Automated Covid Detection
VoCOVID

User Manual

Team Members: Rodrigo Alarcon, Emma Conti, Lamine Deen, Audrey Eley

Advisor: Dr. Nematzadeh

Written By: Emma Conti

Project Goal and Motivation

The COVID-19 pandemic has underscored the need for effective and innovative diagnostic tools. A web-based application that analyzes cough audio to predict COVID-19 infection can offer a convenient and non-invasive screening method, potentially aiding early detection and reducing healthcare burdens. Although at-home testing for COVID-19 is effective, it can be expensive and inconvenient. Therefore, a convenient COVID-19 screening tool to assess the need for testing can help people make an informed decision.

Our Approach

The user can record their coughs and receive predictions on their COVID-19 infection status. This feature not only provides real-time feedback but also aids in maintaining a history of the user's infection status, making it a non-invasive and cost-effective tool for early screening. By tracking this data over time, the user can monitor their health status without the immediate need for a healthcare provider.

The web app design prioritizes ease of use, ensuring anyone can navigate it effortlessly and check their COVID-19 status at any time. Other aspects of the website will include details about the research and development of the ML model

The user can view a continuous progress chart that updates with every entry, making it easier to visualize changes in their infection status over time. This feature helps determine when a user is recovering and no longer symptomatic, or still infected. The user-friendly layout ensures that navigating the web app is straightforward.

Users can access their data at any time, providing continuous access to their COVID-19 status history. This eliminates the need to wait for a healthcare provider for early detection and offers users a convenient and effective way to monitor their health

What It Does

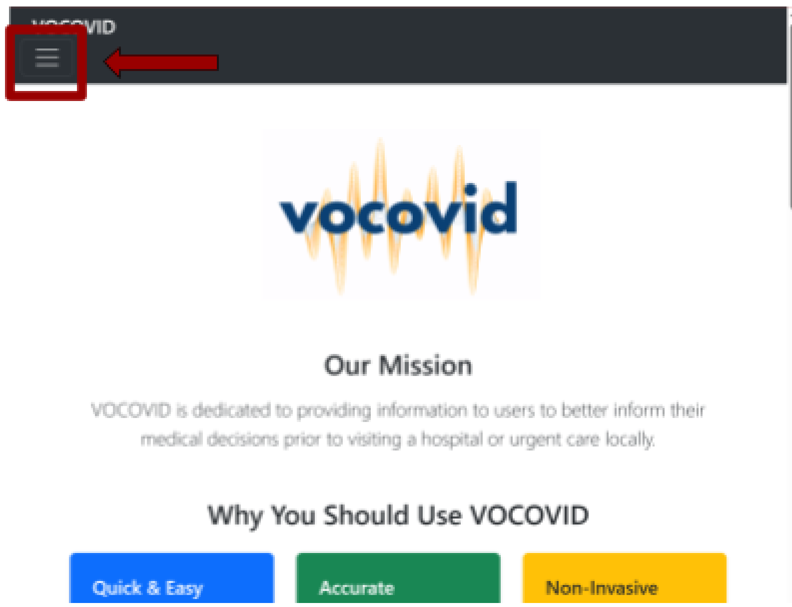
The web app not only predicts COVID-19 infections based on cough recordings but can also track the progression of the user's condition over time. By analyzing daily recordings, the user can observe trends in your symptoms, making it easier to determine when medical intervention is necessary. This continuous monitoring feature offers a more personalized health-tracking experience.

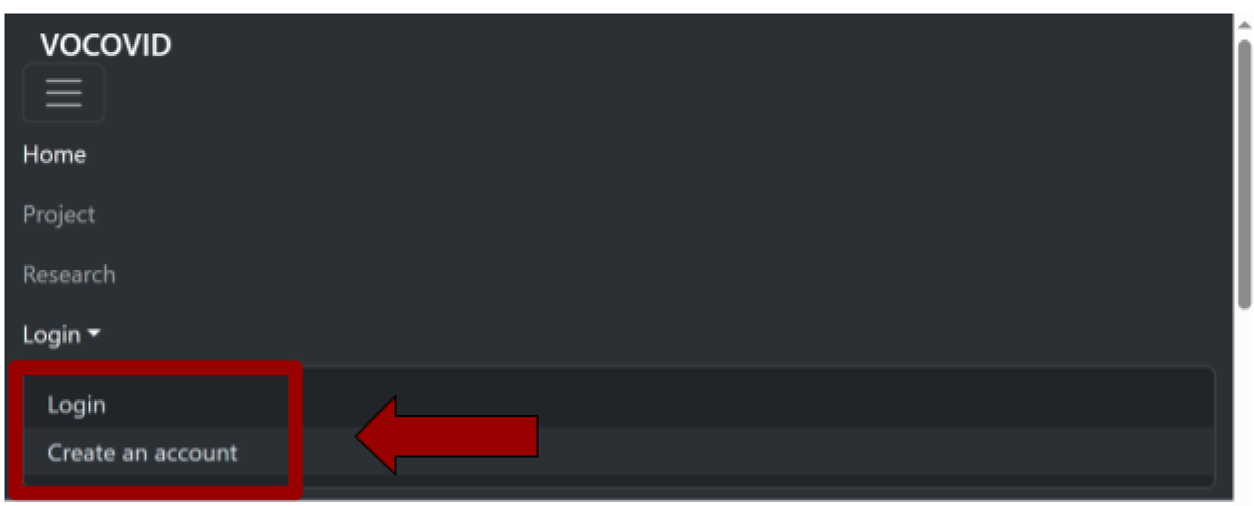
How to Use

Below you can find step by step instructions for navigating and using the VoCOVID Web Application!

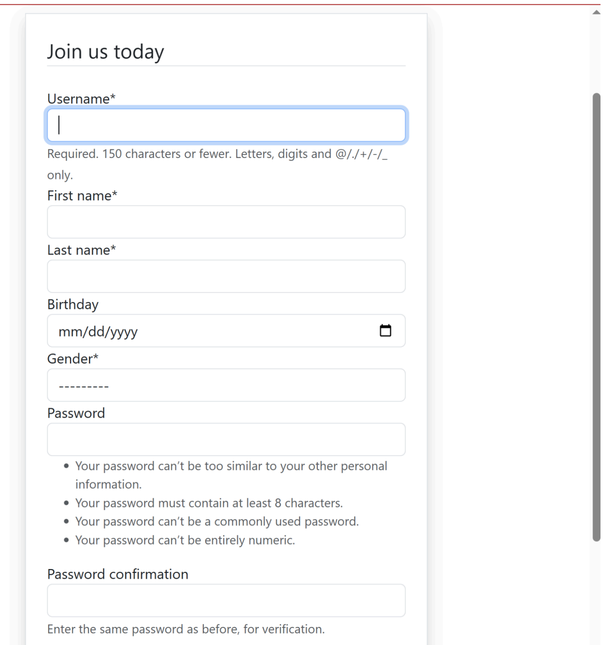
Creating an Account

Step 1: Navigate to the Login/Create an Account Page

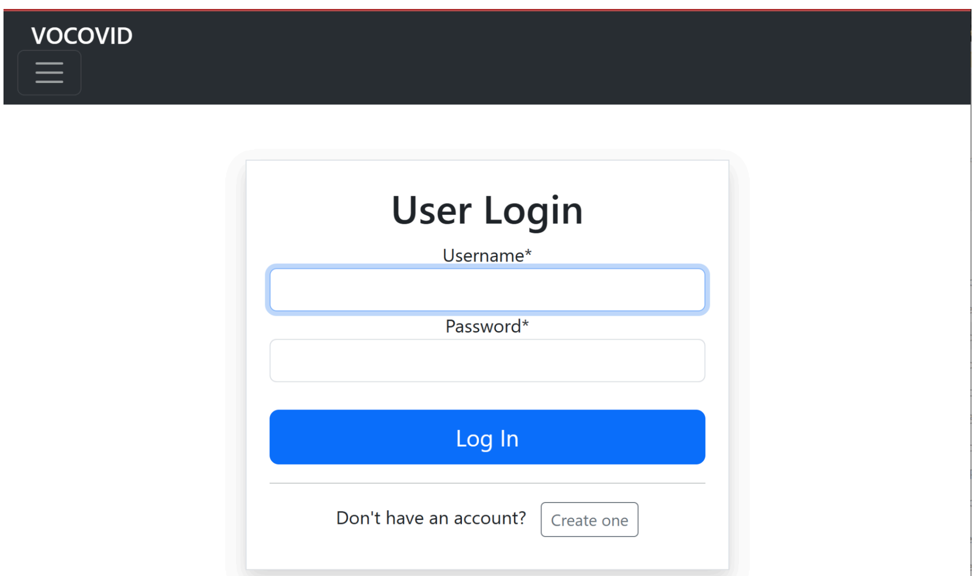




Step 2: Fill out all questions as per the instructions



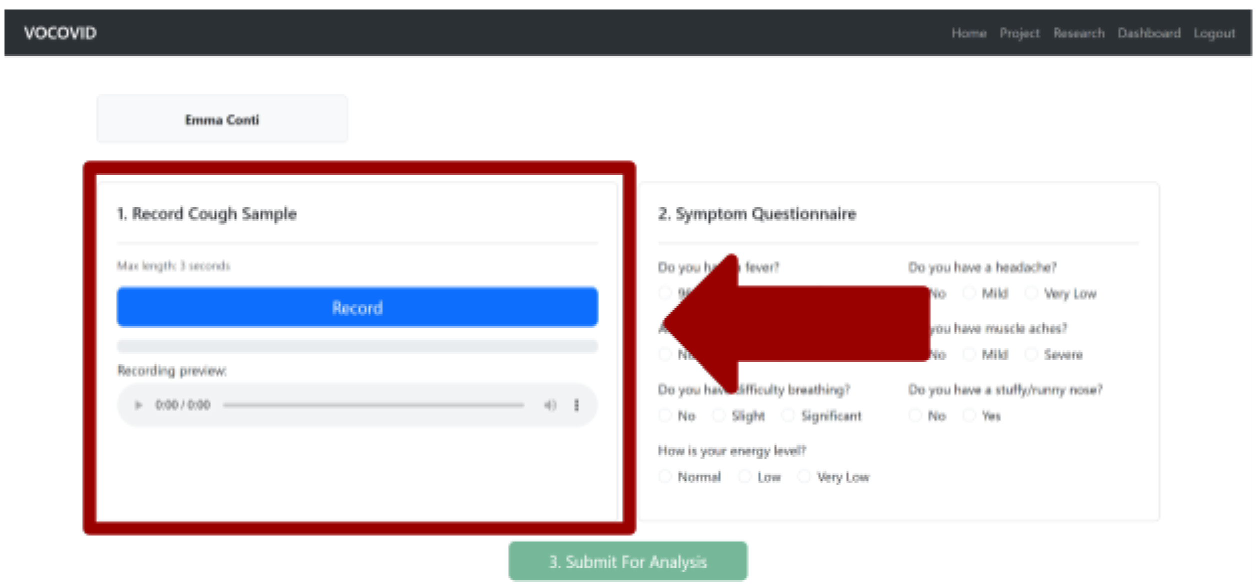
Step 3: Log In!



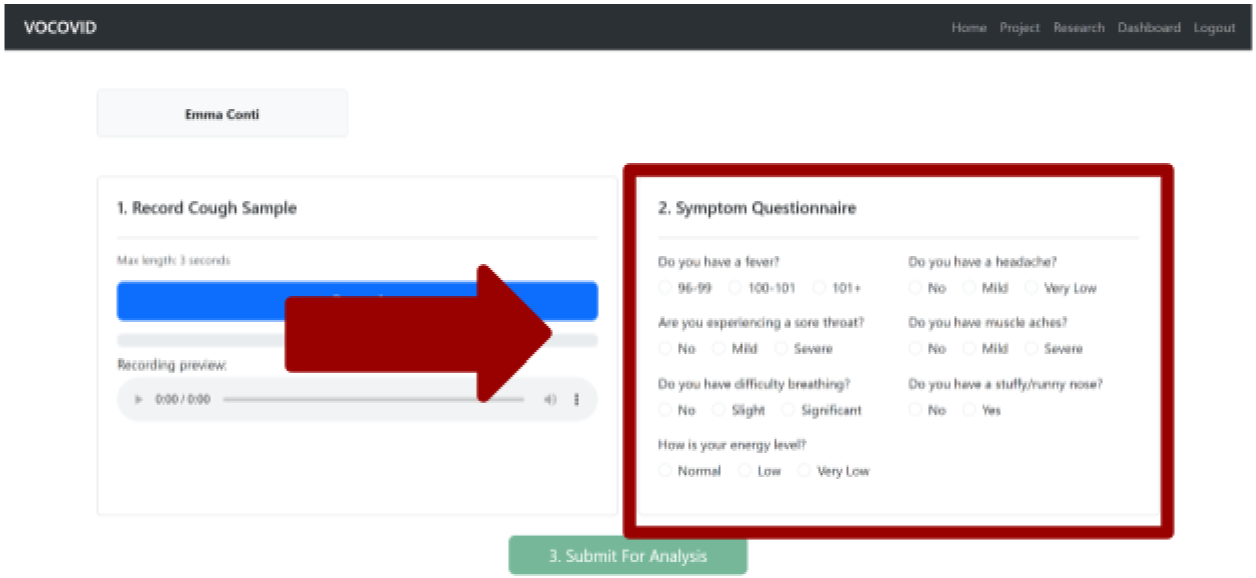
Uploading a Recording

Step 1: Navigate to the Dashboard

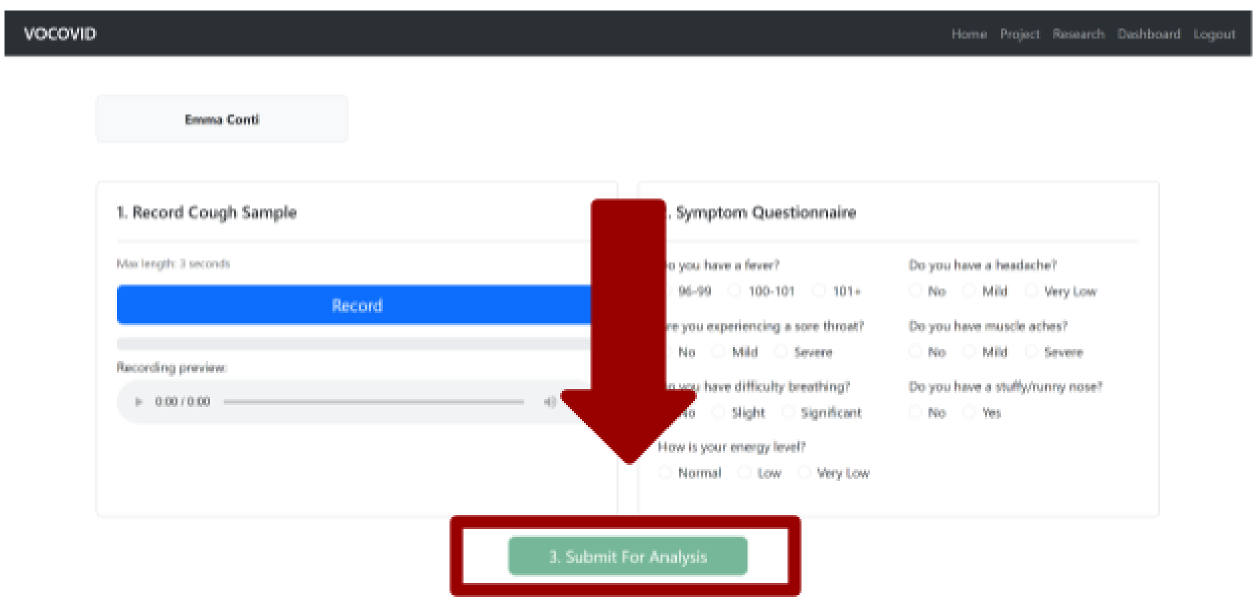
Step 2: Use the Record Button and record yourself coughing



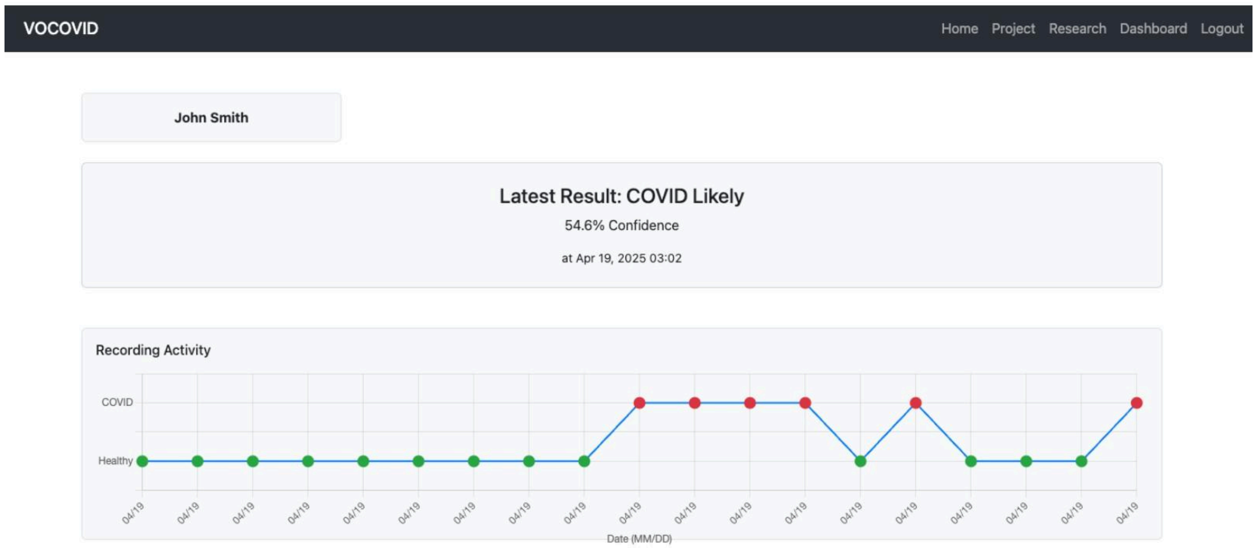
Step 3: Fill out the Symptoms Questionnaire



Step 4: Submit your Results for Analysis



Step 5: View Your Results

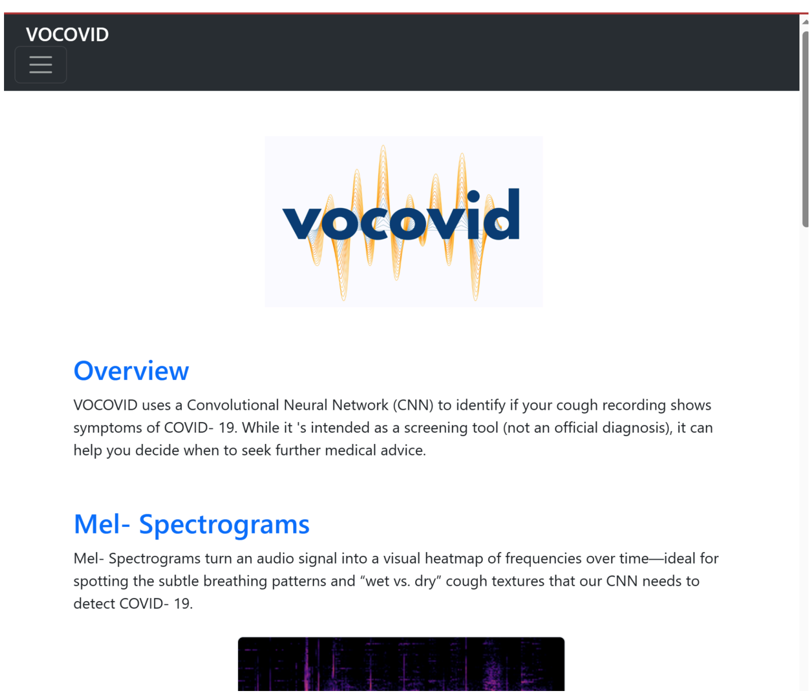


Web Application Development

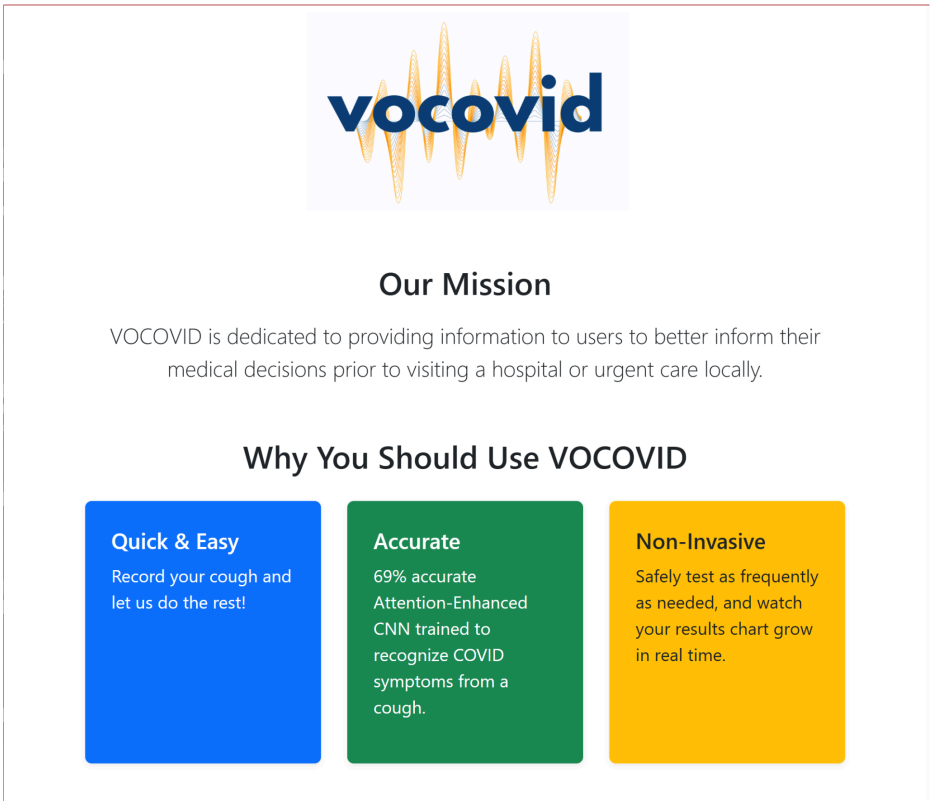
Implemented multiple models as precision of the Attention Enhanced CNN increased. Alongside the model implementation a few features were implemented as well. An initial questionnaire asking about symptoms, a time limit for how long a cough can be recorded for, and a graph to show users their results. Once a final model was implemented, it was sent to users for their review and to see how the user experience could be improved. This led to changes being made to the user dashboard included: additional information provided in graphs, further questions added to the questionnaire users answered when they recorded their coughs, and different page redirects after a user logs into their account or creates a new account to improve the ease of use.

Trained and saved the weights of Attention Enhanced CNN for web integration. Saved weights and model architecture to web application. Implemented data pipeline which retrieves an audio sample from the user model for classification, classifies audio, and displays the classification result and the prediction confidence. As in previous milestones, working to implement a classification pipeline was challenging, and many errors were encountered, but were eventually resolved. We have only had the opportunity to test the web application with non-COVID-19 users, but in these tests, the model was able to correctly classify the users as not having COVID-19.

Added Research and Home Pages to keep website as easy to navigate as possible, while still providing users on background information about how our Attention Enhanced CNN Model works.



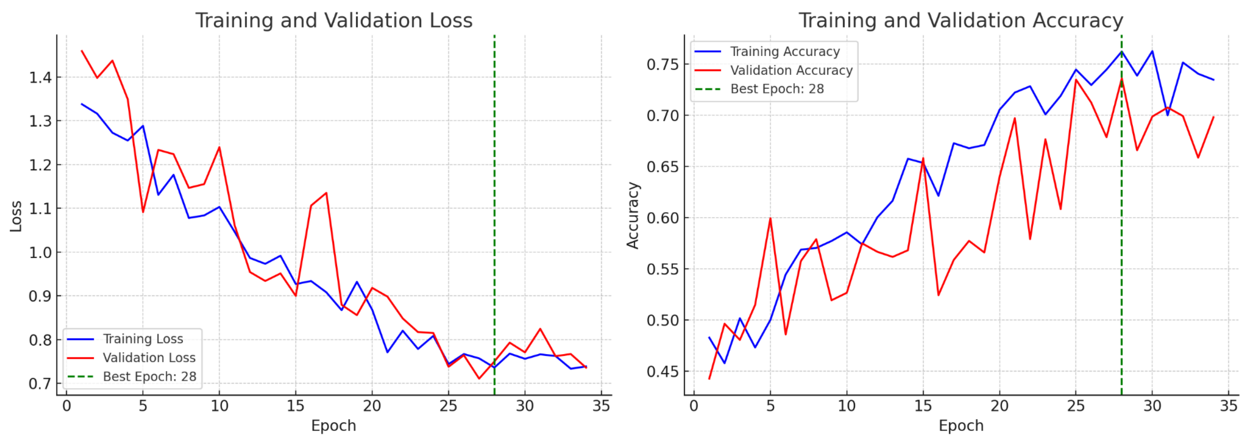
VoCOVID Research Page



VoCOVID Home Page

Machine Learning Model Development

The ResNet50 model was receiving similar precision to the CNN, so the team pivoted to focus entirely on the development of the CNN. It was switched from a model with 3 evaluation results to 2 evaluation results: 'healthy' and 'COVID'. The final model architecture for the developed Attention Enhanced CNN can be seen below. It was determined through various tests that 60 epoch training runs were receiving the best results, so the number of epochs was not increased further, as it did not improve the precision and accuracy. The 69% accuracy achieved will also help diminish the number of false positives and negatives received by a user during their use of the webapp. The model and its weights were passed off to then be implemented into the webapp.



	precision	recall	f1-score	support
healthy	0.73	0.81	0.77	368
COVID-19	0.64	0.53	0.58	234
accuracy			0.70	602
macro avg	0.68	0.67	0.67	602
weighted avg	0.69	0.70	0.69	602

Convolutional Neural Network

CNNs and RNNs are generally heralded as the best choices for a cough detection algorithm. Additionally, because mel Spectrograms are two dimensional, they are primarily paired with CNNs. A downside of CNNs is that they are a training model for a local space and can result in the loss of information based on how the information is prepared and filtered. This cannot be solved entirely, only mitigated.

CNNs are ultimately more useful over RNN's because of they are able to produce a feature map. This allows for further development on dataset augmentation to be based on which features are most important to be identified within the CNN. Mel Spectrograms can therefore be enhanced to more clearly show the features needed specifically to identify COVID-19 during testing.

Our Model

Based on the research done, a CNN was the obvious choice for VOCOVID. As we continue testing, we have found the need to augment our data. The current model features a dataset that is utilizing:

- K-fold cross validation
- Pre-processing on Mel-Spectrograms

Our current model has three convolutional blocks for processing the data during testing, and a final head layer, as can be seen in our model architecture below.

Model Overview Architecture:

- Block 1: Conv 3×3 (1⟶32) → BN → ReLU → MaxPool 2×2 → Channel-Attention 32 → Spatial-Attention
- Block 2: Conv 3×3 (32⟶64) → BN → ReLU → MaxPool 2×2 → Channel-Attention 64 → Spatial-Attention
- Block 3: Conv 3×3 (64⟶128) → BN → ReLU → MaxPool 2×2 → Channel-Attention 128 → Spatial-Attention
- Head: AdaptiveAvgPool 1×1 → Flatten → FC 128⟶64 → ReLU → Dropout 0.5 → FC 64⟶2 (logits)

Current Model Architecture

ResNet50

ResNet50 was used as a benchmark model for testing our own CNN. As a CNN, ResNet50 is already uniquely equipped for audio processing and testing specifically, and yields fantastic results. A variety of cough based mel-spectrogram models using ResNet50 as a basis have published their findings, making it a strong choice for our testing.

One of the benefits of ResNet50 over other CNNs is the accuracy without needing any kind of preprocessing for imbalance data handling on an audio benchmark set. This

allows for data sets to be used without any kind of data augmentation, and still yields usable results.
Results from testing with our dataset using ResNet50 will be included soon.

Our ResNet50 model used a different pre-trained ResNet50 model and at the last layer implemented our data in order to improve the accuracy.

Data Preprocessing

Mel-Spectrograms

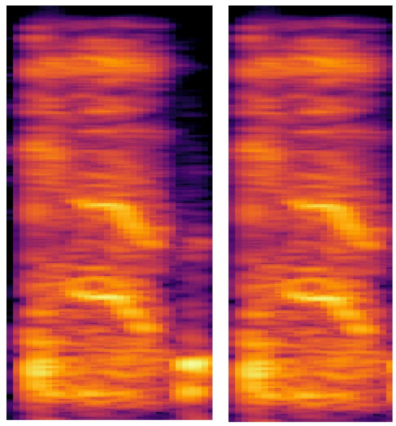
Mel Spectrograms are being used because the way to determine whether or not a cough is COVID-19 is based on the way the user is breathing. What a Mel Spectrogram does is makes a visual representation of the types of noise happening in an audio file. This will allow for the specific types of noise, in our case restricted breathing prior to the cough, to become more identifiable. Differences between “wet” and “dry” coughs can also be determined using Mel Spectrograms.

Within a Mel Spectrogram the x-axis is time and the y-axis is frequency. They are specifically designed in order to better represent what humans hear in a visual format. The vibrancy of the color represents the amplitude at a given time. The main benefit of using Mel Spectrograms is the ability to compress the frequency scale to help reduce the data’s dimensionality, which assists in bringing down the computational load on a learning model. Capturing important pitches that can then be analyzed further by a CNN.
Our Mel Spectrograms have been shortened to one cough in length to ensure that no white noise interferes with your results. Below you can see an example of what this might look like, along with the other Mel Spectrogram Hyperparameters used for our Attention Enhanced CNN.

Data included a variety of audio records that were inaccurately listed as cough recordings. In order to attain results with any level of accuracy, the dataset needed to be completely cleaned before testing could re-commence. Once all audio files were listened to, it was necessary to take this information and pass it through a script to remove all audio files, mel spectrograms, and stored metadata that are no longer being used so the new final dataset can be used for testing.

After removing all audio files that had been labeled but were not cough recordings they are converted into mel-spectrograms with all white space on either end being removed. That way all recordings were focused entirely on the coughs within the recording. Unfortunately, the accuracy of our developed CNN did not improve at all. The next step for the CNN is to build a model with more layers to see if that improves accuracy, but also to further refine the data in pre-processing. All mel-spectrograms have been shortened to one cough in length to further increase the accuracy of the CNN. This was then tested on both the ResNet50 model as well as our CNN.

Example of Past Data vs. Current Data



Audio 23 Original Image, 1st Attempt at Cutting

Data Preprocessing Steps

1. Initial Data Audit
2. Filtering & Cleaning
3. Waveform to Mel-Spectrogram
4. Cough Segment Extraction
5. Metadata Sync & Class Selection
6. Oversampling Minority Class
7. Augmentation Integration
8. Final Balancing & Shuffle
9. Dataset Splitting
10. PyTorch DataLoader Prep

References

Wolf-Monheim, Friedrich. “Spectral and Rhythm Features for Audio Classification with Deep
Convolutional Neural Networks.” *Spectral and Rhythm Features for Audio Classification with Deep Convolutional Neural Networks*, 2024, arxiv.org/html/2410.06927v1#S4.

Kexin Luo, Guanci Yang, Yang Li, Shangen Lan, Yang Wang, Ling He, Binqi Hu,
Croup and pertussis cough sound classification algorithm based on channel attention and multiscale Mel-spectrogram, Biomedical Signal Processing and Control, Volume 91, 2024, 106073, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2024.106073.
(https://www.sciencedirect.com/science/article/pii/S1746809424001319)

Ghrabli, Syrine, Mohamed Elgendi, and Carlo Menon. 2022. "Challenges and Opportunities of Deep Learning for Cough-Based COVID-19 Diagnosis: A Scoping Review" *Diagnostics* 12, no. 9: 2142. https://doi.org/10.3390/diagnostics12092142

Awan, Mazhar & Rahim, Mohd & Salim, Naomie & Mohammed, Mazin & Zapirain, Begoña & Abdulkareem, Karrar. (2021). Efficient Detection of Knee Anterior Cruciate Ligament from Magnetic Resonance Imaging Using Deep Learning Approach. Diagnostics. 11. 105. 10.3390/diagnostics11010105.

Troy, Mazerolle. “Using VGG16 to Classify Spectrograms.” *Kaggle*, Kaggle, 22 Dec. 2023, www.kaggle.com/code/troymazerolle/using-vgg16-to-classify-spectrograms.

Yadav, Jyoti & Varde, Aparna & Liu, Hao & Antoniou, George & Xie, Lei. (2024). Audiovisual Multimodal Cough Data Analysis for Tuberculosis Detection.

Vincenzo Dentamaro, Paolo Giglio, Donato Impedovo, Luigi Moretti, Giuseppe Pirlo, AUCO ResNet: an end-to-end network for Covid-19 pre-screening from cough and breath, Pattern Recognition, Volume 127, 2022, 108656, ISSN 0031-3203, https://doi.org/10.1016/j.patcog.2022.108656.
(https://www.sciencedirect.com/science/article/pii/S0031320322001376)