Few-Shot Lung Cancer Classification Using Prototypical Networks

Nimesha Hansani Amarasinghe
Department of Computer Science and Engineering
University of Moratuwa
Moratuwa, Sri Lanka
nimesha.21@cse.mrt.ac.lk

Thanuja D. Ambegoda
Department of Computer Science and Engineering
University of Moratuwa
Moratuwa,Sri Lanka
thanuja@cse.mrt.ac.lk

***Abstract*—Lung cancer, a leading cause of mortality globally, demands early and accurate detection to improve patient outcomes. Current diagnostic methods, primarily relying on CT scans, face challenges in lung cancer subtype identification, particularly due to the scarcity of extensive medical image datasets for each subtype. This study addresses this limitation by leveraging prototypical networks, an innovative few-shot learning approach, which excels in scenarios with limited data. The proposed method capitalizes on a small number of samples per category, integrating a pre-trained model for feature extraction from lung CT scans. We rigorously evaluated the model's performance, focusing on its accuracy relative to the sample size per category. Remarkably, the method achieved a 98% accuracy rate after 15 epochs, showcasing its efficacy. This research not only confirms the feasibility of using prototypical networks for lung cancer subtype classification but also opens new avenues for applying few-shot learning techniques in medical imaging. Our findings hold significant potential for enhancing lung cancer diagnostics, thereby contributing to improved patient care and survival rates.**

***Keywords*—CT Image, U-net Architecture, VGG16, ResNet50, DenseNet**

I. Introduction

Lung cancer stands as a critical global health crisis, responsible for the deaths of hundreds of people daily worldwide. It is one of the most lethal and prevalent cancers, typically developing in the lung tissues. Accurate identification of the type of lung cancer is crucial, as it significantly influences the patient's treatment and prognosis. This underscores the urgent need for more precise diagnostic methods and therapeutic interventions.

The American Cancer Society categorizes lung cancer into two primary types: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) [1]. NSCLC is further divided into three major subtypes: Adenocarcinoma, Squamous cell carcinoma, and Large cell carcinoma [1]. SCLC, however, is a separate classification altogether. Differentiating these types is vital for developing tailored treatment plans and enhancing patient survival rates.

In practical applications, encountering cases for each of these types is rare, often presenting as singular instances or just a handful in training datasets. The feasibility of collecting extensive medical images for each subtype is low. Challenges also arise due to privacy and data security concerns, hindering the acquisition of the large datasets typically required for thorough research.

979-8-3503-8486-4/24/$31.00 ©2024 IEEE

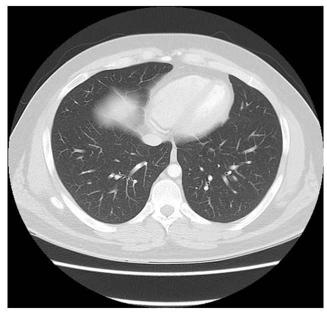


Fig. 1: CT scan of a healthy person. Adapted from [2]

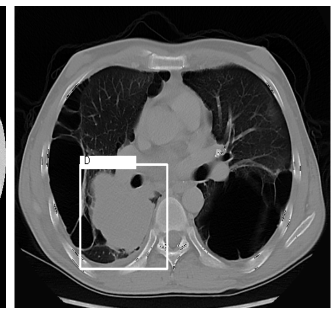


Fig. 2: CT scan of a patient with Large Cell Carcinoma. Adapted from [3]

Computer-Aided Detection/Diagnosis (CADe/CADx) systems are increasingly utilized in the diagnosis of lung cancer due to their precision, efficiency, and rapid processing abilities. These systems analyze detailed cross-sectional images of the chest, facilitating the identification of small nodules and anomalies within the lung tissue. However, a challenge with many CADe/CADx systems is that they are often trained on datasets predominantly featuring common types of lung cancer. While these systems are effective for identifying and classifying these common cancer types, they may not encompass the full clinical and radiological spectrum of lung cancer, potentially limiting their effectiveness in diagnosing less common or atypical lung cancer types.

Recent advancements in image processing and deep learning have shown promise in medical image analysis for disease diagnosis using CT scans. Various deep learning methods for lung cancer detection using CT images have been proposed, leveraging large labeled datasets [4]–[8].

However, applying deep learning for lung cancer type classification is a daunting task. To address this, the emerging field of few-shot learning in machine learning shows potential.

Few-shot learning aims to achieve effective learning outcomes with limited labeled data in the training dataset, which includes instances of inputs paired with their corresponding outcomes.

In few-shot classification, the support set and query set are key components. The support set, containing a small number of examples per class, aids in training the model. For instance, a 4-way 5-shot setting indicates five examples for each of the four classes. The query set is used for testing the model's ability to generalize. It includes new examples not seen during training, and the model must classify them accurately based on the knowledge gained from the support set.

Prototypical Networks, introduced by Snell et al [9], have significantly advanced the field of few-shot learning. These networks are adept at classifying small sample sizes by forming a class prototype through the mean of features within a class, and then making predictions based on the similarity to new instances. Prototypical Networks have been applied in various domains, such as image recognition, natural language processing, and medical image analysis [10], [11], [12], [13]. In this study, the Prototypical Network was employed to effectively tackle the challenge of handling a limited number of examples per class.

II. Background

Few-shot learning, an emerging domain within machine learning, has gained prominence for its ability to tackle the challenges posed by limited data availability and enhance model generalization across various fields. This approach is primarily divided into two categories: non-meta learning and meta-learning methods. Among the latter, Prototypical Networks stand out as a notable meta-learning technique for few-shot learning. These networks utilize a nearest neighbor strategy, which simplifies the meta-testing stage by reducing the need for extensive hyper-parameter tuning, thereby enabling rapid inference.

Prototypical Networks are designed to facilitate model generalization with only a few examples available per category or class. This methodology has shown potential in medical image analysis, as illustrated by [13], who applied it to melanoma detection using a limited dataset of dermoscopy images. This underscores the utility of few-shot learning in scenarios where medical image data is scarce.Another notable application was presented by Yifan Jian et al. [11], who developed a novel approach for CT-based Coronavirus diagnostics using controlled domain adaptation techniques. This method is particularly advantageous in situations where only a limited number of named CT images are available, a frequent limitation in medical imaging studies.

Ahuja et al. [10] explored a P-shot n–ways Siamese network, combining deep learning principles with prototypical nearest neighbor classifiers. Their research focused on classifying COVID-19 infection in lung CT scans, examining how different pre-trained network CNN models influence the performance of Siamese-based multi-class classification networks.

TABLE I: Results from prior research involving few-shot learning applied to medical image analysis

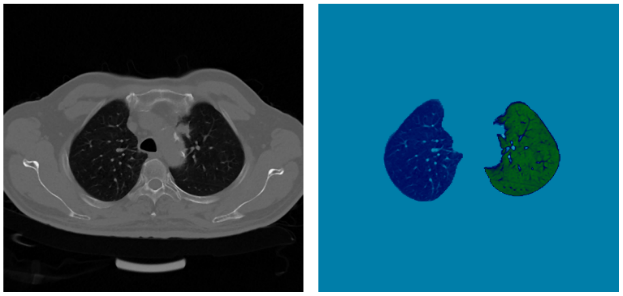
Domain	Dataset	Performance
Siamese network based model [11]	Covid-19 CT segmentation from https://medicalsegmentation.com/covid19	Accuracy 0.8040±0.0356 F1-score 0.7998±0.0384
Prototypical closest neighbors' classifiers combined with a P-shot N-ways Siamese network [10]	Chest CT scans from 1110 patients in medical hospitals from Moscow, Russia	Accuracy 98.07% F1-Score 95.10%
A model based on a few-shot U-Net architecture [14]	Lung-PET-CT-DX dataset in TCIA database. PET/CT scans from 87 patients	Accuracy 99% Precision 70.62%
Comparison using both Zero-shot learning and Few Shot Learning [15]	LC25000 dataset. 25,000 color images in 5 classes	99.87% of accuracy from few-shot setting

Nicholas et al. [14] innovated in dynamic few-shot learning for lung cancer lesion segmentation, merging few-shot learning's strengths with the U-Net architecture. Their approach diverged from traditional methods by integrating global neighborhood PET/CT method fusion, enhancing lung cancer detection and classification.

Meldo et al. [7] developed a Computer-Aided Diagnosis system using a Siamese neural network to differentiate various lung conditions and establish a comprehensive lung cancer classification. This network was trained on a dataset of segmented and labeled lung abnormalities, including tumors, categorized into distinct groups based on CT image patterns. The dataset exclusively contained confirmed tissues, validated through meticulous examinations, ensuring high data integrity and reliability.

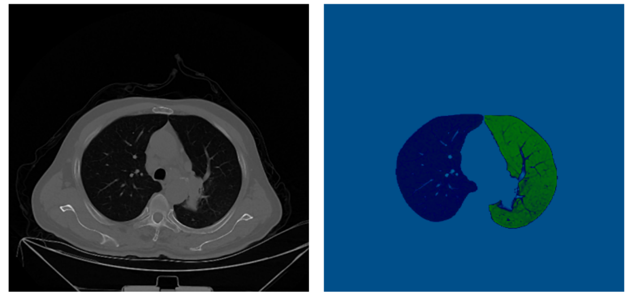
Fu-Ming Guo et al. [15] utilized a pre-trained Vision Transformer (ViT) for classifying lung cancer on histologic slices, employing both Zero-Shot and Few-Shot settings. Their study demonstrated the ViT model's remarkable performance, achieving an impressive 99.87% accuracy in the Few-Shot setting with limited epochs, showcasing the effectiveness of pre-trained models in handling complex classification tasks.

The research highlighted thus far illustrates the significant strides made in applying few-shot learning to lung cancer identification. However, these studies have not fully explored the potential of prototypical networks or the impact of different feature extraction methods and the integration of U-Net architecture for segmentation purposes. Addressing these gaps is crucial to advancing the understanding of lung cancer type identification through few-shot learning.This paper aims to delve deeper into the role of prototypical networks in lung cancer classification. It will also evaluate the efficacy of three pre-trained feature extraction methods and their synergy with convolutional neural networks within the context of few-shot



(a) Before Segmentation (b) After Segmentation

Fig. 3: The CT scan image depicting Squamous Cell Carcinoma before the application of the U-Net (R321) architecture for segmentation, followed by the post-segmentation result after applying U-Net. Adapted from [3]



(a) Before Segmentation (b) After Segmentation

Fig. 4: The CT scan image depicting Adenocarcinoma before the application of the U-Net (R321) architecture for segmentation, followed by the post-segmentation result after applying U-Net. Adapted from [3]

learning. This comprehensive analysis is expected to enrich the understanding of the nuances involved in lung cancer type identification, paving the way for more effective and nuanced diagnostic approaches in the future.

III. Methodology

A. Dataset

The study utilized CT scan images from the Lung-PET CT-Dx dataset housed in The Cancer Imaging Archive (TCIA), generously provided by the National Cancer Institute [16]. Accompanying XML Annotation documents provided crucial information on tumor locations via bounding boxes. This dataset, curated retrospectively, includes patients under lung cancer suspicion, all of whom underwent standard lung biopsies and PET/CT imaging.

The dataset comprises CT images of 355 subjects, each represented as a series of DICOM images. For class labeling within this dataset, a straightforward naming convention was adopted, based on the initial letter of the patient's name. Patients labeled with 'A' were classified as having adenocarcinoma, 'B' for small cell carcinoma, 'E' for large cell carcinoma, and 'G' for squamous cell carcinoma.

B. Image Pre-processing

The initial step involved extracting pixel data from DICOM images. These pixels, the fundamental elements of digital imaging, represent the luminosity of the captured scene. The subsequent step was converting these pixel values into Hounsfield Units, a standardized scale in medical imaging, which enhances image clarity and diagnostic accuracy.Noise reduction in CT images is crucial to prevent misinterpretation. Median filtering was applied to diminish noise. This process involves arranging neighboring pixel values in numerical order and replacing the central pixel with the median value.

C. Image Segmentation

Image segmentation involves dividing an image into distinct regions or classes based on specific characteristics.The U-net architecture was utilized for this study, known for its efficacy in medical image segmentation [14], [17], [18]. Johannes et al. [19] developed a modified U-net (R-231) model, which excelled in the LOLA11 challenge, achieving high scores in metrics like Dice similarity coefficient (DSC), Hearty Hausdorff distance (HD95), Mean Surface Distance (MSD), and cancer cross-over. Due to its superior performance, the U-net (R-231) model was chosen for segmenting preprocessed CT images. Figures 3 and 4 illustrate the output before the U-net(R-231) architecture was applied.

D. Feature Extraction

Feature extraction in medical imaging is crucial for the automated acquisition of complex image features, enhancing training representations. We employed three pre-trained deep learning models: VGG16 [20], ResNet50 [21], DenseNet [22], and a custom Convolutional Neural Network with three convolutional layers, ReLU activation, and a pooling layer. These models facilitated the extraction of nuanced features from lung CT images.

E. Prototypical Network

In the few-shot classification, a small support set of N labeled examples, each represented by a D-dimensional feature vector from a pre-trained model, was used. A prototype for each lung cancer type was calculated by averaging the feature vectors within a class. These prototypes represented the respective classes during inference.

A key element of the prototypical network is the distance matrix, which measures the similarity between the support and query sets. Euclidean distance, as suggested by Snell et al. [9], proved more effective than cosine similarity for this purpose. When employing Euclidean distance, the model effectively functions as a linear model with a specific parameterization.The overall architecture of the model is shown in figure 5.

F. Episodic Learning

The performance of the model was improved by implementing episodic learning. In this method, the model is directly tuned for few-shot classification tasks within each episode, defined as an iteration cycle or an epoch.

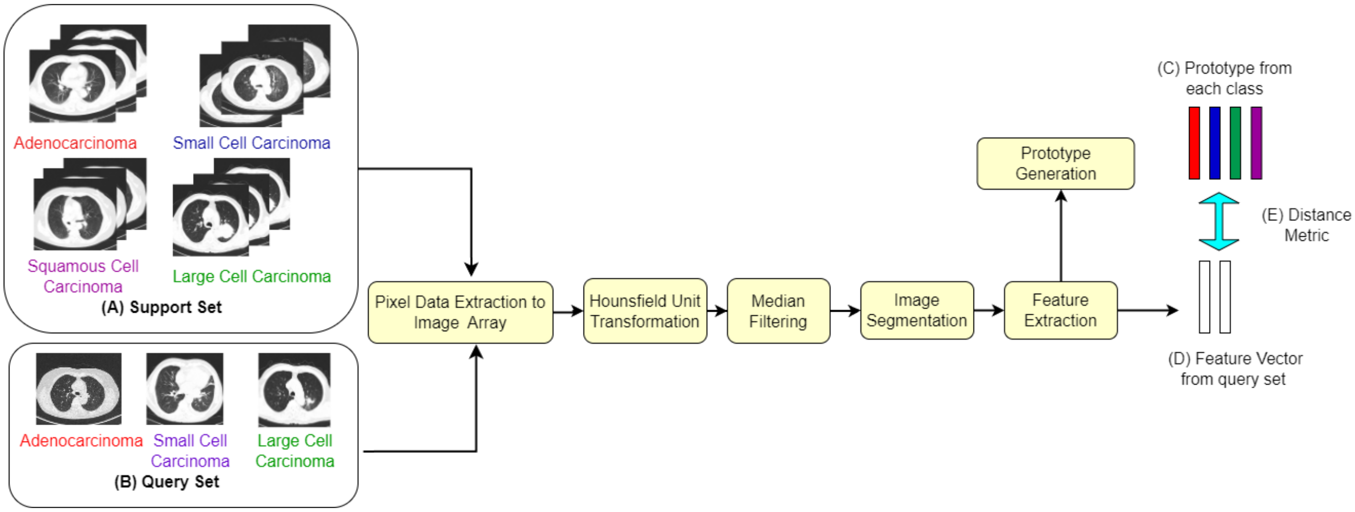


Fig. 5: Overall architecture of the prototypical network.(A) Support set has four classes with a limited number of shots.(B) Query set consists of the samples from the classes.(C) Prototype feature vector for each class is generated.(D) Feature vector from each query images.(E) Similarity between each prototype and the query image is calculated using the Euclidean distance.

During each epoch, the training set undergoes forward and backward propagation and is divided into several batches. The number of iterations per epoch corresponds to the number of batches processed. The support set consists of N-way k-shot random samples, and the query set comprises q random samples for each of the N classes in the support set. These datasets are generated anew in each episode. The model's meta-learner is trained over a series of episodes, learning from the limited dataset, a process termed meta-learning.For this study, the CrossEntropyLoss function was used to define the loss criterion, and the Adam optimizer was employed for updating the model parameters. The learning rate was set at 0.001, dictating the step size in the optimizer's parameter space.

IV. RESULTS & DISCUSSION

The experiment involved the utilization of three pre-trained deep learning models and a convolutional neural network as the feature extractors. The performance evaluation of each pre-trained model was conducted by assessing metrics such as Precision, Recall, F1-Score, and Accuracy score. This evaluation was conducted under varying conditions where the number of shots per class in the support set was altered, aiming to gauge how well the models perform under different scenarios of data availability for each class. Table II presents the Precision, Recall, and F1-Score for four different models, when applied to the prototypical network.

The results demonstrate how the model's performance changes with the number of labeled examples per class. Figure 6 displays the accuracy for each model with the number of lung CT images per class, indicating that the highest accuracy scores are attained for all models with more than 3-shot settings.

Notably, the highest accuracy observed was 98% for the ResNet50 model with a 5-Shot setting for each class. This superior performance can be attributed to several factors inherent to the ResNet50 architecture and its compatibility with few-shot learning. ResNet50 is known for its deep structure and residual connections, which allow it to learn complex features effectively, an advantageous trait for nuanced medical imaging tasks. Furthermore, ResNet50's ability to generalize from smaller datasets, crucial in few-shot settings, combined with the Prototypical Network's approach of learning a metric space for classification, likely contributed to its outstanding performance.

TABLE II: PERFORMANCE METRICS FOR INDIVIDUAL FEATURE EXTRACTION MODEL IN A K-SHOT 4-WAY SETTING.(EPOCHS=3)

Feature Extraction Model	**Shots Per Class = K**	**Precision**	**Recall**	**F1-Score**
VGG16	1-Shot	0.74	0.68	0.68
	2-Shot	0.82	0.81	0.81
	3-Shot	0.86	0.81	0.82
	4-Shot	0.77	0.75	0.74
	5-Shot	0.69	0.66	0.65
CNN	1-Shot	0.6	0.56	0.54
	2-Shot	0.88	0.87	0.87
	3-Shot	0.68	0.68	0.68
	4-Shot	0.95	0.93	0.93
	5-Shot	0.87	0.83	0.82
DenseNet	1-Shot	0.79	0.68	0.66
	2-Shot	0.95	0.93	0.93
	3-Shot	0.88	0.87	0.87
	4-Shot	0.87	0.83	0.82
	5-Shot	0.9	0.9	0.9
ResNet50	1-Shot	0.33	0.43	0.33
	2-Shot	0.22	0.31	0.25
	3-Shot	0.76	0.68	0.69
	4-Shot	0.71	0.62	0.6
	5-Shot	0.98	0.98	0.98

Comparative studies, such as those by Shivan et al. [23] and Shital et al [24], have also employed ResNet50, demonstrating its efficacy in similar tasks. Table III compares the results

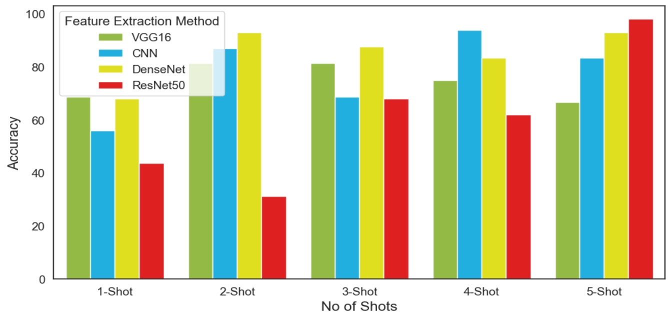


Fig. 6: Accuracy comparison for each pre-trained model before meta-training with 5-Shot 4-Way Setting.(Epochs=3)

Fig. 7: The graphs compares the training accuracy over 15 epochs for each feature extraction model. The chosen configuration utilized five shots per class and a learning rate of 0.001.

obtained by these studies with the outcomes from our research.

TABLE III: COMPARISON OF THE ACCURACY BETWEEN EXISTING METHODS UTILIZING RESNET50 FOR FEATURE EXTRACTION AND THE PROPOSED PROTOTYPICAL NETWORK.

Method	**Accuracy**
Convolutional Neural Network with RestNet50 [23]	97.05%
Support Vector Machine with ResNet50 [24]	97.53%
Prototypical Network with ResNet50	98%

Overall, the models achieve their best performance when there are more shots per class, particularly with 4 or 5-Shot classes. This trend underscores the importance of having a sufficient number of training examples to achieve higher accuracy and F1-Scores. Figure 7 shows the training accuracy and loss tracked over multiple episodes for each model. Each model was set to learn for 15 episodes, revealing interesting patterns in their learning processes.

The VGG16 model exhibited fluctuations in training accuracy, suggesting instability in its learning cycle, possibly due to variations in data availability or task complexity. In contrast, the CNN model displayed a consistent increase in training accuracy, indicative of a steady learning process, potentially due to its architectural efficiency and data handling. The ResNet50 model's training accuracy showed initial variability but a notable increase in later episodes, despite some inconsistencies. The DenseNet model initially experienced fluctuating training accuracy, which stabilized as training progressed, indicating effective adaptation to the dataset.

These observations highlight the complex nature of training neural networks, influenced by factors such as model architecture, data distribution, and convergence dynamics. The experiment, encompassing 15 training episodes with 5 support images and 3 query images from each class, demonstrated that both the ResNet50 and CNN models achieved impressive accuracy, reaching 98%.

The research was conducted using a single dataset from TCIA, providing valuable insights into the application of prototypical network architecture for lung cancer classification. However, the reliance on a single dataset may affect the generalizability of the results. Future studies could benefit from incorporating diverse lung CT image datasets labeled with specific lung cancer types, enhancing the reliability and broader applicability of the findings.

Incorporating datasets with varied characteristics and representations of lung cancer types would offer a more robust validation of the prototypical network's effectiveness across different imaging conditions and patient demographics. Such an approach would also help to ascertain the scalability and adaptability of the proposed model in real-world clinical settings.

Moreover, the fluctuating performance of models like VGG16 and ResNet50 in certain shot settings prompts further investigation. It raises questions about the models' robustness and their ability to handle data scarcity, which is a common challenge in medical imaging. Future research could explore the underlying reasons for these fluctuations, potentially through more granular data analysis or feature visualization techniques. This could lead to improved model designs or training strategies that are more resilient to variations in data availability and complexity.

Additionally, the consistent performance improvement of the CNN model across different episodes suggests a possible avenue for optimizing few-shot learning models for medical imaging tasks. Investigating the factors contributing to this consistency, such as the model's architecture or the nature of the data it handles best, could provide valuable insights for developing more effective learning algorithms in the field.

The study also underscores the importance of choosing the right feature extraction model and the amount of training data in achieving optimal results in few-shot learning scenarios. This finding has practical implications for developing efficient diagnostic tools, especially in situations where acquiring large datasets is challenging. It emphasizes the need for a balanced approach that considers both the computational strengths of the model and the characteristics of the available data.

In conclusion, while the study demonstrates the promising potential of using Prototypical Networks with pre-trained models like ResNet50 for lung cancer classification, it also highlights the complexities involved in machine learning applications in medical imaging. The exploration into different model behaviors and their training dynamics offers a valuable contribution to the field. However, the reliance on a single dataset suggests the need for caution in generalizing the results. Future research should aim to expand the dataset diversity to further validate and enhance the applicability of these findings.

V. CONCLUSION & FUTURE WORK

This study focused on the classification of four main types of lung cancer using CT scan images, leveraging the potential of few-shot learning in scenarios with limited training data. The core methodology employed the prototypical network, a well-regarded model in few-shot learning, complemented by various pre-trained models and a custom CNN for effective feature extraction from lung CT images.

The findings of this research lay a foundational framework for further advancements in lung cancer detection using CT scans and few-shot learning techniques. Such advancements could extend to automated medical report generation for radiology images [25], enhancing diagnostic efficiency and accuracy.

A promising direction for future research is the integration of Siamese Neural Networks with Prototypical Networks. Siamese Networks excel in identifying subtle differences [26], a crucial aspect in medical imaging, while Prototypical Networks are effective in classification by creating representative models of each class. The synergy of these networks could potentially overcome individual limitations and lead to a more robust and precise system for lung cancer classification. The challenge lies in effectively combining these networks to harness their complementary strengths while addressing potential integration issues, presenting an intriguing problem for future exploration.

REFERENCES

[1] "American cancer society," https://www.cancer.org/cancer/lung-cancer.

[2] T. Foster and A. Dixon, "Normal ct chest," May 2015. [Online]. Available: http://dx.doi.org/10.53347/rID-36676

[3] P. Li, S. Wang, T. Li, J. Lu, Y. HuangFu, and D. Wang, "A large-scale ct and pet/ct dataset for lung cancer diagnosis," 2020. [Online]. Available: https://www.cancerimagingarchive.net/collection/lung-pet-ct-dx/

[4] P. Chaturvedi, A. Jhamb, M. Vanani, and V. Nemade, "Prediction and classification of lung cancer using machine learning techniques," *IOP Conference Series: Materials Science and Engineering*, vol. 1099, p. 012059, 03 2021.

[5] S. Makaju, P. P.W.C, A. Alsadoon, A. Singh, and A. Elchouemi, "Lung cancer detection using ct scan images," *Procedia Computer Science*, vol. 125, pp. 107–114, 01 2018.

[6] L. Wang, "Deep learning techniques to diagnose lung cancer," *Cancers*, vol. 14, no. 22, 2022. [Online]. Available: https://www.mdpi.com/2072-6694/14/22/5569

[7] A. A. Meldo and L. V. Utkin, "A new approach to differential lung diagnosis with ct scans based on the siamese neural network," *Journal of Physics: Conference Series*, vol. 1236, no. 1, p. 012058, jun 2019. [Online]. Available: https://dx.doi.org/10.1088/1742-6596/1236/1/012058

[8] G. Xiuhua, S. Tao, W. huan, and L. Zhigang, "Prediction models for malignant pulmonary nodules based-on texture features of ct image," in *Theory and Applications of CT Imaging and Analysis*, N. Homma, Ed. Rijeka: IntechOpen, 2011, ch. 4. [Online]. Available: https://doi.org/10.5772/14766

[9] J. Snell, K. Swersky, and R. Zemel, "Prototypical networks for few-shot learning," in *Advances in Neural Information Processing Systems*, vol. 30. Curran Associates, Inc., 2017.

[10] S. Ahuja, B. K. Panigrahi, N. Dey, A. Taneja, and T. K. Gandhi, "Mcs-net: Multi-class siamese network for severity of covid-19 infection classification from lung ct scan slices," *Applied Soft Computing*, vol. 131, p. 109683, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494622007323

[11] Y. Jiang, H. Chen, H. Ko, and D. Han, "Few-shot learning for ct scan based covid-19 diagnosis," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2021-June, pp. 1045–1049, 2021.

[12] D. Liu, X. Gao, and Q. Shen, "Prototypical network for radar image recognition with few samples," *Journal of Physics: Conference Series*, vol. 1634, p. 012116, 09 2020.

[13] A. Hussaindeen, S. Iqbal, and T. D. Ambegoda, "Multi-label prototype based interpretable machine learning for melanoma detection," *International Journal Of Advances In Signal And Image Sciences*, vol. 8, no. 1, pp. 40–53, 2022.

[14] N. Protonotarios, I. Katsamenis, S. Sykiotis, N. Dikaios, G. Kastis, S. Chatziioannou, M. Metaxas, N. Doulamis, and A. Doulamis, "A few-shot u-net deep learning model for lung cancer lesion segmentation via pet/ct imaging," *Biomedical Physics & Engineering Express*, vol. 8, 02 2022.

[15] F. Guo and Y. Fan, "Zero-shot and few-shot learning for lung cancer multi-label classification using vision transformer," *ArXiv*, vol. abs/2205.15290, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:249192263

[16] "The cancer imaging archive," Year. [Online]. Available: https://www.cancerimagingarchive.net

[17] W. Nimalsiri, M. Hennayake, K. Rathnayake, T. D. Ambegoda, and D. Meedeniya, "Cxlseg dataset: Chest x-ray with lung segmentation," in *2023 International Conference On Cyber Management And Engineering (CyMaEn)*, 2023, pp. 327–331.

[18] S. Dasanayaka, V. Shantha, S. Silva, D. Meedeniya, and T. Ambegoda, "Interpretable machine learning for brain tumour analysis using mri and whole slide images," *Software Impacts*, vol. 13, p. 100340, 2022.

[19] J. Hofmanninger, F. Prayer, J. Pan, S. Röhrich, H. Prosch, and G. Langs, "Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem," *European Radiology Experimental*, vol. 4, p. 50, 08 2020.

[20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd International Conference on Learning Representations (ICLR 2015)*, pp. 1–14, 2015.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[22] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269.

[23] S. Mohammed and A. Çinar, "Lung cancer classification with convolutional neural network architectures," *Qubahan Academic Journal*, vol. 1, pp. 33–39, 02 2021.

[24] S. D. Bhatt, H. B. Soni, H. R. Kher, and T. D. Pawar, "Automated system for lung nodule classification based on resnet50 and svm," in *2022 3rd International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*, 2022, pp. 1–5.

[25] W. Nimalsiri, M. Hennayake, K. Rathnayake, T. D. Ambegoda, and D. Meedeniya, "Automated radiology report generation using transformers," in *2023 3rd International Conference on Advanced Research in Computing (ICARC)*. IEEE, 2023, pp. 90–95.

[26] Y. Li, C. Chen, and T. Zhang, "A survey on siamese network: Methodologies, applications and opportunities," *IEEE Transactions on Artificial Intelligence*, vol. PP, pp. 1–21, 12 2022.