

# A Multi-Modal Respiratory Disease Exacerbation Prediction Technique Based on a Novel Spatio-Temporal Machine Learning Architecture and Local Environmental Sensor Networks

Rohan Tan Bhowmik<sup>1</sup>

**Abstract**—Chronic respiratory diseases, such as the chronic obstructive pulmonary disease and asthma, are a serious health crisis, affecting a large number of people globally and inflicting major costs on the economy. Current methods for assessing the progression of respiratory symptoms are either subjective and inaccurate, or complex and cumbersome, and do not incorporate environmental factors. Lacking predictive assessments and early intervention, unexpected exacerbations can lead to hospitalizations and high medical costs.

This work presents a multi-modal solution for predicting the exacerbation risks of respiratory diseases, such as COPD, based on a novel spatio-temporal machine learning architecture for real-time and accurate respiratory events detection, and tracking of local environmental and meteorological data and trends. The proposed new machine learning architecture blends key attributes of both convolutional and recurrent neural networks, allowing extraction of both spatial and temporal features encoded in respiratory sounds, thereby leading to accurate classification and tracking of symptoms. Combined with the data from environmental and meteorological sensors, and a predictive model based on retrospective medical studies, this solution can assess and provide early warnings of respiratory disease exacerbations.

This research will improve the quality of patients' lives through early medical intervention, thereby reducing hospitalization rates and medical costs.

## I. INTRODUCTION

Chronic respiratory diseases affect a large fraction of the world population, with Chronic Obstructive Pulmonary Disease (COPD) affecting 235 million and asthma affecting 339 million people worldwide, according to the World Health Organization [1]. Lacking effective early intervention, COPD and asthma cost over \$130 Billion annually in the U.S. alone [2].

Existing methods of diagnosis and tracking of these disease conditions in clinical practice, including widely-used patient questionnaires, are highly variable due to the subjectivity of definition, perception, and reporting of respiratory events. In fact, many respiratory diseases are often over- or under-diagnosed. Based on the study by Diab. et al., approximately 70 percent of COPD cases worldwide may be underdiagnosed, while 30 to 60 percent of those diagnosed with COPD may not have the disease at all [3]. As the treatment of respiratory diseases often requires the prescription of steroids, misdiagnosis can cause serious problems.

<sup>1</sup>Rohan T. Bhowmik is a student at the Harker School, 500 Saratoga Ave, San Jose CA, USA 95138. He can be contacted at: 23rohanb@students.harker.org

Currently, no passive monitoring method exists for accurately predicting the exacerbation of respiratory conditions. A number of cough detection methods have been reported, but no accurate real-time tracking technique exists for passive and continuous monitoring. Commonly used methods involve subjective reporting, often leading to frequent and dangerous misdiagnosis [4-6]. Besides the respiratory conditions of the patient, environmental factors such as pollen, humidity, air quality, etc., also play a significant role in the disease progression, exacerbations, and hospitalizations [7]. However, currently there is no multi-modal predictive technique that incorporates the trends of both respiratory events and local environmental factors in order to assess the progression of the patient's conditions.

Thus, the development of an accurate and real-time predictive solution for respiratory disease exacerbation that is easily accessible is highly needed, based on monitoring of patient's respiratory events as well as the local environmental and meteorological parameters. The recent advances in connected devices, sensors, data technologies, and machine learning techniques present a significant opportunity to develop respiratory telehealth capabilities, allowing for accurate remote monitoring of patient conditions as well as assessing potential exacerbations with predictive Artificial Intelligence (AI) models.

This work presents a multi-modal solution for real-time COPD exacerbation prediction that includes a novel spatio-temporal artificial intelligence architecture for cough detection, real-time cough-count and frequency monitoring, analytics of the local environmental and meteorological factors utilizing data from sensor networks, and exacerbation prediction using both respiratory event tracking and environmental conditions based on retrospective medical studies. The goal of this research is to develop an early-warning system based on AI and multi-factor analysis to reduce hospitalizations and medical costs, and demonstrate the feasibility of deploying a passive, continuous, remote patient monitoring and telehealth solution for chronic respiratory diseases.

## II. PRIOR RESEARCH

Researchers have previously identified that monitoring a patient's respiratory events can be utilized to assess the patient's condition [8]. In order to automate this process, a number of cough detection solutions have been proposed [9-15]. A survey of previously reported techniques, performances and limitations are listed in Fig. 1. Earlier methods

used relatively simpler techniques, such as probabilistic statistical models on waveform data [9], but also yielded low accuracies. On the other hand, more recent studies have used specialized equipment and complex setups, such as wireless wearable patch sensors [13] or spirometers [15], to achieve relatively better results. However, no single technique simultaneously meets all of the following requirements: highly accurate, efficient, passive and continuous monitoring, and does not need extra equipment.

With the recent advancements in the field of artificial intelligence, researchers have moved towards exploring solutions based on Deep Neural Networks (DNN). Several researchers have demonstrated detection of cough with either Convolutional Neural Networks (CNN) or Recurrent Neural Networks (RNN). Traditional CNN models are based on learning and detecting spatial features in the data and are typically used for image-based analysis, whereas RNN models are based on extracting temporal features and are often used for time-sequenced tasks such as speech processing. Since respiratory sounds, when converted to spectrograms, encode key spatial and temporal signatures, neither of the traditional models is well suited for respiratory event classification.

Some researchers have recently reported combined Convolutional-Recurrent Neural Networks (CRNN) for acoustic analysis [16-20]. These CRNN models have been shown to work better than CNN and RNN in both image processing and sequence-related tasks [16, 19], but these frameworks do not fully utilize the spatial/temporal feature extraction capabilities of CNN/RNN architectures as they are created by simply stacking RNN layers after CNN layers in a sequential manner. The development of machine learning architecture based on deeply meshed spatio-temporal feature learning for respiratory sound classification has not been previously explored.

Medical researchers have also shown that several key environmental and meteorological factors are related to the exacerbations of COPD [7]; however, this research has not been combined with real-time monitoring of respiratory events to develop predictive models for exacerbations.

### III. METHODS

#### A. Proposed Multi-Modal System Architecture

In this project, a novel multi-modal COPD patient monitoring and exacerbation prediction system has been developed based on real-time analysis and tracking of both respiratory events and environmental factors. As shown in Fig. 2, the system architecture consists of three stages: i) a detection module, ii) an environmental module, and finally, iii) a prediction module.

The detection module utilizes a new spatio-temporal machine learning algorithm for accurately detecting coughs from real-time audio and tracking the patient's cough count and frequency. Simultaneously, the environmental module acquires local environmental and meteorological data from nearby weather stations and sensor networks to calculate the percentage increase of exacerbation risks in any location around the world based on the results of retrospective

medical studies. Finally, the prediction module combines the historical cough count data and trends from the detection module and the calculated exacerbation risk increase from the environmental module in order to forecast the progression of the patient's conditions, and alert the patients and caregivers for early interventions.

#### B. Detection Module

The detection module, as shown on the left-hand side of the system architecture diagram in Fig. 2, consists of a new AI model for real-time detection and tracking of cough. As described earlier, previously reported models for respiratory sound analysis are based on the traditional convolutional, recurrent, or the more recent convolutional-recurrent structures. In this project, a new machine learning algorithm has been developed that incorporates a novel hybrid framework by deeply meshing convolutional and recurrent architectures, enabling more efficient extraction and analysis of spatio-temporal features, leading to better accuracies for classifying and tracking respiratory events.

The following subsections describe the new spatio-temporal machine learning framework for classifying and tracking respiratory events, creation of the dataset to train and test the model, the results of benchmarking the proposed model with traditional neural network architectures, and a live demonstration application showcasing the capability of real-time classification of respiratory sounds.

*1) A New Machine Learning Architecture for Respiratory Sound Analysis:* The new AI model, henceforth referred to as the Spatio-Temporal Artificial Intelligence Network (STAIN), interweaves convolutional neural network models within a recurrent neural network architecture, allowing for sequential image analysis over the time domain. The architecture of the STAIN framework is shown in Fig. 3. First, the respiratory sound files are converted to corresponding spectrogram images by performing Fast Fourier Transforms. The resulting spectrogram is split into 200 millisecond slices, which are used as inputs for the machine learning model.

As illustrated in Fig. 3, the machine learning model architecture incorporates a hybrid network based on a deep mesh integration of convolutional and recurrent architectures to learn spatio-temporal features. The STAIN framework consists of a CNN model which evaluates the corresponding audio slices and outputs its predicted confidence. The CNN architecture is a variation of Yann Lecun's seminal LeNet model [21], which can flexibly adapt to any image dataset. Specifically, it consists of two groupings of Convolutional Layers of 2x2 kernels and 2x2 Maximum Pooling Layers followed by Rectified Linear Unit (ReLU) activation function. Then, the resulting data is flattened into a one-dimensional array before feeding it into two Fully Connected (Dense) Layers to reduce the number of neurons down to just one. The final output is then passed through a Sigmoid Layer to obtain a value between (0, 1).

The CNNs analyzing separate parts of the input image enable spatial feature extraction, while the Encoders passing down compressed inputs as RNN's hidden variables enable

Ref.	Detection Technique	Technical Merits	Equipment Used	Technique Achievements	Limitations
Barry et al., 2006 [9]	PNN on raw waveform	96% SP 80% SE	Audio Microphone	An early method for cough detection using neural network on raw waveform	Achieved high specificity but low sensitivity (cannot identify "No Cough" well)
Liu et al., 2015 [10]	Cough event classification by pretrained DNN	90% SP 85% SE ~89% ACC	Audio Microphone	Utilized spectrograms for cough detection	Accuracy is mediocre, not suitable for precise monitoring
Li et al., 2015 [11]	CNN (AlexNet) on spectrogram features	98.6% ACC 0.977 F1	Audio Microphone	Used transfer learning on audio, achieved high accuracy	Large network creates high latency, not suitable for real-time monitoring
Amoh et al., 2016 [12]	CNN and RNN on spectrograms	92.7% SP for CNN 87.7% SE for RNN	Audio Microphone	Comparative analysis of CNN vs. RNN for cough detection on audio, found that CNN performs better with higher specificity, RNN had higher sensitivity	Both models have mediocre accuracy, not suitable for precise monitoring
Elfaramawy et al., 2018 [13]	Wearable Patch Sensors	N/A	Proprietary Patch Sensors	Using gyro sensors on abdomen and chest to monitor breathing pattern with no AI, abnormal disruption of pattern is identified as a cough	Requires patch sensors, supporting equipment need to be worn at all times, inconvenient for everyday use
Drugman et al., 2019 [14]	Evaluating the efficacy of 6 sensors to detect coughs	94.5% SP 94.4% SE	Audio and Contact Microphones	Comparative analysis of several sensors for cough detection, found that the method using sound and contact gave best results	Requires contact microphones and supporting equipment to be worn at all times, inconvenient for everyday use
Soliński et al., 2020 [15]	DNN on Spirometry Curves	91% SP 86% SE 91% ACC	Spirometer	Analyzed airflow through spirometer to differentiate normal breath vs. cough	Can only detect coughs through spirometer, impractical for broad deployment

Fig. 1. A survey of previously reported techniques for automatic cough detection (PNN = Probabilistic Neural Network; SP = Specificity; SE = Sensitivity; ACC = Accuracy). The above techniques follow these observations: i) generally higher accuracies were achieved with more complex models utilizing spectrograms; ii) techniques aided with extra equipment produced better results; iii) no single technique simultaneously meets all of the following requirements: highly accurate, efficient, passive and continuous monitoring, does not need extra equipment.

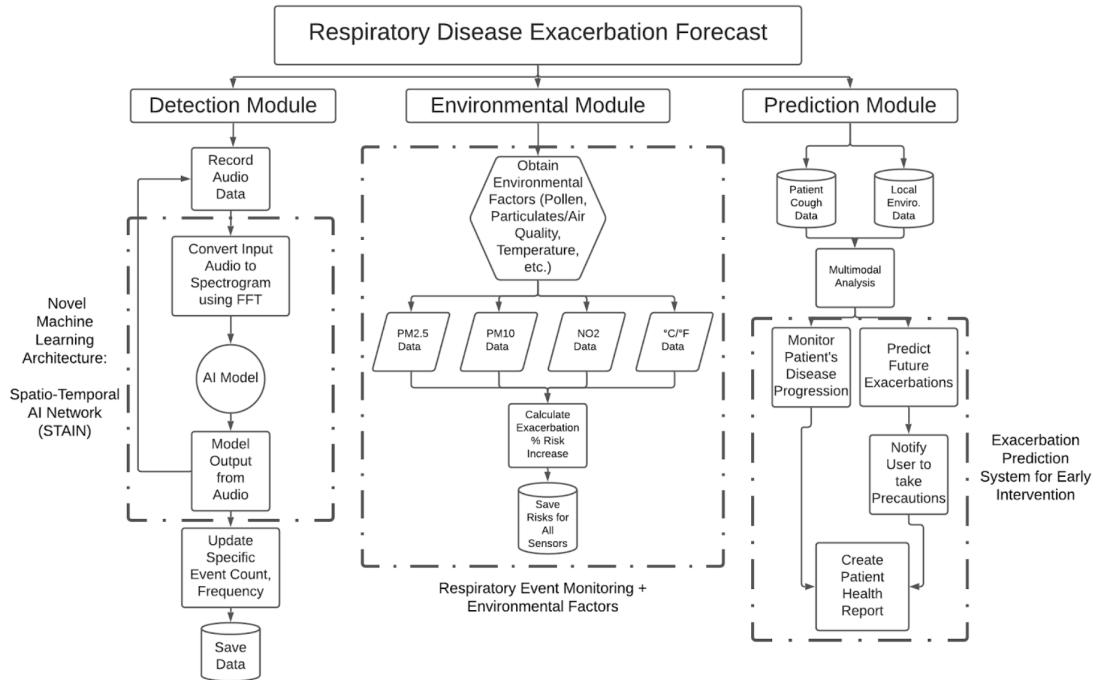


Fig. 2. This flowchart represents the proposed system architecture for real-time multi-modal exacerbation prediction. The detection module depicts the respiratory event analysis system using a novel spatio-temporal artificial intelligence neural network. The prediction module depicts the disease exacerbation prediction system using the output of the machine learning model and environmental factors. The prediction module takes the respiratory event data and trends from the detection module, forecasts the progression of the patient's conditions, and provides necessary alerts for early intervention.

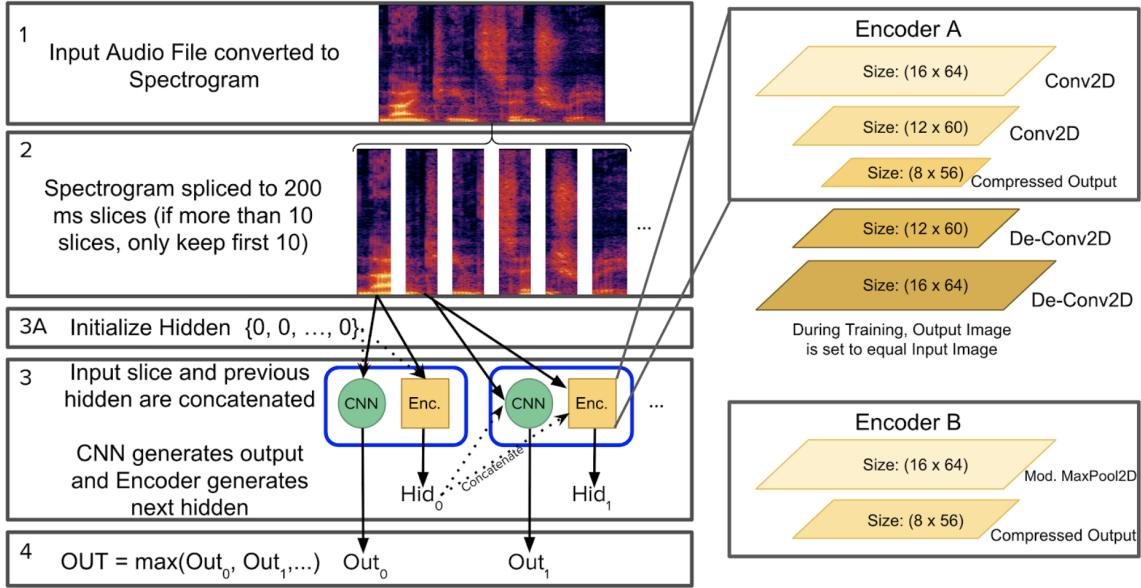


Fig. 3. Architecture for the new machine learning model, which is referred to as the spatio-temporal artificial intelligence network (STAIN). This proposed AI model deeply blends the elements of both convolutional and recurrent neural networks, and effectively learns both spatial and temporal features encoded within the respiratory sound spectrograms for accurate classifications.

temporal feature extraction. Various designs for the Encoder have been explored, starting with a simple architecture consisting of a single Maximum Pooling layer, shrinking the input into a hidden variable. A simple Variational Auto-Encoder (VAE) has also been created, consisting of two Deconvolutional Layers followed by Convolutional Layers.

Effectively, each slice of the spectrogram image is assigned to an RNN unit, wherein a CNN generates an output and the Encoder generates the hidden data. Each output represents the probability of a cough during that slice. The hidden outputs carry on information from previous slices and are concatenated to the next slice. The final output is the maximum of all the outputs from all slices. All the codes in this project were written in Python, and the machine learning models were implemented using the PyTorch Libraries.

2) *Creation of the Dataset:* In order to train and evaluate the proposed STAIN machine learning model as well as benchmark with other state-of-the-art models including CNN, RNN and CRNN, an augmented dataset of audio segments were created and partitioned into 10,000 training files with coughs, 10,000 training files without coughs, 1,000 testing files with coughs, and 1,000 testing files without coughs. The models were trained only on the 20,000 training files and tested only on the 2,000 testing files in order to objectively evaluate and compare the performance of various models.

First, roughly 500 cough sound files were downloaded from the Free Sound Database (FSD) from Kaggle's audio tagging competition [22] and every file was adjusted to only contain either a cough burst or coughing fit. The cough files were sufficiently diverse, containing many variations of coughs from individuals of both genders and from a wide range of ages (from babies to elderly). Each file also has its

unique recording quality, mimicking the varying degrees of audio quality from different devices.

In order to augment the data, the rest of the audio files from Kaggle's FSD were utilized. To create an augmented audio file, an empty audio file is created with a duration randomly chosen between 2 seconds and 5 seconds. Then, using the PyDub Library, a randomly chosen number of non-cough files from the FSD are superimposed on the targeted augmented file. Each of the added audio files are placed at a randomly chosen timestamp, with audio exceeding the augmented files trimmed off. The result of this process creates an augmented audio file categorized as "No Cough". To turn it into a "Cough" file, one of the cough files from the FSD is added in a similar fashion. Additionally, each added file's decibel gain is randomized to simulate sounds from varying distances.

3) *Benchmarking and Results:* Using the dataset described in the previous section, rigorous evaluations of the four different AI models were performed. The results of these analyses are shown in Fig. 4 and Fig. 5, which present the following performance metrics: sensitivity, specificity, accuracy, Matthews Correlation Coefficients, and the confusion matrices.

As these results illustrate, compared to RNN's temporal feature analysis, CNN's spatial analysis was better suited for classifying spectrograms. CRNN, created by simply stacking the CNN and RNN components, could not bring out the best of both architectures, performing worse than CNN. In contrast, the proposed new machine learning model, STAIN, performed better than all other models using its architecture for deeply meshed spatio-temporal feature analysis.

4) *Demonstration of the Detection Module:* A live demo application for the real-time cough detection module has

Performance Metrics	Convolutional Neural Network (CNN)	Recurrent Neural Network (RNN)	Convolutional Recurrent Neural Network (CRNN)	Spatio-Temporal AI Network (STAIN)
Sensitivity	91.6%	83.4%	89.8%	92.7%
Specificity	93.8%	89.2%	92.1%	94.2%
Accuracy	92.7%	86.3%	91.0%	93.4%
Matthews Correlation Coefficient	85.42%	72.72%	81.92%	86.91%
Performance	Spatial feature analysis, works well on spectrograms	Temporal feature analysis, less than optimal	Spatial and temporal feature analysis, not specialized	Spatio-temporal feature analysis, functions best

Fig. 4. Summary of results of the comparative study of CNN, RNN, CRNN, and the proposed STAIN machine learning models for cough detection. The sensitivity, specificity, accuracy, and Matthews Correlation Coefficient metrics were obtained for all four models using the same datasets. As can be seen, the STAIN model outperforms all the other traditional AI models with its deeply meshed spatio-temporal feature extraction architecture, which is more advantageous for effectively classifying respiratory events.

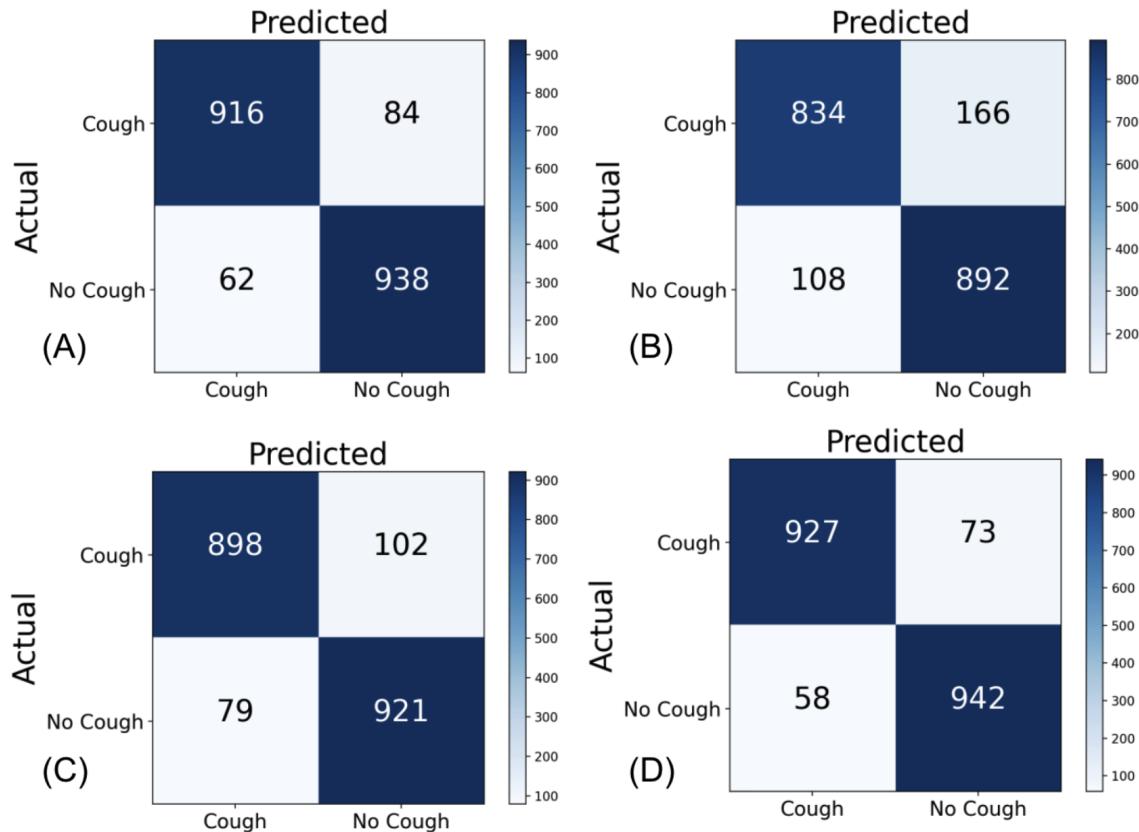


Fig. 5. Confusions Matrices, created using Matplotlib, for (A) CNN, (B) RNN, (C) CRNN, and (D) the proposed STAIN machine learning models. The new STAIN architecture outperforms the traditional neural network architectures for more accurate cough detection.

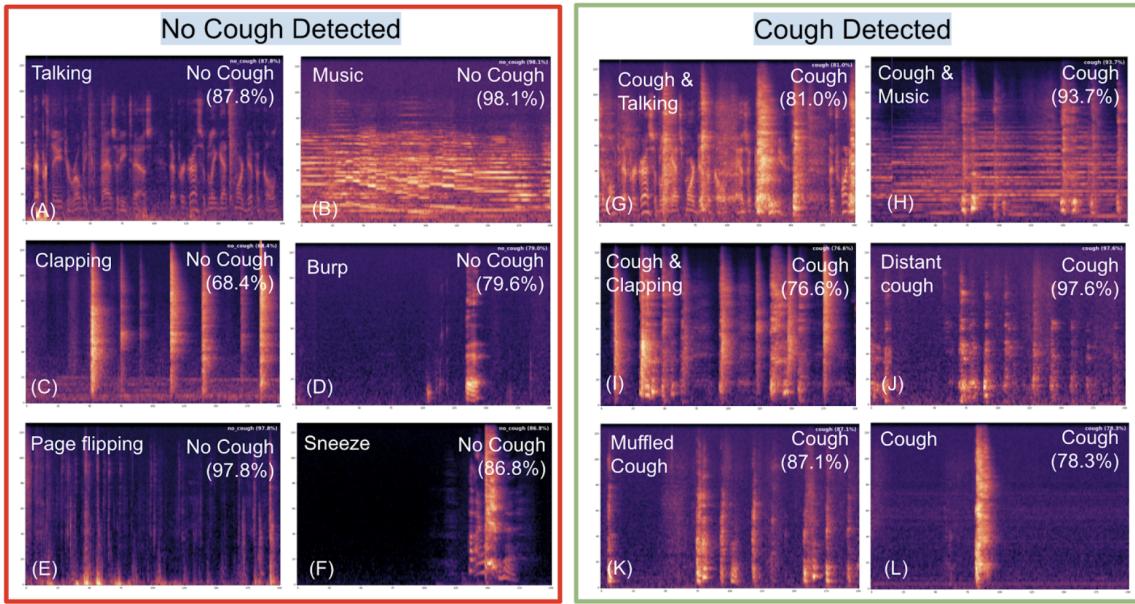


Fig. 6. Screenshots of the live demonstration of the cough detection module based on the new spatio-temporal machine learning model. The real-time application, implemented on a laptop computer, captures user-generated sounds using its integrated microphones, converts the sound into spectrogram images, processes through the STAIN model to detect the presence of cough, and displays the results on the screen.

Parameters	Risk Increase Coefficient [7, 23]	Safety Standards	Rate [7, 23]
PM 2.5	1%	12 $\mu\text{g}/\text{m}^3$ [24]	10. $\mu\text{g}/\text{m}^3$
PM 10	0.8%	7 $\mu\text{g}/\text{m}^3$ [24]	10. $\mu\text{g}/\text{m}^3$
$\text{NO}_2$	2%	101.23 $\mu\text{g}/\text{m}^3$ [25]	10. $\mu\text{g}/\text{m}^3$
Temperature	2.61%	68.0 °F	-1.8 °F

Fig. 7. Correlations between the degradation of the environmental and meteorological factors and the increase in COPD exacerbation risks, derived from retrospective medical studies [7, 23]. As an example, these studies demonstrated that an increase in  $\text{NO}_2$  concentration by 10  $\mu\text{g}/\text{m}^3$  resulted in about 2% increase in the risk. These correlations were used to estimate the overall risk trends based on the real-time data from local sensors.

$$1\% * \frac{([PM_{2.5}] - 12 \frac{\mu\text{g}}{\text{m}^3})}{10. \frac{\mu\text{g}}{\text{m}^3}} + 0.8\% * \frac{([PM_{10}] - 7 \frac{\mu\text{g}}{\text{m}^3})}{10. \frac{\mu\text{g}}{\text{m}^3}} + 2\% * \frac{([NO_2] - 101.23 \frac{\mu\text{g}}{\text{m}^3})}{10. \frac{\mu\text{g}}{\text{m}^3}} + 4.7\% * \frac{68.0^\circ F - T_F}{1.8^\circ F}$$

Fig. 8. Equation to estimate the increase in COPD exacerbation risks as a function of environmental and meteorological factors ( $PM_{2.5}$ ,  $PM_{10}$ ,  $NO_2$ , and  $T_F$  for Temperature), derived based on the retrospective medical studies [7, 23].

been developed. This application, running on a laptop computer, captures user-generated sounds using the built-in microphones of the computer, converts the sound files into spectrogram images, processes the data through the STAIN machine learning model, classifies and tracks the cough count and cough frequency over time. The results are presented on the computer screen with a live display of the spectrogram images corresponding to the sound, superimposed with the classification results of the cough events. Fig. 6 shows the representative screenshots of the application running real-time, and correctly classifying talking, clapping, page flipping, music, burp, and sneezes as “No Cough” (left-hand side of Fig. 6), whereas successfully detecting cough events superimposed with the same background sound environments (right-hand side of Fig. 6).

### C. Environmental Module

While the detection module presented in the previous section tracks real-time cough frequency for patient-specific analysis, the environmental module offers local area-wide environmental and meteorological factor analysis. By examining certain environmental indicators, a patient’s increase of COPD exacerbation likelihood can be determined.

Breathing air quality is one of the most crucial factors in human health; poor air quality can cause any person’s health to significantly deteriorate and is an increasingly important issue following the advent of rapid industrialization. Especially since their lungs are compromised due to inflammation, COPD patients are extremely susceptible to exacerbations caused by bad air quality. A seminal retrospective study analyzed hospitalization and exacerbation rates for COPD patients as functions of the local environmental

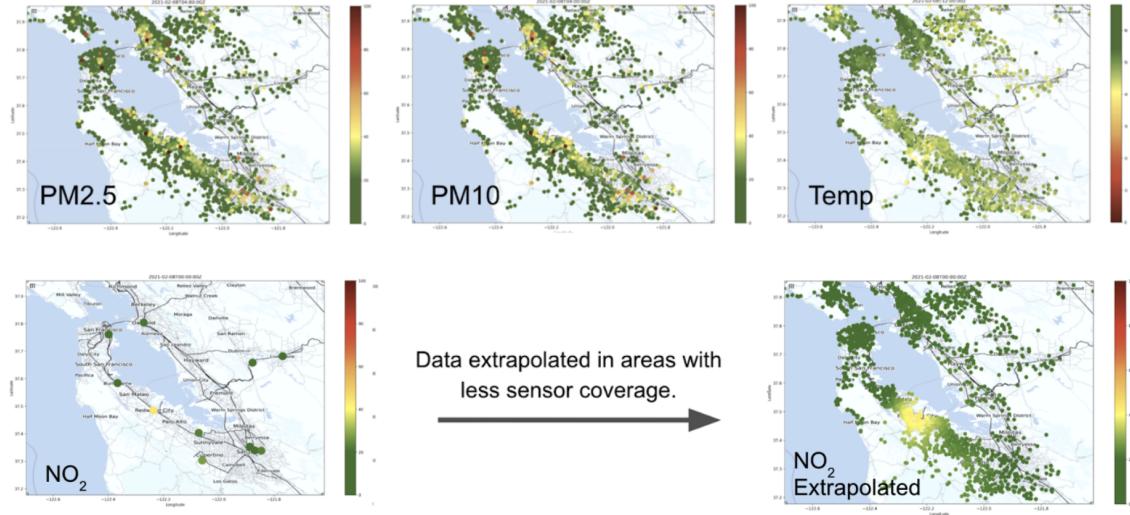


Fig. 9. Data maps for the relevant environmental and meteorological factors ( $\text{PM}_{2.5}$ ,  $\text{PM}_{10}$ ,  $\text{NO}_2$ , and Temperature), obtained from the sensors deployed by PurpleAir and the WAQI data platform. An extrapolation method was used to estimate the data in areas with sparse sensor coverage.

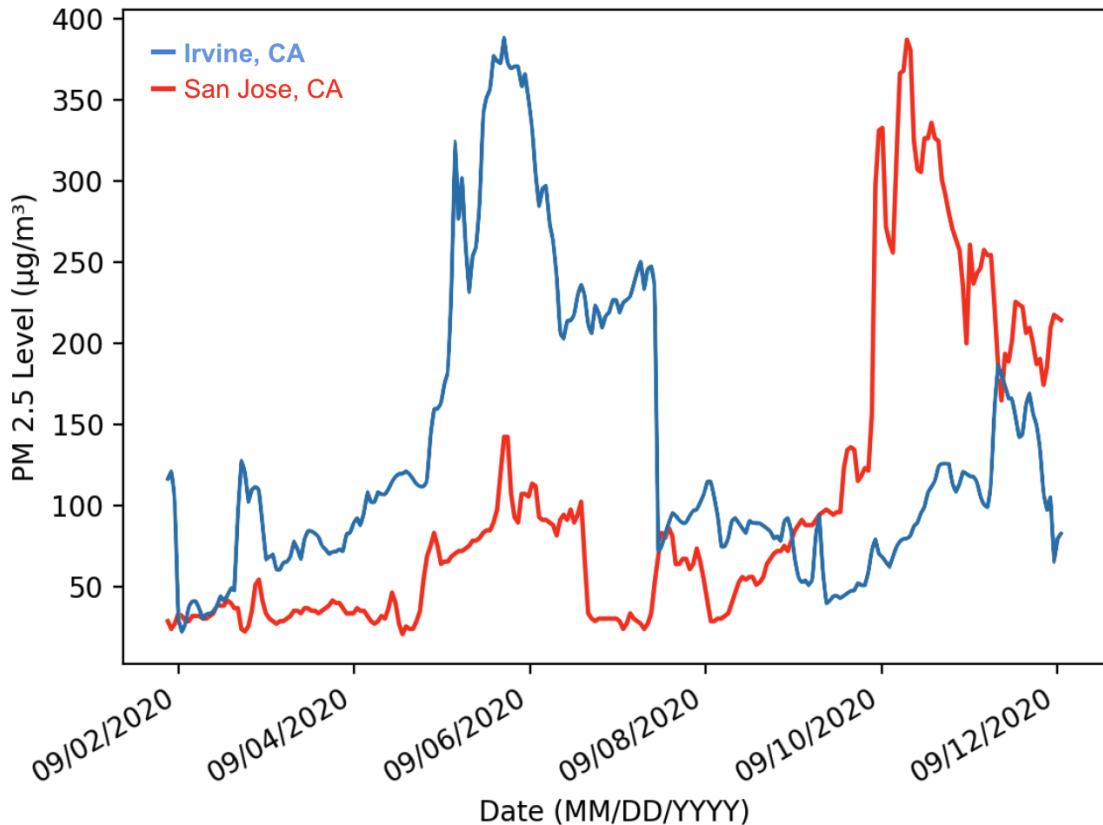


Fig. 10. Sensors deployed by PurpleAir in Irvine and San Jose showed that the  $\text{PM}_{2.5}$  concentration spiked to dangerous levels during Sept. 2-13, 2020, fire season. The onsets of spikes on Sept 6 and Sept 10 correspond to the El Dorado Fire and the SCU Lightning Fire events.

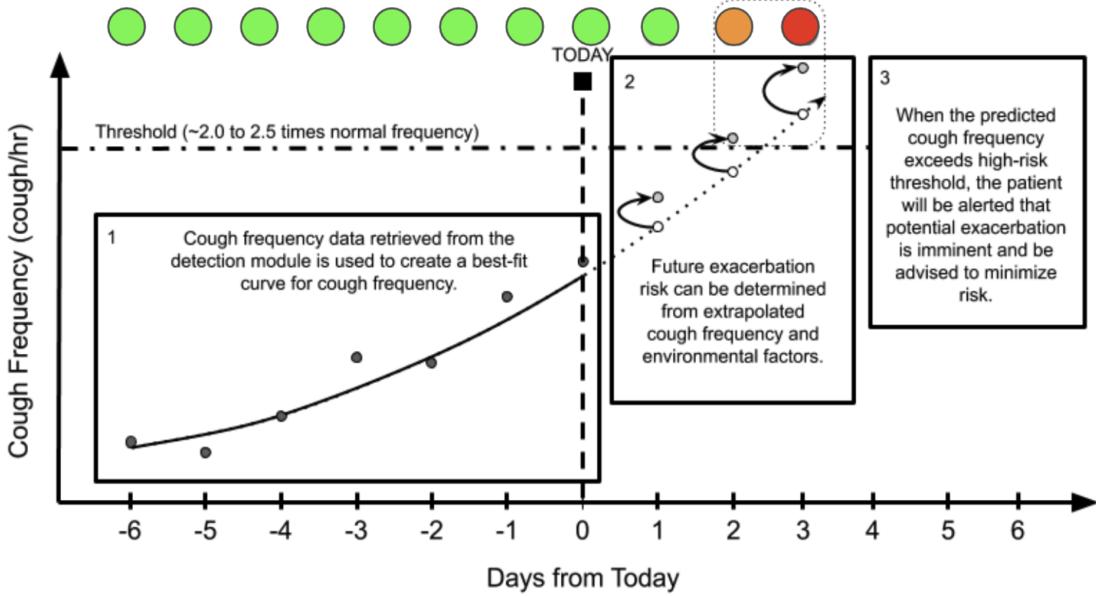


Fig. 11. Illustration of the procedures implemented within the prediction module that forecasts the expected progression of the condition of the patient in the days ahead. This final step in the multi-modal architecture combines the results from the respiratory sound analysis performed by the machine learning model of the detection module, and the environmental and meteorological factors and trends analysis conducted by the environmental module. By extrapolating the cough frequency trends, along with the predicted exacerbation risks due to the environmental and meteorological data, the system can alert the patient and caregivers of the imminent risks and preempt medical interventions to potentially reduce hospitalization costs.

and meteorological factors, including the concentration of fine particulate matters (where  $PM_x$  refers to particles or droplets present in the air that are  $x$  micrometers or less in width),  $NO_2$ , and temperature variations [7, 23]. These medical studies established that the percentage exacerbation risk increases are directly proportional to  $PM_{2.5}$  and  $PM_{10}$  levels,  $NO_2$  concentrations, and temperature variations. The details of the findings are outlined in Fig. 7, with each increase/decrease of the “Rate” from “Safety Standards” constituting an additional “Risk Increase Coefficient” for exacerbations.

Based on the results of these retrospective medical studies, an equation has been formulated in this project to estimate the percentage exacerbation risk increase using the four environmental and meteorological parameters in the patient’s location, as shown in Fig. 8. If a factor falls below the threshold standard, its contribution to the final risk percentage is zero; otherwise, it follows the formula outlined in Fig. 8.

In order to generate a real-time risk map that would represent the exacerbation risk increase for an individual given the environmental factors in the patient’s location, the environmental and climatological data measured by sensors deployed by PurpleAir which are accessible via an open-source database [26], and  $NO_2$  readings from the World Air Quality Index (WAQI) data platform [27], have been incorporated into the above equation and overlaid on the geographical map of the region. Moreover, an extrapolation method has been developed to estimate the data at a specific location using the data from the sensors deployed in adjacent areas. As an example, Fig. 9 shows the data map for  $PM_{2.5}$ ,

$PM_{10}$ , Temperature, and  $NO_2$  from over 6000 sensors in the San Francisco Bay Area. As a spot check for the data, Fig. 10 shows the  $PM_{2.5}$  concentrations recorded by the PurpleAir sensors in Irvine and San Jose areas during the first half of September, 2020. The onsets of spikes on Sept 6 and Sept 10 correspond to the El Dorado Fire and the SCU Lightning Fire events, respectively.

#### D. Prediction Module

Finally, the prediction module combines the results of the respiratory sound analysis from the detection module and the environmental and meteorological factors analysis from the environmental module to forecast a patient’s expected conditions.

Previously reported medical research studies have determined average cough frequencies for COPD-affected smokers, affected ex-smokers, healthy smokers, and healthy non-smokers [24, 25]. Thus, by extrapolating the progression in cough frequency as determined by the STAIN machine learning model and exacerbation risk increase from environmental factors from the data trends, a patient’s expected condition is determined.

This method is illustrated in Fig. 11. First, based on the continuous respiratory event classifications performed by the STAIN machine learning model within the detection module, a best-fit curve is created to determine the patient’s cough frequency trend. Next, the future exacerbation risks are derived based on the extrapolated cough frequency data and the increased risks due to environmental and meteorological factors as determined by the correlations established by the

retrospective medical studies, as explained in the previous section. If the prediction module forecasts exceeding the threshold levels that are acceptable, the patient and caregivers would be alerted of the imminent exacerbations for necessary early medical interventions, thereby improving the patient's quality of life and potentially saving hospitalization costs.

#### IV. CONCLUSION

In summary, a multi-modal technique has been developed for predicting the exacerbation risks for respiratory diseases such as COPD, based on a new artificial intelligence model for respiratory sound analysis and retrospective medical studies correlating key environmental parameters to exacerbations. The proposed solution includes a novel spatio-temporal machine learning model for accurate real-time classification and monitoring of respiratory conditions, tracking of local environmental and meteorological factors with commercially deployed sensors, and forecasting the patient's progression of conditions by combining the trends derived from these two modules.

The proposed new spatio-temporal artificial intelligence network architecture deeply meshes the salient structures of both convolutional and recurrent neural networks, and as a result outperforms both traditional CNN and RNN models, as well as the more recent CRNN models, in extracting the spatial and temporal features that are inherent in spectrograms of respiratory sounds. Extensive comparative tests have been performed to demonstrate that the new model achieves better sensitivity, specificity, accuracy, and Matthews Correlation Coefficient metrics than the traditional machine learning models.

A telehealth solution based on this work can assess the exacerbation risks and alert patients and doctors of early medical intervention, medication, and impending hospitalization. Thus, this technique can conveniently and cost-effectively help minimize and mitigate the impact of respiratory exacerbations, therefore improving patients' quality of life and potentially reducing hospitalization costs.

The future work will include collaboration with medical research institutions to further validate and deploy a remote patient monitoring solution into the real-world.

#### ACKNOWLEDGMENT

I would like to express sincere gratitude to my advisor, Chris Spenner, for providing insightful advices. I would also like to thank Kailas Vodrahalli, Archelle Georgiou, Sridhar Nemala, and Krishna Vastare for their inputs and guidance. Additionally, I would like to thank my parents for their continuous support.

#### REFERENCES

- [1] Forum of International Respiratory Societies (2017). The Global Impact of Respiratory Disease – Second Edition. Sheffield, European Respiratory Society.
- [2] G. Syamlal, A. Bhattacharya, and K.E. Dodd (2020). Morbidity and Mortality Weekly Report. 69 (26), 809–814. doi: 10.15585/mmwr.mm6926a1
- [3] N. Diab, A.S. Gershon, D.D. Sin, W.C. Tan, J. Bourbeau, L.P. Boulet, and S.D. Aaron (2018). American Journal of Respiratory and Critical Care Medicine. 198 (9), 1130-1139. doi: 10.1164/rccm.201804-0621CI
- [4] I. Tomasic, N. Tomasic, R. Trobec, M. Krpan, and T. Kelava (2018). Medical and Biological Engineering and Computing. 56 (4), 547–569. doi: 10.1007/s11517-018-1798-z
- [5] S.B. Bentsen, T. Rustøen, and C. Miaskowski (2012). International Journal of Chronic Obstructive Pulmonary Disease. 7, 137–143. doi: 10.2147/COPD.S28994
- [6] T. Ho, R.P. Cusack, N. Chaudhary, I. Satia, and O.P. Kurmi (2019). Breathe (Sheffield, England). 15 (1), 24–35. doi: 10.1183/20734735.0346-2018
- [7] J. de Miguel-Díez, J. Hernández-Vázquez, A. López-de-Andrés, A. Álvaro-Meca, V. Hernández-Barrera, and R. Jiménez-García (2019). PLOS One. 14 (5), e0217143. doi: 10.1371/journal.pone.0217143
- [8] J. Smith and A. Woodcock (2006). International Journal of Chronic Obstructive Pulmonary Disease. 1 (3), 305–314. doi: 10.2147/copd.2006.1.3.305
- [9] S.J. Barry, A.D. Dane, A.H. Morice, and A.D. Walmsley (2006). Cough. 2, 8. doi: 10.1186/1745-9974-2-8
- [10] J.M. Liu, M. You, Z. Wang, G.Z. Li, X. Xu, and Z. Qiu (2015). BMC Medical Informatics and Decision Making. 15 (Suppl 4), S2. doi: 10.1186/1472-6947-15-S4-S2
- [11] H.H. Wang, J.M. Liu, M.Y. You and G.Z. Li (2015). IEEE International Conference on Bioinformatics and Biomedicine (Washington, DC, USA). 442-445. doi: 10.1109/BIBM.2015.7359724.
- [12] J. Amoh and K. Odame (2016). IEEE Transactions on Biomedical Circuits and Systems. 10 (5), 1003-1011. doi: 10.1109/TB-CAS.2016.2598794
- [13] T. Elfaramawy, C. Latyr Fall, M. Morissette, F. Lelouche and B. Gosselin (2017). 15th IEEE International New Circuits and Systems Conference (Strasbourg, France). pp. 197-200, doi: 10.1109/NEW-CAS.2017.8010139.
- [14] T. Drugman, J. Urbain and T. Dutoit (2011). 19th European Signal Processing Conference (Barcelona, Spain). pp. 1289-1293.
- [15] M. Soliński, M. Lepek, and L. Koltowski (2019). Automatic cough detection based on airflow signals for portable spirometry system. arXiv:1903.03588.
- [16] E. Çakır, G. Parascandolo, T. Heittola, H. Huttunen and T. Virtanen (2017). IEEE/ACM Transactions on Audio, Speech, and Language Processing. 25, doi: 10.1109/TASLP.2017.2690575
- [17] J. Sang, S. Park and J. Lee (2018). 26th European Signal Processing Conference (Rome, Italy). pp. 2444-2448, doi: 10.23919/EUSIPCO.2018.8553247
- [18] S. Deshmukh, B. Raj and R. Singh (2020). Multi-Task Learning for Interpretable Weakly Labelled Sound Event Detection. arXiv:2008.07085.
- [19] Sorana (2020). Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2020/11/a-short-intuitive-explanation-of-convolutional-recurrent-neural-networks/>
- [20] K. Sanjeevan and T. Hung (2020). <https://github.com/ksanjeevan/crnn-audio-classification>
- [21] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner (1998). Proceedings of the IEEE. 86 (11), 2278-2324. doi: 10.1109/5.726791
- [22] Kaggle FSD (2018). <https://www.kaggle.com/c/freesound-audio-tagging>
- [23] E.J. Jo and W.J. Song (2019). Asia Pacific Allergy. 9 (2), e16. doi: 10.5415/apallergy.2019.9.e16
- [24] The United States Environmental Protection Agency. National Ambient Air Quality Standards for Particle Pollution. [https://www.epa.gov/sites/production/files/2016-04/documents/2012\\_aqi\\_factsheet.pdf](https://www.epa.gov/sites/production/files/2016-04/documents/2012_aqi_factsheet.pdf)
- [25] The United States Environmental Protection Agency. National Ambient Air Quality Standards. <https://www.epa.gov/criteria-air-pollutants/naqs-table>
- [26] PurpleAir. Real-time Air Quality Monitoring. <https://www2.purpleair.com/>
- [27] World Air Quality Index: Real-Time Air Pollution. <https://waqi.info/>