



e-ISSN: 2147-8228

www.dergipark.org.tr/ijamec

Volume 08  
Issue 04

December, 2020

**Research Article****Lung disease classification using machine learning algorithms****Murat Aykanat<sup>a,\*</sup> , Özkan Kılıç<sup>a</sup> , Bahar Kurt<sup>b</sup> , Sevgi Saryal<sup>c</sup>** <sup>a</sup>Computer Engineering, Yıldırım Beyazıt University, Ankara, 06010, Turkey<sup>b</sup>Clinic of Chest Diseases, Ankara Dışkapı Yıldırım Beyazıt Education and Research Hospital, Ankara, 06110, Turkey<sup>c</sup>Department of Chest Diseases, Ankara University, Ankara, 06620, Turkey

## ARTICLE INFO

*Article history:*

Received 24 September 2020

Accepted 9 October 2020

*Keywords:*

Gaussian Bayes

K-nearest neighbor

Lung diseases

Lung sounds

Support vector machines

## ABSTRACT

In this study we compared support vector machines (SVM), k-nearest neighbor (k-NN), and Gaussian Bayes (GB) algorithms in classification of respiratory diseases with text and audio data. An electronic stethoscope and its software are used to record patient information and 17930 lung sounds from 1630 subjects. SVM, k-NN and GB algorithms were run on 6 datasets to classify patients into; (1) sick or healthy with text data, (2) sick or healthy with audio MFCC features, (3) sick or healthy with the text data and audio MFCC features, (4) 12 diseases with text data, (5) for 12 disease with audio MFCC features, (6) for 12 disease with the text data and audio MFCC features. Accuracy results in SVM were %75, %88, %64, %73, %63, %70; for k-NN %95, %92, %92, %67, %64, %66; for GB %98, %91, %97, %58, %48, %58 respectively. In 12 class classification of lung diseases, the most accurate algorithm was SVM with text data. In classifying via audio data, k-NN was the most accurate. Using both audio and text data, SVM was the most accurate. When we classify healthy versus sick via text, audio and combined data, GB was always the most accurate with very high accuracy, closely followed by k-NN. We can infer from here that when we have large number of features but limited amount of samples, SVM and k-NN are best in classifying the dataset in more than two classes. However GB is best when it comes to classifying into two classes.

This is an open access article under the CC BY-SA 4.0 license.  
(<https://creativecommons.org/licenses/by-sa/4.0/>)

**1. Introduction**

Millions of people around the world suffer from pulmonary disease. The most common of these diseases are chronic obstructive pulmonary disease (COPD), asthma, pneumonia, lung cancer and tuberculosis [1]. Physicians need a medical history and physical examination in the diagnosis phase of patients with respiratory system disorders [2]. But this information is not available for computer processing [2]. Classifying respiratory diseases by taking a look at the drawn curves of the respiratory impedance or derived parameters is a difficult work for untrained physicians because it depends on the experience and ability of the physician [3]. Inductive learning systems have been used in different medical fields such as oncology, liver pathology, prognosis of the survival in hepatitis, urology, diagnosis of thyroid diseases, rheumatology, diagnosing craniostenosis syndrome,

dermatoglyphic diagnosis, cardiology, neuropsychology, gynecology, and perinatology. Automatically created diagnostic rules have increased the diagnostic correctness of specialist doctors [2].

To diagnose or classify anything, patterns have to be identified. However, if the data we have is too large, it is hard to find these patterns. In addition, traditional methods cannot be used to find patterns or create mathematical models because gathered data is generally not linear [4].

Several successful machine learning algorithms have been developed in recent years and now the error rate has become very small with deep learning algorithms [4]. Recently, machine learning, particularly in computer vision and speech recognition, almost approaches human perception level [4]. Even if expert systems are used in practice in clinical settings, machine learning systems are still being used more experimentally today [5].

\* Corresponding author. E-mail address: aykanatm@gmail.com  
DOI: 10.18100/ijamec.799363

In classification and regression analysis, one of the most commonly used supervised machine learning models is the Support vector machines (SVMs) [6], [7]. The SVM training algorithm creates a model that appoints new samples as a binary linear classifier that is not based on a likelihood [8]. In the SVM model, samples are shown as points in separate categories that are divided by a gap in the space [9]. New samples are paired into the same space and are categorized according to which side of the gap it is based on [9].

In order for SVMs to perform nonlinear classifications efficiently, data must be labeled to apply the supervised learning. When data are not labeled, an unsupervised learning approach is required that naturally clusters data to groups and then tries to match new data to these groups. SVM has a clustering algorithm that provides an improvement [10].

SVM is such a powerful algorithm it has been widely used in the biological and other sciences [11]–[13].

In previous studies, diagnosis classification was predicted by manually selected text data or audio data. As for classification methods, they used traditional machine learning algorithms such as multi-layer perceptron (MLP), multilayer neural network (MLNN), k-nearest neighbor (k-NN), probabilistic neural network (PNN). Their results can be seen in Table 3.

In this study our aim is to classify respiratory diseases, based on the collected data. This data will consist of patient demographic information, preliminary questions, symptoms, lung function test results, blood test results, X-ray results, final diagnosis and audio recordings of lung sounds by chest physicians. Since our experiments include combined text and audio data, our results may also include connections between seemingly unrelated data and conditions which may help the field of medicine with new insights.

## 2. Materials and Methods

Since we need a device to record the breathing sound, we first examined all the electronic stethoscopes on the market and found that two types of electronic stethoscopes were used today. These are the Littman 2100 Electronic Stethoscope and Thinklabs One Electronic Stethoscope. In these devices, a microphone and a series of electronic circuits are used to convert the analog signals coming from the head of the stethoscope into digital signals. This digital signal is then transmitted to the computer via a 3.5 mm microphone jack, which is common in computers and mobile devices. However, the main difference between Littman 2100 and Thinklabs One electronic stethoscope was that while Littman 2100 required proprietary software, Thinklabs One transmits the audio signal to any device utilizing any software [14]. Since these devices didn't suit to our needs, we built a custom electronic stethoscope.

The first prototype was a large device with audio out for headphones and a microphone input for the stethoscope with microphone. However, this device captured too much environmental noise which suppressed the respiratory sounds. It was also too big to carry around in a hospital environment.

The second prototype was a smaller version of the first one which had two inputs: one for stethoscope microphone signal and one for recording. It also had audio output for headphones. The device recorded stereo audio, one channel for respiratory sounds and the other channel for environmental noise. The idea behind the device was to record both audio and extract the noise from the respiratory signal. However, we found that the noise in the respiratory signal was not equivalent to the noise signal coming from the second channel, hence when it is extracted, there was a huge data loss on the signal due to the low frequency nature of the respiratory audio signals. So we decided not to use the second one either.

We found that the environmental noise contains electronic noise from the components of the device, so the more complex the device gets the more electronic noise in the final signal. So we removed the signal enhancing hardware and the device with a small and directional microphone strapped inside the head of the stethoscope with a 3.5mm microphone jack.

However, there was still noise in the recorded audio because:

- Hospital environments naturally have variety of noises such as people talking, phones, noisy medical devices, ambulance and police sirens etc.
- A scratching noise occurs when the diaphragm of stethoscope comes in contact with skin and body hair during recording.

The first problem there was not much we could do because it is impossible to provide perfect silence in hospital rooms. However, we solved the second problem simply by lubrication of the contact area.

### 2.1. Software for Data Acquisition

We developed an application that creates patient records and record, play and modify audio. It has 8 main sections:

- Patient information: First name, last name, age, gender, smoking habits, sport habits (Figure 1).
- Preliminary questions: Shortness of breath, cough, color of mucus, coughing of blood, chest pains (Figure 2).
- Symptoms: High fever, weight loss, swelling in legs, night sweating, palpitation (Figure 3).
- Audio recording: Audio recordings from 11 areas of patient's chest (Figure 4).
- Lung function test results: Forced vital capacity (FVC), forced expiratory volume in 1st second (FEV1), and FEV1 / FVC (Figure 5).

- Blood test results: White blood cell count, C-reactive protein count, and neutrophils count (Figure 6).
- X-ray results: X-ray results from 6 regions of lungs (Figure 7).
- Final diagnosis (Figure 8).

Figure 1. Patient information

Figure 2. Preliminary questions

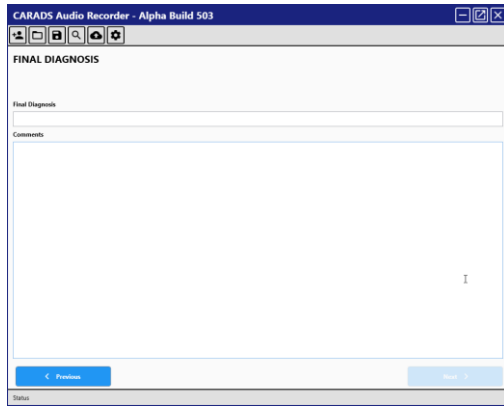
Figure 3. Symptoms

Figure 4. Audio recording

Figure 5. Lung function test results

Figure 6. Blood test results

Figure 7. X-ray results



**Figure 8.** Final diagnosis

## 2.2. Data Acquisition

Three hospitals agreed to host our research in their respiratory diseases department: Ankara University, Yıldırım Beyazıt University and Yıldırım Beyazıt Education and Research Hospital.

**Table 1.** Disease Frequencies in Dataset

Disease Name	ICD-10	Frequency
Normal		805
COPD	J44.9	211
Pneumonia	J18.9	134
Asthma	J45.9	75
Bronchiectasis	J47.0	30
IPF	J84.9	40
PTE	I26.9	43
COPD + Bronchiectasis	J44.9+ J47.0	28
COPD + Pneumonia	J44.9+ J18.9	85
Lung Cancer	D44.3	42
COPD + Emphysema	J44.9+J43.9	14
Pleural Effusion	J90	34
Pneumonia + PTE	J18.9+ I26.9	9
COPD + PTE	J44.9+ I26.9	8
Bronchitis	J41.0	53
Pneumonia + Lung Cancer	J18.9+44.3	19
Total		1630

COPD: Chronic Obstructive Pulmonary Disease  
 ICD-10: International classification of diseases [15]  
 IPF: Interstitial Pulmonary Failure  
 PTE: Pulmonary Thromboembolism

In this study we used Lenovo ThinkPad E550 Laptop for recording respiratory audio and patient data. We recorded patient information, preliminary questions, symptoms, audio recording, lung function test results, blood test results, X-ray results, final diagnosis and respiratory audio from 1630 subjects as can be seen in Table 1 and 11 positions from each patient, totaling to 17930 audio clips, each 10 seconds long.

## 2.3. Experiments

Apart from the manually selected features (age, gender, smoking habits, sport habits, shortness of breath, cough, color of mucus, coughing of blood, chest pains, high fever, weight loss, swelling in legs, night sweating, palpitation, FVC, FEV1, FEV1 / FVC, white blood cell count, c-reactive protein count, neutrophils count and x-ray results from 6 regions of lungs), since Mel Frequency Cepstral Coefficient (MFCC) features are widely used in audio detection systems, we also used MFCC features.

SVMs can help solve problems such as classifying text and hypertext and improve image classification. SVMs can provide higher search accuracy than traditional query improvement schemes after only three or four relevant feedback rounds [11].

One of the most frequently used prospective statistical classification algorithms is k-NN. It is a method utilized to classify objects based on the nearest training instances in the property area [4].

The GB is a probabilistic model. It is supposed that all data points are produced from a combination of a few Gaussian distributions with unknown parameters. It can be considered that the mixture models are a universalization of the k-means cluster, which contains data about the covariance structure of the data and hidden Gaussian centers [4].

Because of the advantages of these models, we used SVM, k-NN and GB algorithms to process the following datasets that were built with 1630 subjects:

- Dataset to predict whether the subject is ill or healthy with data that is collected by physicians manually
- Dataset to predict whether the subject is ill or healthy with MFCC features extracted from combined audio data from each subject's 11 locations on their chest
- Dataset to predict whether the subject is ill or healthy with combining data that is collected by physicians manually with MFCC features extracted from combined audio data from each subject's 11 locations on their chest
- Dataset for 12 class diagnosis classification with data that is collected by physicians manually
- Dataset for 12 class diagnosis classification with MFCC features extracted from combined audio data from each subject's 11 locations on their chest
- Dataset for 12 class diagnosis classification with combining data that is collected by physicians manually with MFCC features extracted from combined audio data from each subject's 11 locations on their chest

## 3. Results and Discussion

Our results are in Table 2. Several studies have been reported that demonstrate the benefit of computerized lung disease analysis [11], [16], [17]. However, there are by the small number of available studies for the diagnosis of lung diseases as shown in Table 3.

As shown in Table 3 the studies in the literature had limited or pre-recorded datasets [3], [18]–[30]. Because of the low number of samples and very low or distinct features, their results were not consistent. They were either very high or very low. Also pre-recorded datasets provide clean

samples which may not be the case in real life, hence producing incorrect results. To overcome this issue, we collected 17930 audio recordings from 1630 healthy and sick subjects.

In the previous studies, they did diagnosis classification with 2 classes, and one study with 3 classes [25], [29]. Also, most studies, classified subjects as healthy and ill while

some of them classified a subject group with two different illnesses. The problem with using low number of classes is that it does not really measure the performance and effectiveness of a given machine learning algorithm. In our study we classified 1630 patients into 12 disease classes as can be seen from Table 1.

**Table 2.** Experiment Results

<b>Lung Disease Classification Method</b>	<b>Accuracy</b>	<b>Precision</b>	<b>Recall</b>	<b>Sensitivity</b>	<b>Specificity</b>
<b>12 classes using text data (SVM)</b>	73%	73%	91%	91%	-
<b>12 classes using text data (k-NN)</b>	67%	67%	100%	100%	-
<b>12 classes using text data (GB)</b>	58%	58%	64%	64%	-
<b>12 classes using audio data (SVM)</b>	63%	63%	96%	96%	-
<b>12 classes using audio data (k-NN)</b>	64%	64%	99%	99%	-
<b>12 classes using audio data (GB)</b>	48%	48%	53%	53%	-
<b>12 classes using text and audio data (SVM)</b>	70%	70%	99%	99%	-
<b>12 classes using text and audio data (k-NN)</b>	66%	66%	100%	100%	-
<b>12 classes using text and audio data (GB)</b>	58%	58%	69%	69%	-
<b>Healthy versus sick using text data (SVM)</b>	75%	100%	55%	55%	52%
<b>Healthy versus sick using text data (k-NN)</b>	95%	94%	98%	98%	96%
<b>Healthy versus sick using text data (GB)</b>	98%	98%	99%	99%	98%
<b>Healthy versus sick using audio data (SVM)</b>	88%	89%	88%	88%	88%
<b>Healthy versus sick using audio data (k-NN)</b>	92%	94%	94%	92%	88%
<b>Healthy versus sick using audio data (GB)</b>	91%	98%	85%	85%	85%
<b>Healthy versus sick using text and audio data (SVM)</b>	64%	100%	43%	43%	30%
<b>Healthy versus sick using text and audio data (k-NN)</b>	92%	90%	96%	96%	95%
<b>Healthy versus sick using text and audio data (GB)</b>	97%	97%	95%	95%	97%

In the literature, diagnosis classification was made either by manually selected text data [3], [20], [23]-[27], [31], [32] or audio data [19], [31] as can be seen in Table 3. In our study we ran our experiments using text, audio and text and audio combined. This provided an insight into which features are more important and if results could be improved with text and audio data combined.

In previous studies, they used traditional machine learning algorithms such as MLP, MLNN, k-NN, PNN. In our study, we used MFCC features of audio data in SVM algorithm for classification.

Our study has three advantages over the state-of-art studies:

- Our data set (1630 subjects and 17930 audio clips) is much bigger compared to the studies done on this field.
- Respiratory audio clips in the data set are not amplified, modified, cleaned or pre-recorded by a third party which is not the case with many of the studies we looked into.
- We tested our algorithms on 6 datasets and obtained consistent results across the board which was not done in any of the state-of-art study so far.

**Table 3.** Machine Learning in Computerized Respiratory Sound Analysis Systems

Author	Subjects	Classified Items	Number of Features	Feature Extraction Method	Classification Method	Accuracy
Kahya 1997 [18]	51p 18n	COPD, restrictive lung disease, normal	14 (audio)	AR Model	k-nn	69.59%
Ashizawa 1999 [19]	110s	ILD	26 (text)	Manual selection	MLNN	90%
El-Solh 1999 [31]	682p	TB	21 (text)	Manual selection	GRNN	92.30%
Santos 2004 [20]	136p	SNPT	26 (text)	Manual selection	MLP	77%
Heckerling 2004 [32]	1160s	Pneumonia	35 (text)	Manual selection	MLP	82.8%
Barua 2004 [21]	131s	Pulmonary diseases	12 (text)	IOS	MLP	61.53%
Barua 2005 [22]	361s	Asthma	12 (text)	IOS	MLP	95.01%
Er 2008 [23]	100n 50p	TB	38 (text)	Manual selection	MLNN, GRNN	95.08%
Er 2008 [24]	100n 55p	COPD	38 (text)	Manual selection	MLNN	96.08%
Er 2009 [25]	100n 101p	Pneumonia COPD	38 (text)	Manual selection	MLNN, PNN, LVQ, AIS	94%
Temurtas 2009 [26]	150n 65p	Thyroid disease	5 (text)	Manual selection	MLNN, PNN, LVQ	94.81%
Er 2010 [27]	100n 257p	TB, COPD, Pneumonia, asthma, lung CA	38 (text)	Manual selection	MLNN, PNN, LVQ, GRNN, RBF	TB vs Others 90% COPD vs Others 88% Pneumonia vs Others 91.67% Asthma vs Others 90.91% Lung CA vs Others 93.75% Normal vs Others 99%
Yamashita 2011 [28]	101p 39n	Normal or emphysema	Not specified	Segmentation	HMM	87.4% and 88.7%
Amaral 2012 [3]	25p 25n	COPD	7 (text)	Manuel selection	KNN, SVM, MLP	95%
Rao 2018[29]	6p 6n	Normal or Diseased	11 (audio)	MFCC	GMM, SVM, KNN	50%, 75%, 91.7%
S. Jayalakshmy 2020 [30]	RALE Dataset	COPD	Automatic	Bump and Morse Scalograms	CNN	83.78%

p: Patient, c: Control, s: Subject, n: Normal, AIS: Artificial immune system, ANN: Artificial neural network, AR: Autoregressive, CA: Cancer, COPD: Chronic obstructive pulmonary disease, GRNN: Generalized regression neural network, HMM: Hidden Markov model, ILD: Interstitial lung disease, IOS: Impulse oscillometry, k-nn: k-nearest neighbour, MLNN: Multilayer neural network, MLP: Multi-layer perceptron, PNN: Probabilistic neural network, RBF: Radial basis function, SVM: Support vector machine, LVQ: Learning vector quantization, TB: Tuberculosis

#### 4. Conclusions

In this study, our first goal was to build an electronic stethoscope along with a software system that can record respiratory sounds and patient information to a computer. Audio and text datasets created by this system were used in SVM, k-NN and GB machine learning algorithms for purposes of automated analysis and diagnosis.

SVM, k-NN and GB algorithms were run on 6 datasets to classify patients into; (1) ill or healthy with text data,

(2) ill or healthy with audio MFCC features, (3) ill or healthy with the text data and audio MFCC features, (4) 12 diseases with text data, (5) for 12 disease with audio MFCC features, (6) for 12 disease with the text data and audio MFCC features. Accuracy results in SVM were %75, %88, %64, %73, %63, %70; for k-NN %95, %92, %92, %67, %64, %66; for GB %98, %91, %97, %58, %48, %58 respectively.

As a result, we have determined that, in these experiments, for the number of patients we had (1630

subjects), it was observed that the best results were found in healthy versus sick classification. The reason for that is our dataset does not have equal number of samples for each disease. Some classes are represented by just a few samples. Therefore, the classification accuracy drops as we have more classes. Also, the total number of samples affects the classification results. We have enough samples to classify 2 classes but for more accurate classification of more classes we need more samples.

In 12 class classification of lung diseases, the most accurate algorithm was SVM with text data. In classifying via audio data, k-NN was the most accurate. Using both audio and text data, SVM was the most accurate.

However when we classify healthy versus sick via text, audio and combined data, GB was always the most accurate with very high accuracy, closely followed by k-NN.

We can infer from here that when we have large number of features but limited amount of samples, SVM and k-NN are best in classifying the dataset in more than two classes. However GB is best when it comes to classifying into two classes.

Also, we can see from the results that when it comes to disease diagnosis, text and combined data produces better results than just audio data. This is also primarily true for deciding if the patient is healthy or sick. However, in deciding if the patient is healthy or sick, pure audio data can also be used as we found it to be highly accurate as well.

In addition, this system will enable to record and store patient information, especially audio data, to be shared with other physicians and to compare the new data recorded later to follow the prognosis of the patient. We believe that our method of diagnosis classification using patient data and respiratory sounds can lead the way for even more advanced computerized analysis techniques in the future.

### Author's Note

Abstract version of this paper was presented at 9th International Conference on Advanced Technologies (ICAT'20), 10-12 August 2020, Istanbul, Turkey with the title of "Lung Disease Classification using Machine Learning Algorithms".

### Acknowledgment

I would like to thank the chest physicians Prof. Dr. Sevgi Saryal, Prof. Dr. Turan Acıcan, Prof. Dr. Banu Gülbay, Assoc. Prof. Dr. Bülent Bozkurt, Assist. Prof. Dr. Gülbahar Yorulmaz Yüce and Dr. Şilan Işık for their help in auscultation, helping me collect patient and diagnosis data.

I would like to thank Dr. Dt. Banu Aykanat for coordinating all the chest physicians contributed in the

study and helping me gather patient data with the electronic stethoscope.

### References

- [1] MedHelp. (2017, January 26). [Online]. Available: <http://www.edhelp.org/Medical-Dictionary/Terms/2/8964.htm>
- [2] I. Kononenko, "Inductive and Bayesian learning in medical diagnosis", *Applied Artificial Intelligence*, vol. 7, no. 4, pp. 317-337, 1993. DOI:10.1080/08839519308949993
- [3] J. L. M. Amaral, A. J. Lopes, J. M. Jansen, A. C. D. Faria, and P. L. Melo, "Machine learning algorithms and forced oscillation measurements applied to the automatic identification of chronic obstructive pulmonary disease", *Computer Methods and Programs in Biomedicine*, vol. 105, pp. 183-193, 2012.
- [4] M. Aykanat, Ö. Kılıç, B. Kurt, and S. Saryal, "Classification of lung sounds using convolutional neural networks", *EURASIP Journal on Image and Video Processing*, vol. 65, pp. 1-9, 2017. DOI: 10.1186/s13640-017-0213-2.
- [5] E. Coiera, "Guide to Health Informatics", 2nd ed., London: CRC Press, 2003.
- [6] R. Palaniappan, K. Sundaraj, and N. U. Ahamed, "Machine learning in lung sound analysis: A systematic review", *Biocybernetics and Biomedical Engineering*, vol. 33, no. 3, pp. 129-135, 2013.
- [7] K. M. Sindhu, and H. S. Suresha, "Hand Gesture Recognition using DTW and Morphological Feature Extraction", *International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)*, vol. 5, no. 5, pp. 10171-10175, 2017.
- [8] Z. S. Huang, C. C. Chuang, C. W. Tao, M. Y. Hsieh, C. X. Zhang, and C. W. Chang, "iOS-based people detection of multi-object detection system" In 2016 Joint 8th International Conference on Soft Computing and Intelligent Systems (SCIS) and 17th International Symposium on Advanced Intelligent Systems (ISIS, IEEE), Sapporo, Japan, pp. 868-873, August 2016.
- [9] G. Serbes, C. O. Sakar, Y. P. Kahya, and N. Aydin, "Feature extraction using time-frequency/scale analysis and ensemble of feature sets for crackle detection", 33rd Annual International Conference of the IEEE EMBS, Boston, Massachusetts USA, pp. 3314-3317, 2011.
- [10] Ng. Andrew, (2017, January 16), Support Vector Machines. [Online]. Available: <http://cs229.stanford.edu/notes/cs229-notes3.pdf>
- [11] B. Flietstra, N. Markuzon, A. Vyshedskiy, and R. Murphy, "Automated analysis of crackles in patients with interstitial pulmonary fibrosis", *Pulm Med.*, vol. 2010, pp. 1-7, 2011.
- [12] A. Gouda, S. El Shehaby, N. Daa, and M. Abougabal, "Classification Techniques for Diagnosing Respiratory Sounds in Infants and Children", In 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, pp. 0354-0360, January 2019.
- [13] S. Don, "Random Subset Feature Selection and Classification of Lung Sound", *Procedia Computer Science*, vol. 167, pp. 313-322, 2020.
- [14] S. H. Ah, and S. Lee, "Hierarchical representation using NMF neural information processing", Heidelberg, Berlin: Springer, 2013.
- [15] T.C. Sağlık Bakanlığı Temel Sağlık Hizmetleri Genel Müdürlüğü, "Türkiye kronik hava yolu hastalıklarının (Astım - KOAH) önleme ve kontrol programı (2009 - 2013)", Ankara: T.C. Sağlık Bakanlığı Eylem Planı, 2009.
- [16] H. Pasterkamp, S. S. Kraman, and G. R. Wodicka, "Respiratory sounds, advances beyond the stethoscope", *Am J Respir Crit Care Med*, vol. 156, pp. 974-987, 1997.
- [17] J. E. Earis, and B. M. G. Cheetham, "Current methods used for computerized respiratory sound analysis", *Eur Respir Rev.*, vol. 10, no. 77, pp. 586-590, 2000.
- [18] Y. P. Kahya, E. C. Guler, and S. Sahin, "Respiratory disease diagnosis using lung sounds", *Engineering in Medicine and Biology Society, Proceedings of the 19th Annual International*

- Conference of the IEEE, Chicago, IL, USA, pp. 2051-2053, 1997.
- [19] K. Ashizawa, T. Ishida, H. MacMahon, C. J. Vyborny, S. Katsuragawa, and K. Doi, "Artificial neural networks in chest radiography: application to the differential diagnosis of interstitial lung disease", *Academic Radiology*, vol. 6, no. 1, pp. 2-9, 1999.
- [20] A. M. Santos, B. B. Pereira, J. M. Seixas, F. C. Mello, and A. L. Kritski, "Neural Networks: An Application for Predicting Smear Negative Pulmonary Tuberculosis", *Advances in Statistical Methods for the Health Sciences Statistics for Industry and Technology*, pp. 275-287. DOI:10.1007/978-0-8176-4542-7\_18.
- [21] M. Barua, H. Nazeran, P. Nava, V. Granda, and B. Diong, "Classification of pulmonary diseases based on impulse oscillometric measurements of lung function using neural networks", In the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Francisco, CA, USA, vol. 2, pp. 3848-3851, September 2004.
- [22] M. Barua, H. Nazeran, P. Nava, B. Diong, and M. Goldman, "Classification of impulse oscillometric patterns of lung function in asthmatic children using artificial neural networks", In 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, pp. 327-331, January 2006.
- [23] O. Er, F. Temurtas, and A. Ç. Tanrikulu, "Tuberculosis Disease Diagnosis Using Artificial Neural Networks", *Journal of Medical Systems*, vol. 34 no. 3, pp. 299-302, 2008. DOI:10.1007/s10916-008-9241-x
- [24] O. Er, and F. Temurtas, "A study on chronic obstructive pulmonary disease diagnosis using multilayer neural networks", *Journal of Medical Systems*, vol. 32, no. 5, pp. 429-432, 2008.
- [25] O. Er, C. Sertkaya, F. Temurtas, and A. C. Tanrikulu, "A comparative study on chronic obstructive pulmonary and pneumonia diseases diagnosis using neural networks and artificial immune system", *Journal of Medical Systems*, vol. 33, no. 6, pp. 485-492, 2009.
- [26] F. Temurtas, "A comparative study on thyroid disease diagnosis using neural networks", *Expert Systems with Applications*, vol. 36, pp. 944-949, 2009.
- [27] O. Er, N. Yumusak, and F. Temurtas, "Chest diseases diagnosis using artificial neural networks", *Expert Systems with Applications*, vol. 37, no. 12, pp. 7648-7655, 2010.
- [28] M. Yamashita, S. Matsunaga, and S. Miyahara, "Discrimination between healthy subjects and patients with pulmonary emphysema by detection of abnormal respiration", In 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, pp. 693-696, May 2011.
- [29] A. Rao, S. Chu, N. Batlivala, S. Zetumer, and S. Roy, "Improved detection of lung fluid with standardized acoustic stimulation of the chest", *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 6, pp. 1-7, 2018.
- [30] S. Jayalakshmy, and G. F. Sudha, "Scalogram based prediction model for respiratory disorders using optimized convolutional neural networks", *Artificial Intelligence in Medicine*, vol. 103, pp. 101809, 2020.
- [31] A. A. El-Solh, C. B. Hsiao, S. Goodnough, J. Serghani, and B. J. B. Grant, "Predicting active pulmonary tuberculosis using an artificial neural network", *Chest*, vol. 116, pp. 968-973, 1999..
- [32] P. S. Heckerling, B. S. Gerbera, T. G. Tapeç, and R. S. Wigton, "Use of genetic algorithms for neural networks to predict community-acquired pneumonia", *Artificial Intelligence in Medicine*, vol. 30, pp. 71-84, 2004.