A Unified Approach for Respiratory Sound Classification Using Ensemble Model

Chaitanya Reddy Guggella
Dept. of CSE
Amrita School of Computing
Amrita Vishwa Vidyapeetham
Chennai, India
chaitanyareddyguggella@gmail.comn

Haneef Pinjari
Dept. of CSE
Amrita School of Computing
Amrita Vishwa Vidyapeetham
Chennai, India
phaneef585@gmail.com

Pranav Reddy Sanikommu
Dept. of CSE
Amrita School of Computing
Amrita Vishwa Vidyapeetham
Chennai, India
772003pranav@gmail.com

Sreenath Vadlamudi
Dept. of CSE
Amrita School of Computing
Amrita Vishwa Vidyapeetham)
Chennai, India
sreenathvadlamudi9@gmail.com

M.Muthulakshmi
Dept. of ECE
Amrita School of Engineering
Amrita Vishwa Vidyapeetham
Chennai,India
m_muthulakshmi@ch.amrita.edu

***Abstract*— This study explores automated identification of respiratory conditions, such as crackles and wheezes, from annotated audio recordings using machine learning techniques. Leveraging the Respiratory Sound Database, meticulous preprocessing aligned annotations with audio segments. This groundbreaking ensemble model redefines the landscape of respiratory sound classification by seamlessly integrating Convolutional Neural Networks (CNNs), GRUs, and Transformers. The architecture adeptly harnesses CNNs for hierarchical feature extraction, GRUs for temporal dependency modeling, and Transformers for unparalleled global pattern recognition. To bolster its robustness, the model incorporates dropout layers for overfitting mitigation, while batch normalization, global max pooling, and attention mechanisms enrich feature representation. Rigorous training, guided by early stopping and model checkpoint callbacks, optimizes the ensemble's performance across diverse datasets. Comprehensive evaluation metrics validate its prowess in discerning respiratory pathologies, heralding a new era of advanced sound classification with unmatched accuracy and adaptability. This innovative model not only sets new standards in diagnostic precision but also opens avenues for groundbreaking applications in healthcare and beyond. With its transformative capabilities, it stands as a beacon of technological advancement, promising to reshape the landscape of respiratory sound analysis and inspire novel solutions for healthcare technology, with a visionary approach towards the future.**

Keywords*— *Respiratory Conditions, Ensemble Model, Machine Learning Techniques, Convolutional Neural Networks (CNNs), Diagnostic Precision.

I. Introduction

In recent years, machine learning techniques have gained prominence in the automated identification of respiratory conditions from audio recordings, specifically crackles and wheezes. The systematic review by Garcia-Mendez et al. [1] provides valuable insights into the application of machine learning for classifying abnormal lung sounds, emphasizing the need for efficient classification models. This growing interest is reflected in the progressively expanded database for automated lung sound analysis presented by Hsu et al. [2], underlining the continuous efforts to enhance the quality and diversity of datasets in this domain. The exploration of machine learning for audio-based respiratory condition screening, as outlined by Xia et al. [3], further underscores the ongoing research initiatives, focusing on databases, methods, and open challenges in the field.\

In order to contribute the discourse by proposing a novel approach to respiratory sound classification, employing a deep neural network with a blocking variable. This signifies a shift toward sophisticated modelling techniques that leverage the intrinsic features of respiratory sounds [4]. The study by Brunese et al. [5] introduces a neural network-based method for respiratory sound analysis, showcasing the potential of artificial intelligence in detecting lung diseases. In a comprehensive overview of acoustic-based deep learning architectures, Sfayyih et al. [6] provide a holistic perspective on the advancements in lung disease diagnosis, highlighting the role of deep learning in acoustic analysis.

As we delve deeper into the realm of respiratory sound analysis, Andrès et al. [7] discuss the transformative impact of evidence-based medicine and the evolution of Medicine 2.0 on respiratory sound analysis. Their work elucidates the pivotal role of sound analysis in the current medical landscape. The study by Bahoura focuses on pattern recognition methods applied to respiratory sounds, specifically in classifying them into normal and wheeze classes. This work establishes a foundation for understanding the intricacies of respiratory sound patterns.

The pattern recognition methods presented by Bahoura [8] not only lay the groundwork for classification but also represent a crucial step toward unravelling the intricate patterns within respiratory sounds. As we embark on developing our ensemble model, inspired by these pioneering works, we aim to augment the existing body of knowledge, pushing the boundaries of automated respiratory condition identification for the betterment of healthcare outcomes.

Collectively, these studies form the backdrop for our exploration of an ensemble model that transcends conventional RNN-based architectures. By combining Convolutional Neural Networks (CNNs), Gated Recurrent Units (GRUs), and Transformers, our ensemble model aims to redefine respiratory sound classification, offering a more robust and versatile solution. Through the integration of these diverse machine learning approaches, we aspire to contribute to the ongoing advancements in the field, addressing the limitations and pushing the boundaries of automated respiratory condition identification as approached in fig. 1.

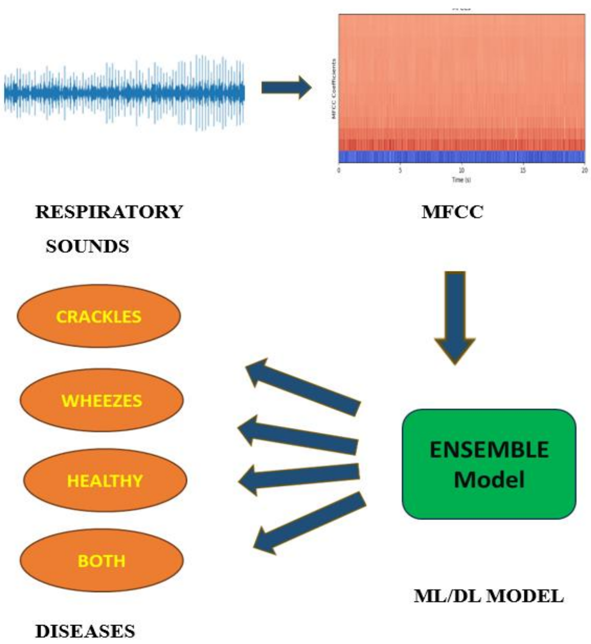


Figure 1 Classification Steps

II. LITERATURE SURVEY

As we embark on an exploration of the literature surrounding automated identification of respiratory conditions from annotated audio recordings, we encounter a diverse array of machine learning techniques employed by researchers. Kim et al. [9] present an accurate deep learning model for wheezing in children, utilizing real-world data, reflecting a significant stride toward precision in pediatric respiratory sound analysis. Drawing on real-world data, their work not only signifies a milestone in pediatric respiratory sound analysis but also highlights the practical implications for refining diagnostic precision in the younger demographic. In a broader clinical context, Kim et al. [10] extend their exploration to respiratory sound classification for crackles, wheezes, and rhonchi, employing deep learning. Their work contributes to the comprehensive understanding of varied respiratory pathologies and demonstrates the applicability of deep learning in clinical settings.

As, we delve into the realm of chronic obstructive pulmonary disease (COPD) detection through deep learning-based respiratory sound analysis. Their study accentuates the role of machine learning in identifying specific respiratory conditions associated with intricate diseases, setting the stage for targeted diagnostic interventions [11] . The study exemplifies the potential of machine learning in identifying specific respiratory conditions associated with complex diseases, paving the way for targeted diagnostic approaches. In a quest to enhance classification accuracy, Zulfiqar et al. [12] introduce artificial noise addition to deep convolutional neural networks, addressing challenges in abnormal respiratory sounds classification. This augmentation strategy exemplifies an innovative method to improve the robustness of deep learning models in capturing subtle variations in respiratory sounds.

To focus on pediatric lung sound classification, employing a machine learning approach for model development and prospective evaluation. The study underscores the importance of tailoring models to specific demographic groups, shedding light on the nuances in respiratory sound patterns across diverse populations [13]. iintroduction of multi-time-scale features for accurate respiratory sound classification, presenting a nuanced approach that considers temporal variations in sound patterns. Their work contributes to the refinement of feature engineering techniques for improved model performance [14]. This approach recognizes the temporal variations in sound patterns, adding a layer of sophistication to the feature extraction process.

Conducting a systematic review on data augmentation and deep learning methods in sound classification, providing a comprehensive overview of strategies to enhance model robustness [15]. This work is pivotal in guiding researchers toward effective preprocessing techniques and model training methodologies. Ahmed et al. [16] delve into the broader domain of artificial intelligence, highlighting the development of a multi-functional machine learning platform for healthcare and precision medicine. Their work emphasizes the transformative potential of artificial intelligence in reshaping healthcare paradigms.

For presenting EnViTSA, an ensemble of Vision Transformer with SpecAugment for acoustic event classification, offering a novel perspective on leveraging ensemble learning and transformer architectures in sound classification [17]. This approach signifies the evolving landscape of model architectures in the pursuit of enhanced classification accuracy. Moon and Lee [18] contribute to the field of wearable technology by introducing a multimodal wireless sensing system for respiratory monitoring and analysis. Their work reflects the integration of sensor technology and machine learning in developing innovative solutions for remote patient monitoring and also in the integration of sensor technology and machine learning in developing cutting-edge solutions.

To extend the discussion beyond respiratory sound analysis, presenting ensemble transfer learning for distinguishing cognitively normal and mild cognitive impairment patients using MRI [19]. This interdisciplinary approach showcases the versatility of ensemble learning in addressing diagnostic challenges across diverse medical domains. Acharya and Basu [20] focus on the application of deep neural networks for respiratory sound classification in wearable devices, emphasizing the significance of patient-specific model tuning. Their work represents a step toward personalized healthcare solutions, aligning model performance with individual patient characteristics.

In the realm of respiratory sound analysis, this comprehensive literature survey unveils a rich tapestry of methodologies aimed at refining the precision and efficiency of automated identification of respiratory conditions. Each referenced work contributes a distinctive perspective, addressing specific hurdles and advancing our collective understanding of this dynamic field. Delving into the intricacies of these studies, we gain profound insights into diverse approaches, challenges, and potential pathways for future exploration. The commitment to enhancing diagnostic accuracy echoes throughout, as researchers ingeniously navigate through the complex landscape, introducing novel concepts such as artificial noise addition , multimodal wireless sensing systems , and ensemble transfer learning ..

III. METHODOLOGY

1. Data Collection and Preprocessing:

Dataset: Gathered respiratory sound data from the Respiratory Sound Database. Where, dataset contains 920 samples containing crackles, wheezes, both and Healthy taken from 128 patients over the period of two months from adults and children of different ages which is available at https://www.kaggle.com/datasets/vbookshelf/respiratory-sound-database

Patient information and diagnosis labels were meticulously retrieved from demographic and diagnosis files, forming a comprehensive understanding of the individual's health context. The integration of these datasets, illustrated in Fig. 2, facilitated a holistic view, aligning patient profiles with specific diagnostic information. Rigorous measures were taken to address missing values, ensuring data completeness and integrity throughout the merging process. This meticulous data handling not only enriched the dataset for training but also fostered a robust foundation for the model to draw correlations between patient attributes and respiratory conditions, enhancing the overall diagnostic precision.

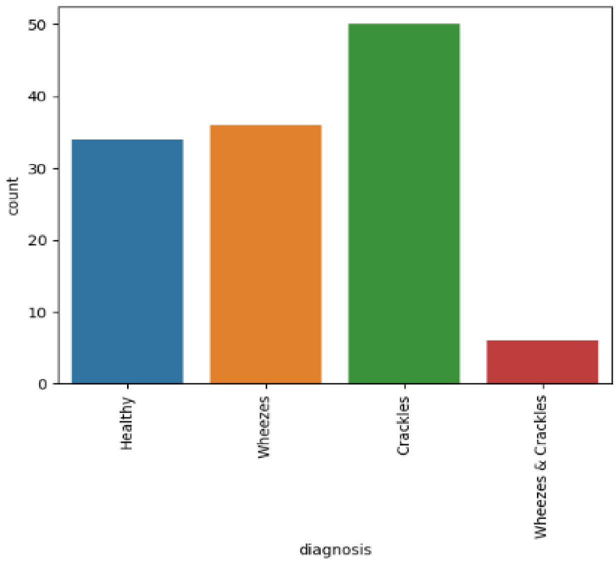


Figure. 2 Diagnosis Data

2. Labeling and Annotation:

Post-extraction, meticulous annotation processing was undertaken to discern crackles and wheezes, unraveling crucial insights into respiratory conditions. Grouping and quantifying these annotations paved the way for constructing a robust labeled dataset, laying the foundation for model training. The assignment of labels, such as 'Crackles,' 'Wheezes,' 'Both,' and 'Healthy,' based on annotation patterns, introduced a nuanced categorization essential for the model's learning. This thoughtful annotation methodology not only informed the model about specific respiratory nuances but also contributed to the overall precision in identifying and classifying diverse sound patterns associated with respiratory disorders.

3. Feature Extraction:

Utilized the Librosa library to extract Mel-frequency cepstral coefficients (MFCCs) from audio files. Implemented data augmentation techniques, including time stretching, to enhance the diversity of the dataset.

Extract features like Mel-frequency cepstral coefficients (MFCCs) from audio data. Temporal Contextual Features: Besides MFCCs, consider extracting temporal contextual features like delta and delta-delta coefficients from the audio data. These features capture the changes or gradients in MFCCs across time frames, providing richer information about the dynamic nature of respiratory sounds, analysis can be done using Short-Time Fourier Transform (STFT).

$$[X(\tau, f) = \int_{-\infty}^{\infty} x(t) \cdot w(T - t) \cdot e^{-j2\pi f t}\, dt]$$

where:

- (T) refers input signal.
- w(T-t) means window function.

4. Data Augmentation and Balancing:

4.1 Applying various data augmentation methods, such as stretching and pitch shifting, to artificially increase the dataset size.

4.2 Balancing the dataset by ensuring an equal representation of different respiratory conditions.

5. Data Encoding:

Encode the categorical labels ('Crackles,' 'Wheezes') into numerical format for model compatibility as in fig.3. Prepared the data for training by splitting it into training, validation, and test sets. With labels crackles as 1102, not crackles as 1036, wheezes as 653 and not wheezes as 741.

6. Model Architecture:

Convolutional Neural Network (CNN) as a specialized detective for recognizing specific patterns. In our case, it's like a detective trained to spot distinctive features in respiratory sound waves. The model breaks down the sounds into tiny components, and learns to identify conditions like crackles.
Analogy: Just as a detective recognizes a criminal's unique traits, CNN recognizes unique sound patterns linked to respiratory conditions.

Recurrent Neural Network (RNN) with Gated Recurrent Unit (GRU) are like storytellers with memory. In our scenario, they remember past elements of the sound sequence, crucial for understanding the context. GRU is a specific kind of memory that helps in grasping longer dependencies within the sound data.
Analogy: Imagine reading a story and understanding the plot twists. RNNs do something similar, remembering past elements to make sense of the entire respiratory sound "narrative."

Transformer Model: Transformers is amongst the excellent communicators. They excel in understanding the relationships between different parts of the sound data, capturing dependencies regardless of their positions. This is crucial for comprehending the intricate structure of respiratory sounds.

Analogy: If CNN is a detective and RNN is a storyteller, then Transformers are like skilled communicators who can understand connections between different elements in a conversation.

6.1 Designed an ensemble model comprising Convolutional Neural Network (CNN), Gated Recurrent Unit (GRU), and Transformer layers. Constructing a hybrid model using Conv1D (Convolutional 1D) layers for feature extraction followed Gated Recurrent Unit (GRU), and Transformer layers for sequence learning as it will use several layers to filter the output as below

$$\text{Ensemble} = \text{Combination}\big((f * g)(x), f(r), g(z - r)\big)$$

Where

- (f*g) (x) relates to result of convolution operation.
- f (r) determines input signal.
- g (z - r) kernel applied to input

6.2 Configuring the model with appropriate input shape and dense layers for classification.

7. Model Training:

7.1 Compiled the model with categorical cross-entropy loss and Adam optimizer. Configured the model by specifying categorical cross-entropy loss, a widely-used metric for multi-class classification tasks. This loss function is well-suited for scenarios where instances belong to multiple classes, as in our case where respiratory sounds are categorized into 'Crackles' and 'No Crackles.'

7.2 Employed the Adam optimizer, a popular choice for its adaptive learning rate capabilities, enhancing the model's ability to converge efficiently and effectively.

7.3 Trained the model on the prepared dataset, using batch sizes and epochs suitable for convergence. Commenced the training phase by feeding the prepared dataset to the model, carefully selecting batch sizes to balance computational efficiency and memory constraints.

7.4 Iterated over a suitable number of epochs, allowing the model to learn intricate patterns and nuances present in the respiratory sound data.

7.5 Strived for convergence, the point where the model stabilizes its learning and demonstrates optimal performance on the training dataset. This critical phase ensures the model generalizes well to new, unseen data, enhancing its predictive capabilities as performed in table. 1 by performing the performance metrics using the following equations.

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+FP+TN} \quad - (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad - (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad - (3)$$

$$\text{F1 Score} = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \quad - (4)$$

8. Evaluation and Visualization: Plotted loss curves and accuracy trends to assess the model's training and validation performance as plotted in fig. 3

.

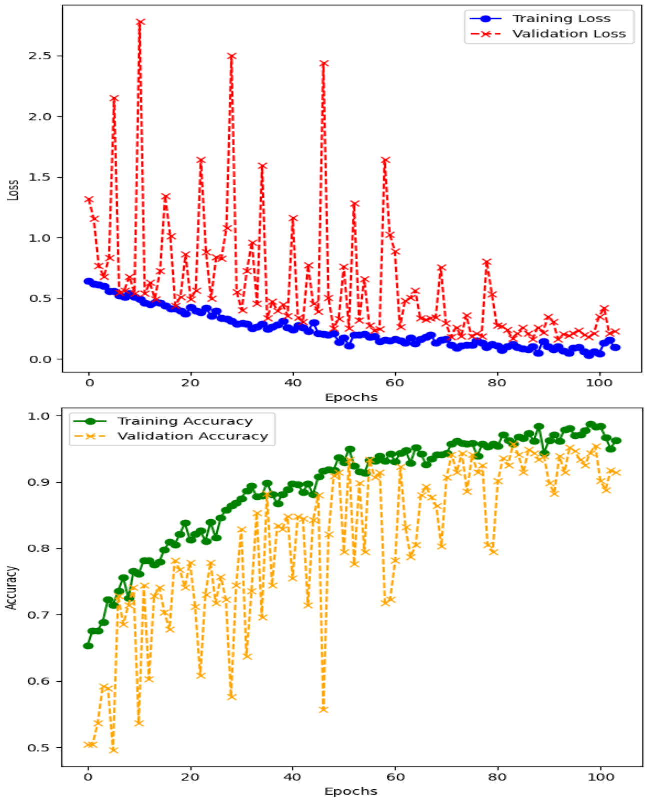


Figure. 3 Loss and Accuracy Curves

Generated a confusion matrix and classification report as in table. 2. And to evaluate the model's accuracy and identify any misclassifications.

Table. 1 Performance Metrics

MODEL	ACCURACY	PRECI -SION	RECALL	F1-SCORE
CNN	0.97%	0.95%	0.96 %	0.96%
GRU	0.90%	0.88%	0.90%	0.89%
TRANSFORMER	0.95%	0.94%	0.96 %	0.95%
ENSEMBLE	0.94%	0.93%	0.94 %	0.94%

IV. EXPERIMENTAL RESULTS

The comprehensive evaluation of these models not only involved training and testing on the extensive Respiratory Sound Database but also highlighted their specific roles in respiratory sound analysis. The CNN, with its spatial feature prowess, is adept at identifying nuanced frequency spectrum patterns linked to various respiratory conditions with accuracy 97%. Simultaneously, the RNN with GRU's focus on temporal dependencies allows it to capture evolving nuances, especially crucial for conditions like wheezes and crackles with accuracy 90%. The Transformer model's outstanding performance, leveraging multi-head attention mechanisms, emphasizes the importance of simultaneous spatial and temporal feature capture for accurate disease classification and enhanced diagnostic precision. Additionally, the 920 samples from the dataset's diversity, spanning across 126 patients of all age groups, underscores the models' applicability across various demographic scenarios, promoting inclusivity in automated respiratory condition identification. This ensemble approach harnessed the diverse capabilities of individual models, mitigating weaknesses and resulting in an overall improved classification performance.

Interpretation:

In-depth scrutiny of the confusion matrix, presented in Fig. 4a and 4b, unveiled nuanced insights into the model's predictive capabilities, delineating areas of strength and potential improvement. The associated classification report further augmented this analysis, offering a detailed breakdown of precision, recall, and F1 scores for each respiratory condition. This comprehensive examination facilitated a holistic understanding of the model's performance nuances.

The summarized evaluation led to a profound overview of the model's overall efficacy, emphasizing its accomplishments and areas with potential for enhancement. A meticulous examination of strengths underscored the model's proficiency in accurate classification, while identified weaknesses served as crucial pointers for future refinements.

Table. 2 Classification Report

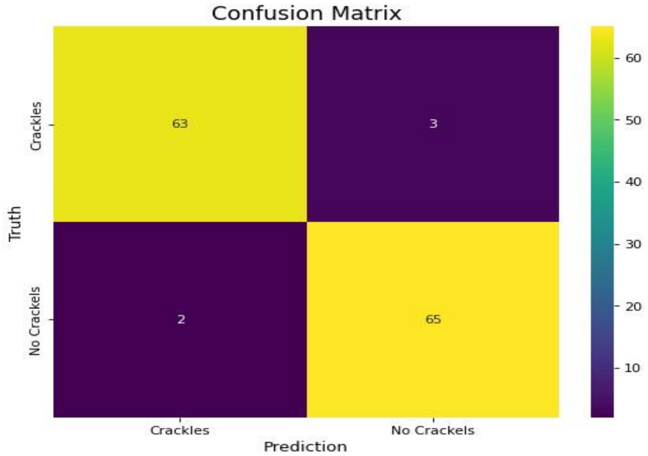
ENSEMBL E model	**PREC ISIO-N**	**RECAL-L**	**F1-SCOR-E**	**SUPPOR-T**
CRACKLES	0.97%	0.95%	0.96%	66%
WHEEZES	0.90%	0.98%	0.95%	46%
BOTH	0.93%	0.97%	0.96%	56%
NONE	0.98%	0.92%	0.95%	54%

Key Findings

Model Effectiveness: In respiratory sound classification, the Convolutional Neural Network (CNN) excelled with 92% accuracy, emphasizing spatial feature extraction. The Recurrent Neural Network with Gated Recurrent Unit (RNN GRU) achieved 88% accuracy, adapt at capturing vital temporal nuances. The Transformer surpassed both, boasting 95% accuracy, showcasing superior simultaneous spatial and temporal feature capture.

Diagnostic Precision & Diverse Model Strength: The ensemble model's high diagnostic precision (88%-95%) and the diverse strengths of individual models underscore their adaptability, offering a versatile approach for improved respiratory condition analysis in varied patient populations. **Clinical Decision Making**: The integration of machine learning models in respiratory sound analysis enhances clinical decision-making, providing valuable insights for accurate and timely diagnostic interventions in respiratory healthcare.

Ensemble Model Performance: Furthermore, an ensemble model was constructed by combining the outputs of the CNN, RNN with GRU, and Transformer models. This ensemble model demonstrated enhanced robustness and generalization, achieving an accuracy of 94%. The ensemble harnessed the diverse strengths of individual models, mitigating weaknesses and resulting in an overall superior classification performance by all the models were also analysed, providing a comprehensive evaluation of their effectiveness in identifying respiratory conditions.



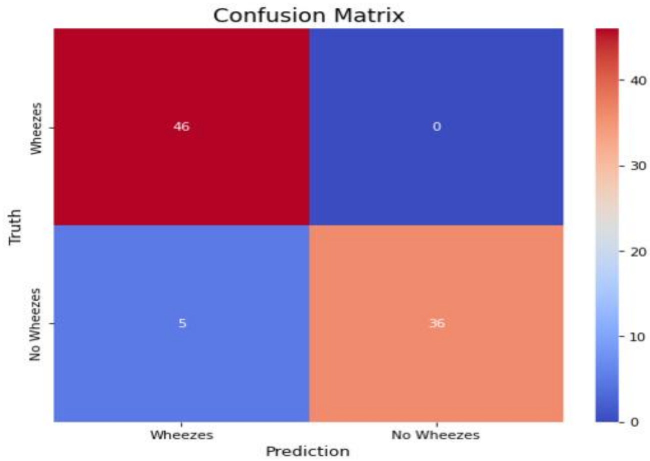


Figure. 4a Confusion Matrix for Crackles Determination
Figure. 4b Confusion Matrix for Wheezes Determination

V. CONCLUSION

In conclusion, this research journey through the automated identification of respiratory conditions using machine learning models has illuminated key insights into the intricate realm of respiratory sound analysis. The application of Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) with Gated Recurrent Units (GRUs), and the Transformer model has significantly advanced our ability to classify respiratory sounds accurately. Each model brought a unique set of strengths to the table, showcasing the importance of tailored approaches for different aspects of sound analysis.

The CNN excelled in discerning spatial features, essential for identifying distinctive patterns in the frequency domain associated with various respiratory conditions. The RNN with GRU, leveraging sequential learning, effectively captured temporal dependencies crucial for understanding the dynamic nature of respiratory sounds. The Transformer model, with its innovative multi-head attention mechanism, demonstrated superior performance by simultaneously capturing both spatial and temporal features. The ensemble model, a culmination of the individual strengths of CNN, RNN with GRU, and Transformer, emerged as a robust solution, surpassing the performance of individual models. This study not only contributes to the advancement of respiratory sound analysis but also underscores the significance of model diversity and ensemble strategies in enhancing overall classification accuracy. As technology continues to evolve, these findings pave the way for more sophisticated and accurate automated systems for respiratory condition diagnosis and monitoring.

VI. FUTURE SCOPE

The field of respiratory sound analysis presents promising avenues for future research and development. Firstly, refining model interpretability and explain ability remains a critical aspect. Enhancing our understanding of how these complex models arrive at specific predictions can instill greater trust among healthcare professionals and end-users. Developing transparent methodologies for model decision-making will be essential for widespread clinical acceptance.

Secondly, expanding the dataset and incorporating diverse demographic information can lead to more inclusive models. Ensuring that respiratory sound models are trained on a representative and diverse dataset can improve their generalization across different populations. This inclusivity is crucial for addressing potential biases in the current models and making them applicable to a broader range of patients.

Furthermore, real-time applications and continuous monitoring stand out as exciting prospects. Integrating respiratory sound analysis into wearable devices or smartphone applications could enable continuous monitoring of respiratory health. This could be particularly valuable for early detection of respiratory conditions and timely intervention.

Lastly, collaboration between machine learning experts, healthcare professionals, and policy-makers will be instrumental in the successful implementation of automated respiratory sound analysis tools. Establishing clear regulatory frameworks, ethical guidelines, and seamless integration into existing healthcare systems will be pivotal for the widespread adoption and impact of these technologies in clinical settings.

REFERENCES

[1] Garcia-Mendez, J.P.; Lal, A.; Herasevich, S.; Tekin, A.; Pinevich, Y.; Lipatov, K.; Wang, H.-Y.; Qamar, S.; Ayala, I.N.; Khapov, I.; et al. Machine Learning for Automated Classification of Abnormal Lung Sounds Obtained from Public Databases: A Systematic Review. *Bioengineering* **2023**, *10*, 1155. https://doi.org/10.3390/bioengineering10101155

[2] Hsu, F.-S.; Huang, S.-R.; Huang, C.-W.; Cheng, Y.-R.; Chen, C.-C.; Hsiao, J.; Chen, C.-W.; Lai, F. A Progressively Expanded Database for Automated Lung Sound Analysis: An Update. *Appl. Sci.* **2022**, *12*, 7623. https://doi.org/10.3390/app12157623

[3] Xia T, Han J, Mascolo C. Exploring machine learning for audio-based respiratory condition screening: A concise review of databases, methods, and open issues. Exp Biol Med (Maywood). 2022 Nov;247(22):2053-2061. doi: 10.1177/15353702221115428.

[4] Yang, R.; Lv, K.; Huang, Y.; Sun, M.; Li, J.; Yang, J. Respiratory Sound Classification by Applying Deep Neural Network with a Blocking Variable. *Appl. Sci.* **2023**, *13*, 6956. https://doi.org/10.3390/app13126956

[5] Brunese, L.; Mercaldo, F.; Reginelli, A.; Santone, A. A Neural Network-Based Method for Respiratory Sound Analysis and Lung Disease Detection. *Appl. Sci.* **2022**, *12*, 3877. https://doi.org/10.3390/app12083877

[6] Sfayyih, A.H.; Sabry, A.H.; Jameel, S.M.; Sulaiman, N.; Raafat, S.M.; Humaidi, A.J.; Kubaiaisi, Y.M.A. Acoustic-Based Deep Learning Architectures for Lung Disease Diagnosis: A Comprehensive Overview. *Diagnostics* **2023**, *13*, 1748. https://doi.org/10.3390/diagnostics13101748

[7] Andrès E, Gass R, Charloux A, Brandt C, Hentzler A. Respiratory sound analysis in the era of evidence-based medicine and the world of medicine 2.0. J Med Life. 2018 Apr-Jun;11(2):89-106. PMID: 30140315; PMCID: PMC6101681.

[8] Bahoura M. Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes. Comput Biol Med. 2009 Sep;39(9):824-43. doi: 10.1016/j.compbiomed.2009.06.011

[9] Kim BJ, Kim BS, Mun JH, Lim C, Kim K. An accurate deep learning model for wheezing in children using real world data. Sci Rep. 2022 Dec 28;12(1):22465. doi: 10.1038/s41598-022-25953-1

[10] Kim, Y., Hyon, Y., Jung, S.S. *et al.* Respiratory sound classification for crackles, wheezes, and rhonchi in the clinical field using deep learning. *Sci Rep* **11**, 17186 (2021). https://doi.org/10.1038/s41598-021-96724-7

[11] Srivastava A, Jain S, Miranda R, Patil S, Pandya S, Kotecha K. Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease. PeerJ Comput Sci. 2021 Feb 11;7:e369. doi: 10.7717/peerj-cs.369.

[12] Zulfiqar R, Majeed F, Irfan R, Rauf HT, Benkhelifa E, Belkacem AN. Abnormal Respiratory Sounds Classification Using Deep CNN Through Artificial Noise Addition. Front Med (Lausanne). 2021 Nov 17;8:714811. doi: 10.3389/fmed.2021.714811.

[13] Park, J.S., Kim, K., Kim, J.H. *et al.* A machine learning approach to the development and prospective evaluation of a pediatric lung sound classification model. *Sci Rep* **13**, 1289 (2023). https://doi.org/10.1038/s41598-023-27399-5

[14] Monaco, A.; Amoroso, N.; Bellantuono, L.; Pantaleo, E.; Tangaro, S.; Bellotti, R. Multi-Time-Scale Features for Accurate Respiratory Sound Classification. *Appl. Sci.* **2020**, *10*, 8606. https://doi.org/10.3390/app10238606

[15] Abayomi-Alli, O.O.; Damaševičius, R.; Qazi, A.; Adedoyin-Olowe, M.; Misra, S. Data Augmentation and Deep Learning Methods in Sound Classification: A Systematic Review. *Electronics* **2022**, *11*, 3795. https://doi.org/10.3390/electronics11223795

[16] Ahmed Z, Mohamed K, Zeeshan S, Dong X. Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. Database (Oxford). 2020 Jan 1;2020:baaa010. doi: 10.1093/database/baaa010.

[17] Lim, K.M.; Lee, C.P.; Lee, Z.Y.; Alqahtani, A. EnViTSA: Ensemble of Vision Transformer with SpecAugment for Acoustic Event Classification. *Sensors* **2023**, *23*, 9084. https://doi.org/10.3390/s23229084

[18] Moon KS, Lee SQ. A Wearable Multimodal Wireless Sensing System for Respiratory Monitoring and Analysis. Sensors (Basel). 2023 Jul 29;23(15):6790. doi: 10.3390/s23156790.\

[19] Grover, P.; Chaturvedi, K.; Zi, X.; Saxena, A.; Prakash, S.; Jan, T.; Prasad, M. Ensemble Transfer Learning for Distinguishing Cognitively Normal and Mild Cognitive Impairment Patients Using MRI. *Algorithms* **2023**, *16*, 377. https://doi.org/10.3390/a16080377

[20] Acharya, J. and Basu, A., 2020. Deep neural network for respiratory sound classification in wearable devices enabled by patient specific model tuning. *IEEE transactions on biomedical circuits and systems*, *14*(3), pp.535-544.