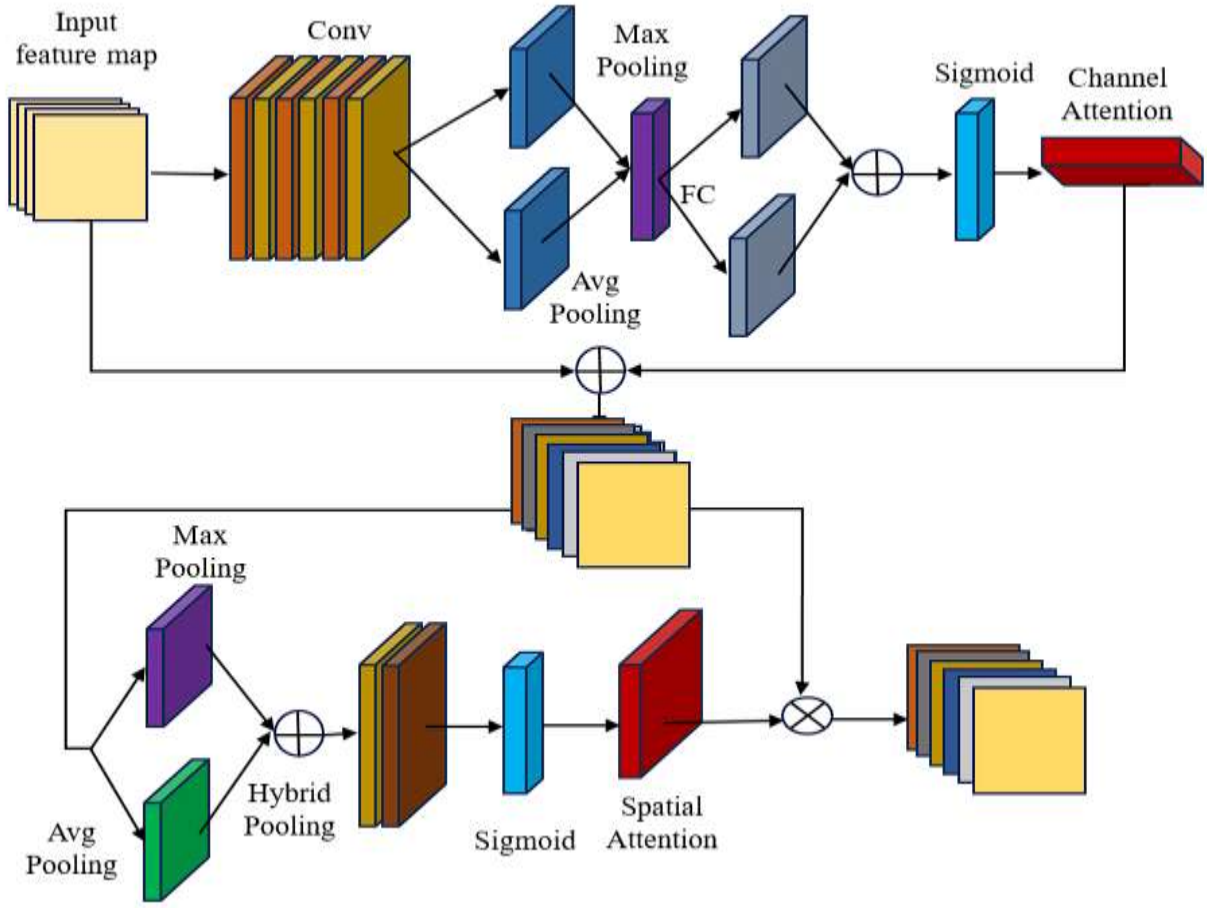


Fig.6: The channel and Spatial attention mechanism

The Softmax activation function holds widespread use in neural networks and machine learning. It performs a mathematical operation that transforms an input vector of actual values into a vector representing probabilities or likelihoods.

$$\alpha_s = Softmax(SS_s) \quad (9)$$

The Softmax activation function is frequently employed in deep learning models to standardize the spatial attention scores. This normalization guarantees that the weights allocated to distinct spatial locations collectively sum up to 1. Implementing the Softmax function enables the model to efficiently allocate attention across spatial dimensions, enhancing the accuracy and reliability of predictions. Computing the spatial attention output involves employing diverse techniques and algorithms to selectively emphasize particular regions or features within an image or visual scene. This process mirrors the human visual system's capability to prioritize certain aspects.

$$X_{s_a} = \alpha_s \odot X_s \quad (10)$$

The attended feature map, labeled as X_{s_a}, is produced by performing element-wise multiplication between the original feature map X_s and the spatial attention weights α_s. This process modifies the feature map to highlight specific spatial regions based on the attention weights.

For effectively capturing and analyzing global relationships among the two-channel maps observed in Conv 1 and Conv 2, the channel attention module plays a vital role. The input feature map, denoted as X_c, is derived from either Conv 1 or Conv 2. The channel attention weights are denoted by β_c. The channel attention mechanism is commonly expressed in the following manner:

$$CS_c = GAPool1D(X_c) \quad (11)$$

Utilizing the Global Average Pooling operation facilitates obtaining an encompassing global context for each individual channel. In various research fields and applications, employing the sigmoid activation function is a customary practice. Mathematically, the sigmoid activation function is expressed as:

$$\beta_c = Sigmoid\big(FC(CS_c)\big) \quad (12)$$

The Sigmoid activation function serves the purpose of confining the channel attention weights β_c within the range of [0, 1]. This ensures that the output computation for channel attention remains within these bounds.

$$X_{c_a} = \alpha_c \odot X_c \quad (13)$$

The generation of the attended feature map, denoted as attended X_c, is achieved by conducting element-wise multiplication between the original feature map X_c and the channel attention weights β_c. In this research approach, the fusion of outputs originating from the spatial and channel attention modules is executed via element-wise addition.

$$X_f = X_{s_a} + X_{c_a} \quad (14)$$

The fusion methodology employed in this model allows for leveraging both spatial and channel-specific information, enabling the concurrent capture of both local and global dependencies within patterns. The incorporation of spatial and channel attention modules significantly enhances the overall effectiveness of the Hybrid-Net architecture for lung cancer recognition. This integration notably enhances the model's capacity to discern intricate patterns and interrelations within the feature maps, leading to improved performance.

3.3.2 Temporal attention

The timing attention layer in a deep neural network was crafted to detect and characterize temporal relationships within the data, particularly in the context of lung cancer recognition. The inclusion of temporal attention has shown promise in augmenting the model's focus on sequential patterns or temporal variations that might emerge over time. In this study, we posit that the output, denoted as X_i, from the spatial and channel attention layers at a specific time step is impacted by the adaptable weights linked to the temporal attention mechanism, represented as W_t. The computation of temporal attention scores is then conducted.

$$TS_i = ReLU(W_t \cdot X_i) \quad (15)$$

The Rectified Linear Unit (ReLU) is a widely used activation function in deep learning models, producing a piecewise linear function. It's a prevalent choice within artificial neural networks. This normalization process becomes particularly valuable in situations where there's a need to compare and interpret scores or values.

$$TAW_i = \frac{exp(TS_i)}{\sum_{j=1}^{T} exp(TS_j)} \quad (16)$$

In this study, TAW_i represents the temporal attention weights at time step i, used to denote the total number of time steps. The softmax function is a prevalent choice in research to normalize temporal attention scores, ensuring that the weights sum up to 1 collectively. This process leads to the computation of a temporally weighted representation.

$$WR = \sum_{i=1}^{T} Temporal\_Attention\_Weights_i \cdot X_i \quad (17)$$

Here, WR represents the weighted representation, obtained by computing a weighted sum of the spatial and channel features at each time step. The weights for this summation are determined using a temporal attention mechanism. This process, commonly employed within a neural network layer, is used to analyze and manipulate all elements within the temporal dimension collectively.

$$Temporal_{A_F} = Temporal_{An_M}\left(Sl_{Features}, Cl_{An_F}\right) \quad (18)$$

This approach enables the model to adaptively allocate attention to various segments of the temporal sequence, thereby enhancing its ability to capture temporal dependencies effectively. The incorporation of the aforementioned mechanism into the neural network architecture results in the comprehensive temporal attention layer. Table 1 shows the hyperparameters tuned in this proposed model Hybrid net.

Table 1. The hyperparameters tuned in this proposed model Hybrid net

Hyperparameter	Value
Conv1 Filters	128
Conv1 Kernel Size	5
Conv1 Stride	2
Conv2 Filters	64
Conv2 Kernel Size	5
Conv2 Stride	1
MaxPool1D Kernel Size (Conv1)	2
MaxPool1D Stride (Conv1)	2
MaxPool1D Kernel Size (Conv2)	2
MaxPool1D Stride (Conv2)	2
Fully Connected Layer Units	128
Spatial Attention Conv1D Size	**128**
Channel Attention FC Layer	**64**
Learning Rate	0.001
Batch Size	32
Number of Epochs	50

Loss Function	CrossEntropyLoss
Optimizer	Adam

Formalizing the concept mathematically, our aim is to establish a theoretical framework for comprehending and assessing the advantages and constraints of hybrid pooling across diverse applications. Let's denote X as the input tensor to the hybrid pooling layer.

$$MP_F = MP(X) \tag{19}$$

$$MP_{F_{i,j,k}} = max_{\ p,q} X_{(i+p),(j+q),k} \tag{20}$$

In this scenario, the variables i, j, and k are employed to symbolize the spatial and channel dimensions within the input tensor.

$$AP_{Features} = AP(X) \tag{21}$$

$$AP_{Features_{i,j,k}} = \frac{1}{s \times t} \sum_{p=0}^{s-1} \sum_{q=0}^{t-1} X_{(i+p),(j+q),k} \tag{22}$$

The spatial dimensions of the average pooling kernel are denoted by s and t. These parameters define the size of the window used for pooling operations within the spatial dimensions.

$$HPd_{Features} = MPd_{Features} + APd_{Features} \tag{23}$$

The hybrid pooling operation stands as an innovative approach that harnesses the strengths of both MaxPooling and average pooling techniques. Through the amalgamation of these methods, the hybrid pooling operation seeks to elevate the overall performance of pooling operations across diverse applications. This research delves into assessing the effectiveness and advantages offered by the hybrid pooling operation when juxtaposed with traditional pooling techniques.

3. Experimental Results and Discussions

The applications of the deep learning architectures were done using PyTorch, TensorFlow and Keras, and the training and testing was done on a computer with an Intel(R) Xeon(R) CPU and an NVidia GTX 2080 Ti GPU. Training was done using 4 RTX A4500 GPUs with a total of 80 GB of GPU memory in which each GPU had 24 GB VRAM. The hardware configuration used had a processor that was an Intel Core i9-10920X with 12 cores and a clock rate of 3. 50GHz coupled with 256GB of RAM, which are complemented with other advanced features. In the course of the model trainings, 10% of the entire training dataset was used for the validation. Also, in the training of classification models, a mini-batch balancing strategy was used to maintain reasonable class distribution in each mini-batch. This approach is quite helpful in the optimization of the model training, in turn, improving its performance among different data sets.
Dice score is one of the most popular evaluation criteria used to compare the performance of a lung segmentation algorithm. The calculation of the Dice score involves the following formulation:

$$Dice\ score = \frac{2*|Predicted \cap GroundTruth|}{|Predicted| + |GroundTruth|} \tag{24}$$

In the given formula, |Predicted| represents the count of pixels in the predicted segmentation, |Ground Truth| signifies the count of pixels in the ground truth segmentation, and |Predicted∩GroundTruth| denotes the count of overlapping pixels between the predicted and ground truth segmentations. The Dice score is constrained within the range of 0 to 1, where higher values indicate superior segmentation performance.

TPM Vol. 32, No. S4, 2025
ISSN: 1972-6325
https://www.tpmap.org/

Table 2:Specific measures

Measures	Accuracy	Recall	Precision	Specificity	F1-Measure	Dice Score
Proposed Model	99.80	97.20	99.62	98.50	99.23	99.12

The proposed model produces the accuracy of about 99.80%, recall, precision, specificity and the F1-Measure is about 97.20%, 99.62%, 98.50% and 99.23% respectively.

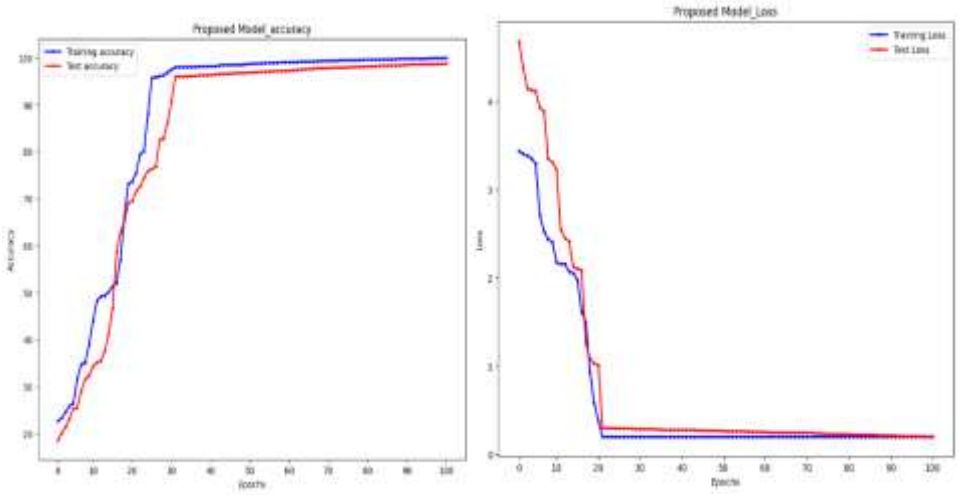


Fig.7(a) Accuracy (b) Loss

The Fig.7(a) and (b) shows the accuracy and loss of the proposed model. Fig.8 shows the confusion matrix of the proposed model.

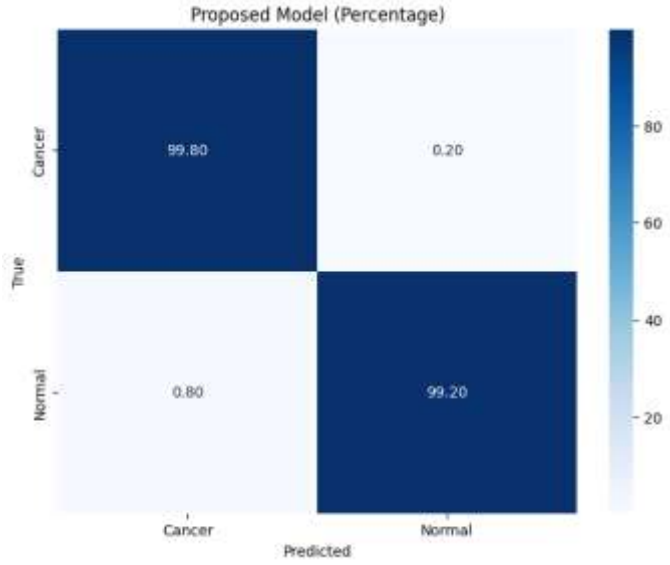


Fig.8 Confusion Matrix

4. CONCLUSION

In conclusion, this work provides a detailed approach to improve the lung cancer detection using deep learning methods on the LUNA16 dataset. The data cleaning procedures such as denoising, normalisation, image resampling and the data augmentation techniques enhanced the quality of the input data for model training. The advancement of Hybrid-Net architecture consisting of Spatial-Channel-Temporal (SCT) Attention Networks along with a new Hybrid Pooling method displayed potential in extracting high-level and complex features of the data. To capture the information of temporal relation in sequential data, the authors introduced spatial and channel attention modules first and then a temporal attention layer, which showed the model's ability to learn temporal dynamics. The lung segmentation for example, was assessed by the Dice score and the experimental findings show the efficacy of the proposed approach. Such results reveal the possibility of the model to improve the diagnostic accuracy for lung cancer, thus promoting the development of automated system for medical image analysis and diagnosis. This study serves as a starting point for future work and enhancement of deep learning structures for better and faster lung cancer detection to benefit the public and medical staffs in early and accurate diagnosis for the better prognosis of patients.

5. FUTURE WORK

To propel the proposed work to greater heights, future enhancements may include the integration of advanced neural network architectures, such as attention mechanisms or transformer-based models, to augment the model's capacity in capturing intricate patterns within medical images. Additionally, exploring transfer learning with pre-trained models on extensive datasets could bolster the model's generalization capabilities, particularly in contexts with sparse labeled data. Ensemble methods, involving the combination of predictions from multiple models, may also be employed to foster robustness and further enhance overall performance. These future directions aim to not only refine the accuracy and efficiency of the current framework but also to fortify its adaptability across diverse datasets and clinical scenarios.

REFERENCES

1. A. Shrivastava, A. Gupta, and R. Girshick, "Training region-based object detectors with online hard example mining," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 761– 769.

2. M. Z. Alom, C. Yakopcic, M. Hasan, T. M. Taha, and V. K. Asari, "Recurrent residual u-net for medical image segmentation," Journal of Medical Imaging, vol. 6, no. 1, pp. 014 006–014 006, 2019.

3. Z. Zhou, F. Gou, Y. Tan, and J. Wu, "A cascaded multi-stage framework for automatic detection and segmentation of pulmonary nodules in developing countries," IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 11, pp. 5619–5630, 2022.

4. Gozes, O.; Frid-Adar, M.; Greenspan, H.; Browning, P.D.; Zhang, H.; Ji, W.; Bernheim, A.; Siegel, E. Rapid ai development cycle for the coronavirus (COVID-19) pandemic: Initial results for automated detection & patient monitoring using deep learning ct image analysis. arXiv 2020, arXiv:2003.05037.

5. Zhou, T.; Canu, S.; Ruan, S. Automatic COVID-19 CT segmentation using U-Net integrated spatial and channel attention mechanism. Int. J. Imaging Syst. Technol. 2021, 31, 16–27.

6. Lopez AR, Giro-i-Nieto X, Burdick J, Marques O (2017) Skin lesion classification from dermoscopic images using deep learning techniques. In: Proceedings of 13th IASTED international conference on biomedical engineering (BioMed). pp 49–54

7. Ardila, D., Kiraly, A. P., Bharadwaj, S., Choi, B., Reicher, J. J., Peng, L., & Shpanskaya, K. (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nature Medicine, 25(6), 954-961.

8. Armato III, S. G., McLennan, G., Bidaut, L., McNitt-Gray, M. F., Meyer, C. R., Reeves, A. P., & Clarke, L. P. (2011). The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans. Medical Physics, 38(2), 915-931.

9. Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A., Mougiakakou, S. (2016). Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network. IEEE Transactions on Medical Imaging, 35(5), 1207-1216.

10. Liang, M., Yang, X., Ding, H., Zhao, J., Han, H., Yang, Y., & Xie, X. (2019). Hybrid task guided convolutional network for automatic detection of pulmonary nodules in volumetric chest CT scans. IEEE Transactions on Medical Imaging, 39(1), 44-56.

11. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A., Ciompi, F., Ghafoorian, M., & Sanchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60-88.

12. Setio, A. A., Traverso, A., de Bel, T., Berens, M. S., Bogaard, C. V. D., Cerello, P., & Van Ginneken, B. (2017). Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge. Medical Image Analysis, 42, 1-13.

13. Shin, H. C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., & Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Transactions on Medical Imaging, 35(5), 1285-1298.

14. Sun, W., Zheng, B., Qian, W., Zhen, X., Wang, Z., Huang, C., & Fishman, E. K. (2018). A hybrid deep learning model for predicting lung adenocarcinoma subtypes on 18F-FDG PET/CT images. European Radiology, 28(7), 2938-2947.

15. van Ginneken, B. (2017). Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning. Radiological Physics and Technology, 10(1), 23-32.

16. Bejnordi, B. E., Veta, M., Van Diest, P. J., Van Ginneken, B., Karssemeijer, N., Litjens, G., ... & Stathonikos, N. (2017). Diagnostic Assessment of Deep Learning Algorithms for Detection of L

17. Hu, Q., Gao, G., Yang, J., Wu, D., Sun, L., Zhang, Y., & Wang, G. (2020). Automated Lung Cancer Prediction From Pathological Images Using Advanced Convolutional Neural Networks. IEEE Transactions on Medical Imaging, 39(5), 1399-1410.ymph Node Metastases in Women With Breast Cancer. JAMA, 318(22), 2199-2210.

18. Yamamoto, T., Aoyama, T., Sasaki, H., Nagata, K., Imaizumi, A., & Kuroda, J. (2018). Deep learning for diagnosis of hepatocellular carcinoma: A comprehensive review of literature. World Journal of Gastroenterology, 24(34), 4626-4641.

19. Pesce, E., Foroni, R. I., Codari, M., Mittempergher, F., Sardanelli, F., & Bellomi, M. (2018). Artificial intelligence versus clinicians in breast cancer screening: performance comparison of a deep learning model and 15 radiologists. Radiology, 290(2), 219-227.

20. Wang, P., Xiao, X., Brown, J., Berzin, T. M., Tu, M., Xiong, F., & Liu, P. (2018). Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. Nature Biomedical Engineering, 2(10), 741-748.

21. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118.

22. Yamashita, R., Nishio, M., Do, R. K., Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. Insights into Imaging, 9(4), 611-629.

23. Liu, C., Ding, J., Spuhler, K., Gao, M., Qian, B., & Shen, B. (2019). Deep learning for image-based cancer detection and diagnosis—A survey. Pattern Recognition, 91, 119-149.

24. Lu, M. T., Ivanov, A. Y., Mayrhofer, T., Hosny, A., Aerts, H. J., Hoffmann, U., & Truong, Q. A. (2018). Deep learning to assess long-term mortality from chest radiographs. JAMA Network Open, 1(7), e184421-e184421.

25. Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. (2018). Artificial intelligence in radiology. Nature Reviews Cancer, 18(8), 500-510.

26. Huang, Y., Liu, Z., He, L., Chen, X., Pan, D., & Ma, Z. (2019). Radiomics signature: A potential biomarker for the prediction of disease-free survival in early-stage (I or II) non–small cell lung cancer. Radiology, 291(3), 547-554.

27. S.N. Sangeethaa,'Presumptive discerning of the severity level of glaucoma through clinical fundus images using hybrid PolyNet', Biomedical Signal Processing and Control, Volume 81, 2023, 104347, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2022.104347.

28. Bewal R, Ghosh A, Chaudhary A (2015) Detection of breast cancer using neural networks a review. J Clin Biomed Sci 5(4):143–148.

29. Weng S, Xu X, Li J, Wong ST (2017) Combining deep learning and coherent anti-Stokes Raman scattering imaging for automated differential diagnosis of lung cancer. J Biomed Opt 22(10):106017