International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 17 (2025) pp. 294–309

Article

Optimized Lung Cancer Identification in CT Imaging: A Synergistic Deep Learning Approach with Residual U-Net Segmentation and Swin Transformer Feature Extraction

Sunil Kumar [1,*], Amit Virmani [2], Ajay Tiwari [3], Nidhi [4], Abhishek Dwivedi [2], Amit Kumar Katiyar [3]

[1] Department of Information Technology, School of Engineering and Technology (UIET), CSJM University, Kanpur 208024, India
[2] Department of Computer Application, School of Engineering and Technology (UIET), CSJM University, Kanpur 208024, India; amitvirmani@csjmu.ac.in (A.V.); abhishekdwivedi@csjmu.ac.in (A.D.)
[3] Department of Electronics and Communication Engineering, School of Engineering and Technology (UIET), CSJM University, Kanpur 208024, India; ajaytiwari@csjmu.ac.in (A.T.); amitkatiyar@csjmu.ac.in (A.K.T.)
[4] Department of Computer Science & Engineering, COER University, Roorkee 247667, India; nidhiamit1529@gmail.com
* Correspondence author: sunilymca24@gmail.com

Received date: 27 November 2024; Accepted date: 25 February 2025; Published online: 27 March 2025

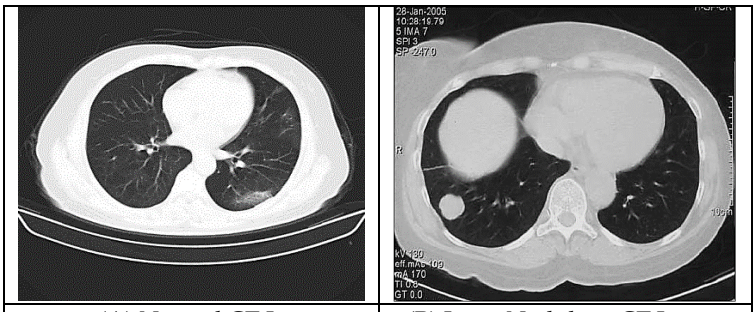
Abstract: Lung cancer remains a leading cause of cancer-related mortality globally, with high fatality rates often due to late-stage diagnosis. This research explores the efficacy of computer tomography (CT) imaging in the early detection of lung cancer. CT imaging, known for its high-resolution capabilities, facilitates the early identification of small nodules and abnormalities, providing detailed visualization of lung structures. This allows for the detection of minute changes that may indicate cancer. The investigation addresses the critical health issue, aiming to reduce the significant mortality and morbidity associated with lung cancer through improved early detection methods. The suggested ensemble method starts with segmentation, using residual-UNet and U-Net with DenseNet models to choose the region of interest (ROI). Following this, the Swin Transformer is employed to extract intricate features from the segmented CT images, leveraging its state-of-the-art capabilities. Principal component analysis (PCA) is implemented as a dimensionality reduction technique to optimize computational efficiency and improve feature selection. Furthermore, the DenseNet-121 and ResNet-101 models are employed to precisely identify lung nodule patterns. The investigation found that the residual U-Net model, with a dice coefficient and accuracy of 0.912 and 93.64%, respectively, is a superior method for segmenting lung nodule regions in CT images. The Swin Transformer successfully identified and extracted 215 distinct features from the segmented data obtained from the segmented lung regions. The PCA decreases the number of features extracted by the swin transformer. With an accuracy of 98.01%, an F1 score of 93.71%, and a dice coefficient of 0.938, the residual U-Net + ResNet-101 ensemble model did the best job of finding lung nodules. The outcomes demonstrate the superior performance of the proposed ensemble models compared to each other, making them the most suitable choice for lung cancer identification.

Keywords: convolutional neural networks; computer tomography; lung nodule; machine learning; medical imaging; transformers

1. Introduction

Lung cancer is a significant global health issue, being the leading cause of cancer-related mortality globally [1]. Lung cancer treatment differs from small-cell lung cancer due to its slower growth and spread, necessitating early detection for effective treatment and improved patient outcomes. However, the diagnostic challenges of this type often lead to late-stage diagnoses and poor prognosis. There are numerous subtypes of lung cancer that account for approximately 85% of all types of cancer [2,3].

Radiological examinations such as X-rays, CT scans, and MRIs often identify lung cancer. CT scans are widely used to detect lung cancer because they produce high-resolution images of the body. A CT scan can identify tiny lung nodules and designate their dimensions, morphology, and position efficiently [4,5]. Figure 1 displays CT scan images with and without lung nodules.



(A) Normal CT Image | (B) Lung Nodulous CT Image

Figure 1. Instances of CT scans [5].

The preprocessing pipeline for lung cancer, i.e., tumor/nodule segmentation, necessitates standardization and enhancement of CT images to improve their accuracy [6]. Contrast enhancement methods, such as Contrast Limited Histogram Equalization (CLAHE), increase the visibility of potential nodules and subtle features [7]. Data augmentation fixes class imbalance by randomly rotating, rotating, magnifying, and shifting the intensity of the data. It makes the model more robust and adds more types of data to the training set. Its purpose is to concentrate on the relevant lung regions to provide more effective input to the subsequent phases [6,8].

Computer-aided diagnosis (CAD) systems using deep learning (DL) algorithms have demonstrated encouraging outcomes in identifying lung cancer [9]. Progress in artificial intelligence (AI) has facilitated the creation of machine/deep learning and convolutional neural networks (CNNs) for the automated identification of lung cancer. These methods have found extensive use in image classification applications, particularly lung cancer categorization in medical imaging [10,11].

Segmentation techniques eliminate extraneous regions, while initial normalization guarantees uniformity for the processing. In lung nodule segmentation, U-Net architectures have shown significant promise in improving the accuracy and precision of defining nodule borders [12]. The residual U-Net incorporates residual connections to address the issue of vanishing gradients and enhance the learning process of deep networks. This approach ultimately improves the model's capacity to accurately divide intricate structures inside the lung [13]. In addition, the U-Net with DenseNet integrates DenseNet's strong feature propagation capabilities with the U-Net's localization strength [14]. The model efficiently collects both micro-level and macro-level information, leading to accurate and consistent segmentation of lung cancer. This is essential for enhancing the precision of treatment planning by offering dependable and thorough segmentation of malignant regions [13,14].

The Vision Transformer (ViT) model combines self-attention techniques with transformer structure to extract global context and spatial relationships from images. This technique is well-suited for processing high-resolution CT images and utilizes the Swin Transformer as the central network [15,16].

The Swin Transformer offers a powerful methodology for lung nodule detection by leveraging its hierarchical feature extraction, attention mechanisms, and global context modeling [17]. The Swin Transformer is a significant advancement in feature extraction, especially for jobs that involve intricate visual data, such as lung cancer [18]. The Swin Transformer is different from other CNN-based methods because it uses a hierarchical structure and shifting windows instead of convolutional kernels of a constant size. This enables it to effectively capture long-range relationships and contextual information across several scales. The capacity to represent both global and local characteristics makes it particularly efficient in obtaining intricate information from images, such as the marginal fluctuations seen in lung nodules or other medical abnormalities [19]. Moreover, the self-attention mechanism effectively focusses on the most pertinent areas of the image, leading to a more precise and flexible feature representation. The Swin Transformer demonstrates superior performance compared to alternative methods in the

processing of high-resolution CT images while utilizing fewer computational resources. This approach offers a more effective and dependable method for extracting features from CT images for segmentation and classification purposes [20].

The research uses a separate ensembling of each DL architecture with a Swin Transformer, where the Swin Transformer extracts feature vectors from CT images and the DL architecture performs classification [17]. Through its hierarchical design, the Swin Transformer captures intricate features, identifying patterns and structures that are important for classification. The DL architecture transforms the resulting output into a suitable format for input, utilizing the extracted features to enhance classification. Traditionally, these models undergo pre-training on an extensive CT image dataset before being fine-tuned to exclusively detect lung cancer [21,22].

The proposed ensemble approach combines feature from Swin Transformers, dimensionality reduction using PCA, and identification from DL architectures. The ensemble approach concatenates the features to create a unified feature vector, then integrates it into a suitable feature matrix. We use PCA to reduce dimensionality of the feature vector, identify the most significant components, and then send the resulting feature matrix to the DL architectures, which process it as input and learn to identify the data. This hierarchical processing enhances the model's ability to derive significant representations from combined features, capturing local and global data.

The substantial contributions of the investigation are listed below:

- The study uses CT images to accurately find lung cancer using the DL architecture identification method and Swin Transformer feature vector extraction.
- The use of modern residual U-Net and U-Net with DenseNet methods improves segmentation precision.
- We propose a unique ensemble method that combines the Swin Transformers, PCA, and DL architectures.
- The investigation carefully assesses at DL architecture s such as DenseNet-121, ResNet-101, and Swin transformers, focusing on how well they work at finding lung cancer and looking at key ways to improve the diagnosis process.

The research work is organized in the following manner: Section II specifically addresses the pertinent research that substantiate the investigation. Section III of our work presents a comprehensive explanation of our proposed ensemble system, including the methodology, datasets, segmentation, ensemble network, and identification mechanism. Section IV of the work summarizes the findings and discusses the experimental conditions used. Section V concludes the study and indicates potential next possibilities.

2. Related Works

This section discusses the analysis of several approaches or procedures now used for the diagnosis of lung cancer by researchers. An analysis was conducted on research articles concerning the identification and prediction of lung cancer.

The 3D Trans-DenseUnet++ model divided the acquired CT scans into segments using a novel loss function. The Multi-Scale Dilated 3D DenseNet uses Atrous Spatial Pyramid Pooling to accurately detect lung cancer by analyzing segmented images. Quantitative performance measures evaluate the accuracy and precision of the model. The proposed approach achieves an accuracy of 92.925 and precision of 93, validating the diagnosis [10]. According to the research, ViT and CNNs distinguish lung tumors better. Convolution and matching blocks capture key encoder-decoder network features. Self-attention helps higher transformer blocks record complicated global feature maps. The network had average dice coefficients of 0.7468 and 0.6847 and Hausdorff distances of 15.336 and 17.435 on a public NSCLC-Radiomics dataset and a local hospital dataset [16]. The Swin Transformer model classifies and segments lung cancer well. After pre-training, the Swin-B model achieved top-1 classification accuracy of 82.26%, exceeding ViT by 2.529%. The Swin-S model outperformed other segmentation methods. According to the data, pre-training in these actions may improve Swin Transformer model accuracy [20]. CT images were preprocessed using guided bilateral filtering to eliminate noise, a transformer-aided generative adversarial network (T-GAN) was used to detect lung cancer kinds, and DyLF-CO was utilized to modify the network model. The model is developed in Python and processes a chest CT picture. The approach has 0.997 accuracy, 0.996 precision, 0.998 specificity, 0.104 RMSE, and 120 s time complexity [23]. The study presented a systematic approach for constructing based-transformer models to automatically generate reports. Empirical measurements demonstrate that the pre-trained ViT feature extractor surpasses the CNN-based encoder (DensNet121) in terms of performance. Furthermore, the use of dual-view input (frontal and lateral images) produces superior outcomes compared to a single input [24].

Investigators created a CNN capable of identifying lung cancer with a sensitivity of 94.4% and a specificity of 93.9%. The researchers trained the CNN on a dataset of over 42,000 CT images, showcasing its superior accuracy in identifying lung lesions that radiologists often overlooked [25]. Investigators developed an automated method for lung segmentation using histograms to determine thresholds and detect lung nodules in CT scan images. After eliminating the threshold-detected outside region, the researchers extracted the lungs from the image. The lung segmentation process included morphological and connected component analysis, resulting in the identification of zones of interest. The Otsu technique was used to identify parenchyma nodules by distinguishing between blood vessels, bronchi, and internal characteristics of the nodules [26]. The development of three mutually dependent deep-fusion learning algorithms aims to identify lung nodules from CT images. MPF, SFMPF, and MFMPF are examples of deep-fusion hierarchical structures. Multi-perspective deep fusion classifies the MPF model into three distinct levels. The SFMPF model is a deep fusion learning architecture that uses features in a hierarchical fashion. The performance of four distinct feature-image-based hierarchical deep-fusion learning models is evaluated by testing these models using bilateral, trilateral, Gabor, and LOG-filtered images. The integration of these four SFMPF models leads to the development of a multi-feature, multi-perspective hierarchical deep fusion learning model [27]. CNN analyzed the CT image after manual segmentation. While DeepLab v3 and VGG-19 outperform artificial segmentation results, the testing clearly showed that both SegNet and artificial segmentation outputs are almost identical. Within the same timeframe, SegNet detected 120 cases of benign lung nodules and 120 cases of early lung cancer [28].

3. Methodology

The approach for lung cancer identification uses segmentation to find nodules in the lungs and separate lung regions from CT scan images. Figure 2 presents the whole process:

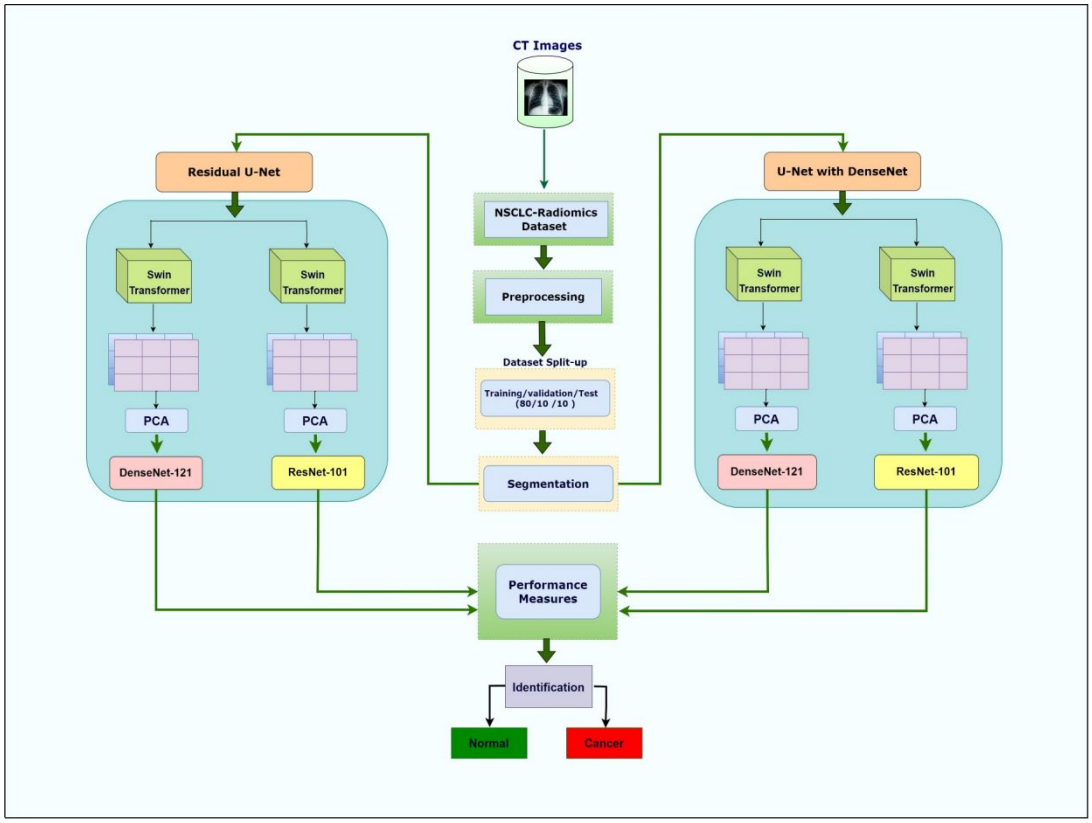


Figure 2. The Proposed Methodology for the lung cancer identification.

The methodology starts with acquiring CT images from the NSCLC-Radiomics dataset. These images undergo a series of preprocessing steps designed to enhance their quality and suitability for further analysis. Following preprocessing, segmentation is performed to identify and isolate regions of interest (ROIs) within the CT images. This step is crucial as it focuses the analysis on relevant areas, potentially containing nodules of lung cancer. The segmented images are then processed using two distinct model

architectures: Residual U-Net and U-Net with DenseNet. The Residual U-Net architecture incorporates a Swin Transformer for feature extraction, which captures intricate details and patterns within the images. PCA is applied for dimensionality reduction, simplifying the data while retaining its most significant features. DenseNet-121 and ResNet-101 models further process these reduced feature sets to enhance the feature representation. Similarly, the U-Net with DenseNet architecture was employed for the task. Finally, the performance of both model architectures is rigorously evaluated based on their performance measures; the input images are classified as either "Normal" or "Cancer." This comprehensive approach combines multiple advanced deep learning techniques, ensuring accurate and reliable detection of lung cancer in CT images.

3.1. Dataset

The NSCLC-Radiomics Dataset contains 422 CT scans obtained from individuals diagnosed with NSCLC and can be accessed at https://doi.org/10.7937/K9/TCIA.2015.PF0M9REI. The dataset consists of 7,368 scans included the lung nodule. We trained the model using 80% lung nodule CT images, and we evaluated its performance using 10% CT images. We preserve a separate test set consisting only of 10% CT images for the purpose of assessing the trained model [29]. Table 1 shows the number of CT scan instances for the investigation.

Table 1. CT Instances in the investigation.

Dataset Split-up	**Instances**
Training	5,896
Validation	736
Testing	736
Total	7,368

3.2. Preprocessing

To achieve effective identification, noisy regions in the raw lung CT image obtained from the data set must be eliminated. Therefore, methods built on screening have been developed to remove clutter from the image and improve the precision of detection. The preprocessing pipeline includes several key steps designed to normalize and enhance the input CT image, ensuring the focus remains on relevant lung regions and potential nodules [6,8]. Normalization is a step in the CT image processing process for finding lung nodules. This standardizes the image intensities to a range of −1000 to 1000 Hounsfield Units (HU) while keeping the contrast and illumination constant with CLAHE [7]. Data augmentation increases the variety of the dataset to address the issue of class imbalance. We achieve this by randomly rotating, flipping, zooming, and shifting the intensity of the data. These techniques ensure that the model is resilient in dealing with changes in lung nodule size and appearance [8].

3.3. Segmentation

Prior to lung segmentation, a collection of CT scans undergone for preprocessing. The process of lung segmentation entails the isolation of the lung from the anatomical structures of the body, such as ribs, arteries, and blood veins [30]. Accurate identification and delineation of lung sections along with prospective nodules in CT imaging depend on lung and nodule segmentation, therefore supporting early detection and diagnosis of lung cancer. Accurate segmentation enables the differentiation between cancerous nodules and benign structures, enhancing the precision of diagnosis and medical care planning [13,14].

3.3.1. Residual U-Net

Residual U-Net is a modified version of the original U-Net architecture. It uses leftover blocks in both the encoder and decoder stages to make segmentation more reliable and effective. This approach facilitates the acquisition of intricate features with greater efficacy, making it well-suited for the precise segmentation of lung nodules. The encoder has many residual blocks, each of which employs skip connections to include the input feature maps into the output of a sequence of convolutional layers. This approach improves feature learning and addresses the vanishing gradient issue. By integrating skip connections from the encoder method, the decoder upsamples the feature maps to merge high-level features with detailed information. Furthermore, the decoder blocks use residual connections to enhance

the feature maps and optimize segmentation [13]. The final output is a segmentation map that emphasizes the designated ROI.

It is possible to express the output $F(x)$ for a residual block that is included in the Residual U-Net as follows:

$$y = F(x, \{Wi\}) + x \tag{1}$$

where x: Input feature map, $F(x, \{Wi\})$: Residual mapping (output after convolution, batch normalization, and activation), y: Output feature map after the residual connection [31].

3.3.2. U-Net with DenseNet

A DenseNet-integrated U-Net design improves feature reuse and network performance in medical image segmentation tasks such as lung cancer diagnosis by including DenseNet blocks into the U-Net structure. This hybrid model incorporates DenseNet blocks in the encoder route of U-Net, interconnecting each layer with every other layer in a dense block. In order to maximize feature reuse and promote the network to learn more compact representations, this connection structure is specified. The decoder upsamples these densely linked features and then uses skip connections to combine them with matching encoder features to make a full segmentation map [14]. Efficiency in feature propagation and reduction of parameter requirements for precise segmentation are achieved by allowing each layer in the DenseNet block to receive the feature maps from all preceding layers as input.

$$x_\ell = H_\ell([x_0, x_1, \dots, x_{\ell-1}]) \tag{2}$$

where x_ℓ: Output of the ℓth layer in the dense block, H_ℓ: Composite function of batch normalization, ReLU, and convolution, and $[x_0, x_1, \dots, x_{\ell-1}]$: Concatenation of feature maps from all preceding layers [32].

3.4. Feature Extraction using Swin Transformer

The Swin Transformer is a groundbreaking advancement in ViTs that enhances the efficacy of image comprehension. The Swin Transformer is a hierarchical vision transformer that operates by processing images using a window-based approach. The system analyzes local variables and gradually combines them to provide global context. The use of both the hierarchical approach and the dynamic shift-based window system in the Swin Transformer has garnered it significant recognition [33]. The Swin Transformer has demonstrated promising results in the detection of nodules. Due to its robust attention mechanism and hierarchical feature representation, it is capable of accurately detecting intricate patterns and anomalies in CT images [17]. Its ability to analyze images at a high level of detail and its incorporation of both local and global contexts allow it to detect subtle indications of infection, such as infiltrates and consolidation. After extensive training on datasets, the Swin Transformer has shown remarkable reliability in distinguishing between healthy lungs and lung nodules [18]. Figure 3 depicts the process of feature extraction using a Swin transformer for the task.

To get localized processing and modest computing in hierarchical methodology, an image *I* of size H × W with three channels is represented into patches (*p*) of size s as:

$$P = \{p1\ , p2\ , \dots, p_N\} \tag{3}$$

The number of patches (*N*) in Eq 3 is represented as:

$$N = \frac{h * w}{s^2} \tag{4}$$

The input image $I \in \mathbb{R}\text{^}(h*w*c)$, where $\mathbb{R}$ represents the computed value and *h, w*, and *c* are the height, width, and channels, respectively, they are fragmented into patches that do not overlap, resulting in a grid of patches with dimensions ($h' \times w'$), where $h' = h/s$ and $w' = w/s$.

Then linearly project each patch with a weight matrix $W_p \in \mathbb{R}\text{^}(D_{in} \times D_{emb})$ to get the patch embeddings [33,34].

3.4.1. Self-Attention Mechanism

The Self-Attention Mechanism allows the model to focus on different aspects of the CT image, for a set of feature vectors $F = \{f_1, f_2, f_3, \dots\dots f_n\}$, and then needs to compute the queries, keys, and values in associated feature vectors:

$$Queries\ (Q)\ =\ F * W_Q \tag{5}$$

$$Keys\ (K)\ =\ F * W_k \tag{6}$$

$$Values\ (V)\ =\ F * W_V \tag{7}$$

The Swin Transformer employs Multi Head Self-Attention (MHSA) mechanisms to capture spatial dependencies across different patches. By computing multiple attention heads, MHSA allows the model to focus on different parts of the image at the same time. Each head captures different aspects of the relationships between patches, enabling the model to learn complex patterns and dependencies [33,34]. The model computes the attention scores for each input token.

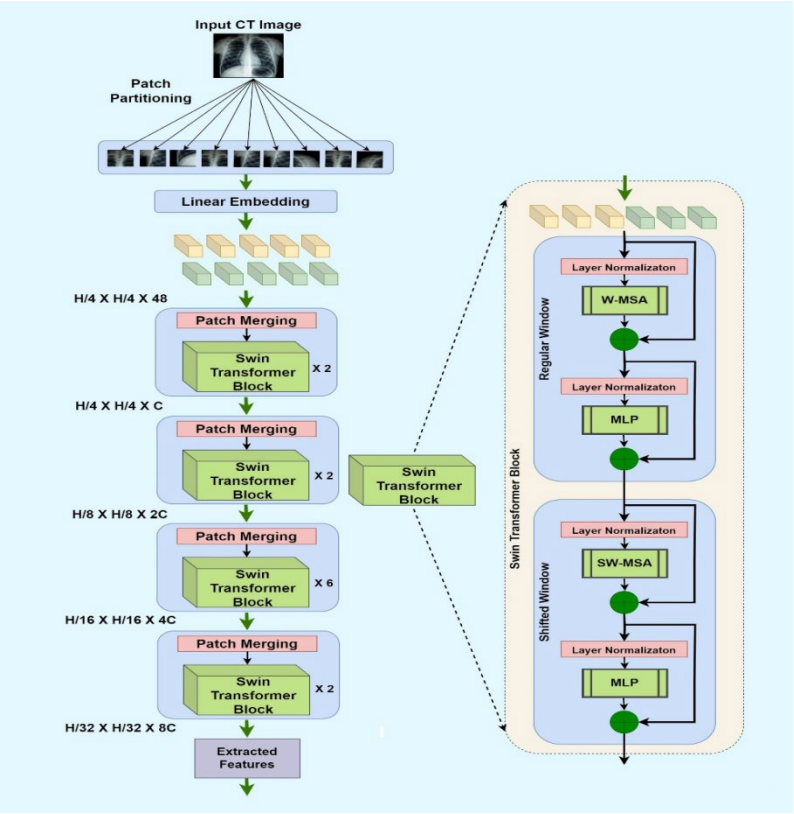


Figure 3. Feature extraction process performed by the Swin Transformer.

3.4.2. Self-Attention Mechanism

The Self-Attention Mechanism allows the model to focus on different aspects of the CT image, for a set of feature vectors $F = \{f_1, f_2, f_3, \ldots\ldots f_n\}$, and then needs to compute the queries, keys, and values in associated feature vectors:

The Swin Transformer employs Multi Head Self-Attention (MHSA) mechanisms to capture spatial dependencies across different patches. By computing multiple attention heads, MHSA allows the model to focus on different parts of the image at the same time. Each head captures different aspects of the relationships between patches, enabling the model to learn complex patterns and dependencies [33,34]. The model computes the attention scores for each input token.

$$Attention\ (Q,K,V) = SoftMax\left(\frac{QK^T}{\sqrt{D_{emb}}}\right) V \tag{8}$$

3.4.3. Shifted Window Partitioning

This mechanism ensures that the model can learn relationships between patches that were not initially in the same window. It enhances global context comprehension while maintaining computational efficiency. Standard window partitioning divides the image into fixed-size windows, computing attention within each one. To capture cross-window interactions, the windows are shifted by a certain number of patches in the next stage [20,33]. This can be represented as:

$$Shifted\_New_{Window} = Old_{Window} + Shift \quad (9)$$

To efficiently capture local and global features, the Swin Transformer introduces a shifted window mechanism. The image is partitioned into windows, and the self-attention is computed within each window. After each stage, the windows are shifted by a predetermined number of patches, allowing cross-window interaction [19].

3.4.4. Feature Hierarchies

The Swin Transformer constructs hierarchical feature representations by progressively merging neighboring patches. This hierarchical structure enables the model to build increasingly complex features from the initial raw patches, akin to how convolutional layers in CNNs build higher-level features. At each stage, the resolution of the image is reduced while increasing the number of channels:

$$Feature_i = MHSA(Feature_i - 1) + FFN(Feature_i - 1) \quad (10)$$

The term FFN refers to a feed-forward neural network that is applied after the attention mechanism. This hierarchical approach enables the extraction of relevant and multi-scale features [19,33].

3.5. PCA

PCA, an advanced dimensionality reduction method, reduces high-dimensional data while maintaining most of it. After the Swin Transformer drives significant features from segmented lung regions, we decrease their dimensionality using PCA. This simplifies data, decreases computational complexity, and reduces classification overfitting. PCA finds the data's principal components, or directions of highest variation [35,36].

3.6. Identification using Deep Learning Methods

3.6.1. DenseNet-121

DenseNet-121 (a 121-layer Dense Convolutional Network) facilitates the reuse of features and expedites the development of deep learning models. The method generates feature maps for succeeding layers by utilizing feature maps from previous layers. The network's direct node connections result in an increase in feature distribution and a reduction in parameters. The convolutional and transition layers constrain the complexity and dimensionality of the architectural design. This approach rectifies gradient distortion in deep neural networks [37,38]. The dense layer's operation is illustrated in Figure 4.

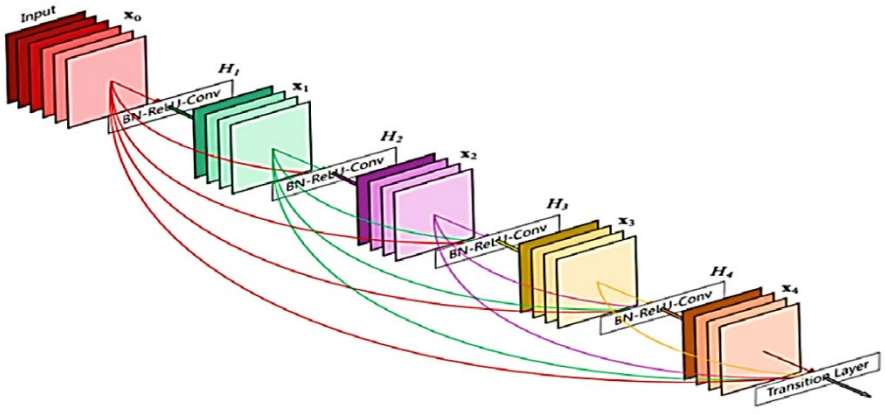


Figure 4. DenseNet Structure.

3.6.2. ResNet-101

Residual network 101 (ResNet-101) is a well-known deep learning architecture that leverages shortcut connections between layers for residual learning to improve accuracy and minimize degradation by using greater depth and identity mapping or projection shortcuts. Using residual mapping instead of input data speeds up dense network training. Shortcut connections allow ResNet-101 to map input to layers while skipping layers. ResNet-101 uses 101 residual blocks, convolutional layers, and shortcut connections to interpret complicated input data. A convolutional block arrangement with distinct weights, batch normalization, and ReLU activations are employed. Design assurances correspond to typical

network parameters, depth, breadth, and computation cost [39]. In appropriate cases, ResNet-101 may get the identity mapping (i.e., Y = X). When the optimal block transition approaches the identity mapping, adjusting the weights of convolutional layers may help to approximate it. This modification makes the optimization process easier during training [40]. Figure 5 presents the skip connection approach of the ResNet architecture.

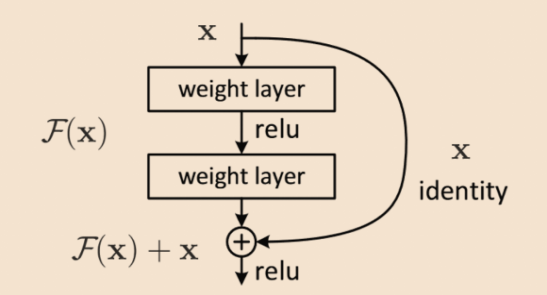


Figure 5. ResNet skip connection approach

3.6.3. Hyper-Parameters

DenseNet-121 and ResNet-101 used certain hyper-parameters that improved the identification. An optimal learning rate of 0.001, accompanied by a decay factor of 0.1 per 10 epochs, guarantees a steady and progressive learning process. Implementing a batch size of 32 achieves a harmonious equilibrium between computing efficiency and model performance. The Adam Optimizer is highly recommended due to its exceptional adaptive learning rate capabilities. Training for 50 epochs produces enough iterations to achieve convergence [36]. Implementing a weight decay of 1e−4 serves to regularize the model, while a dropout rate of 0.5 effectively mitigates overfitting. Binary cross-entropy and dice loss were combined and used to balance pixel-wise accuracy with overall shape accuracy. BCE assesses pixel-level accuracy by comparing predicted and true labels, while Dice Loss focuses on the overlap between predicted and real labels to ensure overall shape correctness. The hyper-parameters that have been specified are shown in Table 2, which includes the employed image size with channels, learning rate, decay, batch size, optimizer, dropout rate, loss function, and epochs.

Table 2. Hyper-parameters.

Terms	**Instances**
Image size	(224*224*3)
Learning rate	0.001
Decay	0.1 per 10 Epochs
Batch size	32
Optimizer	Adam
Dropout rate	0.5
Loss function	Binary Cross-Entropy + Dice Loss
Number of epochs	50

3.7. Performance Measures

Performance measures assess and quantify the effectiveness of a research model. We utilize different metrics for performance, such as the Dice coefficient ($Dice_{co}$) [10], accuracy, precision, recall, and F1 score [41,42]. To specifically assess the effectiveness of the task's segmentation, $Dice_{co}$ measures the overlap between the predicted segmentation and the ground truth. Equation (11) presents $Dice_{co}$, with A representing the predicted pixels and B representing the ground truth pixels. Table 3 provides an overview of all the significant metrics:

Table 3. Performance measures.

Measure = Equation	
$\text{Dice}_{co} = \frac{2 \times \mid A \cap B \mid}{\mid A \mid + \mid B \mid}$	(11)
$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)}$	(12)
$Recall = \frac{TP}{TP + FN}$	(13)
$Precision = \frac{TP}{TP + FP}$	(14)
$F1\ Score = 2 * \left(\frac{Precision * Recall}{Precision + Recall}\right)$	(15)

Legends- *TP*: True Positive, *TN*: True Negative, *FP*: False Positive, *FN*: False Negative.

4. Results and Discussion

4.1. Implementation details

We conduct the model training on a 32 GB NVIDIA GEFORCE RTX 4060 Graphical Processing Unit (GPU) and 128 GB RAM environment. After resizing, the input CT image dimensions are configured as 224 × 224.

4.2. Preprocessing

In our implementation, we rigorously analyzed the CT scans to ensure precise and reliable identification of lung nodules. Our approach to addressing the problem of class imbalance included the use of a variety of data augmentation methods. The methods used included random rotations, flips, zooms, and intensity adjustments. Through this approach to enhancing the dataset, we effectively increased the variety of the training data, thereby improving our models' resilience. Furthermore, we adjusted the CT scans' brightness to a precise range of −1,000 to 1,000 HU. The normalization process was important in ensuring uniform contrast and brightness in all pictures, a critical factor for precise detection and segmentation of lung nodules. The methodology we used extensively depended on these preprocessing procedures to train the models on data of superior quality and diversity.

4.3. Segmentation and Identification

We evaluated the effectiveness of two segmentation models, Residual U-Net and U-Net with DenseNet, using a dataset that included CT scan images. The research aimed to assess the accuracy and reliability of each model in segmenting lung cancer areas. Figure 6 shows the segmented images of the lungs. CT scans in their original state (top row), and lung nodules segregated using Residual U-Net (middle row) and U-Net with DenseNet (bottom row). Residual U-Net exhibits remarkable border delineation, while U-Net with denseNet achieves successful segmentation with somewhat lower precision. The original images provide a reference for evaluating the accuracy of the segmentation results.

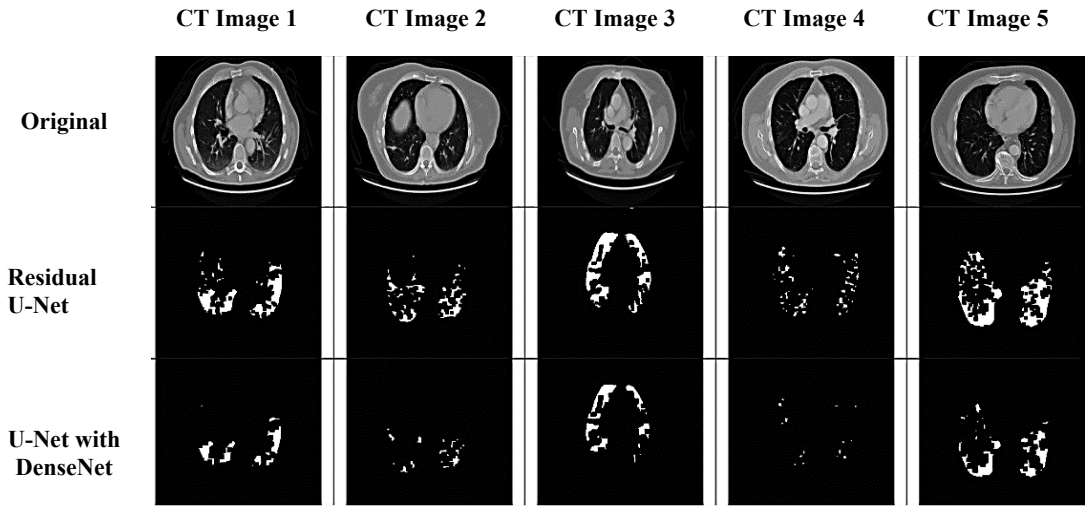
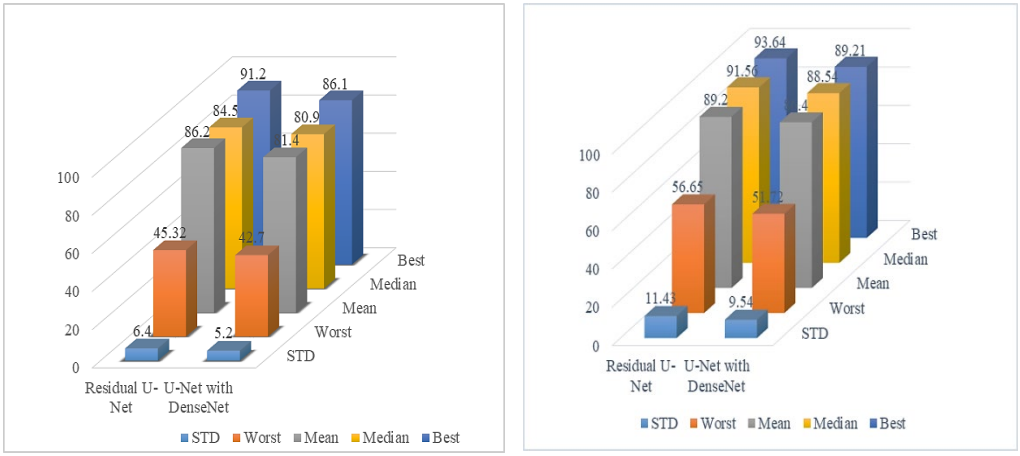


Figure 6. Segmented CT images by the Residual U-Net and U-Net with DenseNet.

We presented the statistical analysis using standard deviation (STD), mean, median, worst, and best performance measures. Figure 7 displays the segmentation results of the residual U-Net and U-Net with DenseNet models.



(a) Dice coefficient (b) Accuracy

Figure 7. Statistical investigation of Residual U-Net and U-Net with DenseNet models.

The residual U-Net achieved the best Diceco and accuracy of 0.912 (91.20%) and 93.64%, surpassing the values of 0.861 (86.11%) and 89.21% observed for the U-Net with DenseNet. We attribute the residual U-Net's superior performance to its ability to preserve spatial information through skip connections, thereby facilitating more effective reconstruction of the segmented areas. Conversely, the U-Net with DenseNet, while successful, may have suffered from a certain degree of spatial information loss caused by the connections, resulting in somewhat worse segmentation.

The residual U-Net model did a great job of capturing the complicated and unique features of lung cancer regions, which led to more accurate segmentation results. The segmented data obtained from the residual U-Net was then subjected to additional processing using the Swin Transformer. An analysis of the segmented lung cancer areas was conducted using the Swin Transformer, which is renowned for its strong feature extraction capabilities.

This study investigated the comprehensive range of features extracted by the Swin Transformer from the segmented data. The Swin Transformer computed 215 distinctive features from the segmented areas, including shape characteristics, texture features, intensity features, edge features, spatial features, and other supplementary features. These features provided a thorough and intricate depiction of the regions affected by lung cancer, therefore enabling more precise analysis and categorization. The extracted features were exceptional, demonstrating the outstanding potential of this integrated method in improving the diagnosis and analysis of lung cancer.

We used PCA on the retrieved features to improve efficiency and refine the feature set. PCA identified the most relevant features that significantly contribute to data variance, simplifying the model, lowering computing complexity, and mitigating the risk of overfitting. We kept the principle components that accounted for 95% of the total variance, thereby conserving the maximum amount of information from the original feature set. The principle component research procedure decreased the number of features to 31 principle components, resulting in a more manageable and efficient attribute set for future research.

For lung nodule identification, we processed the features using ResNet-101 and DenseNet-121 models to further validate the effectiveness of the reduced feature set. A promising way to find lung cancer is to use the residual U-Net and U-Net with densenet for segmentation, the Swin Transformer for feature extraction, and PCA for feature reduction. We assessed the performance of these models in conjunction with the previously mentioned integrated architectures. The accuracy, precision, recall, and F1 scores showed that the features, along with the ResNet-101 and DenseNet-121 models, worked very well together to find lung nodules. This method not only enhances the precision of lung cancer diagnosis but also improves the efficiency of the computational process, making it a valuable tool in diagnostics. Table 4 presents the outcome of the investigation.

Table 4. Identification Outcomes.

Model	Accuracy	Precision	Recall	F1 Score	$Dice_{co}$
Residual U-Net + DenseNet-121	97.23	89.42	96.21	92.89	0.926
Residual U-Net + ResNet-101	98.01	90.39	97.37	93.71	0.938
U-Net with DenseNet + DenseNet-121	96.39	87.62	94.43	91.21	0.911
U-Net with DenseNet + ResNet-101	96.55	88.14	95.07	91.54	0.913

We observed significant variations in performance measures. Among the models that were examined, the Residual U-Net + ResNet-101 model had the most superior performance in all assessed criteria. The accuracy, F1 score, and dice coefficient were calculated to be 98.01%, 93.71%, and 0.938, respectively. The findings demonstrate that this model is quite efficient in detecting lung nodules, exhibiting a well-balanced performance. Additionally, the strong scores for recall and accuracy indicate that the model exhibits both sensitivity and specificity.

The second-best performer was the Residual U-Net + DenseNet-121 model. It achieved an F1 score of 92.89%, an accuracy of 97.23%, and a dice coefficient of 0.926. Despite being slightly inferior to the best performer, this model still exhibits robust performance and reliability in the identification process. Conversely, among the evaluated models, the U-Net with DenseNet + DenseNet-121 model proved to be the least effective. Figure 8 illustrates the confusion matrix for the Residual U-Net + ResNet-101 model.

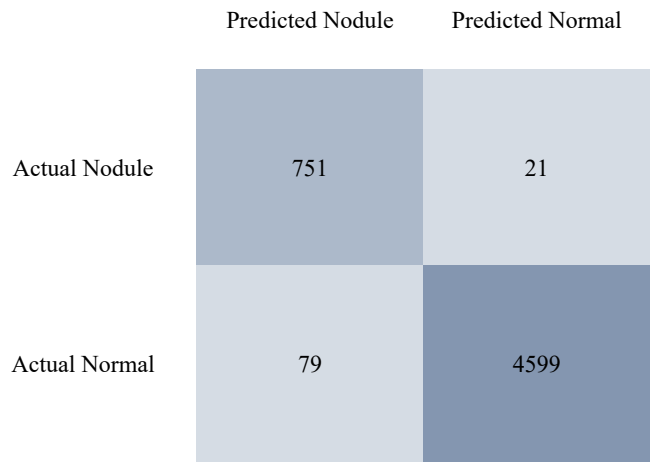


Figure 8. Confusion Matrix for Residual U-Net + ResNet-101.

The accuracy of the Residual U-Net + ResNet-101 model throughout training and validation, in relation to the training and validation loss, is shown in Figure 9.

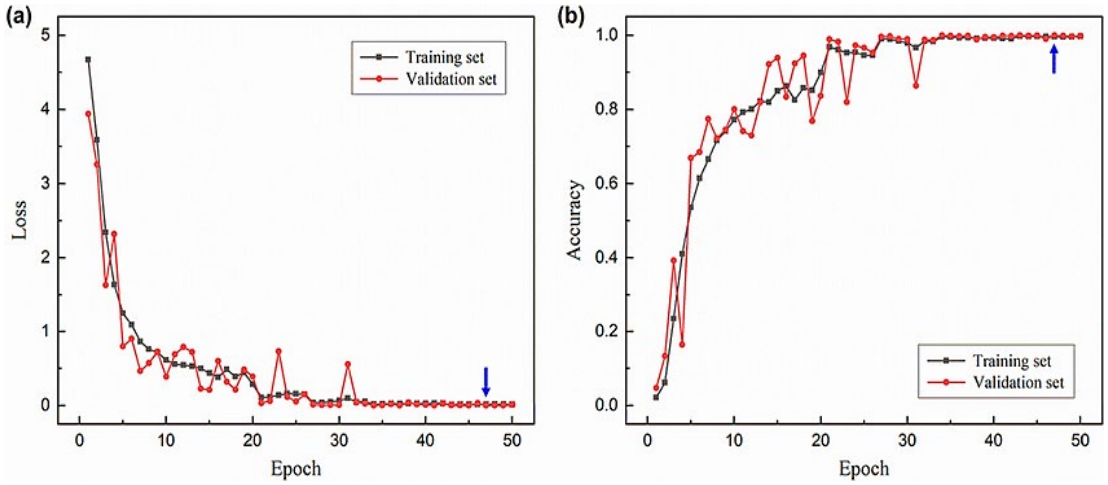


Figure 9. Training and validation loss and accuracy of the Residual U-Net + ResNet-101 model.

Figure 10 presents the identification results of lung nodules from test images of distinct CT scans, highlighting a clear comparison between the input, ground truth, and predicted outcomes by the residual U-Net + ResNet-101 model. The assessment employed the original CT scan image in the first row. The second row presents the ground truth lung nodule, which serves as the accuracy reference. The third row highlights the predicted lung nodule, demonstrating the model's nodule detection ability in comparison to the actual findings.

The first row displays the original CT images, providing a comprehensive view of the lung structures. The second row shows the original lung nodules, isolated and highlighted against a black background. The third row presents the predicted lung nodules, also highlighted against a black background. The research findings indicate that the predicted nodules closely match the original nodules, demonstrating the effectiveness of the Residual U-Net + ResNet-101 model.

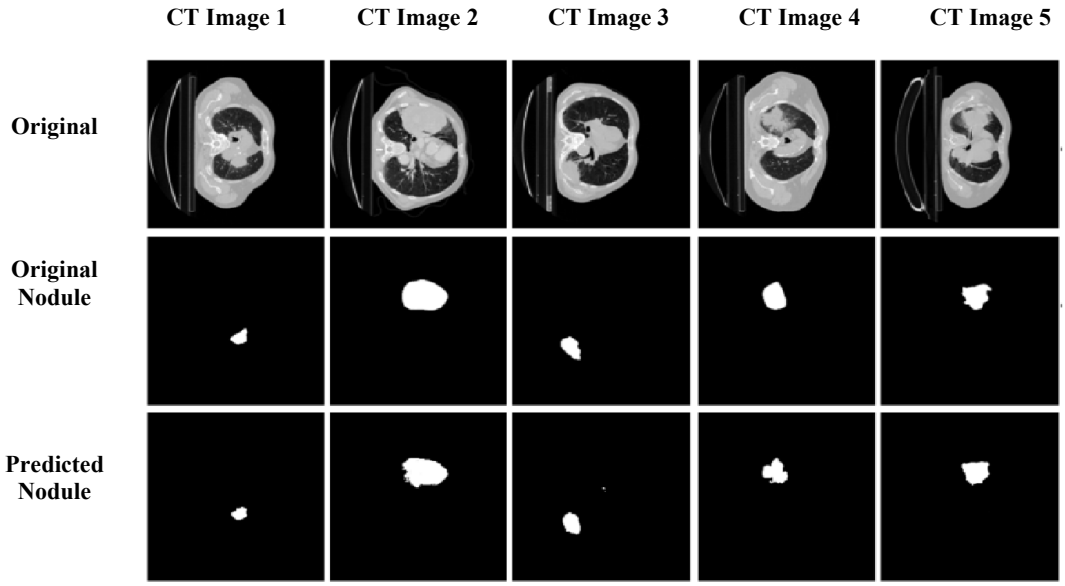


Figure 10. Identification results of the lung nodules.

4.4. Ablation Study

The ablation research aims to tackle the substantial health issue of lung cancer by improving early detection techniques. The suggested ensemble methodology starts with segmentation, using residual-UNet and U-Net models combined with DenseNet to identify the ROI in CT images. Subsequent to segmentation, the Swin Transformer is used to extract intricate features from the segmented images. PCA is used as a dimensionality reduction method to optimize computing performance and improve feature selection. Furthermore, DenseNet-121 and ResNet-101 models are used to precisely recognize lung nodule patterns.

The research indicated that the residual U-Net model, with a dice coefficient of 0.912 and an accuracy of 93.64%, was notably proficient at segmenting lung nodule areas in CT scans. The Swin Transformer effectively recovered 215 distinct features from the segmented data. PCA substantially decreased the quantity of features derived from the Swin Transformer, guaranteeing maximum efficiency. The integration of the residual U-Net and ResNet-101 models demonstrated superior efficacy in lung nodule detection, achieving an accuracy of 98.01%, an F1 score of 93.71%, and a dice coefficient of 0.938.

The results demonstrate the enhanced efficacy of the proposed ensemble models compared to others, emphasizing their viability as the optimal option for lung cancer detection. The research shows a significant improvement in early identification via the use of these new approaches, possibly decreasing the mortality and morbidity linked to lung cancer.

4.5. Comparative Analysis

Researchers preferred accuracy, F1 Score for the classification while $Dice_{co}$ is for the segmentation as their primary performance metric. The comparison table entailed a comprehensive assessment of various techniques for categorizing CT images, primarily focusing on lung cancer diagnosis. Table 5 compares the proposed ensemble approach (Residual U-Net + ResNet-101) to specific alternatives.

Table 5. Comparative analysis.

Model	Accuracy	F1 Score	$Dice_{co}$	Ref.
(3D Trans-DenseUnet++ with Novel Loss Function) + (MDDNetASPP)	92.92	86.79	0.872	[10]
ViT + CNN	--	--	0.746	[16]
UPerNet + Swin Transformer	82.26	--	47.93	[20]
Residual U-Net + ResNet-101 (Ours)	98.01	93.71	0.938	--

The results revealed that residual U-Net + ResNet-101 achieved outstanding effectiveness, showcasing its proficiency in accurately identifying CT instances of lung cancer. The models that were examined, the Residual U-Net + ResNet-101 model has the greatest accuracy of 98.01%, an amazing F1 score of 93.71, and a $Dice_{co}$ of 0.938. These results indicate that the Residual U-Net + ResNet-101 model outperforms the other models in both classification and segmentation tasks for lung cancer detection using CT scans. The (3D Trans-DenseUnet++ with Novel Loss Function) + (MDDNetASPP) model has robust performance, with an accuracy of 92.92% and a Dice coefficient of 0.872. However, it does not reach the same level of performance as the leading model. Both the ViT + CNN model, which has a $Dice_{co}$ of 0.746, and the UPerNet + Swin Transformer, which has an accuracy of 82.26% and a $Dice_{co}$ of 47.93, exhibit moderate to poor performance. This indicates that both models may need more tuning in order to enhance their segmentation accuracy. In summary, the Residual U-Net + ResNet-101 model demonstrates superior efficacy in accurately detecting lung cancer.

5. Conclusions

The late-stage detection of lung cancer is a significant factor in its continued prevalence as a primary cause of cancer-related fatalities worldwide. High-resolution CT imaging has demonstrated effective early identification of lung nodules and abnormalities, significantly improving diagnostic outcomes. The study aims to develop a diagnostic system for image identification that accurately forecasts lung cancer cases through CT images. The residual U-Net model is found to be more effective in segmenting lung cancer areas in CT images, providing more accurate results and clear regions of cancerous areas. It captures complex features of lung cancer areas, leading to more accurate segmentation. Combining the residual U-Net with Swin Transformer for relevant feature extraction and PCA for feature reduction appears to be a promising method for identification. The reduced feature set, which includes more informative principal components, preserves essential information for accurate analysis. We employed the residual U-Net model and U-Net with DenseNet to identify pulmonary nodule patterns, and the DenseNet-121 and ResNet-101 models effectively combined to produce highly precise outcomes. Among the evaluated models, the residual U-Net + ResNet-101 ensemble outperformed other ensemble models in lung nodule detection. This emphasizes the model's potential as the most appropriate method for the precise identification of lung cancer. This research underscores the potential of combining deep learning models for improving early lung cancer detection. We evaluated the models on a constrained dataset, potentially limiting their generalizability. Furthermore, the computational complexity and resource requirements for training these deep learning models may limit their usability and scalability in clinical contexts. Future studies should improve models' speed and computational needs, test them on a

variety of datasets, use advanced methods like attention mechanisms or hybrid models, and look into how easy it is for clinicians to understand predictions in order to get clearer insights and improve the accuracy of detection.

Author Contributions
S.K. wrote, designed, and implemented the research; A.V. and A.D. contributed to the data acquisition and interpretation; N. contributed to the manuscript analysis and visualization; and A.T. and A.K.T. contributed to the manuscript review. All authors have read and agreed to the published version of the manuscript.

Funding
This research received no external funding.

Conflict of Interest Statement
The authors declare no conflicts of interest.

Data Availability Statement
The dataset used to train the models is publicly available for research at the provided URL.

References

1. World Health Organization: WHO, “The top 10 causes of death,” Aug. 07, 2024. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death
2. S. B. Clark and S. Alsubait, “Non–Small Cell Lung Cancer,” StatPearls - NCBI Bookshelf, Sep. 04, 2023. https://www.ncbi.nlm.nih.gov/books/NBK562307/
3. American Lung Association, “Types of Lung Cancer.” https://www.lung.org/lung-health-diseases/lung-disease-lookup/lung-cancer/basics/lung-cancer-types
4. K. J. Ackah, M. Diab, K. Elbrow, E. Lewis, and A. Marchbank, “203 Should interactive 3D-CT models be used as an alternative to repeat contrast computed tomography in lung cancer screening programme patients for surgical planning?” *Lung Cancer*, vol. 190, p. 107764, Apr. 2024, doi: 10.1016/j.lungcan.2024.107764.
5. M. Sherigar, J. Finnegan, D. McManus, T. F. Lioe, and R. A. J. Spence, "CT Scan of chest showing one of the lung nodules," figshare, Figure, 2011. doi: https://doi.org/10.6084/m9.figshare.16069.v1.
6. J. Wang et al., “Preparing CT imaging datasets for deep learning in lung nodule analysis: Insights from four well-known datasets,” *Heliyon*, vol. 9, no. 6, p. e17104, Jun. 2023, doi: 10.1016/j.heliyon.2023.e17104.
7. A. A. Mary and K. K. Thanammal, “Lung cancer detection via deep learning-based pyramid network with honey badger algorithm,” *Measurement Sensors*, vol. 31, p. 100993, Feb. 2024, doi: 10.1016/j.measen.2023.100993.
8. X. Zhou et al., “Customized T-time inner sampling network with uncertainty-aware data augmentation strategy for multi-annotated lesion segmentation,” *Computers in Biology and Medicine*, vol. 180, p. 108990, Sep. 2024, doi: 10.1016/j.compbiomed.2024.108990.
9. M. Alamgeer et al., “Deep Learning Enabled Computer Aided Diagnosis Model for Lung Cancer using Biomedical CT Images,” Computers, Materials & Continua/Computers, *Materials & Continua (Print)*, vol. 73, no. 1, pp. 1437–1448, Jan. 2022, doi: 10.32604/cmc.2022.027896.
10. S. Sridevi and N. ARajivKannan, “Development of 3DTDUnet++ with novel function and multi-scale dilated-based deep learning model for lung cancer diagnosis using CT images,” *Biomedical Signal Processing and Control*, vol. 94, p. 106243, Aug. 2024, doi: 10.1016/j.bspc.2024.106243.
11. G. Mohandass, G. H. Krishnan, D. Selvaraj, and C. Sridhathan, “Lung Cancer Classification using Optimized Attention-based Convolutional Neural Network with DenseNet-201 Transfer Learning Model on CT image,” *Biomedical Signal Processing and Control*, vol. 95, p. 106330, Sep. 2024, doi: 10.1016/j.bspc.2024.106330.
12. T.-W. Wang, J.-S. Hong, J.-W. Huang, C.-Y. Liao, C.-F. Lu, and Y.-T. Wu, “Systematic review and meta-analysis of deep learning applications in computed tomography lung cancer segmentation,” *Radiotherapy and Oncology*, vol. 197, p. 110344, Aug. 2024, doi: 10.1016/j.radonc.2024.110344.
13. R. Bbosa, H. Gui, F. Luo, F. Liu, K. Efio-Akolly, and Y.-P. P. Chen, “MRUNet-3D: A multi-stride residual 3D UNet for lung nodule segmentation,” *Methods*, vol. 226, pp. 89–101, Jun. 2024, doi: 10.1016/j.ymeth.2024.04.008.
14. F. T. J. Faria, M. B. Moin, P. Debnath, A. I. Fahim, and F. M. Shah, “Explainable Convolutional Neural Networks for Retinal Fundus Classification and Cutting-Edge Segmentation Models for Retinal Blood Vessels from Fundus Images,” *arXiv*, no. 2405.07338v1, [Online]. Available: https://arxiv.org/pdf/2405.07338v1
15. H. Ali, F. Mohsen, and Z. Shah, “Improving diagnosis and prognosis of lung cancer using vision transformers: a scoping review,” *BMC Medical Imaging*, vol. 23, no. 1, Sep. 2023, doi: 10.1186/s12880-023-01098-z.
16. S. Tyagi, D. T. Kushnure, and S. N. Talbar, “An amalgamation of vision transformer with convolutional neural network for automatic lung tumor segmentation,” *Computerized Medical Imaging and Graphics*, vol. 108, p. 102258, Sep. 2023, doi: 10.1016/j.compmedimag.2023.102258.
17. F. Cui, Y. Li, H. Luo, C. Zhang, and H. Du, “SF2T: Leveraging Swin Transformer and Two-stream networks for lung nodule detection,” *Biomedical Signal Processing and Control,* vol. 95, p. 106389, Sep. 2024, doi: 10.1016/j.bspc.2024.106389.
18. J.-X. Ren, Y.-J. Xiong, X.-J. Xie, and Y.-F. Dai, “Learning Transferable Feature Representation with Swin Transformer for Object Recognition,” *Neural Processing Letters*, vol. 55, no. 3, pp. 2211–2223, Aug. 2022, doi: 10.1007/s11063-022-11004-3.

19. J.-H. Kim, N. Kim, and C. S. Won, “Global–local feature learning for fine-grained food classification based on Swin Transformer,” *Engineering Applications of Artificial Intelligence*, vol. 133, p. 108248, Jul. 2024, doi: 10.1016/j.engappai.2024.108248.
20. R. Sun, Y. Pang, and W. Li, “Efficient Lung Cancer Image Classification and Segmentation Algorithm Based on an Improved Swin Transformer,” *Electronics*, vol. 12, no. 4, p. 1024, Feb. 2023, doi: 10.3390/electronics12041024.
21. Z. Tan, H. Madzin, B. Norafida, R. W. O. Rahmat, F. Khalid, and P. S. Sulaiman, “SwinUNeLCsT: Global-local spatial representation learning with hybrid CNN-transformer for efficient tuberculosis lung cavity weakly supervised semantic segmentation,” *Journal of King Saud University - Computer and Information Sciences, vol. 36, no. 4, p. 102012, Apr. 2024, doi: 10.1016/j.jksuci.2024.102012.*
22. A. Lin, B. Chen, J. Xu, Z. Zhang, and G. Lu, “DS-TransUNet:Dual Swin Transformer U-Net for Medical Image Segmentation,” arXiv (Cornell University), Jan. 2021, doi: 10.48550/arxiv.2106.06716.
23. S. V. S. N. Murthy and P. M. K. Prasad, “Adversarial transformer network for classification of lung cancer disease from CT scan images,” *Biomedical Signal Processing and Control*, vol. 86, p. 105327, Sep. 2023, doi: 10.1016/j.bspc.2023.105327.
24. A. Elaanba, M. Ridouani, and L. Hassouni, “Transformer-Based Model for Radiology Text Reports Generation from Frontal and Lateral Chest X-ray Images,” Jul. 10, 2024. https://cspub-ijcisim.org/index.php/ijcisim/article/view/732
25. D. Ardila et al., “End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography,” *Nature Medicine*, vol. 25, no. 6, pp. 954–961, May 2019, doi: 10.1038/s41591-019-0447-x.
26. N. Khehrah, M. S. Farid, S. Bilal, and M. H. Khan, “Lung Nodule Detection in CT Images Using Statistical and Shape-Based Features,” *Journal of Imaging*, vol. 6, no. 2, p. 6, Feb. 2020, doi: 10.3390/jimaging6020006.
27. K. Sekeroglu and Ö. M. Soysal, “Multi-Perspective Hierarchical Deep-Fusion Learning Framework for Lung Nodule Classification,” *Sensors*, vol. 22, no. 22, p. 8949, Nov. 2022, doi: 10.3390/s22228949.
28. X. Chen, Q. Duan, R. Wu, and Z. Yang, “Segmentation of lung computed tomography images based on SegNet in the diagnosis of lung cancer,” *Journal of Radiation Research and Applied Sciences*, vol. 14, no. 1, pp. 396–403, Dec. 2021, doi: 10.1080/16878507.2021.1981753.
29. “NSCLC-Radiomics,” The Cancer Imaging Archive. *NIH*. doi: 10.7937/K9/TCIA.2015.PF0M9REI.
30. C. F. J Kuo et al., “Automatic lung nodule detection system using image processing techniques in computed tomography,” *Biomedical Signal Processing and Control*, vol. 56, p. 101659, Feb. 2020, doi: 10.1016/j.bspc.2019.101659.
31. Z. Zhang, Q. Liu, and Y. Wang, “Road Extraction by Deep Residual U-Net,” IEEE Geoscience and Remote Sensing Letters, vol. 15, no. 5, pp. 749–753, May 2018, doi: 10.1109/lgrs.2018.2802944.
32. S. He et al., “Intelligent Mapping of Urban Forests from High-Resolution Remotely Sensed Imagery Using Object-Based U-Net-DenseNet-Coupled Network,” *Remote Sensing*, vol. 12, no. 23, p. 3928, Nov. 2020, doi: 10.3390/rs12233928.
33. Z. Liu et al., “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,” *arXiv* (Cornell University), Jan. 2021, doi: 10.48550/arxiv.2103.14030.
34. D. Li, “Attention-enhanced architecture for improved pneumonia detection in chest X-ray images,” *BMC Medical Imaging*, vol. 24, no. 1, Jan. 2024, doi: 10.1186/s12880-023-01177-1.
35. S. Wold, K. Esbensen, and P. Geladi, “Principal component analysis,” *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1–3, pp. 37–52, Aug. 1987, doi: 10.1016/0169-7439(87)80084-9.
36. S. Kumar and H. Kumar, “Classification of COVID-19 X-ray images using transfer learning with visual geometrical groups and novel sequential convolutional neural networks,” *MethodsX*, vol. 11, p. 102295, Dec. 2023, doi: 10.1016/j.mex.2023.102295.
37. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely Connected Convolutional Networks,” *arXiv* (Cornell University), Jan. 2016, doi: 10.48550/arxiv.1608.06993.
38. C. C. Ukwuoma et al., “A hybrid explainable ensemble transformer encoder for pneumonia identification from chest X-ray images,” *Journal of Advanced Research*, vol. 48, pp. 191–211, Jun. 2023, doi: 10.1016/j.jare.2022.08.021.
39. K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” *arXiv* (Cornell University), Jan. 2015, doi: 10.48550/arxiv.1512.03385.
40. S. Ioffe and C. Szegedy, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,” *arXiv* (Cornell University), Jan. 2015, doi: 10.48550/arxiv.1502.03167.
41. S. Kumar and H. Kumar, “Efficient-VGG16: A Novel Ensemble Method for the Classification of COVID-19 X-ray Images in Contrast to Machine and Transfer Learning,” *Procedia Computer Science*, vol. 235, pp. 1289–1299, Jan. 2024, doi: 10.1016/j.procs.2024.04.122.
42. S. Kumar and H. Kumar, “Lung Cancer Diagnosis Using X-Ray and CT Scan Images Based on Machine Learning Approaches,” in *Lecture notes in networks and systems,* 2023, pp. 399–412. doi: 10.1007/978-981-99-1479-1_30.