DL GROUP PROJECT - Covid-19 detection based on breathing sounds

Vitaliia Marina
vitaliia.marina@tu-ilmenau.de

Ye Yun Khor
ye-yun.khor@tu-ilmenau.de

Edwin G. Carreño
edwin-carreno.lozano@tu-ilmenau.de

Ralina Safina
ralina.safina@tu-ilmenau.de

Project Mentor:
M. Sc. Dominik Walther

December 5, 2021

ABSTRACT

COVID-19 has caused a pandemic of respiratory illness that has negatively affected the economic and social stability of nations worldwide. On the other hand, the impossibility of obtaining updated information on the number of infections and their presence aggravates the situation for effective decision making by governments and health organizations. For those reasons, this work aims to propose a non-invasive strategy that allows a rapid identification of COVID-19 through Convolutional Neural Networks and the use of spectrograms based on audio samples taken from coughing patients (with and without COVID-19).

1 Introduction

COVID-19 causes distinct pathomorphological alterations in the respiratory system, thus, we are interested to find out whether COVID-19 can be detected based on breathing sounds using our own neural network. The input to our algorithm are the spectrogram images of breathing sounds from kaggle. We plan to use a convolutional neural network to output the classification of the breathing sound, i.e., whether it belongs to a COVID-19 or a non COVID-19 patient.

2 Related Work

Deep learning methods have been used successfully to solve many complex medical problems in the past, so it is not surprising that many researchers have turned to them again in the face of a new challenge. In [1], a group of scientists tackles a problem of detecting COVID-19 from a patient's breathing and coughing patterns through usage of an LSTM model which consists of 3 LSTM layers. It provided a validation accuracy of 80%. In order to complete their task, they used a dataset from Pfizer Digital Medicine Challenge which consists of raw audio files, and then utilized Mel-frequency cepstral coefficients (MFCCs) to extract relevant features.

The modification of VGG-13 architecture is suggested by the authors in [2]. To prevent overfitting of the model and balance the data two datasets were used: DiCOVA 2021 and augmented positive samples of COUGHVID. By combining cross entropy and focal loss while using an ensemble of two VGG-13 models the validation AUROCs (82.23%) and blind test AUROCs (78.3%) were achieved.

Another interesting approach to tackle the problem of identifying COVID-19 in data samples is related to translating the power-time domain (audio) to frequency-time-power domain (images), in other words using spectrograms as input data for a CNN architecture. In [3] has been tested the usage of spectrograms in the Xception CNN architecture, achieving a precision of 0.75%-0.80% with a fine-tuning approach as the initial step. However, authors argue that the quality of results is highly related to the quality of the dataset used, high quality images in spectrograms could lead to better results in classification. Other than that, the background noise while recording the breathing sound and the microphone's sensitivity also affect the model [4].

3 Dataset and Features

The dataset in [5] contains breathing sounds spectrograms, which are divided into groups according to the origin (real case and generated spectrograms) and the class of breathing sound (healthy and covid-19 case). The total number of samples is 6684, which will be distributed among training, validation and test sets. The main focus will stay on real cases of Covid-19 and non Covid-19 (4424 samples in total).

As spectrograms presented in the dataset have different resolutions, they will be scaled before feeding to the neural network in order to train the parameters properly. Moreover, the data of the training set might be augmented using frequency or time masks in case the overfitting of the model will occur during training.

4 Methods

In order to classify images, we have chosen the convolutional neural network architecture for its visual information processing efficiency provided by local parameter sharing. We are planning on implementing our own neural network architecture without reusing pre-existing models. The precise architecture will be chosen through experimentation.

5 Design of Experiment

A bias/variance analysis for different models such as Xception[6] and custom architectures mentioned in the section "Methods" will be applied. Depending on if the Bias or Variance is high [7][8], different approaches such as training longer, testing different optimization algorithms, activation functions and applying regularization techniques will be applied in order to obtain a better generalization . Additionally, different metrics will be utilized to report results in classification problems, those metrics are: confusing matrix, accuracy, cost-sensitivity accuracy, precision and recall and finally ROC (Receiver Operating Characteristic) [9].

6 Contributions

This section should describe what each team member will work on and contribute to the project.

Vitaliia Marina - Methods section, data flow and CNN model implementation

Edwin Carreño - Hypertuning, results and reporting

Ralina Safina - Data preprocessing, results and reporting

Ye Yun Khor - Reporting and video recording

References

[1] Lazhar Khriji, Ahmed Ammari, Seifeddine Messaoud, Soulef Bouaafia, Amna Maraoui and Mohsen Machhout. *COVID-19 Recognition Based on Patient's Coughing and Breathing Patterns Analysis: Deep Learning Approach*. IEEE, 2021. https://ieeexplore.ieee.org/document/9435454

[2] Sunil Rao, Vivek Narayanaswamy, Michael Esposito, Jayaraman Thiagarajan, Andreas Spanias. *Deep Learning with hyper-parameter tuning for COVID-19 Cough Detection*. IEEE, 2021.

[3] C. R. Rodriguez, D. Angeles, R. Chafloque, F. Kaseng and B. Pandey, "Deep Learning Audio Spectrograms Processing to the Early COVID-19 Detection," *2020 12th International Conference on Computational Intelligence and Communication Networks (CICN)*, 2020, pp. 429-434, doi: 10.1109/CICN49253.2020.9242583.

[4] Unais Sait, Gokul Lal K.V., Sanjana Shivakumar, Tarun Kumar, Rahul Bhaumik, Sunny Prajapati, Kriti Bhalla, Anaghaa Chakrapani, "A deep-learning based multimodal system for Covid-19 diagnosis using breathing sounds and chest X-ray images", Applied Soft Computing, 2021, doi: 10.1016/j.asoc.2021.107522.

[5] Sait, Unais; Kv, Gokul; Shivakumar, Sanjana ; Kumar, Tarun; Bhaumik, Rahul ; Prakash Prajapati, Sunny; Bhalla, Kriti (2021), "Spectrogram Images of Breathing Sounds for COVID-19 and other Pulmonary Abnormalities", Mendeley Data, V1, doi: 10.17632/pr7bgzxpgv.1

[6] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1800-1807, doi: 10.1109/CVPR.2017.195.

[7] Deep Learning.AI , "Deep Learning Specialization Course 2: Lecture Notes," DeepLearning.AI, 2021. [Online]. Available: https://community.deeplearning.ai/t/dls-course-2-lecture-notes/11866. [Accessed: Dec. 06, 2021].

[8] Deep Learning.AI , "Deep Learning Specialization Course 3: Lecture Notes," DeepLearning.AI, 2021. [Online]. Available: https://community.deeplearning.ai/t/dls-course-3-lecture-notes/11868. [Accessed: Dec. 06, 2021].

[9] A. Burkov, The Hundred-Page Machine Learning Book. 2019, p. 13-19.