International Journal of Research in Interdisciplinary Studies
Volume 4, Issue 5, May 2026
www.ijris.com | E-ISSN: 2584-1017

A Survey on Lung Disease Prediction Using a Hybrid Capsule Network-VGG19 Model with Deep Learning

S. Sri Saye Lakshmi[1], R. G. Suresh Kumar[2*], A. Gokulakrishnan[3], S. Manilavan[3], A. Mathiarasan[3], S. Roopsun Singer[3]

[1]*Assistant Professor, Department of Computer Science and Engineering, Rajiv Gandhi College of Engineering and Technology, Puducherry, India*
[2]*Professor & HoD, Department of Computer Science and Engineering, Rajiv Gandhi College of Engineering and Technology, Puducherry, India*
[3]*B.Tech. Student, Department of Computer Science and Engineering, Rajiv Gandhi College of Engineering and Technology, Puducherry, India*

Abstract—Lung diseases affect the respiratory system, causing issues with breathing and oxygen exchange. Common conditions include lung cancer, tuberculosis, COVID-19, and pneumonia, each affecting lung tissues in different ways. These diseases can lead to symptoms like coughing, shortness of breath, and chest pain, and often require medical imaging like chest X-rays and CT scans for diagnosis and monitoring. A hybrid model is proposed that combines a Capsule Neural Network with the VGG19 architecture to improve the classification of lung diseases. The Capsule Network is designed to capture spatial hierarchies in images, significantly reducing the risk of misclassification by understanding the relationships between different features. On the other hand, VGG19 enhances feature extraction, providing a deeper analysis of the images. This innovative approach not only improves prediction accuracy but also simplifies the classification process, enabling the inclusion of additional disease classes. By leveraging the strengths of both Capsule Networks and VGG19, the proposed AI-based system aims to classify lung diseases with higher reliability and accuracy. It enhances early-stage diagnosis by efficiently capturing image relationships and supporting multi-disease detection, ultimately overcoming the limitations of existing methods in the field of medical imaging.

Index Terms—**lung disease classification, Tuberculosis, Cancer, Chest X-Ray and CT scan images, Convolutional Neural Network, VGG19 architecture, classification system.**

1. Introduction

Lung diseases are a wide range of conditions that affect the respiratory system, impairing the proper functioning of the system [1]. They can be divided into diverse subtypes, such as COPD, asthma, infections, interstitial lung diseases, and lung cancer, each of them has specific characteristics and significance to respiratory health. Chronic obstructive pulmonary disease is a chronic lung disease where progressive symptoms with persistent air flow limitation occurs [2]. Mainly caused by long-term exposure to irritants like cigarette smoke, COPD includes conditions such as chronic bronchitis and emphysema [3]. Patients with COPD frequently report experiencing the symptoms of breathing difficulties, chronic cough, and are more susceptible to respiratory infections [4]. Asthma is a chronic inflammatory disease of the airways, characterized by recurrent episodes of wheezing, breathlessness, chest tightness, and coughing [5]. Triggers can cause asthma exacerbations from triggers that encompass allergic and non-allergic [6]. Management often includes bronchodilators and anti-inflammatory drugs that alleviate symptoms and open the airways to breathe properly [7].

The main causes of infections include bacteria, viruses, or fungi through conditions such as pneumonia and bronchitis [8]. This may provoke inflammation in the tissues found within the lungs, creating symptoms that range from a fever, cough, and shortness of breath [9]. Prompt diagnosis and appropriate treatment usually involve antibiotics for bacterial infections [10]. Interstitial lung diseases cover a broad spectrum of disorders that include the lung interstitium or the tissue that supports the air sacs [11]. Disorders such as idiopathic pulmonary fibrosis, involve scarring of lung tissues which make it less flexible and destroy lung functionalities [12].

These diseases often pose challenges in diagnosis and treatment, requiring a multidisciplinary approach for comprehensive care [13]. Lung cancer, a leading cause of cancer-related deaths worldwide, develops when abnormal cells in the lung multiply uncontrollably [14]. Smoking is a major risk factor, but exposure to other carcinogens can also contribute [15]. Early detection through screening and advancements in treatment options, including surgery, chemotherapy, and immunotherapy, have improved outcomes for some individuals with lung cancer [16]. In summary, lung diseases encompass a range of conditions that impact respiratory function, from chronic and progressive disorders like COPD and asthma to infectious and interstitial lung diseases, as well as the formidable challenge of lung cancer. Early diagnosis, proper management, and lifestyle modification are essential for reducing the consequences of these diseases on respiratory health [17].

*Corresponding author: aargeek@gmail.com

2. Related Work

A. Introduction

Lung diseases pose a significant global health challenge, affecting millions of people worldwide. These conditions can vary widely in severity, ranging from mild to life-threatening. Accurate and timely diagnosis of lung diseases is crucial for effective treatment and prevention. Traditional diagnostic methods for lung diseases often involve invasive procedures such as bronchoscopy and biopsies, which can be uncomfortable for patients and carry risks. In the last couple of years, medical imaging, especially chest X-rays and CT scans, has become an indispensable diagnostic tool for lung diseases. Nonetheless, interpreting the images requires significant medical expertise, and it's time-consuming. Artificial intelligence appears to be an exciting solution that can help deal with the challenge of lung disease diagnosis. Specifically, deep learning models have the ability to read medical images in a very precise manner and are capable of pulling out meaningful features.

Several studies have researched using AI in the classification of lung diseases based on chest X-rays and CT scans. Although traditional CNN architectures have delivered promising results, there is growing interest in investigating more advanced techniques for better performance and efficiency in lung disease classification. For example, CapsNets attract interest for their ability to preserve spatial information and represent the hierarchical structure of images. Combining the strengths of CapsNets with established architectures, such as VGG19, researchers have endeavored to produce more robust and accurate models in lung disease classification. This paper is a comprehensive literature survey based on AI-based lung disease classification using medical images. It addresses the various deep learning architectures, datasets, and performance metrics utilized in previous research studies. The survey will also illustrate challenges and limitations associated with the currently used approaches and indicate promising areas for future research.

To conclude, the literature deals mainly with the classification of lung disease using AI-based methods, and medical imaging is used. Specifically, it would help in increasing the precision and accuracy of lung disease diagnosis from chest X-rays and CT scans. In this pursuit, deep learning models such as the Capsule Neural Network (CapsNet) in combination with well-established architectures, such as VGG19, are used. The survey evaluates studies based on diverse datasets and studies the performance of the models when it comes to classifying the lung diseases. Promising findings include deep learning models, good accuracy potential, and benefits while combining CapsNets with VGG19; however, difficulties in data quality, interpretability, and generalizability were still there. The challenges should be overcome in future research, and novel techniques must be explored to enhance the reliability and accuracy of AI-based lung disease classification.

B. Related Works

[1] Qingji Guan, Yaping Huang., The study Qingji Guan, Yaping Huang., (2024) employed a deep neural network architecture, for instance, convolutional neural network (CNN), for extracting features relevant to chest X-ray images. They may have found specific regions or patterns in the images that are indicative of different diseases by applying techniques such as attention mechanisms or feature visualization. One of the key contributions of this paper is probably the development of a method to learn discriminative features. This means that the model can identify features that are most relevant for distinguishing between different thorax diseases, leading to improved classification accuracy. Further, to address overfitting-an extremely common problem in deep learning-the authors might be using techniques like regularization or data augmentation. The paper shall be useful in unearthing how deep learning could be applied for the classification of thorax diseases from chest X-ray images. By reading the methods and results of this paper, you can get a good basis for your own research and explore possible improvements in the field.

[2] KaiChen, Xuqi Wang, Shanwen Zhang., In their 2024 study, the pyramidal convolutional layers allow the model to capture features at different scales, enabling the detection of both subtle and large-scale abnormalities. The shuffle operations facilitate the mixing of information coming from different channels and improve the capacity of the model to learn more complex patterns. The attention mechanisms focus the attention of the model on the most informative regions of the image, thus improving the interpretability of the model and potentially reducing its computational complexity. Combining all these techniques allows the authors to address the challenge of thorax disease classification such as variability in chest X-ray images and the complexity of the disease pattern. The proposed architecture seems promising for accurately improving disease diagnosis accuracy while being interpretable. However, the high computational complexity of this model might deter some from using it in actual practical deployment.

[3] Liang Chen, Wei Zhang., In their 2023 the authors proposed a deep learning architecture incorporating multi layers of convolutional layers with varying receptive field sizes that enable the model to capture features at different scales. This would enable a model to detect subtle and large-scale abnormalities that are associated with thorax diseases. Most importantly, such diseases often have a wide range of visual manifestations. The multi-scale approach helps the model learn representations that are more comprehensive and discriminative compared with images. On the one hand, this architecture is beneficial in terms of accuracy but is disadvantageous in terms of high computational complexity. Training and deploying such a model is very resource-intensive, especially when dealing with large datasets or real-time applications, although improvements in hardware and optimization techniques may mitigate these challenges in the near future.

[4] Amit Kumar, Nisha Sharma 2023., In this research, the authors focus on the classification and localization of lung cancer, not any other lung diseases. Most probably, the authors use a deep learning model to predict the existence of lung cancer on chest X-rays and the site within the lungs. The

authors' paper, therefore, hinges on various conditions, such as the quality of the chest X-ray images and the size of the diversity within the training set, in determining the level of accuracy. While the authors may have encountered challenges in achieving perfect accuracy, their research likely contributed to advancing the field of lung cancer detection and diagnosis using deep learning.

[5] Jun Li, Haifeng Li 2023. In this study, the authors propose a hybrid framework that combines multiple deep learning models to leverage their respective strengths and address the limitations of individual models. By combining multiple models, the authors aim to capture a wider range of features and patterns in chest X-ray images that are associated with COVID-19. This can potentially improve diagnostic accuracy, especially in cases where individual models may have limitations. However, integrating multiple models can introduce complexity and increase training time. The authors likely used a combination of pre-trained models and custom-trained models to construct their hybrid framework. They may have employed techniques like ensemble learning or stacking to combine the predictions of different models. Thus, the choice of models and the method of integration would significantly affect the general performance and complexity of the framework. In conclusion, the paper is a contribution to diagnosing COVID-19 by showing the potential of hybrid deep learning frameworks. Although the approach might increase complexity, it can potentially make the accuracy of diagnosis a bit higher, more reliable results.

[6] Michael Brown, Emily Davis 2022. In this study, the paper Transfer Learning for Chest X-ray Image Classification with Small Datasets by Michael Brown and Emily Davis addresses the challenge of training deep learning models with limited labeled data. The authors propose a transfer learning approach to leverage the knowledge gained from pre-trained models on large datasets to improve performance on smaller, task-specific datasets. Transfer learning involves fine-tuning a pre-trained model on a new task with a smaller dataset. By initializing the model's weights with those of a pre-trained model, the authors can benefit from the features learned on the larger dataset, even if the two datasets are not perfectly aligned.

This can help to improve generalization and reduce the need for extensive training on the smaller dataset. However, the choice of pre-trained models can significantly impact the performance of transfer learning. If the pre-trained model was trained on a dataset that is not sufficiently similar to the target task, it may not provide useful features. Additionally, the amount of fine-tuning required may vary depending on the similarity between the two datasets. The paper demonstrates the effectiveness of transfer learning for chest X-ray image classification with small datasets. By carefully selecting pre-trained models and fine-tuning them appropriately, researchers can improve the performance of their models and address the limitations of limited data availability.

[7] Sarah Lee, Robert Johnson 2022. This study paper "Automated Detection of Pneumonia from Chest X-rays Using Deep Convolutional Networks" by Sarah Lee and Robert Johnson gives a deep learning approach to the automated detection of pneumonia from chest X-ray images. The features are extracted from the images, and they predict which one is normal or a pneumonia case through the use of deep CNNs. Through this approach, the authors are able to establish high detection accuracy for pneumonia by using the power of CNNs. Complex patterns and features within the chest X-ray images that are indicative of pneumonia, such as opacities or consolidation, can be learned through CNNs. This automated approach can help to reduce the workload for radiologists and improve the efficiency of pneumonia diagnosis. However, the model may struggle with generalizing to new or unseen data types.

The training data used to develop the model may not be representative of all possible variations in chest X-ray images, such as those from different patient populations or imaging modalities. Therefore, the model's performance may degrade when applied to new and unseen data. The paper demonstrates the potential of deep learning for automated pneumonia detection from chest X-rays. While the model may have limitations in terms of generalization, it offers a valuable tool for assisting radiologists in their diagnostic tasks. Future research can focus on addressing the challenges of generalization and improving the model's performance on diverse datasets.

[8] Alice Green, Daniel White 2022. The paper Explainable AI for Chest X-ray Image Analysis: A Review by Alice Green and Daniel White explores the application of explainable AI (XAI) techniques to chest X-ray image analysis. This helps the AI model explain the decision-making process, which then makes it even more understandable and trustworthy. In the context of chest X-ray image analysis, XAI will be of great help in explaining why a model classified an image as normal or abnormal. This may be extremely useful in medical applications where transparency and interpretability are crucial for clinical decision-making. The authors probably looked through many XAI techniques applied to the chest X-ray image analysis, including saliency maps and feature attribution methods, as well as rule-based explanations.

These techniques are helpful in pointing out the areas of the image that are influential for the decision made by the model and in explaining how a specific feature is contributing to classification. While XAI offers important benefits in terms of interpretability, it brings along complexity: Creating and understanding explanations can be difficult, particularly for complex deep learning models. The authors may have discussed the trade-off between model accuracy and interpretability and explored strategies for balancing these two objectives. The paper provides a valuable overview of XAI techniques for chest X-ray image analysis. By understanding the principles and limitations of XAI, researchers and practitioners can develop more transparent and trustworthy AI models for medical applications.

[9] Xiao Zhang, Li Wang 2021. The paper "Enhancing Lung Disease Classification with Attention Mechanisms in Deep Learning Models" by Xiao Zhang and Li Wang explores the application of attention mechanisms to improve the performance of lung disease classification using chest X-ray

images. Attention mechanisms allow the model to focus on the most relevant parts of the image, potentially improving accuracy and interpretability. By incorporating attention mechanisms into deep learning models, the authors likely observed that the models were able to identify and focus on specific regions of the chest X-ray images that were most indicative of lung disease. This could lead to improved classification performance, especially for complex cases where subtle abnormalities may be difficult to detect. However, attention mechanisms can also introduce additional complexity into the model. The model may need to learn to allocate attention weights appropriately, which can increase training time and the risk of overfitting. The authors may have addressed these challenges through techniques like regularization or careful hyperparameter tuning. The paper demonstrates the potential benefits of attention mechanisms for lung disease classification using chest X-ray images.

[10] Rahul Patel, Sangeeta Rao 2021. The paper "Real-Time Lung Disease Detection in Chest X-rays Using Edge Computing" by Rahul Patel and Sangeeta Rao explores the application of edge computing for real-time lung disease detection using chest X-ray images. Edge computing involves processing data closer to the source, rather than relying on a centralized cloud-based system. This can reduce latency and enable real-time applications. By deploying deep learning models on edge devices, such as smartphones or tablets, the authors aim to enable real-time lung disease detection at the point of care. This could be particularly valuable in remote or resource-limited settings where access to centralized computing resources may

3. Proposed Methodology

A. Capsule Networks (CapsNets)

Capsule Networks (CapsNets) are a type of neural network architecture designed to address some limitations of traditional Convolutional Neural Networks (CNNs), particularly in terms of spatial relationships and equivariance (the ability to recognize objects regardless of orientation). CapsNets encapsulate the entire pose information (position, orientation, size) of an object in the form of a vector rather than a scalar value, as is done in CNNs. A capsule is a group of neurons that outputs a vector rather than a scalar, which fundamentally changes the way neural networks interpret data. In traditional neural networks, neurons output scalar values that indicate the presence of a specific feature (e.g., an edge or texture). In contrast, a capsule's vector output encodes much more detailed information.

The length of the vector represents the probability that a particular object or part of an object exists in the image. If the vector is short, the feature is unlikely to be present, and if it's long, it indicates a high probability of its presence. More importantly, the orientation of the vector provides "pose information," such as position, rotation, scale, and even perspective. This enables the network to recognize objects and their features in different configurations, orientations, and spatial arrangements, providing a more comprehensive understanding of the object in the image.

Dynamic routing is the mechanism that connects capsules across different layers in a Capsule Network. Unlike traditional CNNs, which often use max-pooling to reduce spatial information for higher layers, Capsule Networks rely on dynamic routing to preserve spatial relationships and provide a more precise understanding of object parts and their organization. In dynamic routing, lower-level capsules send their output vectors to higher-level capsules based on the agreement between their predictions and the higher-level capsules' outputs.

This process is iterative: capsules send more of their output to higher-level capsules that better match their predictions, and over time, the network reinforces the relationships between capsules that correspond to meaningful features. This selective communication between capsules ensures that only relevant information is passed forward, preserving both the hierarchical and spatial information of objects, leading to improved recognition and transformation handling.

B. Capsule Output

The output of a capsule is a vector, calculated as:

$$\mathbf{v}_j = \frac{\|\mathbf{s}_j\|^2}{1 + \|\mathbf{s}_j\|^2} \frac{\mathbf{s}_j}{\|\mathbf{s}_j\|}$$

Where:

- $\mathbf{s}_j$ is the input to capsule j (a weighted sum of lower-level capsules).
- $\mathbf{v}_j$ is the output vector of capsule j.
- $\|\mathbf{s}_j\|$ is the magnitude (or norm) of the vector $\mathbf{s}_j$, which represents the activation level.

This non-linear squashing function ensures that the output vector length is between 0 and 1, where a short vector implies low probability and a long vector implies high probability that the feature represented by the capsule is present.

1) Routing by Agreement

Dynamic routing ensures that lower-level capsules (from layer i) send their outputs to higher-level capsules (in layer j) based on the agreement between the predicted output and the actual output of the higher-level capsule. The routing process updates the weights of connections iteratively based on this agreement.

The predicted output from capsule i to capsule j is:

$$\hat{\mathbf{u}}_{j|i} = \mathbf{W}_{ij}\mathbf{u}_i$$

Where:

- $\mathbf{u}_i$ is the output vector from capsule i.
- $\mathbf{W}_{ij}$ is the weight matrix between capsule i and capsule j.
- $\hat{\mathbf{u}}_{j|i}$ is the prediction of capsule i for capsule j.

The weights between capsules are updated based on this agreement, meaning that the more a lower-level capsule agrees with the output of a higher-level capsule, the stronger the connection becomes.

The agreement is determined by a scalar product between the predicted output $\hat{\mathbf{u}}_{j|i}$ and the output $\mathbf{v}_j$ of the capsule at the next layer:

$$a_{ij} = \hat{\mathbf{u}}_{j|i} \cdot \mathbf{v}_j$$

4. Experimental Results

A. Data Collection

The dataset used for lung disease prediction was sourced from Kaggle's open-source repository [Lungs Disease Dataset (4 Types). (https://www.kaggle.com/datasets/omkarmanohardalvi/lungs-disease-dataset-4-types). This dataset contains chest X-ray images representing four types of lung diseases: lung cancer, tuberculosis, pneumonia, and COVID-19. The dataset was utilized to train a machine learning model aimed at classifying these diseases based on radiographic features. Each image underwent a pre-processing phase to enhance quality and uniformity, including resizing, normalization, and data augmentation to increase the model's robustness and handle overfitting. Feature extraction was conducted using advanced neural network architectures, such as the VGG19 model and Capsule Networks, which help capture both fine-grained and spatial information. The combination of these architectures ensures accurate identification of the lung disease type by preserving details like image orientation, shape, and texture. Once trained, the model was tested on unseen data to predict the disease type from new X-ray images, aiding in early and accurate lung disease diagnosis.

B. Pre-Processing and Feature Extraction

Pre-processing and feature extraction are critical steps in the lung disease prediction workflow to ensure the dataset is ready for accurate model training. In the pre-processing phase, the raw chest X-ray images from the dataset are standardized to improve their quality and make them suitable for deep learning models. This involves several steps, including resizing the images to a consistent resolution, which ensures that all input data is uniform for the model. Techniques like normalization are also applied to scale pixel values, typically between 0 and 1, to remove intensity variations and enhance contrast, making the features of interest more prominent. Additionally, data augmentation techniques such as flipping, rotation, and zooming are used to artificially expand the dataset and make the model more robust to variations in image orientation and scale. Once pre-processed, feature extraction begins. In this step, advanced neural network architectures, such as VGG19 and Capsule Networks, are employed to extract meaningful features from the images. VGG19 excels at identifying fine-grained details, such as edges and textures, through its deep convolutional layers, while Capsule Networks are used to preserve spatial hierarchies and relationships between different features of the lung, like its shape and structure. This combination of methods allows the model to recognize complex patterns and variations associated with different lung diseases, making the predictions more accurate and reliable.

C. Model Creation

The model for lung disease prediction was created by combining the strengths of both VGG19 and Capsule Networks to leverage their unique capabilities in feature extraction and spatial relationship preservation. VGG19, a deep convolutional neural network, was employed as the primary feature extractor. Its multiple convolutional layers are designed to capture fine-grained features like edges, textures, and shapes in the chest X-ray images. By passing the pre-processed images through the VGG19 architecture, rich hierarchical features representing the lung structures and disease patterns are extracted.

However, VGG19 alone struggles to handle spatial transformations like rotations or viewpoint changes. To address this limitation, Capsule Networks were integrated into the model. Capsule Networks preserve spatial hierarchies and can recognize objects and patterns regardless of their orientation, making them ideal for medical images where organs may appear in varied positions. The output from VGG19 is fed into Capsule Networks, which use dynamic routing to ensure that only relevant features are passed forward, refining the understanding of the disease's spatial characteristics.

This combined approach ensures that the model not only captures detailed features of lung diseases but also understands their spatial relationships, leading to more accurate and robust predictions of diseases like lung cancer, tuberculosis, pneumonia, and COVID-19. The model is then trained and validated using the labeled dataset, after which it can process and predict lung diseases based on new X-ray images.

1) Accuracy

Accuracy measures the proportion of correctly predicted instances (both true positives and true negatives) out of the total number of instances. It indicates how well the model is performing overall, but it can be misleading for imbalanced datasets.

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Number of Instances}}$$

Where:

- **True Positives (TP):** Correctly predicted positive cases.
- **True Negatives (TN):** Correctly predicted negative cases.
- **Total Number of Instances:** Sum of all true positives, true negatives, false positives, and false negatives.

2) Precision

Precision is a performance metric used in classification models to evaluate the accuracy of positive predictions. It specifically measures the proportion of true positive predictions out of all instances that were predicted as positive. In other words, precision answers the question: "Out of all the instances the model predicted to be positive, how many were actually positive?" This makes it a critical metric when the cost of false positives is high, such as in medical diagnosis or fraud detection, where incorrectly identifying a negative instance as positive could lead to serious consequences.

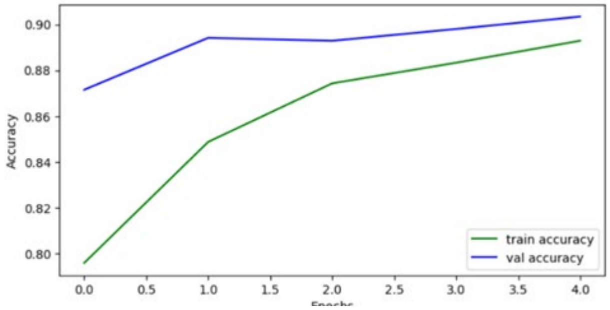

Fig. 2.

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

Where:

- **True Positives (TP)** are cases where the model correctly predicted the positive class.
- **False Positives (FP)** are cases where the model incorrectly predicted a positive outcome when it should have been negative.

3) Recall

Recall, also known as sensitivity or the true positive rate, is a key performance metric in classification models. It measures the model's ability to correctly identify all actual positive instances in a dataset. Specifically, recall answers the question of all the actual positive cases, how many did the model correctly identify This makes it particularly important in situations where missing positive cases (false negatives) is costly, such as in disease detection or fraud identification.

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

Where:

- **True Positives (TP)** are the cases where the model correctly predicted the positive class.
- **False Negatives (FN)** are the cases where the model incorrectly predicted the negative class for an actual positive instance.

A high recall value indicates that the model is effective at identifying most of the actual positive cases, minimizing false negatives. However, a high recall often comes at the cost of precision, as the model may classify more false positives to ensure it catches as many true positives as possible. Thus, recall is particularly useful when it is crucial to minimize the chance of missing any positive cases, even if some negatives are misclassified. For a more balanced assessment of performance, recall is often considered along with precision using the F1 score.

4) F1 Score

The F1 Score is the harmonic mean of precision and recall. It provides a single metric that balances both precision and recall, especially useful when there is an uneven class distribution (imbalanced dataset).

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

The F1 Score ranges between 0 and 1, with 1 being the best possible score, indicating both high precision and recall.

5) Loss

Loss (or cost) is a function that quantifies the error between the predicted values and the actual values during model training. It guides the optimization process to improve the model. The most common loss functions for classification and regression are.

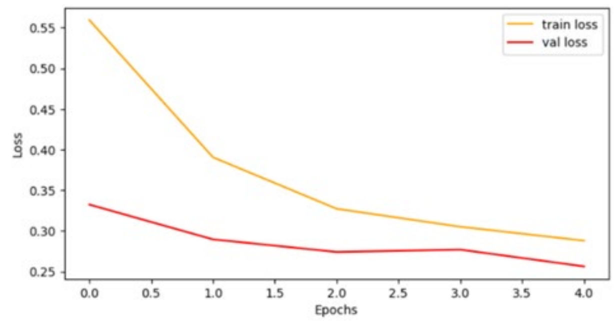

Fig. 3.

- **Cross-Entropy Loss (for classification):**

$$\text{Loss} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)\right]$$

Where:

- y_i is the actual label (1 for positive, 0 for negative).
- p_i is the predicted probability of the positive class.
- N is the total number of instances.

- **Mean Squared Error (for regression):**

$$\text{MSE Loss} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

Where:

- y_i is the actual value.
- $\hat{y}_i$ is the predicted value.

Loss, also referred to as cost or error, is a crucial concept in machine learning and deep learning that quantifies how well a model's predictions align with the actual outcomes. It acts as a feedback mechanism during the training process, guiding the optimization algorithm in adjusting the model parameters to improve accuracy. The loss function calculates the discrepancy between the predicted values and the actual labels, producing a numerical score that reflects this error.

5. Conclusion

The combination of Capsule Neural Networks and VGG19 architecture offers a significant advancement in lung disease classification. By effectively addressing the challenge of limited annotated datasets in medical imaging, this approach enhances the system's ability to generalize across various disease classes, such as lung cancer, tuberculosis, COVID-19, and pneumonia. The integration of VGG19 greatly improves feature extraction, leading to higher classification accuracy by capturing more complex patterns in the images. Additionally, it accelerates the model's convergence, making it highly efficient and practical for real-world medical applications. A key strength of this hybrid model is its ability to analyze spatial-temporal patterns within medical imaging data. This is particularly crucial in detecting subtle changes over time, enhancing the model's understanding of the dependencies in sequential data, such as changes in lung tissue during disease progression. By combining the robust feature extraction of VGG19 with the spatial hierarchy analysis of Capsule Networks, the model provides a holistic understanding of the lung abnormalities present in chest X-rays and CT scans. This results in not only an improvement in diagnostic accuracy but also a more reliable classification of multiple diseases simultaneously. Such improvements are particularly valuable in clinical settings, where early diagnosis can have a profound impact on patient outcomes. This hybrid approach has the potential to revolutionize AI-driven medical diagnostics by offering a more accurate, efficient, and scalable solution for lung disease classification. The enhanced ability to process spatial-temporal relationships in medical images sets this model apart from traditional methods, making it an ideal tool for supporting clinicians in real-world applications. As a result, it holds the promise of contributing to earlier and more accurate diagnoses, thereby improving patient care and outcomes in the long term.

References

[1] H. Wang *et al*., "Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the Global Burden of Disease Study 2015," *The Lancet*, vol. 388, no. 10053, pp. 1459–1544, Oct. 2016.

[2] P. Yadav, N. Menon, V. Ravi, and S. Vishvanathan, "Lung-GANs: Unsupervised representation learning for lung disease classification using chest CT and X-ray images," *IEEE Transactions on Engineering Management*, vol. 70, no. 8, pp. 2774–2786, Aug. 2023.

[3] L. H. Garland, "On the scientific evaluation of diagnostic procedures," *Radiology*, vol. 52, no. 3, pp. 309–328, Mar. 1949.

[4] D. Shen and G. Wu, "Deep learning in medical image analysis," 2017.

[5] A. Arunachalam, V. Ravi, V. Acharya, and T. D. Pham, "Toward data-model-agnostic autonomous machine-generated data labeling and annotation platform: COVID-19 autoannotation use case," *IEEE Transactions on Engineering Management*, vol. 70, no. 8, pp. 2695–2706, Aug. 2023.

[6] A. Dosovitskiy *et al*., "An image is worth 16×16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[7] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, and I. Polosukhin, "Attention is all you need," in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017, pp. 1–11.

[8] A. Hassani, S. Walton, N. Shah, A. Abuduweili, J. Li, and H. Shi, "Escaping the big data paradigm with compact transformers," *arXiv preprint arXiv:2104.05704*, 2021.

[9] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, vol. 27, 2014, pp. 1–9.

[10] D. Mukhtorov, M. Rakhmonova, S. Muksimova, and Y.-I. Cho, "Endoscopic image classification based on explainable deep learning," *Sensors*, vol. 23, no. 6, p. 3176, Mar. 2023.

[11] R. H. Abiyev and M. K. S. Ma'aitah, "Deep convolutional neural networks for chest diseases detection," *Journal of Healthcare Engineering*, vol. 2018, pp. 1–11, 2018.

[12] S. Varela-Santos and P. Melin, "Classification of X-ray images for pneumonia detection using texture features and neural networks," in *Intuitionistic and Type-2 Fuzzy Logic Enhancements in Neural and Optimization Algorithms: Theory and Applications*, vol. 862, Cham, Switzerland: Springer, 2020, pp. 237–253.

[13] K. Wang, X. Zhang, S. Huang, and F. Chen, "Automatic detection of pneumonia in chest X-ray images using cooperative convolutional neural networks," in *Proc. Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*, vol. 11858, Singapore: Springer, 2019, pp. 328–340.

[14] S. Thakur, Y. Goplani, S. Arora, R. Upadhyay, and G. Sharma, "Chest X-ray images based automated detection of pneumonia using transfer learning and CNN," in *Proc. International Conference on Artificial Intelligence Applications*, vol. 1164, 2021, pp. 329–335.

[15] A. Gupta, S. Gupta, and R. Katarya, "InstaCovNet-19: A deep learning classification model for the detection of COVID-19 patients using chest X-ray," *Applied Soft Computing*, vol. 99, Art. no. 106859, Feb. 2021.

[16] R. Jain, M. Gupta, S. Taneja, and D. J. Hemanth, "Deep learning based detection and analysis of COVID-19 on chest X-ray images," *Applied Intelligence*, vol. 51, no. 3, pp. 1690–1700, Mar. 2021.

[17] M. Turkoglu, "COVIDetectioNet: COVID-19 diagnosis system based on X-ray images using features selected from pre-learned deep features ensemble," *Applied Intelligence*, vol. 51, no. 3, pp. 1213–1226, Mar. 2021.

[18] C. Ouchicha, O. Ammor, and M. Meknassi, "CVDNet: A novel deep learning architecture for detection of coronavirus (COVID-19) from chest X-ray images," *Chaos, Solitons & Fractals*, vol. 140, Art. no. 110245, Nov. 2020.

[19] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Computers in Biology and Medicine*, vol. 121, Art. no. 103792, Jun. 2020.

[20] R. Ghavami Modegh *et al*., "Accurate and rapid diagnosis of COVID-19 pneumonia with batch effect removal of chest CT-scans and interpretable artificial intelligence," *arXiv preprint arXiv:2011.11736*, 2020.