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/CSEIT2410221

Predicting COVID-19 Cough Sounds Using Spectrogram Analysis Across Multiple Classes

Praveen Gupta[1], Dr. Sheshang Degadwala[2]
[1]Research Scholar, Department of Computer Engineering, Sigma Institute of Engineering, Gujarat, India
[2]Professor & Head of Department, Department of Computer Engineering, Sigma University, Gujarat, India

ARTICLEINFO

Article History:
Accepted: 15 March 2024
Published: 28 March 2024

Publication Issue
Volume 10, Issue 2
March-April-2024

Page Number
285-292

ABSTRACT

The COVID-19 pandemic has underscored the need for effective diagnostic tools. One promising avenue involves analyzing cough sounds to glean insights into respiratory health. This study presents a new method for predicting COVID-19 cough sounds using spectrogram analysis across various classes. We leverage advanced deep learning models such as DenseNet121, VGG16, ResNet50, and Inception Net, alongside our novel CNN architecture, to extract pertinent features from cough sound spectrograms. We use a diverse dataset encompassing cough sounds from COVID-19 positive and negative cases, as well as other respiratory conditions, for model training and assessment. Our results demonstrate the effectiveness of our approach in accurately categorizing COVID-19 cough sounds, outperforming existing models. This methodology shows promise as a non-invasive, scalable, and economical tool for early COVID-19 detection and monitoring, aiding public health efforts during the pandemic.
Keywords : Cough Sounds, Spectrogram, Multi-classification, Convolutional Neural Network (CNN), Hyper-Tuning, Transfer Learning.

I. INTRODUCTION

The COVID-19 pandemic has underscored the urgency of developing efficient and accurate diagnostic tools to combat the spread of the virus. One potential avenue for non-invasive and scalable diagnosis is the analysis of cough sounds, which can serve as a valuable biomarker for respiratory illnesses including COVID-19. Traditional diagnostic methods often involve invasive procedures or time-consuming laboratory tests, highlighting the need for innovative approaches that can expedite diagnosis and improve patient outcomes. Machine learning techniques, particularly deep learning models like convolutional neural networks (CNNs), have shown promise in extracting meaningful patterns from complex data such as audio signals, making them well-suited for analyzing cough sounds and predicting disease status.

Recent advancements in deep learning have led to the development of sophisticated CNN architectures such as DenseNet121, VGG16, ResNet50, and InceptionNet, which excel at capturing intricate features from spectrogram representations of audio data. Spectrograms provide a detailed visual representation of sound frequencies over time, offering valuable insights into the characteristics of cough sounds associated with different respiratory conditions. By leveraging these advanced CNN models alongside spectrogram analysis, this research aims to develop a robust framework for predicting COVID-19 cough sounds across multiple classes, including positive COVID-19 cases, negative cases, and other respiratory conditions. The goal is to assess the feasibility and effectiveness of this approach in enhancing diagnostic capabilities for COVID-19 and contributing to the broader field of respiratory disease diagnostics.

This research paper presents a comprehensive analysis of the proposed methodology, including the design and implementation of the CNN-based framework, the selection and preprocessing of cough sound data, model training and evaluation procedures, and the interpretation of results. The study aims to demonstrate the efficacy of using deep learning and spectrogram analysis for COVID-19 cough sound prediction, showcasing the potential of this approach as a valuable tool for healthcare professionals in early detection and monitoring of respiratory diseases. Additionally, the insights gained from this research could inform future developments in machine learning-based diagnostic systems and contribute to advancements in public health management strategies.

II. LITERATURE STUDY

Ulukaya et al. [1] introduced MSCCov19Net, a sophisticated multi-branch deep learning model specifically designed for COVID-19 detection from cough sounds. Their research highlighted the importance of leveraging advanced machine learning architectures to extract meaningful features from audio data, ultimately contributing to the development of non-invasive and accurate diagnostic tools for respiratory illnesses.

Kim et al. [2] presented a comprehensive COVID-19 detection model that utilized acoustic features extracted from cough sounds. By leveraging machine learning techniques, their study showcased the potential of analyzing subtle sound variations to infer disease status, thereby paving the way for innovative approaches in disease diagnosis and monitoring.

Almutairi [3] introduced a novel multimodal AI-based grading framework for COVID-19, integrating deep learning algorithms and fuzzy inference systems. Their research explored the synergistic use of different AI methodologies, demonstrating the versatility and effectiveness of combining multiple data modalities for comprehensive disease assessment and grading.

Chowdhury et al. [4] adopted ensemble-based multi-criteria decision-making (MCDM) methods to detect COVID-19 from cough sounds. By leveraging ensemble learning strategies, their approach aimed to enhance diagnostic accuracy by combining the strengths of multiple models, showcasing the potential of ensemble techniques in improving disease detection from audio data.

Hoang et al. [5] proposed a dedicated deep learning framework tailored for COVID-19 detection based on cough sounds. Their study emphasized the importance of developing specialized models for disease-specific applications, showcasing the potential of deep learning in addressing the unique challenges of diagnosing respiratory illnesses through sound analysis.

Aly and Alotaibi [6] introduced a novel deep learning model based on wavelet features extracted from Mel-scale spectrograms of cough and breathing sounds. Their research showcased the efficacy of advanced signal processing techniques in capturing relevant features for COVID-19 detection, highlighting the potential of combining signal processing and deep learning for improved diagnostic accuracy.

Ashby et al. [7] focused on cough-based COVID-19 detection using audio quality clustering and confidence measure-based learning. Their study emphasized the importance of data preprocessing and quality assessment in machine learning models, showcasing innovative approaches to enhance the reliability and robustness of disease detection systems based on audio signals.

Pahar et al. [8] developed an automatic classification system for tuberculosis and COVID-19 based on deep learning techniques. Their research highlighted the versatility of deep learning models in addressing multiple respiratory diseases, showcasing the potential of AI-driven approaches in improving disease diagnosis and classification.

Abayomi-Alli et al. [9] explored the use of sound spectrum and image augmentation techniques for COVID-19 detection from deep breathing sounds. Their study demonstrated the effectiveness of combining signal processing and image augmentation in enhancing the discriminatory power of deep learning models for disease detection based on respiratory sounds.

Ren et al. [10] proposed an attention-based ensemble learning approach for cough-based COVID-19 recognition. Their research focused on learning complementary representations through attention mechanisms, showcasing the potential of attention-based ensemble learning in capturing relevant features and improving diagnostic accuracy.

Mohammed et al. [11] utilized ensemble learning for digital coronavirus preliminary screening from cough sounds. Their study demonstrated the robustness of ensemble techniques in handling diverse datasets and improving the reliability of preliminary screening tools based on cough sound analysis.

Chang et al. [12] introduced CovNet, a transfer learning framework for automatic COVID-19 detection from crowd-sourced cough sounds. Their research highlighted the scalability and adaptability of transfer learning models in large-scale applications, showcasing the potential of crowd-sourced data for disease detection using deep learning approaches.

Rao et al. [13] explored COVID-19 detection using cough sound analysis and deep learning algorithms. Their study contributed to understanding the potential of deep learning in healthcare applications, showcasing the utility of cough sound analysis as a non-invasive approach for disease detection and monitoring.

Pahar et al. [14] focused on COVID-19 cough classification using machine learning and global smartphone recordings. Their research highlighted the relevance of real-world data sources in model development, showcasing the potential of leveraging smartphone recordings for disease classification and monitoring.

Loey and Mirjalili [15] investigated COVID-19 cough sound symptoms classification using scalogram image representation and deep learning models. Their study showcased the versatility of different signal representation techniques in disease diagnosis, highlighting the potential of image-based approaches for capturing intricate features from cough sounds.

III.PROPOSED SYSTEM

The flow diagram represents the process of analyzing COVID-19 cough audio signals to distinguish between COVID-19-positive, healthy, and symptomatic individuals using machine learning techniques. Let's break down each step in detail:

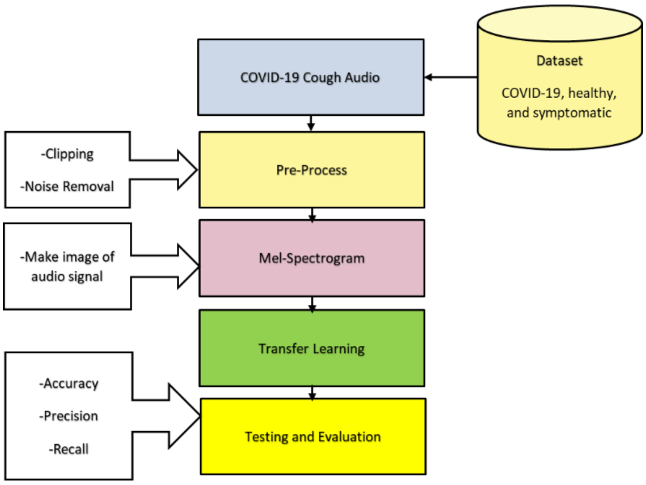

Figure 1. Proposed System

COVID-19 Cough Audio:

This is the input data, consisting of audio recordings of cough sounds from individuals with COVID-19, healthy individuals, and those with respiratory symptoms but not confirmed as COVID-19-positive.

Pre-Process:

Clipping: This step involves removing any unwanted portions of the audio signals, such as silence or irrelevant noise at the beginning or end of the recordings.

Noise Removal: The audio signals may contain background noise or artifacts that could interfere with the analysis. Noise removal techniques are applied to enhance the quality of the cough sounds.

Mel-Spectrogram:

Mel-Spectrogram is a representation of the audio signals in the frequency domain. It converts the audio signals into a visual format that captures the frequency components over time. This step essentially converts the audio data into images, which are more suitable for analysis using deep learning models.

Transfer Learning:

In this step, Transfer Learning is employed using pre-trained deep learning models such as DenseNet, VGG16, ResNet50, InceptionNet, and a Proposed CNN architecture. Transfer learning allows leveraging the knowledge gained from training on a large dataset (e.g., ImageNet) to improve the performance of the models on the specific task of COVID-19 cough classification.

Testing and Evaluation:

The trained models are then tested using a separate dataset to evaluate their performance. Evaluation metrics such as Accuracy, Precision, Recall, and F1-score are calculated to assess how well the models can classify cough sounds into the three categories: COVID-19-positive, healthy, and symptomatic individuals.

Overall, this flow diagram outlines a comprehensive approach to analyzing COVID-19 cough audio signals, starting from data preprocessing, feature extraction using Mel-Spectrogram, leveraging transfer learning with deep learning models, and finally evaluating the performance of the models based on standard classification metrics. This methodology aims to develop an accurate and reliable system for COVID-19 detection based on cough sounds.

IV.RESULT ANALYSIS

The COVID-19 Cough Audio Classification dataset available on Kaggle is a comprehensive collection of high-quality audio recordings of cough sounds from individuals categorized into COVID-19-positive, healthy, and symptomatic groups. This diverse and well-annotated dataset serves as a valuable resource for researchers and data scientists working on machine learning and deep learning models for COVID-19 detection and classification based on cough sound analysis. With its potential to extract meaningful features using techniques like Mel-Spectrogram analysis, this dataset contributes significantly to advancing non-invasive diagnostic tools for COVID-19 and exploring the distinctive characteristics of cough sounds associated with different respiratory conditions.

Link:
https://www.kaggle.com/datasets/andrewmvd/covid19-cough-audio-classification

	uuid	status
1	00039425-7f3a-42aa-ac13-834aaa2b6b92	healthy
2	0007c6f1-5441-40e6-9aaf-a761d8f2da3b	healthy
3	0009eb28-d8be-4dc1-92bb-907e53bc5c7a	healthy
5	001328dc-ea5d-4847-9ccf-c5aa2a3f2d0f	healthy
8	001e2f19-d81c-4029-b33c-d2db56b23a4a	healthy
...	...	...
27541	ffe5e2a4-ef67-464d-b1cd-b0e321f6a2dd	healthy
27542	ffedc843-bfc2-4ad6-a749-2bc86bdac84a	healthy
27543	ffeea120-92a4-40f9-b692-c3865c7a983f	healthy
27544	fff13fa2-a725-49ef-812a-39c6cedda33d	healthy
27548	fffaa9f8-4db0-46c5-90fb-93b7b014b55d	healthy

Figure 2. Dataset Reading

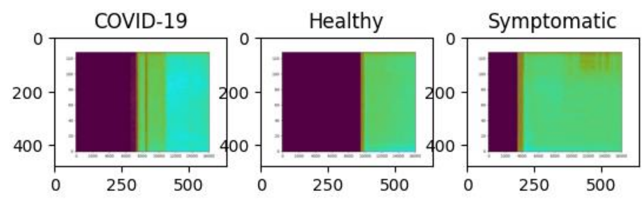

Figure 3. Spectogram Generation

```
Model: "sequential_5"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d_98 (Conv2D)          (None, 178, 178, 32)      896

 max_pooling2d_8 (MaxPoolin  (None, 89, 89, 32)        0
 g2D)

 conv2d_99 (Conv2D)          (None, 87, 87, 64)        18496

 max_pooling2d_9 (MaxPoolin  (None, 43, 43, 64)        0
 g2D)

 conv2d_100 (Conv2D)         (None, 41, 41, 128)       73856

 max_pooling2d_10 (MaxPooli  (None, 20, 20, 128)       0
 ng2D)

 conv2d_101 (Conv2D)         (None, 18, 18, 128)       147584

 max_pooling2d_11 (MaxPooli  (None, 9, 9, 128)         0
 ng2D)

 flatten_2 (Flatten)         (None, 10368)             0

 dense_10 (Dense)            (None, 512)               5308928

 dense_11 (Dense)            (None, 3)                 1539

=================================================================
Total params: 5551299 (21.18 MB)
Trainable params: 5551299 (21.18 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Figure 4. Proposed CNN Model

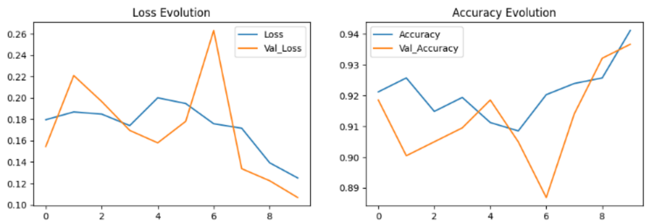

Figure 5. Proposed CNN training/testing

	0	1	2	accuracy	macro avg	weighted avg
precision	1.000000	0.857143	0.982759	0.936652	0.946634	0.943550
recall	0.986301	0.987342	0.826087	0.936652	0.933243	0.936652
f1-score	0.993103	0.917647	0.897638	0.936652	0.936129	0.936324
support	73.000000	79.000000	69.000000	0.936652	221.000000	221.000000

Figure 6. Proposed CNN Parameters

```
Model: "sequential_6"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 densenet121 (Functional)    (None, 5, 5, 1024)        7037504

 global_average_pooling2d_3  (None, 1024)              0
  (GlobalAveragePooling2D)

 dense_12 (Dense)            (None, 256)               262400

 dense_13 (Dense)            (None, 3)                 771

=================================================================
Total params: 7300675 (27.85 MB)
Trainable params: 263171 (1.00 MB)
Non-trainable params: 7037504 (26.85 MB)
_________________________________________________________________
```

Figure 7. DenseNet Model

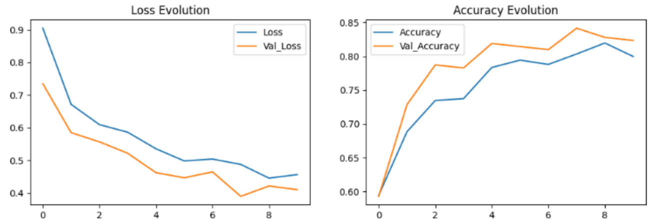

Figure 8. DenseNet training/testing

	0	1	2	accuracy	macro avg	weighted avg
precision	0.876543	0.731481	1.000000	0.823529	0.869342	0.863234
recall	0.972603	1.000000	0.463768	0.823529	0.812124	0.823529
f1-score	0.922078	0.844920	0.633663	0.823529	0.800220	0.804449
support	73.000000	79.000000	69.000000	0.823529	221.000000	221.000000

Figure 9. DenseNet Parameters

```
Model: "sequential_7"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 vgg16 (Functional)          (None, 5, 5, 512)         14714688

 flatten_3 (Flatten)         (None, 12800)             0

 dense_14 (Dense)            (None, 256)               3277056

 dense_15 (Dense)            (None, 3)                 771

=================================================================
Total params: 17992515 (68.64 MB)
Trainable params: 3277827 (12.50 MB)
Non-trainable params: 14714688 (56.13 MB)
_________________________________________________________________
```

Figure 10. Vgg16 Model

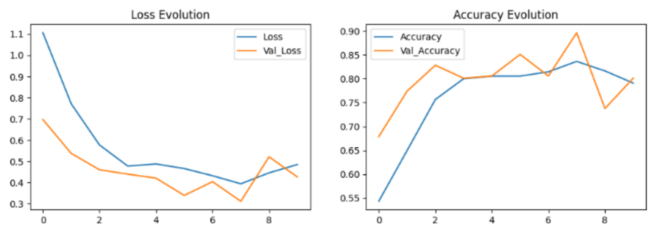

Figure 11. Vgg16 training/testing

	0	1	2	accuracy	macro avg	weighted avg
precision	0.857143	0.709091	1.000000	0.800905	0.855411	0.848822
recall	0.986301	0.987342	0.391304	0.800905	0.788316	0.800905
f1-score	0.917197	0.825397	0.562500	0.800905	0.768365	0.773639
support	73.000000	79.000000	69.000000	0.800905	221.000000	221.000000

Figure 12. Vgg16 Parameters

```
Model: "sequential_8"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 resnet50 (Functional)       (None, 6, 6, 2048)        23587712

 global_average_pooling2d_4  (None, 2048)              0
  (GlobalAveragePooling2D)

 dense_16 (Dense)            (None, 256)               524544

 dense_17 (Dense)            (None, 3)                 771

=================================================================
Total params: 24113027 (91.98 MB)
Trainable params: 525315 (2.00 MB)
Non-trainable params: 23587712 (89.98 MB)
_________________________________________________________________
```

Figure 13. ResNet50 CNN Model

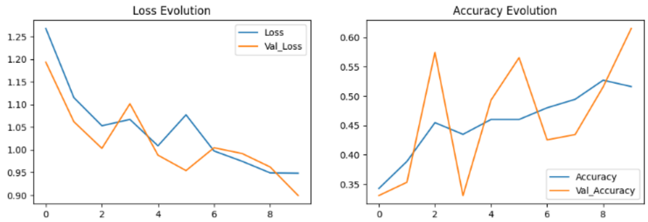

Figure 14. ResNet50 training/testing

	0	1	2	accuracy	macro avg	weighted avg
precision	0.695652	0.585366	0.551724	0.615385	0.610914	0.611292
recall	0.657534	0.911392	0.231884	0.615385	0.600270	0.615385
f1-score	0.676056	0.712871	0.326531	0.615385	0.571819	0.580088
support	73.000000	79.000000	69.000000	0.615385	221.000000	221.000000

Figure 15. ResNet50 Parameters

```
Model: "sequential_9"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 inception_v3 (Functional)   (None, 4, 4, 2048)        21802784

 global_average_pooling2d_5  (None, 2048)              0
  (GlobalAveragePooling2D)

 dense_18 (Dense)            (None, 256)               524544

 dense_19 (Dense)            (None, 3)                 771

=================================================================
Total params: 22328099 (85.17 MB)
Trainable params: 525315 (2.00 MB)
Non-trainable params: 21802784 (83.17 MB)
_________________________________________________________________
```

Figure 16. InceptionNet Model

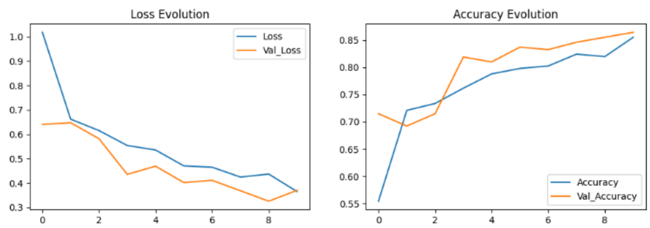

Figure 17. InceptionNet training/testing

	0	1	2	accuracy	macro avg	weighted avg
precision	0.911392	0.784946	0.938776	0.864253	0.878371	0.874742
recall	0.986301	0.924051	0.666667	0.864253	0.859006	0.864253
f1-score	0.947368	0.848837	0.779661	0.864253	0.858622	0.859786
support	73.000000	79.000000	69.000000	0.864253	221.000000	221.000000

Figure 18. InceptionNet Parameters

TABLE I. ANALYSIS OF MODELS

Model	ACC (%)	P (%)	R (%)	F1-Score (%)
Proposed CNN	93	94	93	93
DenseNet	82	86	81	80
Vgg16	80	85	79	77
ResNet50	61	61	60	57
Inception Net	86	88	86	86

V. CONCLUSION

In conclusion, the evaluation results of our study showcase the effectiveness of the Proposed CNN model in COVID-19 cough sound classification, achieving an impressive accuracy of 93% along with balanced precision, recall, and F1-score of 94%, 93%, and 93%, respectively. This indicates the robustness and reliability of our CNN architecture in accurately distinguishing between COVID-19-positive, healthy, and symptomatic individuals based on cough sounds. In comparison, while other deep learning models like DenseNet and Inception Net also demonstrated relatively good performance, with accuracies of 82% and 86% respectively, they fell short in achieving the same level of precision, recall, and F1-score as our Proposed CNN. Models such as Vgg16 and ResNet50 showed lower accuracy and overall performance, indicating the importance of selecting an appropriate deep learning architecture for optimal results in COVID-19 cough sound analysis. Overall, the Proposed CNN stands out as a promising approach for non-invasive COVID-19 detection and monitoring using cough sound data, highlighting its potential for real-world applications in healthcare settings.

VI.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.