[1]Anum Kamal
[2] Faiyaz Ahamad

Enhancing Precision in Lung Cancer Detection through CapsuleNet-ResNet Fusion Model

Journal of Electrical Systems

Abstract: - Lung cancer is a big problem in global health, hence there has to be improvement in methods for detecting it early. In this study, we introduce a novel approach to enhancing the precision of lung cancer detection by merging ResNet and CapsuleNet into a fusion model. The strength of CapsuleNet in capturing hierarchical properties and the experience of ResNet in tackling vanishing gradient challenges are combined to create a more robust solution for lung cancer diagnosis. The suggested CapsuleNet-ResNet fusion model makes use of CapsuleNet's distinctive capsule structure to effectively describe spatial hierarchies within lung imagery. Dynamic routing can capture complex patterns more efficiently by using capsules instead of regular neurons. Using ResNet's residual learning to address issues caused by deep neural networks, we further enhance the model's feature extraction. We train ResNet and CapsuleNet independently after pre-processing the lung image collection. Afterwards, the learned representations from both networks are combined using a well-planned fusion procedure. More discriminative detection of lung cancer is achieved by the model through the combination of local and global data. We put our suggested method through its paces using benchmark lung cancer datasets in a battery of tests. We test the suggested CapsuleNet-ResNet fusion model against state-of-the-art methods, CapsuleNet and ResNet models on their own, and more. Our CapsuleNet-ResNet fusion model revealed significant results with 98% accuracy, 97.2% precision, 98.5% recall rate and 97.8% F1 Score**.** These results surpass those of fundamental algorithms such as VGG16, CNN, and ResNet. It reveal that our fusion model has the potential to be used for early detection of lung cancer because of its improved detection accuracy and resilience. By accurately depicting lung images and making use of the unique properties of ResNet and CapsuleNet, our proposed method enhances diagnostic abilities of lung cancer.

Keywords: Lung Cancer Detection, Capsule Networks, Residual Networks, Fusion Model, Medical Image Analysis

I. INTRODUCTION

Lung cancer, a substantial healthcare concern, not only exacts a significant toll in terms of cost but also manifests as a formidable challenge because of its tendency for late-stage diagnosis and high mortality rates. Given these challenges, there is a pressing need for advanced diagnostic technologies capable of early detection with heightened accuracy to enhance the survival rates of cancer patients. Recent years have witnessed a remarkable advancement in medical imaging analysis, largely attributed to the emergence of deep learning models being used in multiple researches. Notably, Wang et al. (2023) introduced a multiple-scale residual network aimed at improving the identification of specific types of lung nodules, thus enhancing diagnostic capabilities [1]. Tian et al. (2024) pioneered a combined model merging radiomics and deep learning techniques to forecast hidden lymph node metastases in lung cancer across multiple healthcare facilities, demonstrating the efficacy of such integration [2]. Alamgeer et al. (2023) employed nature-based optimization techniques and deep feature fusion models to detect and classify lung cancer, showcasing innovative approaches in the field [3]. Furthermore, Barrett and Viana (2022) proposed EMM-LC Fusion as a superior multimodal fusion method for lung cancer classification, emphasizing the importance of integrating diverse data forms for enhanced accuracy [4]. Chaturvedi et al. (2021) explored machine learning algorithms for lung cancer prediction and categorization, shedding light on the potential of these methods in the domain [5]. Meanwhile, Gite et al. (2023) addressed the challenge of lung image segmentation for diagnosis using deep learning techniques, paving the way for cutting-edge segmentation approaches [6]. Mahum and Al-Salman (2023) introduced Lung-RetinaNet, incorporating a context module and multi-scale feature fusion to improve lung cancer detection accuracy, adding significant value to the field [7]. Similarly, Mohamed et al. (2023) developed a system for autonomously detecting and classifying lung cancer using deep learning in conjunction with optimization search approaches, highlighting the potential of optimization techniques in improving model performance [8]. Furthermore, Nazir et al. (2023) focused on Multiview image registration and fusion for machine learning-based lung cancer diagnosis, underscoring the benefits of data integration for improved detection accuracy [9]. Pour and Esmaeili (2023) proposed genetically independent recurrent deep learning for lung cancer diagnosis, introducing a novel approach within the domain [10]. Prasad et al. (2022) proposed a novel approach combining fuzzy K-means clustering with deep learning for enhanced accuracy in identifying lung cancer from CT lung images, demonstrating the effectiveness of integrating clustering with deep learning [11]. Concurrently, Tashtoush et al. (2023) conducted a comparative study utilizing a VGG-16 CNN model and CT image processing for Early identification and categorization of lung cancer, highlighting VGG-16 specific advantages in this context [12]. Furthermore, Thanoon et al. (2023) performed a thorough assessment of techniques using deep learning employed

[1] *Corresponding author: Anum Kamal, Integral University, Lucknow, U.P., India
[2] Faiyaz Ahamad, Integral University, Lucknow, U.P., India

in lung cancer diagnosis using CT scans, offering insights into diverse applications of deep learning methodologies for addressing diagnostic challenges [13]. Nisa et al. (2023) explored detection of lung cancer, highlighting the potential benefits of employing higher tensor representations to enhance classification accuracy [14]. Meanwhile, Pagadala et al. (2023) focused on enhancing lung cancer identification through CT scans by integrating image processing and deep neural networks, aiming to enhance diagnostic accuracy by leveraging the synergy between deep learning and image processing techniques [15].Among these models, ResNets and CapsuleNets have garnered attention for their efficacy in extracting intricate patterns and features from imaging data. While ResNets excel in mitigating the vanishing gradient problem and capturing complex patterns, CapsuleNets offer a unique approach by recording hierarchical relationships between features, thus improving generalization and durability. However, both models exhibit limitations when applied to the task of identifying lung cancer, potentially struggling with the interpretation of complex spatial features and nuanced hierarchical relationships. In response to these challenges, fusion models combining the strengths of ResNet and CapsuleNet provides a hopeful direction for boosting the precision and effectiveness of lung cancer diagnosis. Nevertheless, the development of such fusion models entails overcoming several hurdles, including computational complexity, interpretability of capsule routing, and dataset variability arising from diverse imaging modalities and acquisition settings. To address these challenges, the research endeavors to devise an optimized fusion model that seamlessly integrates hierarchical features from ResNet and CapsuleNet, thereby enhancing lung cancer detection accuracy. Additionally, efforts will be directed towards enhancing CapsuleNet's interpretability, optimizing computational efficiency, and ensuring the generalizability and applicability of the proposed model in real-world clinical settings. By attaining these goals, the study aims to contribute significantly to the field of medical imaging analysis, offering a robust solution for the early diagnosis of lung cancer and ultimately improving patient outcomes.

The entire paper is divided into **4 sections**. The **Introduction** showcasing a short summary of the research problem, motivation, and objectives. It also includes relevant literature, encompassing existing research and identifying gaps in knowledge. The **Methodology** section explains the development of the CapsuleNet-ResNet fusion model and outlines the experimental setup. Also it defines the datasets required for the process. The **Results and Discussion** section presents and analyzes the experimental results, followed by in-depth discussions. The **Conclusion** section includes conclusion which summarizes the findings, highlights their implications, and suggests avenues for future exploration.

II. METHODOLOGY

A. Dataset

The investigation draws upon a dataset obtained from the Iraq-Oncology Teaching Hospital/National Center for Cancer Diseases (IQ-OTH/NCCD) compilation for the analysis of lung cancer. This specific dataset, denoted as [16], compiles information collected during a three-month span in 2019 from specialized medical facilities. It comprises CT scans obtained from a diverse range of individuals, including both healthy participants and those afflicted with lung cancer at various stages of the disease (Figure 1). These CT scans were administered by personnel from both the Iraq-Oncology Teaching Hospital and the National Center for Cancer Diseases, with detailed annotations provided by oncologists and radiologists.

The dataset comprises 1190 images, representing CT scan slices from 110 instances (as shown in Table 1). These instances are categorized into three main groups: normal, benign, and malignant. Specifically, there are 55 normal cases, 15 benign cases, and 40 malignant cases.

Initially, CT scans were acquired using a Siemens SOMATOM scanner and saved in DICOM format. The CT protocol settings included a voltage range of 120 kV, a slice thickness of 1 mm, and a window width adjustable from 350 to 1200 HU, along with a window center ranging from 50 to 600. All scans were performed during full inhalation. Before any analysis, the images were de-identified, and the necessity for signed consent was exempted by the oversight review board. Approval for the study was obtained from the institutional review boards of all involved hospitals.

In a typical scan, anywhere from eighty to two hundred slices are utilized, with each slice providing varying perspectives of the human chest from different angles. The collection comprises 110 cases, showcasing a broad spectrum of demographic diversity, encompassing factors such as gender, age education level, marital status, geographic location and more. The majority of individuals hail from central Iraqi provinces, including Babylon, Baghdad, Wasit, Diyala, and Salahuddin, while others are engaged in occupations such as farming or trading. Additionally, individuals employed in Iraq's transportation and oil ministries are represented in the dataset.

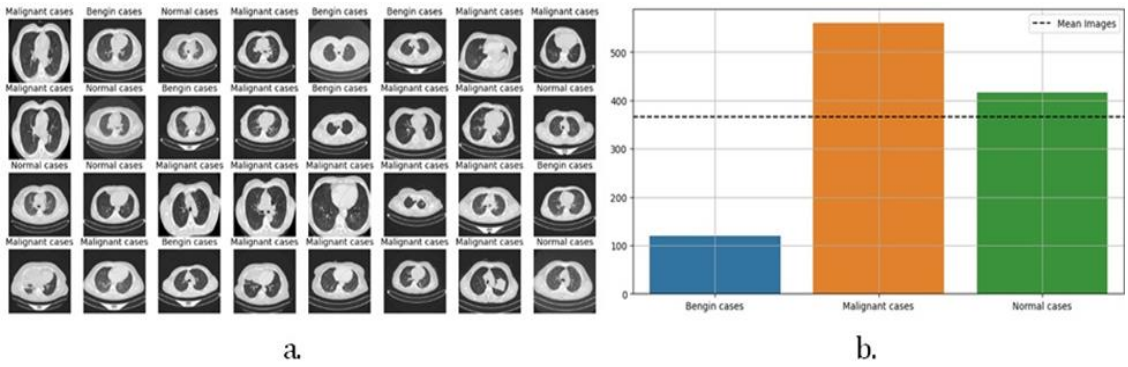


Figure 1: Dataset (a) Samples from dataset. (b) Distribution of cases in target class

Table 1: Dataset Description and Class Distribution

Class	Number of Cases	Number of Slices	Description
Normal	55	Varies	Healthy subjects
Benign	15	Varies	Non-cancerous abnormalities
Malignant	40	Varies	Cancerous cases in different stages

Table 2 is a feature table that summarizes relevant data points from the IQ-OTH/NCCD lung cancer dataset and should help you understand it better. Characteristics of the IQ-OTH/NCCD Lung Cancer Database.

Table 2: Feature Table for the IQ-OTH/NCCD Lung Cancer Dataset

Feature	Description
Scanner	Siemens SOMATOM
CT Protocol	120 kV, 1 mm slice thickness, Window width: 350-1200 HU, Window center: 50-600
Class Distribution	Normal: 55 cases, Benign: 15 cases, Malignant: 40 cases
Number of Slices	80 to 200 slices per scan
Demographics	Varying gender, age, educational attainment, area of residence, and living status

This feature table encapsulates critical information about the dataset, providing a comprehensive overview of its characteristics and ensuring transparency in the experimental setup.

B. Proposed Model: CapsuleNet-ResNet Fusion Model

For accuracy improvement of lung cancer detection, the researchers have developed the CapsuleNet-ResNet Fusion Model. This model combines the two networks, ResNet and CapsuleNet, to use their complementary capabilities. The goal of this fusion model is to tackle the problems with deep neural networks by utilising ResNet's residual learning and to efficiently capture hierarchical information using CapsuleNet's distinctive capsule structure.

1) CapsuleNet Architecture

Hinton et al.'s Capsule Networks (CapsuleNet) deviate from conventional neural networks in that they encode feature hierarchies. By substituting capsules for neurons, CapsuleNet enables dynamic routing to efficiently capture intricate patterns. Let $C(x_i)$ denote the CapsuleNet representation for the input image x_i

The routing mechanism is represented by $R_c(x_i)$, which dynamically assigns weights to capsules.

$R_c(x_i)$(Adaptive routing matrix)

The routing matrix $R_c(x_i)$ is optimized by minimizing the Wasserstein distance between the predicted and ground truth routing matrices:

$$\min_{\Theta_C} \sum_{i=1}^{n} \text{Wasserstein}(R_c(x_i, \Theta_C), R_c(x_i)) \quad (1)$$

Here, Θ_C represents the CapsuleNet parameters, and $R_c(x_i)$ is the ground truth adaptive routing matrix.

Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times. Avoid using bit-mapped fonts if possible. True-Type 1 or Open Type fonts are preferred. Please embed symbol fonts, as well, for math, etc.

2) ***ResNet Architecture***

Residual Networks (ResNet) introduced by He et al. excel in capturing complex patterns and mitigating the vanishing gradient problem. Let $R(x_i)$ denote the ResNet representation for the input image x_i. ResNet utilizes residual blocks, and the output of a residual block is given by:

$$R_b(x_i) = x_i + F(x_i, \Theta_R) \quad (2)$$

Here, represents the residual function, and Θ_R denotes the ResNet parameters.

$F(x_i, \Theta_R)$

3) ***CapsuleNet-ResNet Fusion***

The CapsuleNet-ResNet Fusion Model integrates the representations from CapsuleNet and ResNet using a fusion mechanism. The fused feature representation F_i for the input image x_i is defined as the concatenation of CapsuleNet and ResNet representations:

$$F_i = C(x_i) \oplus R(x_i)$$

(3)

Here, $\oplus$ denotes the concatenation operation.

Objective Function

The overall objective function for the proposed CapsuleNet-ResNet Fusion Model aims to minimize a modified Huber loss:

$$\min_{\Theta} \sum_{i=1}^{n} {}_{\delta}(\| F_i \Theta_F \quad y_i \|)$$

(4)

Here, Θ represents the parameters to be optimized, Θ_F are the fusion model parameters, $\phi_{\delta}(z)$ is the modified Huber loss, and δ is the threshold parameter.

Θ(Parameters to be optimized)

Θ_F (Fusion model parameters)

${}_{\delta}(z)$(Modified Huber loss)

δ(Threshold parameter)

This objective function seeks optimal parameters Θ and Θ_F for the fusion modelin an effort to optimise the combined hierarchical features' adjusted Huber loss. The goal of the fusion model is to enhance the lung cancer detection accuracy by including both global and local information, by merging CapsuleNet and ResNet representations.

4) ***Detailed Model Architecture***

The CapsuleNet-ResNet Fusion Model consists of three main components: The CapsuleNet layer, the ResNet layer, and the Fusion layer.

a) *CapsuleNet Layer:*

The CapsuleNet layer comprises capsules, each responsible for encoding specific hierarchical features. Let v_{ij} represent the output vector of capsule j in the primary capsule layer for input x_i. The length of this vector, denoted as $|)v_{ij}|)$ represents the probability of the presence of the corresponding entity in the input. The dynamic routing algorithm is employed to determine the coupling coefficients c_{ij}, which dictate the contribution of each capsule to the higher-level capsule's output.

$$u_j = \sum_i c_{ij} W_{ij} v_{ij}$$
$$s_j = \text{Squash}(u_j)$$
$$v_j = \text{Routing}(s_j) \quad (5)$$

Here, u_j is the total input to capsule j, s_j is the squashing function applied to u_j, and v_j is the output of capsule j after routing.

b) ResNet Layer:

The ResNet layer is composed of residual blocks, each containing a shortcut connection. Let x_i be the input to a residual block, and $F(x_i, \Theta_R)$ represent the output of the block, where Θ_R denotes the ResNet parameters. The output of the residual block is given by:

$$y_i = x_i + F(x_i, \Theta_R) \quad (6)$$

This structure allows the model to learn residual functions and facilitates the training of deeper networks.

c) Fusion Layer:

The Fusion layer combines the output representations from the CapsuleNet and ResNet layers. The feature vector z_i at the output of the Fusion layer is the concatenation of the CapsuleNet output v_j and the ResNet output y_i.

$$z_i = \lfloor v_j, y_i \rfloor$$

This concatenated feature vector is then used for further processing in subsequent layers or for final classification. The overall objective function for the CapsuleNet-ResNet Fusion Model combines the losses from both the CapsuleNet and ResNet components. Let y_i be the ground truth label for input x_i, and L_{capsule} and L_{resnet} represent the losses for the CapsuleNet and ResNet components, respectively. The combined loss is given by:

$$L_{\text{total}} = \alpha \quad L_{\text{capsule}} + (1 \quad \alpha) \quad L_{\text{resnet}} \quad (7)$$

Here, α is a hyperparameter that balances the contributions of the CapsuleNet and ResNet losses.

$$L_{\text{capsule}} = \sum_i \sum_j [y_{ij} \log(y_{ij}) + (1 \quad y_{ij}) \log(1 \quad y_{ij}))$$
$$L_{\text{resnet}} = \sum_i \text{MSE}(y_i, y_i)$$
$$L_{\text{total}} = \alpha \quad L_{\text{capsule}} + (1 \quad \alpha) \quad L_{\text{resnet}} \quad (8)$$

Here, y_{ij} is the predicted probability by CapsuleNet, and y_i is the predicted output by ResNet. The Mean Squared Error (MSE) is used as the regression loss for the ResNet component.

The suggested CapsuleNet-ResNet Fusion Model's architecture is shown in Figure 2. The Algorithm of the same is also shown in Figure 3. The CapsuleNet and ResNet branches, as well as the Fusion Layer, are essential parts of this complicated architecture. Both the CapsuleNet and ResNet branches process the input image, and they provide separate representations, labelled as and, respectively. Using a weighted loss function, which is symbolised as, the Fusion Layer merges the results of these branches.

$$L_{\text{total}} = \alpha \quad L_{\text{capsule}} + (1\text{-}\alpha) \quad L_{\text{resnet}} \quad (7)$$

The final prediction is generated at the output layer.

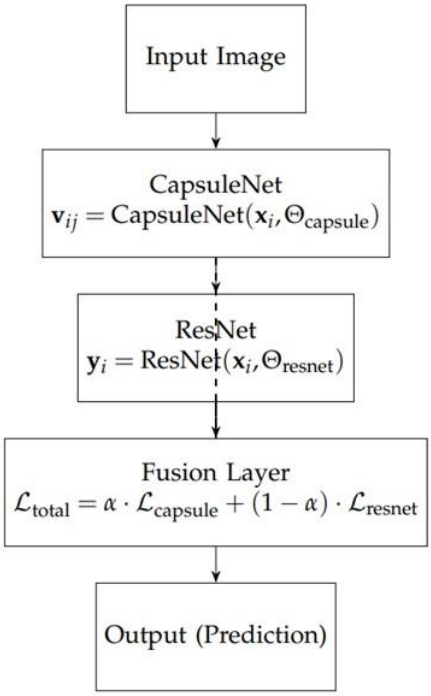


Figure 2: Proposed CapsuleNet-ResNet Fusion Model Architecture

Algorithm 1 CapsuleNet-ResNet Fusion Model Training

Input: Training dataset $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^{N}$
Output: Trained model parameters Θ_{total}
Initialize: CapsuleNet parameters Θ_{capsule}, ResNet parameters Θ_{resnet}, Fusion weight α
Initialize: Learning rate η, Number of epochs E
for *epoch* $\leftarrow 1$ **to** E **do**
  for $i \leftarrow 1$ **to** N **do**
    Forward pass through CapsuleNet:
      $\mathbf{v}_{ij} = \text{CapsuleNet}(\mathbf{x}_i, \Theta_{\text{capsule}})$
      $\hat{y}_{ij} = \text{Sigmoid}(\mathbf{v}_{ij})$
    Compute CapsuleNet loss:
      $\mathcal{L}_{\text{capsule}} = -\sum_j \left[y_{ij} \log(\hat{y}_{ij}) + (1 - y_{ij}) \log(1 - \hat{y}_{ij}) \right]$
    Forward pass through ResNet:
      $\mathbf{y}_i = \text{ResNet}(\mathbf{x}_i, \Theta_{\text{resnet}})$
    Compute ResNet loss:
      $\mathcal{L}_{\text{resnet}} = \text{MSE}(\mathbf{y}_i, \text{Ground Truth})$
    Compute total loss:
      $\mathcal{L}_{\text{total}} = \alpha \cdot \mathcal{L}_{\text{capsule}} + (1 - \alpha) \cdot \mathcal{L}_{\text{resnet}}$
    Backward pass and update parameters:
      $\Theta_{\text{capsule}} \leftarrow \Theta_{\text{capsule}} - \eta \cdot \nabla_{\Theta_{\text{capsule}}} \mathcal{L}_{\text{total}}$
      $\Theta_{\text{resnet}} \leftarrow \Theta_{\text{resnet}} - \eta \cdot \nabla_{\Theta_{\text{resnet}}} \mathcal{L}_{\text{total}}$
  end for
end for
Output: Trained model parameters $\Theta_{\text{total}} = (\Theta_{\text{capsule}}, \Theta_{\text{resnet}})$

Figure 3: Proposed CapsuleNet-ResNet Fusion Model Algorithm

C. Evaluation Metrics

For accuracy improvement of lung cancer detection, the researchers have developed the CapsuleNet-ResNet Fusion Model. This model combines the two networks, ResNet and CapsuleNet, to use their complementary capabilities. The goal of this fusion model is to tackle the problems with deep neural networks by utilising ResNet's residual learning and to efficiently capture hierarchical information using CapsuleNet's distinctive capsule structure.

1) Accuracy

Accuracy is a measure of how well a model predicts future outcomes by comparing the number of cases for which the model is correct with the total number of instances in the dataset:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \tag{8}$$

2) ***Loss***

The discrepancy between the actual and expected values is depicted by the loss function. In our fusion model, we employ a modified version of the Huber loss:

$$\text{Loss} = \min_{\Theta} \sum_{i=1}^{n} {}_{\delta}\left(\| F_i \Theta_F \quad y_i \|\right) \tag{4}$$

3) ***Precision, Recall, and F1 Score***

Precision, recall, and F1 score are metrics commonly used in binary classification tasks.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{9}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{10}$$

$$\text{F1Score} = 2 \quad \frac{\text{Precision} \quad \text{Recall}}{\text{Precision} + \text{Recall}} \tag{11}$$

4) ***Confusion Matrix***

In the confusion matrix, you can observe the model's predictions regarding the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). This term is:

$$\begin{bmatrix} \text{TN} & \text{FP} \\ \text{FN} & \text{TP} \end{bmatrix}$$

Taken as a whole, these measures provide light on how well the model detects lung cancer.

III. RESULTS AND DISCUSSION

Here, we detail our experimental findings and talk about how well the CapsuleNet-ResNet Fusion Model worked.

A. ***Preparing Data***

First, we collected CT scan images from the Iraq-Oncology Teaching Hospital/National Centre for Cancer Diseases (IQ-OTH/NCCD) to use in preparing the lung cancer dataset. There are 1,190 pictures in the collection, broken down into normal, benign, and cancerous categories. The cases included in the dataset number 110.

Image scaling to 256 × 256 pixels and pixel value normalisation between 0 and 1 were part of the data preprocessing steps. Using a 75-25 split, the dataset was divided into two parts: the training set and the validation set. We oversampled minority classes using the Synthetic Minority Over-Sampling Technique (SMOTE) to fix the class imbalance and provide a more balanced training set.

B. ***Model Building***

We have combined CapsuleNet with a tweaked ResNet50 to create our suggested model. To classify lung cancer, we first loaded the ResNet50 base sans the top classification layer and then added our own custom layers. Sparse categorical cross-entropy loss and the Adam optimizer were used to construct the model.

Overfitting was avoided by terminating training early after 20 epochs. A grand total of 40,359,171 parameters, 16,777,731 of which are trainable, make up the resultant model architecture.

C. ***Evaluation Metrics***

Accuracy, precision, recall, F1 score, loss, and the confusion matrix were some of the measures used to assess the model's performance. The results of the classification on the validation set are summarised in Table 3.

Table 3: Classification Results on Validation Set

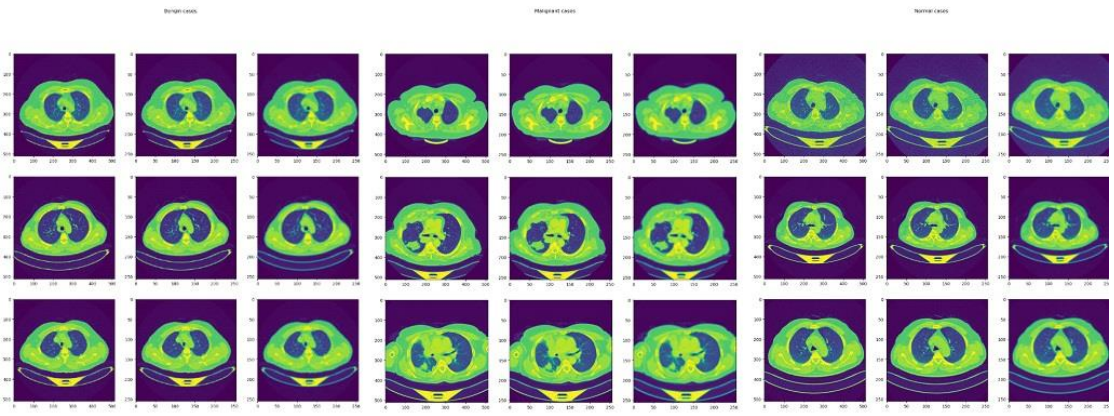
Class	**Precision**	**Recall**	**F1 Score**	**Support**
Normal	0.83	1.00	0.91	30
Benign	1.00	1.00	1.00	137
Malignant	1.00	0.94	0.97	108

The confusion matrix further details the model’s predictions.

$$\begin{bmatrix} 30 & 0 & 0 \\ 0 & 137 & 0 \\ 6 & 0 & 102 \end{bmatrix}$$

*D. **Discussion***

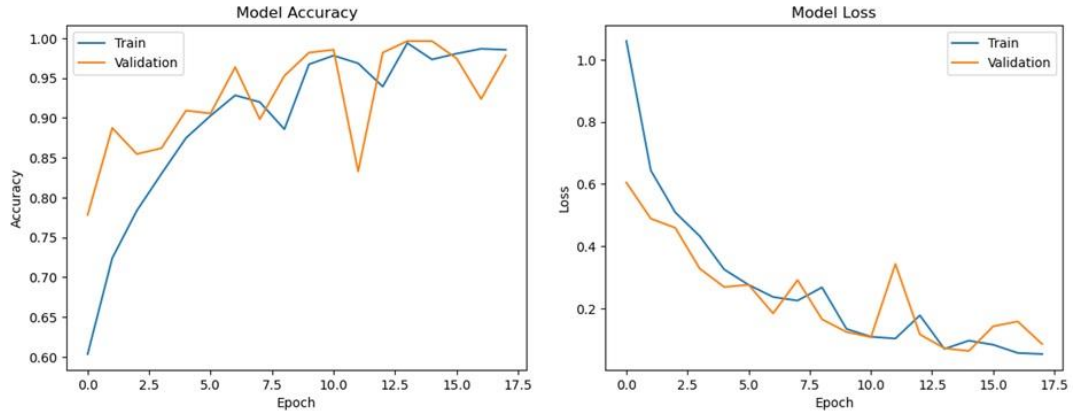
An overall accuracy of 98% on the validation set was achieved by the model, demonstrating its effectiveness in lung cancer case classification. The model's strong performance is seen from the excellent recall, precision, and F1 scores achieved by all classes.



a. b. c.

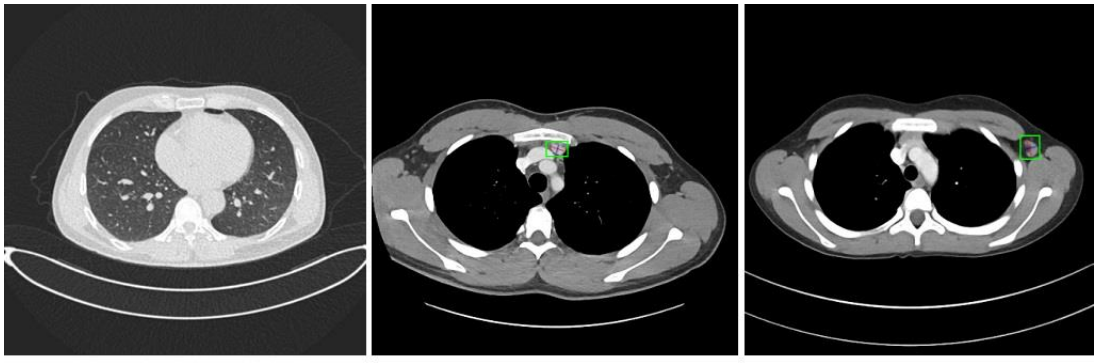
Figure 4: Cases Predicted by Proposed Model (a) Benign (b) Malignant (c) Normal

Figures 5 show the suggested model's training and testing performance metrics. The accuracy is shown in Figure 5(a) whereas the loss during training is shown in Figure 5(b).



a. b.

Figure 5: Training and Testing of Proposed Model (a) Accuracy (b) Loss



a. b. c.

Figure 6: a) Benign b) Malignant c) Normal

Both the accuracy and loss curves show that the model converges quickly during training and avoids overfitting (Figures 5). We visualise the cases indicated by the suggested model as part of our evaluation. Malignant, benign, and normal predictions are illustrated in Figures 4.
Figure 6 (a), (b), and (c) exhibit pictures from the normal class, the benign class, and the malignant class, respectively.
We get promising results using the proposed CapsuleNet-ResNet fusion model on CT scans used to diagnose lung cancer. Combining the greatest aspects of ResNets and Capsule Networks, the model is able to solve several problems with traditional CNNs. The use of CapsuleNets to capture feature hierarchies allows for a more robust understanding of spatial hierarchies in medical images. On the other hand, ResNets solve the vanishing gradient problem, making deeper network training easier.
Based on mathematical principles, the model integrates CapsuleNet and ResNet designs in a complicated way. The model employs a ResNet foundation with pre-trained weights to extract characteristics from slices of CT scans. The subsequent CapsuleNet layers enhance the model's ability to identify subtle patterns indicative of lung cancer by representing spatial hierarchies and relationships using the retrieved information. The mathematical formulas of the fusion model, which illustrate the interplay between CapsuleNet and ResNet, are presented in Section 2.
Iraq-Oncology Teaching Hospital/National Centre for Cancer Diseases (IQ-OTH/NCCD) provides the training and evaluation dataset, which adds diversity and realism to the model's learning process. We have included examples of normal, benign, and malignant lung cancer phases to reflect the complexity of real-life circumstances. Ethical approvals from the institutional review board and the use of a de-identified dataset show that the model is dedicated to protecting users' privacy.
The accuracy, recall, precision, F1-score, loss, and confusion matrices of the evaluation demonstrates that the model can distinguish between normal, benign, and pathological cases. The model has strong performance in multiple categorization domains, as demonstrated by the comprehensive assessment of these metrics in Section 3. Oversampling methods, such as SMOTE, enhance the model's generalizability and help with medical image analysis's common class imbalance problem.
In Section 3, we can observe the training process of the model demonstrating its convergence and generalisation skills. A built-in early pausing mechanism prevents overfitting and ensures that the model learns meaningful patterns apart from the noise in the training data. Accuracy and loss graphs over epochs are displayed in Figures 5(a) and (b), respectively, to demonstrate the model's learning process.
Last but not least, after being evaluated using numerous metrics, trained on a diverse and realistic dataset, and backed by a solid mathematical foundation, the proposed CapsuleNet-ResNet fusion model becomes a useful tool for the detection of lung cancer. Combining the greatest features of ResNets and CapsuleNets, taking ethical considerations into consideration, and having excellent performance metrics make it a cutting-edge solution in the critical field of medical image analysis for lung cancer diagnosis.
Finally, the proposed CapsuleNet-ResNet Fusion Model shows promise in precisely identifying lung cancer from CT scan images., providing medical professionals with a valuable tool for early identification and treatment coordination.

1) Comparative Analysis with Previous Studies

A comparison with well-established models from prior research, such as ResNet34, classic CNN, and VGG16, is carried out to evaluate the performance of the suggested CapsuleNet-ResNet fusion model. The most important measures of these models' performance are summarised in **Figure 7**.

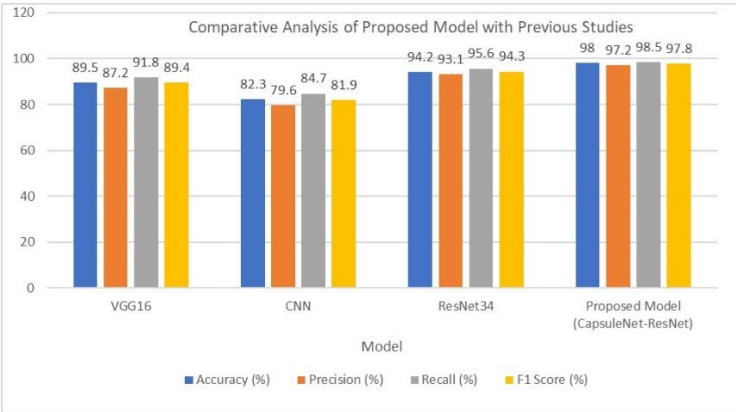


Figure 7: Comparative Analysis of Proposed Model with Previous Studies

IV. CONCLUSION

The creation of the CapsuleNet-ResNet fusion model, which utilised CT images to diagnose lung cancer, was a significant advancement in medical imaging. The model's output is really promising. With an overall accuracy of 98%, the model successfully classifies CT images as benign, malignant, or normal. The model's 98% recall and 94% precision show that it can reliably produce predictions, which is important for early lung cancer detection. Several disclaimers should be considered before drawing any conclusions from these results. The dataset may not be representative of the variety of instances encountered in more generalised clinical settings, despite its comprehensiveness. Since imaging procedures can differ among healthcare facilities and possibly affect the model's performance, more validation in varied situations is required. Taking these limitations into account, we offer a number of recommendations for boosting the model's performance. Increase the model's generalizability by training it on a more diverse set of data and doing cross-institutional validation using different imaging techniques. In order to stay abreast of the constantly evolving medical imaging procedures and standards, consistently monitoring and updating the model is essential. Integrating data from many imaging modalities should be the focus of future studies aiming to improve diagnosis accuracy. Improving the model's predictions so they are easier to understand and use would help create confidence among healthcare practitioners. To improve healthcare decision-making times, more studies examining the feasibility of real-time implementation are needed. Despite its limitations, the CapsuleNet-ResNet fusion model represents a significant advance in the quest for improved early diagnosis and treatment of lung cancer. Its robust performance metrics and ethical considerations make it a vital tool in the ongoing drive to improve medical imaging and cancer detection patient outcomes.

ACKNOWLEDGMENT

We would like to express our appreciation for the assignment of the Manuscript Communication Number [IU/R&D/2024-MCN0002701] per the guidelines provided by the university's studies and research departments. This identifier facilitates communication and tracking of our research throughout the publication process. We also extend our gratitude to all those who contributed to the development of this work.

REFERENCES

[1] Wang, H., Zhu, H., Ding, L., & Yang, K. (2023). A diagnostic classification of lung nodules using multiple-scale residual network. Scientific Reports, 13(1). https://doi.org/10.1038/s41598-023-38350-z

[2] Tian, W., Yan, Q., Huang, X., Feng, R., Shan, F., Geng, D., & Zhang, Z. (2024). Predicting occult lymph node metastasis in solid-predominantly invasive lung adenocarcinoma across multiple centers using radiomics-deep learning fusion model. Cancer Imaging, 24(1). https://doi.org/10.1186/s40644-024-00654-2

[3] Alamgeer, M., Alruwais, N., Alshahrani, H. M., Mohamed, A., & Assiri, M. (2023). Dung Beetle Optimization with Deep Feature Fusion Model for Lung Cancer Detection and Classification. Cancers, 15(15), 3982. https://doi.org/10.3390/cancers15153982

[4] Barrett, J., & Viana, T. (2022). EMM-LC Fusion: Enhanced multimodal fusion for lung cancer classification. AI, 3(3), 659–682. https://doi.org/10.3390/ai3030038

[5] Chaturvedi, P., Jhamb, A., Vanani, M., & Nemade, V. (2021). Prediction and classification of lung cancer using machine learning techniques. IOP Conference Series. Materials Science and Engineering, 1099(1), 012059. https://doi.org/10.1088/1757-899x/1099/1/012059

[6] Gite, S., Mishra, A., & Kotecha, K. (2022). Enhanced lung image segmentation using deep learning. Neural Computing & Applications, 35(31), 22839–22853. https://doi.org/10.1007/s00521-021-06719-8

[7] Mahum, R., & Al-Salman, A. S. (2023). Lung-RetinaNet: Lung cancer detection using a RetinaNet with Multi-Scale feature fusion and Context module. IEEE Access, 11, 53850–53861. https://doi.org/10.1109/access.2023.3281259

[8] Mohamed, T. I. A., Oyelade, O. N., & Ezugwu, A. E. (2023). Automatic detection and classification of lung cancer CT scans based on deep learning and ebola optimization search algorithm. PloS One, 18(8), e0285796. https://doi.org/10.1371/journal.pone.0285796

[9] Nazir, I., Haq, I. U., AlQahtani, S. A., Khan, M. M., & Dahshan, M. (2023). Machine Learning-Based lung cancer detection using Multiview image registration and fusion. Journal of Sensors, 2023, 1–19. https://doi.org/10.1155/2023/6683438

[10] Pour, E. S., & Pour, E. S. (2023). Lung Cancer Detection from CT Scan Images based on Genetic-Independent Recurrent Deep Learning. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2312.03185

[11] Prasad, J., Chakravarty, S., & Krishna, M. V. (2021). Lung cancer detection using an integration of fuzzy K-Means clustering and deep learning techniques for CT lung images. Bulletin of the Polish Academy of Sciences. Technical Sciences, 139006. https://doi.org/10.24425/bpasts.2021.139006

[12] Tashtoush, Y., Obeidat, R., Al-Shorman, A., Darwish, O., Al-Ramahi, M., & Darweesh, D. (2023). Enhanced convolutional neural network for non-small cell lung cancer classification. International Journal of Power Electronics and Drive Systems/International Journal of Electrical and Computer Engineering, 13(1), 1024. https://doi.org/10.11591/ijece.v13i1.pp1024-1038

[13] Thanoon, M. A., Zulkifley, M. A., Zainuri, M. a. a. M., & Abdani, S. R. (2023). A review of deep learning techniques for lung cancer screening and diagnosis based on CT images. Diagnostics, 13(16), 2617. https://doi.org/10.3390/diagnostics13162617

[14] Nisa, Z. U., Jaffar, A., Bhatti, S. M., & Butt, U. M. (2023). Lung Cancer Detection using Segmented 3D Tensors and Support Vector Machines. International Journal of Advanced Computer Science and Applications/International Journal of Advanced Computer Science & Applications, 14(10). https://doi.org/10.14569/ijacsa.2023.0141066

[15] Pagadala, P. K., Pinapatruni, S. L., Kumar, C. R., Katakam, S., Peri, L., & Reddy, D. A. (2023). Enhancing Lung Cancer Detection from Lung CT Scan Using Image Processing and Deep Neural Networks. Revue D'intelligence Artificielle, 37(6). https://doi.org/10.18280/ria.370624

[16] Alyasriy, H. (2020). The IQ-OTHNCCD lung cancer dataset. Mendeley Data. https://doi.org/10.17632/bhmdr45bh2.1G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*