International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307 OPEN ACCESS

Available Online at :www.ijsrcseit.com

doi : https://doi.org/10.32628/CSEIT2361049

A Comprehensive Review on COVID-19 Cough Audio Classification through Deep Learning

Praveen Gupta*1, Sheshang Degadwala2

*1Research Scholar, Dept. of Computer Engineering, Sigma Institute of Engineering, Gujarat, India

praveengupta949@gmail.com1

2Associate Professor & Head of Department, Dept. of Computer Engineering, Sigma University, Gujarat, India

sheshang13@gmail.com 2

ARTICLEINFO

Article History:

Accepted: 10 Oct 2023

Published: 20 Nov 2023

Publication Issue

Volume 9, Issue 10

September-October-2023

Page Number

289-294

ABSTRACT

This review paper provides a comprehensive analysis of the advancements in COVID-19 cough audio classification through deep learning techniques. With the ongoing global pandemic, there is a growing need for non-intrusive and rapid diagnostic tools, and the utilization of audio-based methods for COVID-19 detection has gained considerable attention. The paper systematically reviews and compares various deep learning models, methodologies, and datasets employed for COVID-19 cough audio classification. The effectiveness, challenges, and future directions of these approaches are discussed, shedding light on the potential of audio-based diagnostics in the context of the current public health crisis.

Keywords: COVID-19, Cough audio classification, Deep learning, Convolutional Neural Networks, Recurrent Neural Networks, Feature extraction, Diagnostic tools.

I. INTRODUCTION

In the face of the persisting global COVID-19 pandemic, the quest for effective diagnostic tools has intensified, prompting a closer look at innovative methodologies. Among these, the examination of cough audio signals has emerged as a promising avenue for the non-invasive detection of COVID-19 symptoms [1,3]. This review paper undertakes a comprehensive exploration of the role played by deep learning techniques in the classification of COVID-19 cough audio. As the world grapples with the complexities of the pandemic, the need for accurate and rapid diagnostic solutions becomes increasingly evident, making the intersection of deep learning and audio analysis a focal point for research and development [4,6].

Deep learning models, particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have exhibited remarkable capabilities in discerning nuanced patterns within cough audio associated with COVID-19. This review systematically investigates the diverse methodologies employed for feature extraction and model training, shedding light on the strengths and limitations of these approaches. Beyond technical intricacies, the paper also explores the broader implications of integrating deep learning into COVID-19 diagnostics, emphasizing the potential impact on public health [7,8]. By offering a nuanced analysis of the current landscape, this review aims to guide researchers and practitioners in navigating the evolving field of COVID-19 cough audio classification through deep learning.

The review critically evaluates the deep learning models deployed for COVID-19 cough audio classification, emphasizing the roles of convolutional neural networks (CNNs) and spectrogram features. Beyond technical intricacies, the paper explores the challenges associated with feature extraction and model training, providing insights into the effectiveness and limitations of existing approaches. As the global community navigates the complexities of the pandemic, this review aims to contribute a nuanced understanding of the evolving landscape, guiding future research directions and policy decisions. By addressing both the technical aspects and broader implications, the paper seeks to foster a comprehensive perspective on the potential of deep learning in advancing COVID-19 diagnostics through cough audio analysis [11].

II. LITERATURE STUDY

In [1], Ulukaya et al. introduced MSCCov19Net, a multi-branch deep learning model for COVID-19 detection using cough sounds. Their study provides insights into the application of acoustic features for accurate detection, highlighting the model's potential in enhancing diagnostic capabilities.

Kim et al., in [2], presented a COVID-19 detection model leveraging acoustic features from cough sounds. Their research explores the practical implementation of the model, demonstrating its potential applications in real-world scenarios.

In [3], Almutairi proposed a multimodal AI-based non-invasive COVID-19 grading framework using deep learning, manta ray, and fuzzy inference system. The study provides a comprehensive approach to COVID-19 grading, incorporating diverse data sources for robust analysis.

Chowdhury et al., in [4], employed a machine learning ensemble-based MCDM method for COVID-19 detection from cough sounds. Their research contributes to the ongoing efforts in developing accurate and reliable diagnostic tools.

Hoang et al. proposed a cough-based deep learning framework for COVID-19 detection in [5]. Their study highlights the potential of utilizing cough sounds for effective and non-invasive diagnosis.

Aly and Alotaibi, in [6], introduced a novel deep learning model for COVID-19 detection based on wavelet features extracted from Mel-scale spectrogram of patients' cough and breathing sounds. Their research provides a unique approach to feature extraction and model development.

In [7], Ashby et al. proposed a cough-based COVID-19 detection method with audio quality clustering and confidence measure based learning. Their study emphasizes the importance of audio quality in enhancing the reliability of detection models.

Pahar et al., in [8], presented an automatic classification model for tuberculosis and COVID-19 cough using deep learning. Their research explores the potential of a unified approach to detect both diseases.

Abayomi-Alli et al., in [9], proposed a COVID-19 detection method using deep breathing sounds and image augmentation. Their study explores the synergy of sound spectrum analysis and deep learning techniques for accurate detection.

Ren et al., in [10], introduced an attention-based ensemble learning method for cough-based COVID-

19 recognition. Their research contributes to the development of robust and complementary representations for accurate detection.

In [11], Mohammed et al. presented an ensemble learning approach for digital coronavirus preliminary screening from cough sounds. Their study emphasizes the importance of ensemble methods in enhancing the reliability of screening models.

Chang et al., in [12], introduced CovNet, a transfer learning framework for automatic COVID-19 detection from crowd-sourced cough sounds. Their research demonstrates the potential of leveraging crowd-sourced data for large-scale screening.

Rao et al., in [13], proposed COVID-19 detection using cough sound analysis and deep learning algorithms. Their study explores the application of deep learning in developing reliable diagnostic tools.

Pahar et al., in [14], presented a COVID-19 cough classification model using machine learning and global smartphone recordings. Their research showcases the potential of utilizing widespread data sources for effective disease classification.

In [15], Loey and Mirjalili focused on COVID-19 cough sound symptoms classification using scalogram image representation and deep learning models. Their work likely contributes to the ongoing research on image-based representations for disease classification.

III.METHODOLOGY

A. Dataset [1,3,6]

The dataset utilized in this study, referred to as CoughVid, comprises over 25,000 crowdsourced cough recordings. These recordings cover a diverse range of participant demographics, including various ages, genders, geographic locations, and COVID-19 statuses. To enhance the dataset's reliability, a subset of 2,800 recordings was meticulously labeled by four experienced physicians. This subset serves as one of the most extensive expert-labeled cough datasets available, facilitating a broad spectrum of cough audio classification tasks.

B. Pre-Processing [2,4,8]

Clipping: The audio signals undergo clipping to ensure uniformity and manageability in subsequent processing stages.

Noise Removal: To enhance the signal-to-noise ratio, a noise removal process is applied, mitigating potential interference and improving the overall quality of the cough recordings.

C. Feature Extraction [2,4,7,14]

Acoustic Features: Relevant acoustic features are extracted from the pre-processed cough audio signals. These features may include but are not limited to pitch, intensity, and other characteristics essential for differentiating cough patterns.

Spectrogram Features: Spectrogram features are extracted to capture the frequency and time-domain characteristics of the cough signals, providing a comprehensive representation of the audio data.

D. Transfer Learning Models [1,3,8,12,15]

Utilizing transfer learning, established deep learning models are employed for cough audio classification:

AlexNet: Renowned for its effectiveness in image classification tasks, AlexNet has been adapted to harness its deep neural network architecture for the unique spectrogram-based features extracted from cough audio. This adaptation aims to leverage the model's ability to capture high-level abstractions, facilitating the identification of distinctive patterns and characteristics within cough recordings.

VggNet: Utilizing the VggNet architecture, known for its proficiency in capturing hierarchical features, proves advantageous in handling the intricate and complex patterns present in cough audio. The model's deep structure allows for the extraction of multi-scale representations, enabling a more comprehensive understanding of the diverse acoustic features inherent in cough recordings.

ResNet: The ResNet model, characterized by its innovative residual learning framework, is employed to address the challenges of training deeper networks for cough audio classification. By mitigating vanishing gradient issues, ResNet enhances the model's capacity to discern subtle nuances and intricate characteristics within cough recordings, potentially leading to improved performance.

EfficientNet: Incorporating EfficientNet into the framework capitalizes on the model's reputation for efficiency and scalability. By optimizing resource utilization without compromising performance, EfficientNet is well-suited for cough audio classification tasks, providing a balance between computational efficiency and classification accuracy.

CNN (Convolutional Neural Network): The use of a convolutional neural network is motivated by its ability to exploit local patterns within acoustic features. This approach enables the extraction of meaningful spatial hierarchies, allowing the model to discern relevant information and patterns at different levels of granularity within the audio data, ultimately enhancing the classification of cough recordings.

By combining a diverse dataset, rigorous pre-processing techniques, and powerful transfer learning models, this methodology aims to develop a robust and accurate cough audio classification system for COVID-19 screening. The chosen transfer learning models offer a balance between computational efficiency and model performance, ensuring effective utilization in real-world applications.

TABLE I
COMPARATIVE ANALYSIS

Feature/ Model	Pros.	Cons.
Acoustic Features	- Capture essential characteristics of audio signals. - Widely used in traditional audio processing.	- Limited ability to capture complex patterns in audio. - May not perform well on tasks requiring high-level abstraction.
Spectrogram	- Represents frequency content over time. - Better captures temporal and spectral information.	- Fixed time and frequency resolution. - May not capture fine-grained details.
Mel-Spectrogram	- Emphasizes relevant frequency bands. - Human perception-based representation.	- Loss of fine frequency details. - Fixed time and frequency resolution.
AlexNet	- Pioneering deep convolutional neural network (CNN). - Effective feature extraction in images.	- Relatively large and computationally expensive. - Prone to overfitting on smaller datasets.
VggNet	- Simplified architecture with uniform filter sizes. - Easy to understand and implement.	- Deeper architecture, leading to increased computational complexity. - Memory-intensive.
ResNet	- Introduced	- Complex

	residual learning, easing training of very deep networks. - Mitigates vanishing gradients.	architectures may lead to overfitting. - Increased computational requirements.
EfficientNet	- Achieves high accuracy with fewer parameters. - Efficient scaling across depth, width, and resolution.	- Requires careful balancing of scaling coefficients. - May not perform as well on some specialized tasks.
CNN (Convolutional Neural Network)	- Excellent feature learning from image data. - Effective for image classification tasks.	- Limited ability to capture sequential or temporal patterns. - May struggle with variable-sized inputs.

IV.CONCLUSION

In conclusion, the integration of Mel-Spectrograms with modified layers within Transfer Learning Models demonstrates promising potential for achieving enhanced performance in future applications. The utilization of Mel-Spectrograms, which prioritize relevant frequency bands aligned with human perception, coupled with the adaptability and knowledge transfer capabilities of Transfer Learning Models, offers a robust framework for feature extraction and representation learning. Further exploration and refinement of the modified layers within these models, tailored to the unique characteristics of Mel-Spectrograms, could yield even more optimized results. This avenue of research holds the prospect of advancing the state-of-the-art in audio signal processing, paving the way for improved performance across various domains such as speech recognition, music analysis, and other audio-based applications.

V. REFERENCES

[1] S. Ulukaya, A. A. Sarıca, O. Erdem, and A. Karaali, "MSCCov19Net: multi-branch deep learning model for COVID-19 detection from cough sounds," Medical and Biological Engineering and Computing, vol. 61, no. 7, pp. 1619–1629, 2023, doi: 10.1007/s11517-023-02803-4.

[2] S. Kim, J. Y. Baek, and S. P. Lee, "COVID-19 Detection Model with Acoustic Features from Cough Sound and Its Application," Applied Sciences (Switzerland), vol. 13, no. 4, 2023, doi: 10.3390/app13042378.

[3] S. A. Almutairi, "A multimodal AI-based non-invasive COVID-19 grading framework powered by deep learning, manta ray, and fuzzy inference system from multimedia vital signs," Heliyon, vol. 9, no. 6, p. e16552, 2023, doi: 10.1016/j.heliyon.2023.e16552.

[4] N. K. Chowdhury, M. A. Kabir, M. M. Rahman, and S. M. S. Islam, "Machine learning for detecting COVID-19 from cough sounds: An ensemble-based MCDM method," Computers in Biology and Medicine, vol. 145, no. March, p. 105405, 2022, doi: 10.1016/j.compbiomed.2022.105405.

[5] T. Hoang, L. Pham, D. Ngo, and H. D. Nguyen, "A Cough-based deep learning framework for detecting COVID-19," Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, vol. 2022-July, no. 1, pp. 3422–3425, 2022, doi: 10.1109/EMBC48229.2022.9871179.

[6] M. Aly and N. S. Alotaibi, “A novel deep learning model to detect COVID-19 based on wavelet features extracted from Mel-scale spectrogram of patients’ cough and breathing sounds,” Informatics in Medicine Unlocked, vol. 32, no. June, p. 101049, 2022, doi: 10.1016/j.imu.2022.101049.

[7] A. E. Ashby et al., “Cough-based COVID-19 detection with audio quality clustering and confidence measure based learning Khuong An Nguyen,” Proceedings of Machine Learning Research, vol. 179, no. Ml, pp. 1–20, 2022.

[8] M. Pahar et al., “Automatic Tuberculosis and COVID-19 cough classification using deep learning,” International Conference on Electrical, Computer, and Energy Technologies, ICECET 2022, no. July, pp. 20–22, 2022, doi: 10.1109/ICECET55527.2022.9873469.

[9] O. O. Abayomi-Alli, R. Damaševičius, A. A. Abbasi, and R. Maskeliūnas, “Detection of COVID-19 from Deep Breathing Sounds Using Sound Spectrum with Image Augmentation and Deep Learning Techniques,” Electronics (Switzerland), vol. 11, no. 16, 2022, doi: 10.3390/electronics11162520.

[10] Z. Ren, Y. Chang, W. Nejdl, and B. W. Schuller, “Learning complementary representations via attention-based ensemble learning for cough-based COVID-19 recognition,” Acta Acustica, vol. 6, pp. 0–4, 2022, doi: 10.1051/aacus/2022029.

[11] E. A. Mohammed, M. Keyhani, A. Sanati-Nezhad, S. H. Hejazi, and B. H. Far, “An ensemble learning approach to digital corona virus preliminary screening from cough sounds,” Scientific Reports, vol. 11, no. 1, pp. 1–11, 2021, doi: 10.1038/s41598-021-95042-2.

[12] Y. Chang, X. Jing, Z. Ren, and B. W. Schuller, “CovNet: A Transfer Learning Framework for Automatic COVID-19 Detection From Crowd-Sourced Cough Sounds,” Frontiers in Digital Health, vol. 3, no. August 2021, pp. 1–11, 2022, doi: 10.3389/fdgth.2021.799067.

[13] S. Rao, V. Narayanaswamy, M. Esposito, J. J. Thiagarajan, and A. Spanias, “COVID-19 detection using cough sound analysis and deep learning algorithms,” Intelligent Decision Technologies, vol. 15, no. 4, pp. 655–665, 2021, doi: 10.3233/IDT-210206.

[14] M. Pahar, M. Klopper, R. Warren, and T. Niesler, “COVID-19 cough classification using machine learning and global smartphone recordings,” Computers in Biology and Medicine, vol. 135, no. June, p. 104572, 2021, doi: 10.1016/j.compbiomed.2021.104572.

[15] M. Loey and S. Mirjalili, “COVID-19 cough sound symptoms classification from scalogram image representation using deep learning models Mohamed,” Computers in Biology and Medicine, no. January, 2021.

Cite this article as :

Praveen Gupta, Sheshang Degadwala, "A Comprehensive Review on COVID-19 Cough Audio Classification through Deep Learning", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 10, pp.289-294, September-October-2023. Available at doi : https://doi.org/10.32628/CSEIT2361049

Journal URL : https://ijsrcseit.com/CSEIT2361049