IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

PULMONARY SOUND RECOGNITION USING DEEP LEARNING

[1]T.Veni Priya, [2]S.Jayabharathi, [3]S.Lavanya, [4]S.Mohana Vasanthi

[1]Assistant Professor, [2]Student, [3]Student, [4]Student
[1]Department of Computer Science and Engineering ,
[1]Arasu Engineering College, Kumbakonam, India

Abstract: Pulmonary breathing sounds act as the important indicators of respiratory tract health and respiratory tract disorders. The breathing sound emitted by a person is directly related to air movement, changes within lung tissue and the position of secretions within the lung. For instance a wheezing sound, is a common sign that a patient has an obstructive airway disease like asthma or Chronic Obstructive Pulmonary Disease (COPD). Pulmonary breathing sounds can be recorded using digital stethoscopes and other recording techniques. This digital data provides us with the possibility of using the deep learning technique CNN to automatically diagnose respiratory disorders like asthma, pneumonia and bronchiolitis, etc.

***Index Terms* – Lung Sound, CNN, VGG19, Mel Spectrogram.**

I. INTRODUCTION

The lung sound of the respiratory system is an inexpensive, noninvasive, safe, easy-to-perform, and one of the oldest diagnostic techniques used by the physicians to diagnose various pulmonary diseases. The stethoscope is a well-known and widely available diagnostic instrument. Health care personnel routinely use it to listen for abnormal sounds in the lungs to establish a diagnosis. Even though lung auscultation is a very old technique, recent technological advances in the 2 fields of hardware, acoustics, and digital sound analysis and classification gives new possibilities that should be further explored. This is reflected by the many new commercial solutions for recording and viewing sounds from a stethoscope, such as the Bluetooth solution for smart phones from Eko Devices (https://ekodevices.com/), and the MIT mobile stethoscope [5]. However, these solutions do not have automated approaches for detecting abnormal sounds in lung sound that can easily be integrated with smart devices and phones. Crackles are short, explosive nonmusical sounds heard mostly during inspiration [4]. These sounds are present in lung and heart related diseases like chronic obstructive pulmonary disease (COPD), pneumonia, heart failure, and asbestosis. In these diseases, the presence of crackles helps to stablish a diagnosis [10]–[13]. These diseases represent a major public health problem. In 2012, over 3 million deaths were caused by COPD which represented a 6% of the total deaths in that year [15]. Recent reports [15] state that 23 million people worldwide have a diagnosis of heart failure.Breath sound has three characters; frequency, intensity, and timbre or quality; which helps us to differentiate two similar sounds. Diagnosis or classification requires recognizing patterns. But most of the time, it is very hard to spot these patterns, especially if the data is very large. Data collected from the environment is usually non-linear, so we cannot use traditional methods to find patterns or create mathematical models. In the past decade, various technologies, such as expert systems, have been used to attempt to solve this problem. However, for critical systems, the error rate for the decision was too high [1].The above described problem can be solved using machine learning. In the past decade various successful algorithms were developed for machine learning and now with the help of deep learning algorithms, error rate has been reduce drastically. Particularly in computer vision and speech recognition, deep learning plays a vital role in reaching high level of accuracy.Research in this area attempts to make better representations and create models to learn these representations from large-scale unlabeled data [3]. Some of the representations are inspired by advances in neuroscience and are loosely based on interpretation of information processing and communication patterns in a nervous system, such as neural coding which attempts to define a relationship between the stimulus and the neuronal responses and the relationship among the electrical activities of the neurons in the brain [3, 16].Deep learning is a branch of machine learning based on a set of algorithms that attempt to model high-level abstractions in data by using model architectures, with complex structures, composed of multiple non-linear transformations [3, 3]. An observation (e.g., an image) can be represented in many ways including a vector of intensity values per pixel, or in a more abstract way as a set of edges, regions of a particular shape, and various other features. Some representations make it easier to learn tasks (e.g., face recognition or facial expression recognition) from examples [14, 3, 16]. One of the promises of deep learning is replacing handcrafted features with efficient algorithms for unsupervised or semi-supervised feature learning and hierarchical feature extraction [7].Various deep learning architectures such as deep neural networks, convolutional deep neural networks, deep belief networks, and recurrent neural networks have been applied to fields like computer vision, automatic speech recognition, natural language processing, audio recognition, and bioinformatics where they have been shown to produce state-of-the-art results on various tasks [1,9].The convolutional network architecture is a remarkably versatile yet conceptually simple paradigm that can be applied to a wide spectrum of perceptual tasks. Convolutional networks are trainable, multistage architectures. The input and output of each stage are sets of arrays called feature maps [10]. Convolutional neural networks (CNNs) are designed to process data that come in the form of multiple arrays. There are four key ideas behind CNN that take advantage of the properties of natural signals: local connections, shared weights, pooling, and the use of many layers. The architecture of a typical CNN is structured as a series of stages. The first few stages are composed of two types of layers: convolutional layers and pooling layers. Units in a convolutional layer are organized in feature maps, within which each unit is connected to local patches in the feature maps of the previous layer through a set of weights called a filter bank. Although the role of the convolutional layer is to detect local conjunctions of features from the previous layer, the role of the pooling layer is to merge semantically similar features into one [10]. The CNN has been found highly effective and has been commonly used in computer vision and image recognition. More recently, with appropriate changes from designing CNN for image analysis to taking into account speech-specific properties, the CNN is also found effective for speech recognition [9].In this paper, we aim to improve on this invention by analyzing audio data with deep learning algorithms and by classifying respiratory sounds. The dataset consists of audio recordings of lung sounds that were recorded by chest physicians. We believe, using deep learning, audio data can be analyzed for patterns that will lead to the detection of various pathological lung sounds and help in the diagnosis of pulmonary diseases.

II. METHODS

Lung sound is recorded using a electronic stethoscope. Recorded sound is used to provide diagnosis of various lung related diseases. Deep learning technique is used to classify various lung related diseases based on the Mel Spectogram of lung sound.

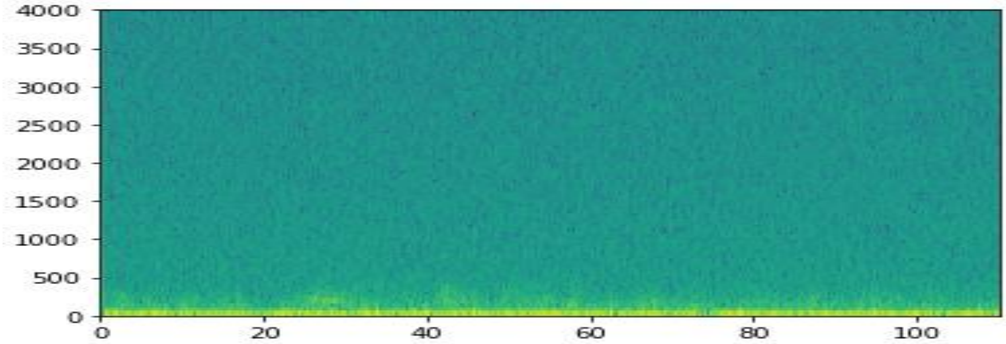

Figure 2.1 Spectrogram For Lung Sounds

III. PROPOSED SYSTEM

Lung breathing sounds act as the important indicators of breathing path health and breathing path disorders. The breathing sound emitted by a person is directly related to air movement, changes within lung tissue and the position of secretions within the lung. In a wheezing sound, is a common sign that a patient has an obstructive airway disease like asthma or chronic obstructive lung disease (COPD).Lung breathing sounds can be recorded using digital stethoscopes and other recording techniques. This digital data provides us with the possibility of using the deep learning technique CNN along with VGG19 to automatically diagnose respiratory disorders like asthma etc.

IV. IMPLEMENTATION

4.1 Pre-Processing

Pre-processing refers to the transformations applied to our data before feeding it to the algorithm. Data Preprocessing is a technique that is used to convert the raw data into a clean data set. In other words, whenever the data is gathered from different sources it is collected in raw format which is not feasible for the analysis.

4.2 Feature Selection

Exploratory Data Analysis or (EDA) is understanding the data sets by summarizing their main characteristics often plotting them visually. This step is very important especially when we arrive at modeling the data in order to apply Deep learning. Plotting in EDA consists of Histograms, Box plot, Scatter plot and many more. It often takes much time to explore the data. A Boxplot I is a standardized way of displaying the distribution of databased on five number summary("minimum",firstquartile(q1),median,thirdquartile(q3),and "maximum") it can also tell you it your data is symmetrical ,how tightly the data is grouped and if ,and how your data is skewed. Padding is a term relevant to convolutional neural networks as it refers to the amount of pixels added to an image when it is being processed by the kernel of a CNN. For example, if the padding in a CNN is set to zero, then every pixel value that is added will be of values zero. In order to slice the audio files, there are things need to consider as well. We need to make sure that they will have the same length (this is in preparation for feeding them into the model for training). If they're not of same length, then we have to pad the audio with silent (or zeroes) sounds. For the length, the optimal length of time that should be used must be known.

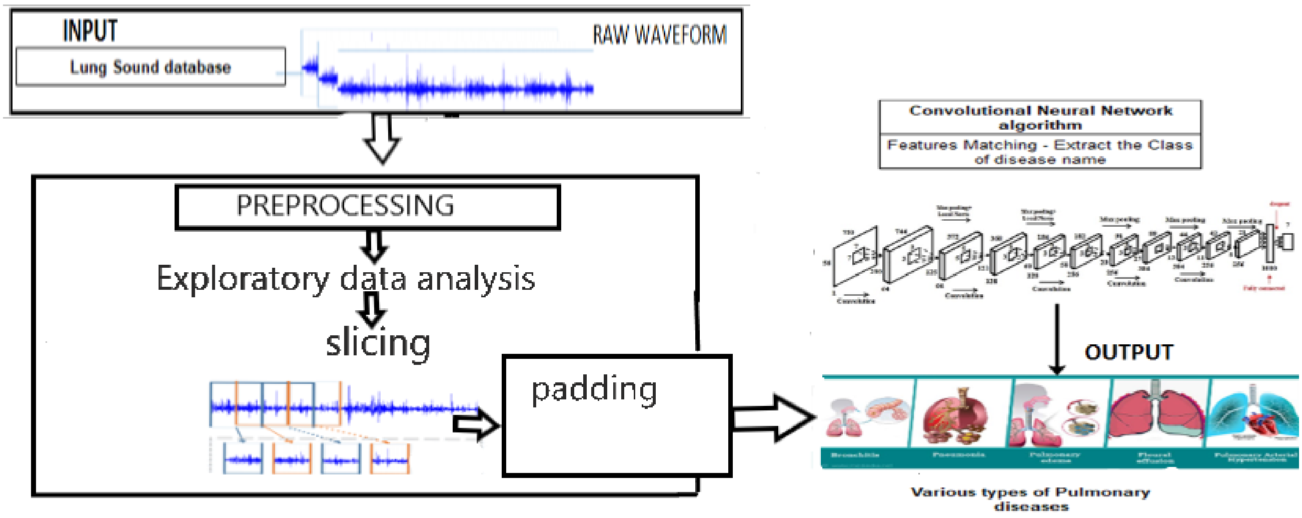

Figure 4.1 System Architecture

4.3 Classification

VGG-19 is a convolutional neural network that is trained on more than a million images from the sound database [1]. The network is 19 layers deep and can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals. A convolutional neural network consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of a series of convolutional layers that convolve with a multiplication or other dot product. The activation function is commonly a RELU layer and is subsequently followed by additional convolutions such as pooling layers, fully connected layers and normalization layers, referred to as hidden layers because their inputs and outputs are masked by the activation function and final convolutional. Convolutional neural network (CNN) is an interconnected assembly of simple parallel processing elements, unit or nodes, whose functionality is loosely based on lung co-efficient values.

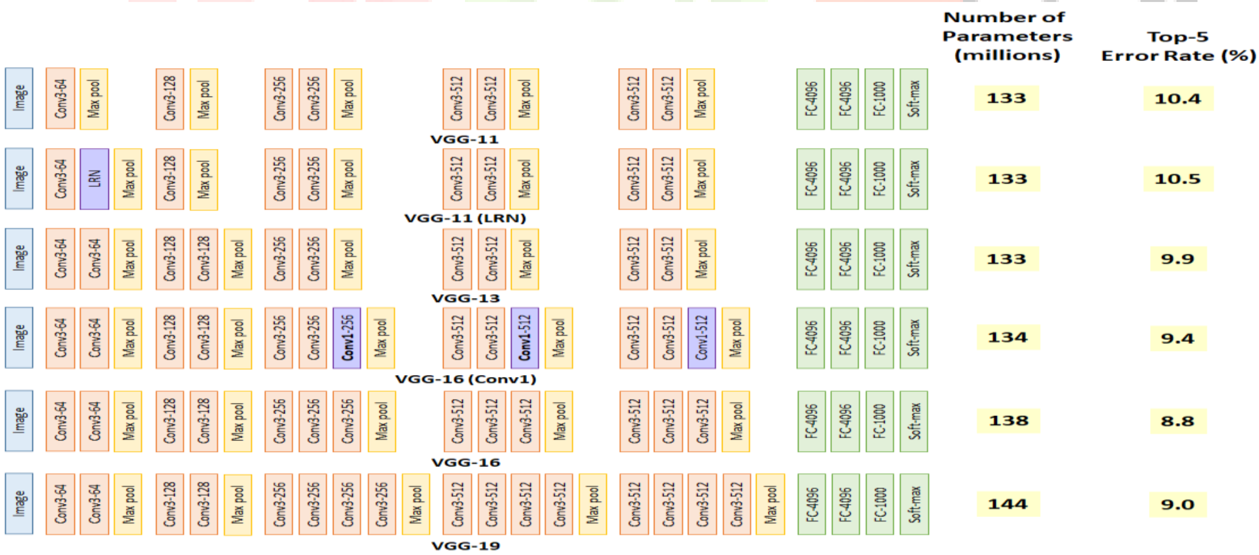

Figure 4.2 VGG19 Architecture

4.4 Data Set

The Respiratory Sound Database used in this paper is created by two research teams in Portugal and Greece. It includes 920 annotated recordings of varying length - 10s to 90s. These recordings were taken from 126patients. There are a total of 5.5 hours of recordings containing 6898 respiratory cycles - 1864 contain crackles, 886 contain wheezes and 506 contain both crackles and wheezes. The data includes both clean respiratory sounds as well as noisy recordings that simulate real life conditions. The patients span all age groups - children, adults and the elderly. This Kaggle dataset includes:920.wav sound files,920 annotation .txt files. A text file listing the diagnosis for each patient

V. RESULTS

We see that slowly but surely our model is learning our data. Validation accuracy is increasing steadily while mean absolute error is decreasing. However, we should be aware that loss is increasing. In this context, but maybe we can just ditch loss since we are already looking at me.

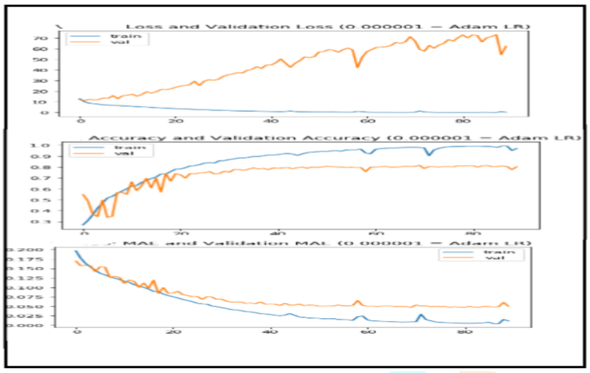

Figure 5.1 VGG19 Loss and Validation, Accuracy and Validation, MAE Accuracy and MAE Validation

VI. Conclusion

The software system is to train VGG19 for the automated analysis and diagnosis. The complete system can also be used for all types of lung sounds. In this study, we experimented using CNN algorithms in audio classification. Since VGG19 features combined with is a generally accepted practice for audio classification, we used it as a benchmark for our CNN algorithm. As a result, we found out that spectrogram image classification with CNN algorithm works as well. CNN can accurately classify and pre-diagnose respiratory audio. This system can be combined with a telemedicine system to store and share information among physicians. We believe our method can improve the results of previous studies and help in medical research.

References

[1]. AB Olshausen, Emergence of simple-cell receptive field properties by learningasparsecodefornaturalimages.Nature381(6583),607–609(1996)

[2].BDCN Prasadl, PESNK Prasad, Y Sagar, An approach to develop expert systems in medical diagnosis using machine learning algorithms (asthma) andaperformancestudy.IJSC2(1),26–

[3].Y Bengio, A Courville, P Vincent, Representation learning: a review and new perspectives. IEEE Trans. Pattern Anal. Mach. Intell. 35, 1–31(2013)

[4].Y Bengio, Learning Deep Architectures for AI. (2009), http://www.iro. umontreal.ca/~bengioy/papers/ftml.pdf.Accessed26Jan2016

[5]. A. Bohadana, G. Izbicki, and S. S. Kraman, "Fundamentals of Lung asultation," *N. Engl. J. Med.*, vol. 370, no. 8, pp. 744–751, Feb. 2014

[6].D.Chamberlain,J.Mofor,R.Fletcher,and R.Kodgule"Mobile stethoscope and signal processing algorithms for pulmonary screening and diagnostics," in *2015 IEEE Global Humanitarian Technology Conference (GHTC)*, 2015, pp.385–392.

[7]. RCollobert ,Deep Learning for Efficient Discriminative Parsing .(2011),http:// www.video.lectures.net. Accessed 26 Jan2016

[8].L Deng, D Yu, Deep Learning: Methods and Applications. (2014), http://research.microsoft.com/pubs/209355/DeepLearning-NowPublishing-Vol7-SIG-039.pdf. Accessed 26 Jan2016

[9] .L Gome, Machine-Learning Maestro Michael Jordan on the Delusions of Big Data and Other Huge Engineering Efforts. (2014), http://spectrum.ieee.org/ robotics/artificialintelligence/machinelearning-maestro-michael-jordan-on- the-delusions-of-big-data-and-other-huge-engineering-efforts.

[10].Y LeCun, K Kavukcuoglu, C Farabet, Convolutional networks and applications in vision. Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on, IEEE, pp. 253–256,2010

[11].Y LeCun, Y Bengio, G Hinton, Deep learning. Nature 521(7553), 436–444 (2015)

[12] A.F.Members et al., "ESC Guidelines for the diagnosis 2012," Eur. Heart J., vol. 33, no. 14, pp. 1787-1847, Jul. 2012

[13] Metlay JP, Kapoor WN, and Fine MJ, "Does this patient have community-acquired pneumonia: Diagnosing pneumonia by history and physical examination," *JAMA*, vol.

[14]. KNasrollahi, T Telve, S Escalera, J Gonzalez, TB Moeslund, P Rasti and G Anbarjafari, Spatio-temporal pain recognition in CNN-based super-resolved facial images. In video analytics. Face and facial expression recognition and audience measurement. Third International Workshop, VAAM 2016, and Second International Workshop, FFER 2016, Cancun, Mexico, Revised SelectedPapers,Springer,Vol.10165,pp.151,December4,2016

[15].V. L. Roger, "Epidemiology of Heart Failure," *Circ. Res.*, vol. 113, no. 6, pp. 46–659, Aug. 2013.

[16]. J Schmidhuber, Deep learning in neural networks: an overview. Neural Net w. 61, 85−117(2015)

s[17]. “WHO | Chronic obstructive pulmonary disease (COPD),” WHO. [Online].Available:http://www.who.int/mediacentre/factsheets/fs315/en/.[Accessed: 05-Aug-2016].