Detection of Pulmonary Disorders with CNN model using Knowledge Distillation technique

Ms.R.Keerthana
Department of Computer Technology
Madras Institute Of Technology
Chennai, India
rskeerthanaravi@gmail.com

Jayaprakash.S
Department of Computer Technology
Madras Institute Of Technology
Chennai, India
jayaprakashmit3@gmail.com

Monisha.R
Department of Computer Technology
Madras Institute Of Technology
Chennai, India
monisha2847@gmail.com

Guru Raman.C
Department of Computer Technology
Madras Institute Of Technology
Chennai, India
gururaman.c@gmail.com

***Abstract*—Chronic respiratory conditions are a major public health concern worldwide and pose a significant burden on individuals and healthcare systems. Early detection of pulmonary diseases play a major role in treatment. However, existing diagnostic techniques can be expensive, time-consuming, and require specialized expertise and any misdiagnosis can lead to delayed treatment. Lung auscultation is a widely used diagnostic technique that involves listening to respiratory sounds through a stethoscope which provide valuable information about the respiratory system and help in the identification of pulmonary diseases. However, interpretation of these sounds requires extensive training and experience. In this paper, we present the "Pulmonet" system, which has been developed to identify pulmonary diseases by analyzing lung auscultation sounds. The methodology employed involves knowledge distillation, where valuable insights from a large and computationally intensive model (referred to as the teacher model) are transferred to a computationally efficient model (referred to as the student model). Our approach involves converting the respiratory sounds into Mel spectrogram images, which can be processed by our system. The model's task is to make predictions about the presence of pulmonary diseases in a patient and identify the specific disease if any (such as Pneumonia, Upper Respiratory Tract Infection, Chronic Obstructive Pulmonary Disease, Bronchitis, Bronchiolitis) or classify them as healthy if no disease is detected. Following that, the insights acquired by the initial model are conveyed to the student model, which possesses lower computational complexity, making it feasible for deployment on mobile devices. Finding the suitable loss function for student model is crucial and challenging in our work. The loss function should be designed to balance both knowledge transfer and preservation of student model. In our work we have achieved the accuracy of 82% in teacher model and 78% in student model.**

***Index Terms*—Deep learning , MelSpectrogram , Knowledge distillation , Teacher model, Student Model**

I. Introduction

Machine learning is a technique that improves computer systems' performance using algorithms, making it crucial in healthcare for analyzing large datasets and identifying patterns and trends that may be missed by human clinicians [1]. Machine learning models handle structured and unstructured data inputs, converting structured data into suitable forms for efficient processing[2]. Machine learning applications in healthcare include image analysis, predictive analytics, drug discovery, and personalized medicine. Covid-19 pandemic impacted lives; machine learning model predicts and reduces loss of lives through supervised, semi-supervised, or unsupervised learning methods [3] .

Deep learning enhances healthcare outcomes through medical image analysis, drug discovery, and personalized medicine [4]. Deep learning technique uses image classification for predicting pulmonary diseases using chest X-ray images. [5].n this study, we utilized the ICBHI dataset [6], which consists of 920 audio samples collected from 226 people, amounting to a total duration of approximately 5.5 hours.

Convolutional neural networks (CNNs) represent a specialized class of deep neural networks meticulously designed to process visual data, such as images and videos [3]. Remarkably transformative for computer vision, CNNs intrinsically possess the ability to autonomously discern and extract key features from images, thus facilitating a wide array of tasks, including object detection, image segmentation, and facial recognition. Accomplished through convolutional layers equipped with adaptive filters, these networks yield discernible output feature maps [7]. Sequentially, network layers synergistically combine the aforementioned feature maps, ultimately executing operations of classification or regression with profound efficacy and precision.

Knowledge distillation encompasses the transference of knowledge from a higher-complexity teacher model to a lower-complexity student model, streamlining learning processes [5]. Employing the One Teacher One Student technique, a solitary teacher model is trained on a sizable dataset through supervised learning [10]. The student model emulates the teacher's output using the identical dataset, minimizing prediction differences and leveraging teacher predictions as soft targets [6]. This amplifies the student model's performance by empowering it to adeptly capture intricate characteristics.

II. Related Work

In our work, pulmonary diseases will be detected using a CNN model and the Knowledge Distillation approach. our CNN model examines spectrogram images generated from the auscultation sounds of the lungs to produce precise classifications. By using knowledge distillation, we move knowledge from a larger student model to a pre-trained instructor model, improving performance and lowering complexity. Our study highlights the potential of automated diagnosis tools to help medical personnel identify lung problems promptly and accurately.

Patel et al.,[11] utilized 920 lung sound recordings from 368 patients, preprocessed with a CNN architecture. Data augmentation techniques were applied to enhance dataset size and performance. The CNN achieved a remarkable 92.5% accuracy in classifying lung sounds into normal, crackle, and wheeze categories.

L. Pham et al.,[9] employed chest radiographs from public datasets, preprocessing them with a combination of a convolutional neural network (CNN) to extract features and a mixture of experts (MoE) model to fuse predictions.

Rab et al.,[5] employs deep CNNs to detect COPD from 3D lung airway tree snapshots. The study includes dataset collection, preprocessing, and training a deep CNN architecture. The model's evaluation reveals its effectiveness in COPD detection, showcasing promising advancements in medical image analysis for respiratory diseases.

Fernando et al.,[10] Fernando et al.,[10] used 1200 lung sound recordings from 267 patients and preprocessed them using a Temporal Convolutional Neural Network (TCN) architecture which is trained using binary cross-entropy and weighted mean squared error loss functions, and improved with data augmentation techniques. It achieved an F1-score of 0.806 for crackle detection and 0.766 for wheeze detection, outperforming other event detection methods.

Chen .Y et al.,[12] propose a dynamic knowledge distillation method that enhances student model performance while preserving teacher model interpretability. They tested the approach on visual analytics tasks, using deep neural networks and distillation loss functions.

J. Song et al. [13] ntroduce a novel approach with two neural networks: a deep teacher model and a smaller student model. Using a distillation loss function, they transfer knowledge and achieve better results in image classification tasks.

K. Zhang et al.,[14] introduce a method using genetic algorithm for optimal student network architecture in knowledge distillation. They use a deep teacher model and a lightweight student model, incorporating a distillation loss function and regularization to prevent overfitting. Their approach excels in image classification tasks, outperforming other methods.

H. Purwins et al. [15] present a comprehensive overview of deep learning techniques for audio tasks, highlighting their significance in speech recognition, music analysis, sound synthesis, and environmental sound classification.

Convolutional neural networks are popular deep learning models for lung disease detection, but cloud deployment requires high computation power and inference time. Knowledge distillation technique compresses large networks into smaller, more efficient models, making them suitable for resource-constrained environments.

III. Proposed Work

Our system Pulmonet aims to provide early and accurate diagnosis of pulmonary diseases, improve patient outcomes, and reduce the burden by reducing the complexity of healthcare systems. In this paper, our main objective is to design and develop a system specialized in classifying respiratory sounds into distinct pulmonary diseases. We have named this model "Pulmonet". In Figure 1, we present a detailed workflow that outlines the steps involved in the Pulmonet system. The workflow illustrates the process of how we preprocess and organize the respiratory sound data, prepare it for training, and create the deep learning architecture for disease classification. We adopt a method that transforms respiratory sounds into Mel spectrogram images, which are then fed into the deep learning model. The model's task is to determine the presence of any pulmonary disease in the patient. To expedite output predictions, we employ Knowledge distillation to reduce the model's computational complexity.

A. Dataset Description

The 2017 International Conference on Biomedical and Health Informatics (ICBHI) dataset offers an all-encompassing repository of meticulously annotated respiratory sounds. This all-encompassing assemblage comprises an impressive repository of 920 audio samples procured from 226 people, with an impressive cumulative duration of 5.5 hours. The audio recordings manifest diverse lengths, spanning from 10 to 90 seconds, and have been acquired at varying sampling frequencies, ranging from 4 kHz to 44.1 kHz. This rich dataset encompasses the contributions of 128 patients, thoughtfully distributed across a diverse spectrum of respiratory conditions. These include 14 instances of Upper Respiratory Tract Infection (URTI), 8 cases of Pneumonia, 26 healthy subjects, 44 instances of Chronic Obstructive Pulmonary Disease (COPD), 12 cases of Bronchiolitis, and 12 cases of Bronchiectasis. Such an extensive and meticulously curated dataset serves as a valuable resource for pioneering research endeavors in the realm of respiratory sound analysis.

B. Data Preprocessing Module

The Audio signals of 226 patients are present in .wav format along with the class labels. First raw audio signals are converted to a fixed length signal by framing and windowing and truncating, filtering and resampling to enhance the signal's quality. Then the processed audio signals are transformed to spectrogram by applying fourier transformation. Then the spectrograms are transformed into Mel-frequency spectrograms using Mel scale.

1) Data Augmentation: Data augmentation techniques like adding noise, stretching, shifting sound, pitching are used to augment the audio data. Then the audio is resampled to fixed sample rate of 320kbps.

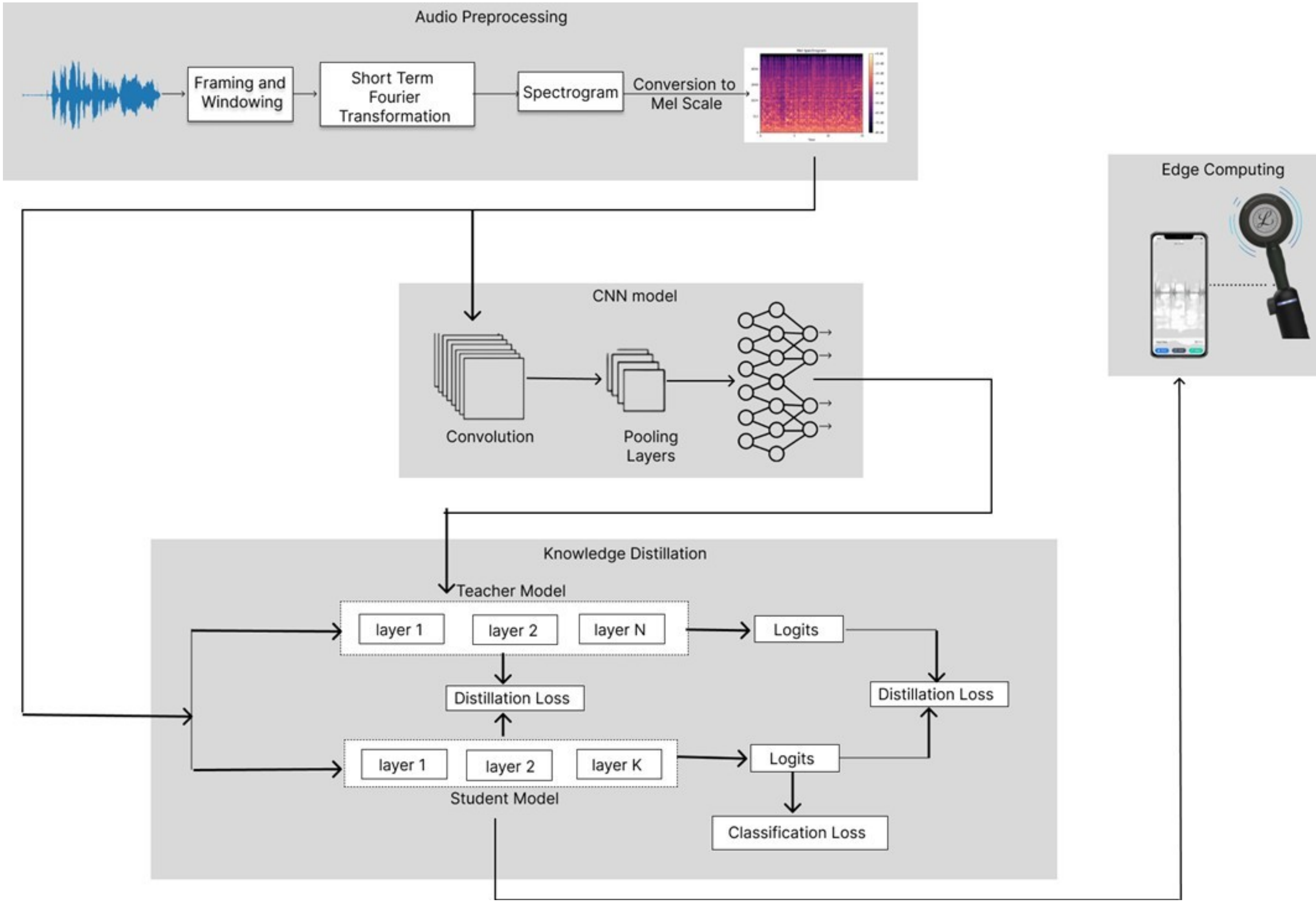

Fig 1 Architecture Diagram

2) Framing: In respiratory audio processing,framing involves segmenting the continuous respiratory signal into shorter frames of a fixed duration, typically ranging from a few milliseconds to a few seconds depending on the application.The audio signal is framed into smaller segments where the crackles or wheezes occurs. The threshold values for the wheeze and crackle is 400 to 2000Hz and 50 to 500 Hz respectively.

3) Windowing: Windowing is meticulously applied to each frame to mitigate spectral leakage and significantly enhance frequency resolution. Specifically, we employ the rectangular function, which attains a value of 1 during the window duration and 0 beyond it. In our experimentation, the window length is set to 10, implying that the initial 10 samples are assigned a value of 1, while the subsequent samples are set to 0. The adoption of the rectangular window bestows unparalleled time resolution, ensuring the absence of any signal distortion in the time domain when the window is enacted. As a cumulative result of augmentation, framing, and windowing, the tabulated outcomes in Table 1 illustrate the number of audio files.

4) Fourier Transformation: The windowed audios are then fourier transformed with a sampling rate of 22050 Hz which results in a spectrogram image. In our proposed work we use Short Term Fourier Transformation(STFT). In STFT the audio signal is splitted into short overlapping windows and Fourier Transformation of each window is computed. The crackles and wheezes from the auscultation sounds are extracted and the

Algorithm 1: Audio to Spectrogram Conversion

Input : Recordings of lung auscultation sounds of patients with pulmonary diseases (pneumonia, URTI, bronchitis, bronchiolitis)

Output: Melspectrogram images of the respective lung auscultation sounds

```
foreach audio in database do
    Cropped_audios ← Extract_crackle_wheeze(audio);
    foreach aud in Cropped_audios do
        Save aud in ‘audiopath/disease_folder‘;

foreach disease_folder in audiopath do
    foreach audio in disease_folder do
        Spectrogram ← STFT(audio);
        Mel_Spec ← melFilterBank(Spectrogram);
        Image ← plot_to_image(Mel_Spec);
        Cropped_image ← crop(Image);
        Save Cropped_image in ‘database/disease_folder‘;
```

Label	Before Augmentation (no.of audio files)	After Augmentation (no.of audio files)	After Framing and Windowing (no.of audio files)
Bronchitis	84	323	1051
Bronchiolitis	97	361	1112
URTI	76	304	1068
COPD	86	340	1136
Pneumonia	90	349	1150

Table 1 Data augmentation

corresponding frames are converted to spectrogram by short term fourier transformation.

$$\text{STFT}[n, k] = \sum_{m=0}^{n-1} (n+m) w[m] e^{-2\pi k m / N} \tag{1}$$

Equation(1) shows Short Term Fourier Transformation, n is the index of the current time window, k is the index of the frequency bin, N is the length of the window (usually a power of 2), m is time, w[m] is the window function , j is the imaginary unit . Here the image is viewed as a 2D matrix. The column represents the time and rows represent the frequency bins. Each element in the matrix corresponds to the magnitude of a specific frequency at a specific time. By applying STFT the audio signals are convertd to spectrogram images. And then by applying n_mfcc of 40 which extracts 40 mfccs from the audio signals. Padding width is calculated, which is the difference between the maximum length of an MFCC sequence and the length of the current MFCC sequence. Then the current MFCC sequence is padded with zeros to the maximum length. The spectrogram images obtained are transformed into Mel scale, resulting in the creation of Mel spectrogram images.

5) Conversion of Mel scale to frequency scale: The spectrogram images of STFT will be frequency domain representation of audio signal in linear. The linear frequency domain is transformed into the Mel scale, representing the frequencies in a logarithmic manner which will be more perceptual to human beings.

$$\text{Mel Spectrogram} = 2595 \times \log_{10}\left(1 + \frac{f}{700}\right) \tag{2}$$

f is frequency, 2595 700 are fixed constants Equation(2) maps the linear frequency scale to the non-linear Mel scale, which is based on human perception of pitch.

All the steps in the data preprocessing module of our system "Pulmonet" is explained in fig 2 .

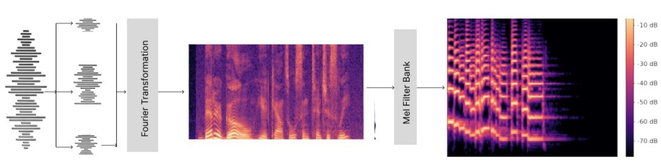

Fig 2: Audio to Mel Spectrogram Conversion

C. Knowledge distillation

Knowledge distillation represents an erudite strategy for transmitting knowledge from a highly intricate teacher model to a relatively less intricate student model. To accomplish this, the student model is not only trained on the conventional hard targets, i.e., class labels, but also exposed to the supple soft targets, manifested as the teacher model's class probabilities. This allows the student model to learn from the teacher's expertise and improve its performance, as shown in Figure 3. The model of having one teacher and one student presents a hopeful solution for implementing complex machine learning models on edge devices such as mobile phones and handheld monitors. The primary benefit of employing this strategy on edge devices lies in the ability to achieve quicker inference times. Looking ahead, our future work involves extending knowledge distillation beyond traditional single-modal scenarios. We plan to explore the exciting realm of multi-modal knowledge distillation, where information from multiple modalities, such as text, images, or audio, is integrated into a single student model. The primary objective of this approach is to augment the student model's performance and endow it with the proficiency to effectively tackle intricate real-world tasks encompassing diverse data types.

1) Teacher Model: Knowledge distillation is predicated upon leveraging a voluminous and highly precise pre-existing model, colloquially known as the teacher model, to serve as the wellspring of knowledge during the distillation process. The objective is to adeptly transfer the acquired knowledge from the teacher model to a more compact student model. In this investigation, the envisaged teacher model is elegantly characterized as a sequential neural network, wherein the output of each layer serves as the input to the subsequent layer. Every layer is abundantly furnished with neurons that diligently execute a linear transformation on the input data, seamlessly complemented by an intricate non-linear activation function.Notably, the convolutional layers make judicious use of the ReLU activation function, which is mathematically defined as elegantly expressed in Eq. (3).

$$f(x) = \max(0, x) \tag{3}$$

The function takes an input x and produces an output that equals the input itself if the input is positive, and it gives zero as the output if the input is negative. During the training phase, backpropagation is employed to learn the weights and biases of the neurons. This iterative algorithm adjusts the parameters to minimize the error between the predicted output and the actual target values.The teacher model is defined using the Keras Sequential API.In the neural network's output layer, the application of the softmax function facilitates multiclass classification, expertly assigning an input to one among several prospective classes. The ultimate predicted class is the one that boasts the highest probability value within the output vector. Fig 4 elegantly showcases the exhaustive structural composition of the teacher model

2) Student Model: The student model is defined using the Keras Sequential API. The complete structural definition of

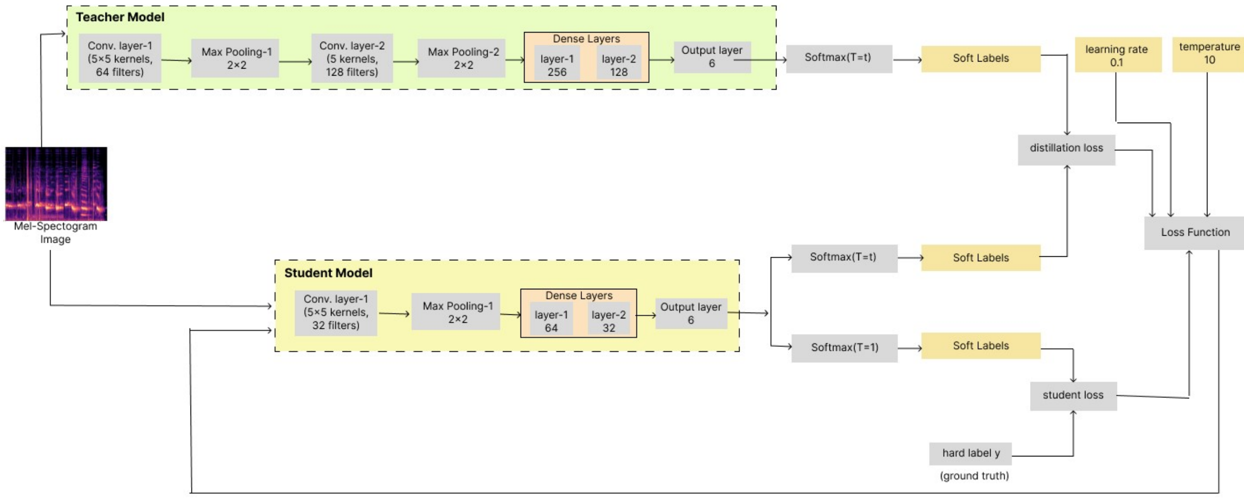

Fig 3 Knowledge Distillation

Architecture	Input	Layers	Striding	Padding	Output
		Input Layer (Melspectrogram Image)			255×255
Conv.Block 01	255×255	BN – Cv [5 × 5] @ 64 -ReLU	1	0	255×255×64
Max.Pooling 01	255×255×64	[2 × 2]	1	0	127×127×64
Conv.Block 02	127×127×64	BN – Cv [5] @ 128 -ReLU	1	0	127×127×128
Max.Pooling 02	127×127×128	[2 × 2]	1	0	63×63×128
Dense.Block 01	63×63×128	FC @ 256 - ReLu	1	0	256
Dense.Block 02	256	FC @ 128 – ReLu	1	0	128
Dense.Block 03	128	FC @ 6 - Softmax	1	0	6

Fig 4 Teacher Model Architecture

the student model is depicted in fig 5. The Mel spectrogram images from the datasets are read and resized to numpy array and then stored in a list along with their labels after normalizing the pixels and the input dataset is divided into two. Subsequently, a portion of the list will be forwarded to the teacher model, and the model's predictions will be recorded as soft labels. In contrast, for the other segment of the list, the hard labels from the dataset will be used. The soft and hard labels will be concatenated into y_train. The Mel spectrogram images along with the labels(y_train) will be passed to the student model.

Architecture	Input	Layers	Striding	Padding	Output
		Input Layer (Melspectrogram Image)			255×255
Conv.Block 01	255×255	BN – Cv [5 × 5] @ 32 -ReLU	1	0	255×255×32
Max.Pooling 02	255×255×32	[2 × 2]	1	0	127×127×32
Dense.Block 01	127×127×32	FC @ 64 - ReLu	1	0	64
Dense.Block 02	64	FC @ 32 – ReLu	1	0	32
Dense.Block 03	32	FC @ 6 - Softmax	1	0	6

Fig 5 Student Model Architecture

3) Loss function: The student model is fitted using the user-defined loss function, which calculates the total loss as the linear combination of the following components:

Algorithm 2: Teacher Model Training

Input : Mel spectrogram images of the lung auscultation sounds of patients with respiratory diseases

Output: A teacher model to predict the class of disease from the lung auscultation sound or to declare them as healthy

```
Create variables to store images and labels;
foreach disease_folder in database do
    foreach image in disease_folder do
        Load the image using cv2;
        Reduce the dimensions to 255*255 pixels;
        Add labels to images;
Create a sequential model (Teacher Model);
Add(Conv2D(64,kernel_size(5,5),activation='relu'))
Add(BatchNormalization);
Add(MaxPooling(2,2));
Add(Conv2D(128,kernel_size = 5,activation='relu'))
Add(BatchNormalization);
Add(MaxPooling(2,2));
Add(Dropout(0.3));
Add(Flatten);
Add(Dense(256, activation='relu'));
Add(Dense(128, activation='relu'));
Add(Dense(6, activation='softmax'));
Compile the model (loss=sparse_categorical_crossentropy,
    metrics=accuracy, optimizer=adam);
Fit the model with loaded dataset_images and class labels;
Train the model with 30 epochs;
```

Distillation loss: This quantifies the disparity between the soft targets and the predictions of the student model.
Student loss: This quantifies the divergence between the true labels and the predictions of the student model.
The combined loss function is formulated by considering both the distillation loss and the student loss.

$$L(x) = (1-\alpha) \times \text{student\_loss} + (\alpha \times \text{temp}^2 \times \text{dist\_loss}) \quad (4)$$

The combined loss, as depicted in Equation (4), is computed by incorporating both the student loss and the distillation loss. The temperature parameter, initially set to 10, is utilized to soften the probability distribution of the teacher's output during knowledge transfer. Throughout the training process of the student model using the distilled knowledge from the teacher model, a learning rate of alpha=0.1 is adopted as a hyperparameter.

$$\mathrm{TL} = (1-\alpha) \times \mathrm{SL} \times \alpha \times \mathrm{temp}^2 \times \mathrm{DL} \tag{5}$$

$$\mathrm{DL} = \mathrm{KL\_Divergence}(\mathrm{soft\_pred}, \mathrm{soft\_true}) \tag{6}$$

$$\mathrm{SL} = \mathrm{categorical\_crossentropy}(y_{\mathrm{true\_student}}, y_{\mathrm{pred\_student}}) \tag{7}$$

$$\mathrm{Soft\_target}(x) = \frac{e^{y_i/\mathrm{temp}}}{\sum_{j=1}^{N} e^{y_j/\mathrm{temp}}} \tag{8}$$

$$\mathrm{KL\_Divergence}(p||q) = \sum p(x) \times \log_{10}\left(\frac{p(x)}{q(x)}\right) \tag{9}$$

$$\mathrm{Categorical\_crossentropy} = -\sum p(x) \times \log_{10}(q(x)) \tag{10}$$

where TL is total loss, SL is student loss, DL is distillation loss, KL Divergence is Kullback Leibler divergence.

Algorithm 3: Creating Student Model

Input : A trained teacher model that can classify classes of diseases or healthy using mel spectrogram images

Output: A student model that has less number of layers and neurons

```
add(Conv2D(32, kernel_size=(5,5), activation='relu'));
add(BatchNormalization);
add(Dropout(0.3));
add(Flatten);
add(Dense(64, activation='relu'));
add(Dense(32, activation='relu'));
add(Dense(6, activation='softmax'));
```

Algorithm 4: Knowledge Distillation

Input : A trained teacher model that can classify classes of diseases or healthy using mel spectrogram images and a newly created student model

Output: A student model that gained knowledge from the teacher model and can predict lung disease

```
Function loss_func(y_true, y_pred):
    y_true_soft, y_true_hard = split(y_true, 2);
    y_pred_soft, y_pred_hard = split(y_pred, 2);
    soft_pred = y_pred_soft/temp;
    soft_true = y_true_soft/temp;
    distillation_loss = kl_divergence(soft_pred, soft_true);
    student_loss = categorical_crossentropy(y_true_hard, y_pred_hard);
    tot_loss = (1 − α) × student_loss + α × temp^2 × distillation_loss;
    return tot_loss;
x_train_soft, x_train_hard = tf.split(x_train, 2);
y_train_soft, y_train_hard = tf.split(y_train, 2);
y_train_soft = teacher_model.predict(x_train_soft);
y_train = tf.concat(y_train_soft, y_train_hard);
student_model.compile(loss=loss_func, metrics=['accuracy'],
    optimizer='adam');
studentmodel.fit(xtrain, ytrain, batchsize=128, epochs=10,
    validationsplit=0.2);
```

In Equations 5,6,7,8,9,10 soft_pred is the student model's predicted soft label. soft_true is the teacher model's predicted soft label. α is the learning rate temp is the temperature hyperparameter. student_loss delineates the loss computation between the student model's output and the ground truth label. distillation_loss entails the meticulous computation of loss between the student model's output and the output yielded by the esteemed teacher model. $y_{\mathrm{true\_student}}$ is the actual class label in the dataset. $y_{\mathrm{pred\_student}}$ is the student model's predicted class label. y is the output of the model. p and q are input variables.

The Mel spectrogram images are normalized, split into two input datasets, and stored in a list. The teacher model is given a subset of the list, with predictions saved as soft labels and hard labels as hard labels. The student model is trained using a customized loss function for 50 epochs, with 20% of the training data used for validation.

IV. Performance Analysis And Results

Model	Accuracy
Teacher Model	**82%**
Student Model	**78%**

Table 2 Model Accuracy

According to Table 2, the suggested teacher model achieved an accuracy of 82%, while the student model, using knowledge distillation, achieved an accuracy of 78%. Notably, the student model maintains a lower complexity compared to the teacher model, all while maintaining a high level of accuracy without significant compromise.

classes	Class label	precision	recall	f1-score
0	Bronchiectasis	0.85	0.68	0.76
1	Bronchiolitis	0.90	0.69	0.82
2	COPD	0.96	0.99	0.98
3	Healthy	0.92	0.66	0.77
4	Pneumonia	0.82	0.76	0.80
5	URTI	0.93	0.57	0.70
accuracy				0.82
Macro avg		0.92	0.73	0.80
Weighted avg		0.95	0.95	0.95

Table 3 Teacher Model Evaluation Metrics

classes	Class label	precision	recall	f1-score
0	Bronchiectasis	0.80	0.56	0.67
1	Bronchiolitis	0.82	0.59	0.77
2	COPD	0.84	0.80	0.85
3	Healthy	0.86	0.76	0.69
4	Pneumonia	0.90	0.67	0.73
5	URTI	0.88	0.63	0.78
accuracy				0.78
Macro avg		0.87	0.65	0.74
Weighted avg		0.89	0.76	0.83

Table 4 Student Model Evaluation Metrics

Table 3 presents the evaluation metrics of the teacher model, including precision, recall, and F1-score.Precision measures accurate positive predictions among all predicted positives. Recall evaluates model's ability to identify all relevant cases.

F1-score combines precision and recall for a comprehensive performance metric.

Table 4 presents the evaluation metrics of the knowledge-distilled student model, including precision, recall, and F1-score. High precision implies minimal false positive errors, while recall denotes few false negative errors. The F1-score showcases a harmonious balance between precision and recall, capturing the model's overall performance adeptly.

V. Conclusion

This project proposes an automated pulmonary disorder detection system using respiratory sound analysis, Mel spectrogram representation, and Convolutional Neural Networks. The system involves data augmentation, short-term fourier transformation, and converting respiratory sounds into Mel spectrogram images. A deep learning model is used to detect pulmonary disorders. To decrease computational complexity, a knowledge distillation technique with one teacher and one student is employed, allowing the compression of larger teacher models into smaller student models. This method achieves greater accuracy on the ICBHI dataset and reduces computing complexity while maintaining excellent accuracy. The system enhances patient outcomes and reduces healthcare expenditures.

References

[1] Erickson, Bradley J. and Korfiatis, "Machine Learning for Medical Imaging", in RadioGraphics journal, vol. 37, no. 2, pp. 505-515, 2017,doi: 10.1148/rg.2017160130.

[2] M. Chen, Y. Hao, K. Hwang, L. Wang and L. Wang, "Disease Prediction by Machine Learning Over Big Data From Healthcare Communities," in IEEE Access, vol. 5, pp. 8869-8879, 2017, doi: 10.1109/ACCESS.2017.2694446.

[3] R. F. Albuquerque Paiva de Oliveira, C. J. A. Bastos Filho, A. C. A. M. V. F. de Medeiros, P. J. Buarque Lins dos Santos and D. Lopes Freire, "Machine learning applied in SARS-CoV-2 COVID 19 screening using clinical analysis parameters," in IEEE Latin America Transactions, vol. 19, no. 6, pp. 978-985, June 2021, doi: 10.1109/TLA.2021.9451243.

[4] Shen, Dinggang and Wu, Guorong and Suk, " Deep Learning in Medical Image Analysis ", in Annual Review of Biomedical Engineering, vol. 19, no. 1, 2017, pp. 221-248, doi: 10.1146/annurev-bioeng-071516-044442.

[5] S. Cannata, A. Paviglianiti, E. Pasero, G. Cirrincione and M. Cirrincione, "Deep Learning Algorithms for Automatic COVID-19 Detection on Chest X-Ray Images," in IEEE Access, vol. 10, pp. 119905-119913, 2022, doi: 10.1109/ACCESS.2022.3221531.

[6] Heang-Ping Chan, Ravi K. Samala, Lubomir M. Hadjiiski, Chuan Zhou, "Deep Learning in Medical Image Analysis" , 2020, Volume 1213, ISBN: 978-3-030-33127-6

[7] B. M. Rocha, D. Filos, L. Mendes, I. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jácome, A. Marques, R. P. Paiva, I. Chouvarda, P. Carvalho, N. Maglaveras, "A Respiratory Sound Database for the Development of Automated Classification", Precision Medicine Powered by pHealth and Connected Health, 2018, Volume 66 ISBN : 978-981-10-7418-9.

[8] L. Pham, H. Phan, R. Palaniappan, A. Mertins and I. McLoughlin, "CNN-MoE Based Framework for Classification of Respiratory Anomalies and Lung Disease Detection," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 8, pp. 2938-2947, Aug. 2021, doi: 10.1109/JBHI.2021.3064237.

[9] T. Fernando, S. Sridharan, S. Denman, H. Ghaemmaghami and C. Fookes, "Robust and Interpretable Temporal Convolution Network for Event Detection in Lung Sound Recordings," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 7, pp. 2898-2908, July 2022, doi: 10.1109/JBHI.2022.3144314.

[10] Aykanat, M., Kılıç, Ö., Kurt, B. et al. "Classification of lung sounds using convolutional neural networks", J Image Video Proc. 2017, 65 (2017). https://doi.org/10.1186/s13640-017-0213-2.

[11] Z. Tu, X. Liu and X. Xiao, "A General Dynamic Knowledge Distillation Method for Visual Analytics," in IEEE Transactions on Image Processing, vol. 31, pp. 6517-6531, 2022, doi: 10.1109/TIP.2022.3212905.

[12] R. Du et al., "Identification of COPD From Multi-View Snapshots of 3D Lung Airway Tree via Deep CNN," in IEEE Access, vol. 8, pp. 38907-38919, 2020, doi: 10.1109/ACCESS.2020.2974617.

[13] J. Song, Y. Chen, J. Ye and M. Song, "Spot-Adaptive Knowledge Distillation," in IEEE Transactions on Image Processing, vol. 31, pp. 3359-3370, 2022, doi: 10.1109/TIP.2022.3170728.

[14] K. Zhang, C. Zhang, S. Li, D. Zeng and S. Ge, "Student Network Learning via Evolutionary Knowledge Distillation," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 4, pp. 2251-2263, April 2022, doi: 10.1109/TCSVT.2021.3090902.

[15] H. Purwins, B. Li, T. Virtanen, J. Schlüter, S. -Y. Chang and T. Sainath, "Deep Learning for Audio Signal Processing," in IEEE Journal of Selected Topics in Signal Processing, vol. 13, no. 2, pp. 206-219, May 2019, doi: 10.1109/JSTSP.2019.2908700.