International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048
Volume 7, Issue 5, pp 99-107, May-2023
https://doi.org/10.47001/IRJIET/2023.705011

# Review "Smoker/Non-Smoker Classification of People Using a Speech Signal"

[1]Ibrahim Khudhur Zaal, [2]Yusra Faisal Mohammad

[1]Computer Science Department, College of Computer and Mathematics & Mosul University, Iraq
[2]Assistant Professor, Computer Science Department, College of Computer and Mathematics &, Mosul University, Iraq

***Abstract* - Speech is a behavioral biometric that can reveal a person's age, gender, race, and emotional state. The speech signal may also be used to ascertain a person's behavior, such as whether or not they smoke or take drugs. One of the topics that is frequently studied in the field of speech technology is the smoking habits of speakers. Over the past years, a lot of research has been done in this area, but little progress has been made in this field. As deep learning techniques have advanced in most machine learning fields, they have replaced earlier research techniques for speech recognition and verification. The most cutting-edge method for confirming and recognizing a speaker's identity is currently deep learning. This study's objective is to analyze research that uses speech signals and artificial intelligence to distinguish smokers from non-smokers. Every speech recognition system uses a variety of algorithms to convert sound waves into information that can be interpreted and processed by the system, which then generates an output that can be used as needed.**

***Keywords:*** Classification, Biometric techniques, Machine learning, Smokers detection, speech signal.

## I. INTRODUCTION

Unquestionably, spoken communication plays a significant part in human life. One of the biometric characteristics that set one individual apart from the other is their voice. By taking characteristics out of the voice stream, the identity of the speaker is discovered. The procedure may be summed up as the classification of a person as a smoker or not based on the information that is extracted from the speaker's voice signal. Because smoking changes the cell components on the vocal cords and has a detrimental impact on sound quality, it has an influence on the human larynx. Moreover, smoking is a major contributor to voice problems and laryngeal lesions. To develop a preventive, treatment, and rehabilitation strategy, it is vital to investigate the impact of smoking on the smoker's voice. The goal of this work was to explore what is associated with the speech signal through a variety of valuable information retrieved from since. Despite several studies that focused on enhancing the extraction of features from speech, the accuracy of classifications remained wholly unmet. And be aware of how much close-to-reality accuracy was achieved [1][14].

All papers relating to earlier deep learning techniques for separating smokers from non-smokers are covered in this study. The study comprises an abstract, an introduction, a voice signal, classification strategies, deep learning, earlier investigations, a study-to-study comparison, and a conclusion.

### 1.1 The Impact of Smoking on Voice

The human larynx is affected by the cigarette smoke because smoking leads to change in the cell components on the vocal cords and has a negative effect on the sound quality. It irritates and dries the vocal cords, causing them to swell and preventing them from working properly. Based on some research show that a short duration of a smoking habit – less than a decade – has a clear effect on some voice parameters. The number of cigarettes smoked per day displayed a linear effect on these parameters in the group [2].

**Figure 1: 1 on the left shows a vocal cord of a smoker who suffers from bulge and inflammation and a vocal cord of a non-smoker on the right**

## II. BIOMETRIC TECHNOLOGY

The automated use of physiological or behavioral traits to establish or confirm an individual's identification is known as biometric technology. Any physiological or behavioral trait of a human that possesses the necessary biometric qualities is also referred to as biometric. They are: universal (everyone should possess that characteristic), unique (no two people should be exactly alike in that characteristic), permanent (invariant with time), collectable (can be measured quantitatively), reliable (must be safe and perform at an

© 2023-2017 IRJIET All Rights Reserved 

ISSN (online): 2581-3048
Volume 7, Issue 5, pp 99-107, May-2023
https://doi.org/10.47001/IRJIET/2023.705011

acceptable level), acceptable (non-invasive and socially tolerable), and non-circum ventable (how easily the system is fooled into granting access to impostors).

In an effort to provide accurate personal identification, biometric technologies create computer models of the physical and behavioral traits of individuals. The aforementioned traits cover both visual cues and other human traits including voice, movement, odor, DNA, and almost anything else that may be used to identify a person specifically.

Since that biometric technology is designed to identify people, they are frequently considered to be applications of pattern recognition algorithms. As a result, it adopts techniques and algorithms from computer vision; however it is not limited to visual imagery. More focus has recently been placed on biometric synthesis, which is the opposite direction of biometrics. The portrayal of biometric phenomena from their associated computer models is known as synthesis. This is in contrast to identification, which calls for signal processing in order to highlight the models required for recognition. But, using the equivalent computer model, we could create a synthetic face. Muscular movements may be included into such a model to simulate face expressions [21][22].

### 2.1 Physiological Biometric techniques

Biometrics refers to a person's bodily traits, such as their fingerprints, hand geometry, iris, face, and DNA. Each biometric characteristic has advantages and disadvantages [21][31]. Below are the most important techniques used in this field:

- Fingerprint
- Hand Geometry
- Iris
- Face
- DNA

### 2.2 Behavioral Biometric techniques

Methods for identifying behavior include paying attention to a person's behaviors and allowing the user to exert control over those actions. As a result of the high amount of internal variations (mood, health, etc.) that these approaches account for, biometrics based on them are only relevant when used consistently. Keystroke, signature, and voice are all included [21][31]. Below are the most important techniques used in this field:

- Keystroke
- Signatures
- Voice

## III. CLASSIFICATION METHODS

Classification is one of the most frequently encountered decision-making tasks in human activity. A classification problem occurs when an object needs to be assigned to a predefined group or class based on a number of observed attributes related to that object. For examples, see stock market prediction, weather forecasting, bankruptcy prediction, medical diagnosis, speech recognition, and character recognition, to name a few. These classification problems can be solved both mathematically and in a nonlinear fashion. The difficulty of solving such problems mathematically lies in the accuracy and distribution of data properties and model capabilities [6].

The purpose of classification methods is to group data into the required structure based on common characteristics. Classification makes it possible to determine those data whose type or affiliated group is unknown. Classification methods application stage of a data mining process a machine learning technique selected at this stage is applied to the data set, and the result is obtained accordingly. Machine learning is not merely a technique used on data; it is an area of artificial intelligence. Classification is based on a learning algorithm. The purpose of learning is to create a classification model. Some data are selected as training data during the process of applying classification, and the algorithm is operated on both this training data set and the other data set with unknown classification so as to determine which group this test data belongs to. Classification in machine learning makes use of decision trees, regression trees, support vector machines, and statistical classification methods. There is a different algorithm for each method. Classification methods are estimator models. Besides these models, classification also uses artificial intelligence techniques such as genetic algorithms and artificial neural networks [13].

## IV. MACHINE LEARNING

Over time, advanced abilities for recognizing patterns such as speech, handwriting, facial traits, and so on have evolved. Machine learning arose from the development of computer programs that allow machines to acquire the aforementioned abilities from prior experience. "A computer program is said to learn from experience E with regard to some class of tasks in T and performance measure P if its performance at tasks in T, as measured by P, increases with experience E," Mitchell noted in the context of machine learning.

The following are some fundamental machine learning terms:

- Features: an attribute set representing the input as a vector or linear array.

© 2023-2017 IRJIET All Rights Reserved

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 5, pp 99-107, May-2023

https://doi.org/10.47001/IRJIET/2023.705011

- Labels: the category or class with which the item is related (e.g., positive or negative in binary classification or a real value in regression).
- During the learning phase, training data is utilized to train the ML algorithm.
- Test data: data used to evaluate the learning algorithm's performance during the generalization phase [1].

### 4.1 Supervised Learning

In supervised learning, the computer is trained using a labeled data set with known output responses or classes for each input data vector. The machine is taught to construct the curved surface that best matches the training data set given a set of input data, and during testing, the machine is expected to appropriately interpolate the new data across the curved surface. This category includes feed-forward neural networks such as perceptrons (which use the delta learning rule or the perceptron learning rule), multilayer perceptrons (which use back propagation), and restricted MLPs [1][18].

### 4.2 Unsupervised Learning

Unsupervised learning assumes that the computer will learn the patterns in the unlabeled input data set on its own, without any assistance from the environment. The task may be characterized as identifying patterns in the input data set in order to divide or cluster the training data into relevant subgroups. Which include creating efficient techniques to categorize data into meaningful groups. Hebb and Hopfield networks (Hebbian learning), Kohonen networks (self-organizing maps), and adaptive resonance theory (ART) networks (ART) are some examples of competitive learning. The gradient descent back propagation unsupervised learning approach is used to train the network to replicate the input [1][18].

### 4.3 Semi-supervised Learning

Semi-supervised learning uses both labeled and unlabeled data to train the system. A small fraction of labeled data is often combined with a large quantity of unlabeled data. This sort of learning strategy is often used in situations where getting tagged data is too costly [1].

## V. PREVIOUS STUDIES

In 2007, the researchers Khalid Saeed and Mohammed K. Nammous presented their study "Automatic detection of smoking habits from spontaneous telephone speech signals". Each utterance is modeled using i-vector and non-negative factor analysis (NFA) frameworks, which yield low-dimensional representations of utterances by applying factor analysis to Gaussian mixture model means and weights, respectively. Each framework is evaluated using different classification algorithms to detect the smoker speakers. Finally, score-level fusion of the i-vector-based and the NFA-based recognizers is considered to improve the classification accuracy. The proposed method is evaluated on telephone speech signals of speakers whose smoking habits are known and drawn from the National Institute of Standards and Technology (NIST) 2008 and 2010 Speaker Recognition Evaluation databases. Experimental results over 1194 utterances show the effectiveness of the proposed approach for the automatic smoking habit detection task [4].

In 2012, the researcher Hiba Adreese Youni and others presented in this study an algorithm for studying speech signal properties for both smokers and non-smokers and then determining whether the person is a smoker or not based on his speech signal. A data base that contains 30 speech signals, of which 15 belong to smokers and 15 belong to non-smokers, is for males only. In this algorithm, formant frequencies such as f1 and f2 were adopted as characteristic properties for speech signals for splitting between two classes, which it calculates using the lpc algorithm. The algorithm consists of two stages: At the data base preparation stage and speaker state classification stage, the absolute, eclideance, and d1 distance were adopted as measures for evaluating the performance of the algorithm, and it gave convergence results [5].

In 2012, Lingying Chai and others used Acoustic dimension and correlation dimension (D2) analyses in research, the study included 73 subjects, 36 nonsmokers and 37 smokers. A segment of sustained vowel production. D2 values for smokers were significantly higher than D2 values for nonsmokers ($P < 0.001$). Jitter and shimmer analysis showed higher values for these parameters among smokers. Logistic regression indicated a higher predictive power with D2, and ROC analysis found no significant difference between the analysis methods [15].

In 2014, the researcher Rosa L. Figueroa presented a system to identify and extract patients' smoking status from clinical narrative text in Spanish. The clinical narrative text was processed using natural language processing techniques. The dataset used for classification had 2,465 documents, each one annotated with one of the four smoking status categories. They used two types of feature representations: single-word tokens and bigrams. The classification problem was divided into two levels. First recognizing smokers (S) and non-smokers (NS); second recognizing current smokers (CS) and past smokers (PS). For each feature representation and classification level, they used two classifiers: support vector machines (SVM) and Bayesian networks (BN). Our results show that SVM combined with the bigram representation

*International Research Journal of Innovations in Engineering and Technology (IRJIET)*

ISSN (online): 2581-3048

Volume 7, Issue 5, pp 99-107, May-2023

https://doi.org/10.47001/IRJIET/2023.705011

performed better in both classification levels. For S vs. NS classification level performance measures, they were: ACC = 85%, Precision = 85%, and Recall = 90%. For CS vs. PS classification level performance measures, they were: ACC = 87%, Precision = 91%, and Recall = 94% [8].

In 2014, the researcher S. Ben Jebar used The software Glottex, a bio-mechanical characterization of speech, is used in order to identify smoker speakers. All voice samples analyzed in the study were extracted from the Massachusetts Eye and Ear Infirmary (MEEI) voice disorders database. A population of 292 speakers, producing the sustained vowel /a/, is considered. For each gender, the subjects are subdivided into two groups: smokers and non-smokers. The distribution of the whole database is as follows: 65 female smokers, 109 female non-smokers, 62 male smokers, and 56 male non-smokers The features used to describe the cover and the body parts of the vocal folds in terms of mass, loss, and stiffness are relatively correlated with the smoking state because the vocal folds are damaged by tobacco. A statistical analysis is carried out to identify relevant features. A classification procedure is carried out to show that four smokers over five are correctly identified [11].

In 2016, the researchers Ouissam Zealouk and others presented an automatic speech recognition system based on Sphinx4 that permits to detect smokers. This research project is carried out using the Amazigh language in order to compare the voices of normal persons and smokers. In this work, they employed the ASR technology to develop a system that is able to automatically detect smokers and non-smokers using the Arabic digits of speech. After the experiences the best recognition rate for smokers is 45.93 % and for non-smokers they have 87.23 %. After this experience, they find that CMU Sphinx4 is able to detect a difference between the voices of smokers and non-smokers [2].

In 2016, Using speech biometrics, researcher Mohammad A. M. Abushariah described in his work "Automatic Identity Recognition Using Speech Biometric", the use of vector quantization (VQ) to build an automatic identity recognizer is heavily stressed in this work. We used speech data from 20 participants—10 men and 10 women—for our initial round of recognizer development. A total of 600 utterances were recorded by all participants, three of each English digit (0–9) being recorded by each participant. We collected speech data from 100 participants—50 males and 50 women—during our second phase in order to build the recognizer. Text-dependent and text-independent voice data are separated. For our phase one automated identity recognizer development project, the feature extraction and feature classification phases have been applied to the 600 utterances The speech-based automatic identifying recognition system (MFCC) is built on the Mel-Frequency Cepstral Coefficient, the most widely used feature extraction technique. 76% accuracy was the best possible score. [19].

In 2016, the researcher Abdurrahman Yldz and others presented in his study “Analysing human voice and classification of voice frequencies according to smoking effect”, the basic frequency of the human voice has been studied, as well as how smoking affects certain frequencies. The voices of the adult male research participants were captured and recorded in the proper setting and circumstances. Signal processing strategies are used to examine the recorded voices. The numerical value of the speech analysis data is compiled into a data set, and this data set is then categorized using machine learning techniques. According to the outcome of this categorization, the voices that are obtained from fresh people are forecasted. In order to identify diseases from the human voice in the future, voices will be examined and subjected to pathological investigation, similar to findings seen in diseases of the human voice using methods based on the results achieved in this study. With between 69.5% and 87.5% of success rates, machine learning and classification algorithms are used in the study to classify the fundamental frequency of the voice in men between the ages of 25 and 35[30].

In 2017, the researchers Dionysus Tafiadis and others designed a system to examine the relationship between voice handicap index and acoustic parameters of university student smokers' voices in Greece. The impact on voice characteristics of different risk factors has been studied and also correlated to cigarette smoking. One hundred and ten female non-dysphonic students (aged 18 to 34) who smoked were recruited. Participants answered the Voice Handicap Index, and their voices were recorded. Acoustic analysis of voice characteristics was performed with the Dr. Speech software system. Results indicated that some measures were predictive of overall, functional, and emotional Voice Handicap Index scores. Other voice parameters had no cohesive or predictable pattern on voice handicap index scores. Significant relationships between Voice Handicap Index individual statements and smokers’ voice characteristics were also observed [12].

In 2017, the researcher Mohammadzadeh presented in his study that this cross-sectional study was performed on 45 smokers (34 male and 11 female, mean age: 38.4 9.7 years) attending the stop smoking center of Beheshti University of Medical Sciences and 32 non-smokers (19 male and 13 female, age range: 28.43  7.8 years). The vowel /a/ was sustained for 5 seconds, and the acoustic characteristics, i.e., F0 and its variation, jitter, shimmer, changes in amplitude, and noise to harmonic ratio, were analyzed with Multi-

© 2023-2017 IRJIET All Rights Reserved 

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 5, pp 99-107, May-2023

https://doi.org/10.47001/IRJIET/2023.705011

Dimensional Voice Program software. The results from this study show that F0 and Fmin decreased in smokers, and there was a significant difference with non-smokers (p 0.05). The NHR, vAm, Atri, and vF0 in smokers are significantly higher than non-smokers (p 0.05). In addition, the relation between the number of cigarettes smoked per day showed that it has a significant correlation with parameters F0 (P = 0.010; r = -0.374) and Fmin (p = 0.004; r = -0.423)[3].

In 2017, the researcher Mohammad Akhtar presented in his study “Text Independent Biometric Authentication System Based On Voice Recognition” This study implements and analyses a text-independent biometric authentication system based on randomly prompted text. Since the prompted text-phrase is not known to the speaker in advance. The system uses Mel Frequency Cepstrum Coefficient (MFCC) to extract speech features and Gaussian Mixture Model (GMM) for speaker modeling. Universal Back Model (UBM) is used as an alternate (impostor) model for likelihood ratio testing in the authentication application. The system uses a fast segmentation and clustering technique for testing. Experiment was conducted on 10 speech samples of 1 minute duration each. An improvement of 57.93% in testing timewas observed by using 'FOLSS’ as compared to traditional segmentation and clustering technique (AHC). Average testing time for 10 speech samples using FOLSS was 1.83 second as compared to 4.35 second for traditional method of STE and AHC[24].

In 2018, the researcher Aws Saad Shawkat presented in his study “Evaluation of Human Voice Biometrics and Frog Bioacoustics Identification Systems Based on Feature Extraction Method and Classifiers”, In this study, the frog species is determined using bioacoustics based on frog cries. The datasets are gathered from two sources, namely online databases. Both databases are wave-shaped. The human database is recorded at a frequency of 32 kHz and a resolution of 16 bits. At 48 kHz and 32 bits, the frog database is captured. In contrast to the 37 speakers used by the human speaker identification method, the frog database uses nine different species. As a result, the well-known features utilized in audio-based biometric systems, i.e., MFCC, are being investigated as features for the frog bioacoustics-based identification system. SVM, k-Nearest Neighbor (k-NN), Local Mean k-Nearest Neighbor (LMk-NN), and Fk-NN classifiers have all had their classification results. The effectiveness of the proposed classifiers-based biometric and frog bioacoustics systems is assessed. The frog bioacoustics identification system and the biometric speaker identification system with 20 training data both had accuracy of 97% and 93.38%, respectively, while utilizing the FkNN classifier to get the best results[25].

In 2018, the researcher Humberto Pérez Espinosa and others presented in his study “Children Age and Gender Classification Based on Speech Using ConvNets”. In this study, presented a work on creating age- and gender-automatic classifiers for kids in the early years of school (between 6 and 11 years old). They produced a speech corpus including 174 kids interacting in a Wizard of Oz-inspired setting with a few robots. Low-level acoustic characteristics were used to manually segregate and describe the recorded speech. They then used a convolutional neural network architecture to train the classification models. They incorporated the use of a mathematical object called covering arrays to construct the set of ideal parameters for neural network design due to the difficulty of the tuning process for the appropriate selection of the parameters used for this form of neural network. They got positive findings considering how difficult it is to categorize children's speech [26].

In 2019, the researcher Jordan J. Bird and others presented in his study “Accent Classification in Human Speech Biometrics for Native and Non-native English Speakers” This preliminary study explores various methods of human accent recognition through classification of locale. Classical, ensemble, time series and deep learning techniques are all explored and compared. A set of diphthong vowel sounds are recorded from participants from the United Kingdom and Mexico, and then formed into a large static dataset of statistical descriptions by way of their MFCC at a sample window length of 0.02 seconds. Using both flat and time series data, various machine learning models are trained and compared to the scientific standard Hidden Markov Model (HMM). Results through 10 fold cross validation show that a vote of average probabilities between a Random Forest and Long Short-term Memory Neural Network result in a classification accuracy of 94.74%, outperforming the speech classification standard Hidden Markov Model by a 5% increase in accuracy[23].

In 2019, the researcher Andrzej Majkowski and others presented their study "Identification of Gender Based on Speech Signal", This study presents a gender identification based on speech signals with supervised machine learning implementation. At first, a database of speech signals in Polish was collected. Next, a set of features from the audio signal were calculated. The features were further used to train a neural network. Audio signal processing and implementation of the neural network were performed in Python, and the calculation of features was done in the R language. The neural network training process was carried out using only CPU, then CPU with GPU, and the times of the programs execution were compared. The obtained accuracy of gender recognition was 92.4%. The use of the GPU accelerated the network learning process several times [29].

© 2023-2017 IRJIET All Rights Reserved

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 5, pp 99-107, May-2023

https://doi.org/10.47001/IRJIET/2023.705011

In 2020, the researcher Zhizhong Maa and others discussed smoking cessation clinical research and practice, that is, smoking abstinence. Speech signals convey important information about a speaker, such as age, gender, body size, emotional state, and health state. They investigated (1) if smoking could measurably alter voice features, (2) if smoking cessation could lead to changes in voice, and therefore (3) if the voice-based smoking status assessment has the potential to be used as an objective smoking cessation. They searched nine scientific databases for original studies involving the effects of smoking on voice features, the effects of smoking cessation on voice features. They found that fundamental frequency, jitter, shimmer, harmonics to noise ratio, and other voice features are affected by smoking and could be used to assess smoking status[9].

In 2020, the researcher Meng Zhao and others applied a support vector machine-based classification method to discriminate between 70 young male smokers and 70 matched nonsmokers using their diffusion tensor imaging (DTI) data. The classification procedure achieved an average accuracy of 88.6% and an average area under the curve of 0.95. The most discriminative features that contributed to the classification were primarily located in the sagittal stratum (SS), external capsule (EC), superior longitudinal fasciculus (SLF), anterior corona radiata (ACR), and inferior fronto-occipital fasciculus (IFOF). The following regression analysis showed a significant negative correlation between the average RD values of the left ACR ($r = -0.247$, $p = 0.039$) and FTND. The average MD values in the right EC ($r = -0.254$, $p = 0.034$) and RD values in the right IFOF ($r = -0.240$, $p = 0.046$) were inversely associated with pack-years. The findings indicate that the discriminative white matter (WM) features as brain biomarkers provide great predictive power for smoking status and suggest that machine learning techniques can reveal underlying smoking-related neurobiology [13].

In 2021, the researcher Yuki Saishu presented and others in his study "A CNN-based approach to identification of degradations in speech signals", the acoustic mismatch in speech-based applications has been addressed in this work through the development of a number of improvement strategies. In this study, the researchers suggest a new convolutional neural network (CNN)-based method for automatically identifying the three main degradation types that are frequently present in speech-based applications: additive noise, nonlinear distortion, and reverberation. This method uses the log-mel spectrogram of audio signals to apply a series of parallel CNNs, each of which detects a specific sort of deterioration. According to experimental results employing two different speech kinds, namely diseased voice and normal running speech. By highlighting the areas of the log-mel spectrogram that have a greater impact on the target degradation, they can visually see how the network decides to distinguish between different forms of deterioration in speech signals using the score weighted class activation mapping[27].

In 2021, the researcher Bhagyalaxmi Jena and others presented in his study "Gender Recognition and Classification of Speech Signal", In this work, speech signal analysis was based on both time and frequency domain. Different speech parameters were generated by short-time, statistical and spectral analysis. The differences in parameters was used as a working principle for the gender model to recognize the gender of the unknown user. The classifier model based on Genetic Algorithm, GMM is known to have an accuracy of 70% with complex training. So, the classifiers used in this work were KNN and SVM. After training and testing, the accuracy of the system was found to be 80 percent on average[28].

In 2022, the researcher Zhizhong Ma and others aim to identify speakers' smoking status from their speech. This study focuses on determining the best acoustic features for smoker identification. They investigate the performance of four acoustic feature sets or representations extracted using three feature extraction or learning approaches: (i) hand-crafted feature sets, including the extended Geneva Minimalistic Acoustic Parameter Set and the Computational Paralinguistics Challenge Set; (ii) the Bag-of-Audio-Words representations; (iii) the neural representations extracted from raw waveform signals by SincNet. Experimental results show that: (i) SincNet feature representations are the most effective for smoker identification and outperform the MFCC baseline features by 16% in absolute accuracy; (ii) the performance of hand-crafted feature sets and the Bag-of-Audio-Words representations rely on the scale of the dimensions of feature vectors [10].

## VI. COMPARISON BETWEEN RELATED WORKS

The most significant prior research in the area of categorizing smokers and non-smokers are included in this table. The table includes the study year, the algorithm that was used, the data that was used, as well as the measurements and findings.

**Table 1: Comparison between related works**

| Ref No. | Year | Algorithm | Type ofdataset | Metrics |
|---|---|---|---|---|
| **4** | 2007 | i-vector and (NFA) frameworks. | (NIST) 2008 | 99.5% Multi speaker mode and 94.5% speaker |

ISSN (online): 2581-3048
Volume 7, Issue 5, pp 99-107, May-2023
https://doi.org/10.47001/IRJIET/2023.705011

| | | | | |
|---|---|---|---|---|
| | | | | independent mode |
| **5** | 2012 | lpc algorithm | 30 speech signals 15 smoker and 15 non-smokers | D1 it gave convergence results. |
| **15** | 2012 | Acoustic dimension and correlation dimension analyses | 36 nonsmokers and 37 smokers | D2 higher than D2 values for nonsmokers (P < 0.001). |
| **8** | 2014 | SVM and BN | 2,465 each one annotated with one of the four smoking status | Measures were: ACC=87%, Precision=91%, and Recall=94%. |
| **11** | 2014 | The software Glottex. | (MEEI) voice disorders database. | A classification procedure is carried to show that 4 smokers over 5 are correctly identified. |
| **2** | 2016 | Sphinx4, ASR | Amazigh digits (0-9). | The best recognition rate for smokers is 45.93 % and for non-smokers they have 87.23 % |
| **19** | 2016 | vector quantization | 10 men and 10 women. Each participant recorded each English digit (0–9) three times. | The highest degree of accuracy was 76%. |
| **30** | 2016 | Decision Table | The voices of the adult male captured and recorded in the proper setting and circumstances. | Between 69.5% and 87.5% of success rates. |
| **12** | 2017 | Dr. Speech software system | One hundred and ten female non-dysphonic students (aged 18 to 34) | Results indicated that some measures were predictive of overall, functional and emotional Voice Handicap Index scores |
| **3** | 2017 | MDVP | 45 smokers<br>32 non-smokers | F0 decreased in smokers and difference with non-smokers (p<0.05).<br>The NHR , vAm, Atri, vF0 in smokers higher than non-smokers (p<0.05). |
| **24** | 2017 | MFCC,GMM) and UBM | Direct record voice person from 5 words | 57.93% in testing time was observed by using 'FOLSS' |
| **25** | 2018 | SVM and Fk-NN | namely online databases. | Accuracy of 97% and identification 93.38% |
| **26** | 2018 | ConvNets | 174 children (78 female, 96 male), aged 6–11 (8.62 mean, 1.73 standard deviation) | For age and gender classification, they got comparable results to the ones reported in the literature. |
| **23** | 2019 | HMM | Person from United Kingdom and Mexico | Accuracy of 94.74% |
| **29** | 2019 | NN with Keras and Tensorflow libraries. | a database of speech signals in Polish was collected. | accuracy of gender recognition was 92.4% |
| **9** | 2020 | A review studies on smoking status based on voice features. | Nine scientific databases. | F0, jitter, shimmer, and HNR are affected by cigarette smoking. |
| **13** | 2020 | SVM | 70 male smokers and 70 nonsmokers. | Accuracy : 83% |
| **27** | 2021 | CNN | mPower mobile Parkinson's disease (MMPD) data set | Baseline Proposed<br>Noise: 0.71±0.00 0.95±0.01<br>Distortion: 0.83±0.01 1.00±0.00<br>Reverberation: 0.84±0.01 0.99±0.00 |
| **28** | 2021 | GMM | 300 different males and females of age 20-22 The sentence | Accuracy of 70% with complex training. |
| **10** | 2022 | SincNet | Online dataset | MFCC 16% |

## VII. CONCLUSION

The significance of human voice cues in automated speech recognition and the impact of smoking on these signals have been reviewed in this work. The most significant research is on using artificial intelligence to distinguish smokers from non-smokers. Numerous researches have produced cutting-edge findings about the capability of differentiating between smokers and non-smokers using a verbal signal that incorporates a variety of elements. Neural networks may be taught to recognize certain qualities via which a classification determination can be achieved.

## REFERENCES

[1] Fadhel, Mazin Ali, et al. " Unsupervised and Semi-Supervised Speech Recognition System: A Review."Al-Rafidain Journal of Computer Sciences and Mathematics (RJCM), Vol. x, No. x, 2022 (x-x).

© 2023-2017 IRJIET All Rights Reserved 

ISSN (online): 2581-3048
Volume 7, Issue 5, pp 99-107, May-2023
https://doi.org/10.47001/IRJIET/2023.705011

[2] Zealouk, Ouissam, Hassan Satori, and Khalid Satori. "Voice comparison between smokers and non-smokers using Amazigh digits." Int. J. Multi-disciplinary Sci 1 (2016): 106-110.

[3] Mohammadzadeh, A., and Z. Mousavi. "The Comparison of Acoustic Voice Features in Smokers and Non-Smokers." Journal of Paramedical Sciences & Rehabilitation 5.4 (2016): 50-58.

[4] Saeed, Khalid, and Mohammad Kheir Nammous. "A speech-and-speaker identification system: Feature extraction, description, and classification of speech-signal image." IEEE transactions on industrial electronics 54.2 (2007): 887-897.

[5] Youni, Hiba Adreese, et al. "An Algorithm for Smoker Detection." AL-Rafidain Journal of Computer Sciences and Mathematics 10.1 (2013): 61-74.

[6] Sathya, Ramadass, and Annamma Abraham. "Comparison of supervised and unsupervised learning algorithms for pattern classification." International Journal of Advanced Research in Artificial Intelligence 2.2 (2013): 34-38.

[7] Rachman, Azril Izha.,"Analysis of Lung Capacity for Smokers and Non-Smokers Based on Age and Gender "Article 2021.

[8] Figueroa, Rosa L., Diego A. Soto, and Esteban J. Pino. "Identifying and extracting patient smoking status information from clinical narrative texts in Spanish." 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2014.

[9] Ma, Zhizhong, et al. "Towards the objective speech assessment of smoking status based on voice features: a review of the literature." Journal of Voice (2021).

[10] Ma, Zhizhong, et al. "Determining the best Acoustic Features for Smoker Identification." ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022.

[11] Jebara, S. Ben. "Bio-mechanical characterization of voice for smoking detection." 2014 22nd European Signal Processing Conference (EUSIPCO). IEEE, 2014.

[12] Tafiadis, Dionysios, et al. "Voice data on female smokers: coherence between the voice handicap index and acoustic voice parameters." AIMS Med Sci 4 (2017): 151-163.

[13] Bulbul, Halil Ibrahim, and Özkan Unsal. "Comparison of classification techniques used in machine learning as applied on vocational guidance data." 2011 10th International Conference on Machine Learning and Applications and Workshops. Vol. 2. IEEE, 2011.

[14] Mohammed, Aalaa Ahmed, and Yusra Faisal Al-Irhayim. "An overview for assessing a number of systems for estimating age and gender of speakers." Tikrit Journal of Pure Science 26.1 (2021): 94-100.

[15] Chai, Lingying, et al. "Perturbation and nonlinear dynamic analysis of adult male smokers." Journal of voice 25.3 (2011): 342-347.

[16] Díaz, José Antonio, A. Antonio Arroyo, and Howard B. Rothman. "Study and proposal of parameters for the objective assessment of voice quality in smokers." Revista Ingenieria UC 21.3 (2014): 7-16.

[17] Pratama, Timothy, and Ayu Purwarianti. "Topic classification and clustering on Indonesian complaint tweets for bandung government using supervised and unsupervised learning." 2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA). IEEE, 2017.

[18] Sathya, Ramadass, and Annamma Abraham. "Comparison of supervised and unsupervised learning algorithms for pattern classification." International Journal of Advanced Research in Artificial Intelligence 2.2 (2013): 34-38.

[19] Abushariah, Mohammad AM, and Assal AM Alqudah. "Automatic Identity Recognition Using Speech Biometric." European Scientific Journal 12.12 (2016).

[20] Joshi, Krishna Kumar, Neelam Joshi, and Ravi Ray Chaudhari. "Machine Learning–Learning Techniques, CNN, Languages and APIs." International journal of scientific research in computer science, engineering and information technology 6.3 (2020).

[21] Wang, Patrick SP, and Svetlana N. Yanushkevich. "Biometric technologies and applications." Artificial Intelligence and Applications. 2007.

[22] Ramesha, K., et al. "Advanced biometric identification on face, gender and age recognition." 2009 International Conference on Advances in Recent Technologies in Communication and Computing. IEEE, 2009.

[23] Bird, Jordan J., et al. "Accent classification in human speech biometrics for native and non-native english speakers." Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments. 2019.

[24] Akhtar, Mohammad. "Text Independent Biometric Authentication System Based On Voice Recognition." Biom. Authentication (2017): 60.

[25] Shawkat, Asst Lecturer Aws Saad. "Evaluation of Human Voice Biometrics and Frog Bioacoustics Identification Systems Based on Feature Extraction Method and Classifiers." Journal of Al-Ma'moon College 31 (2018).

[26] Pérez-Espinosa, Humberto, et al. "Children age and gender classification based on speech using convnets." Research in Computing Science 147.4 (2018): 23-35.

**ISSN (online): 2581-3048**
**Volume 7, Issue 5, pp 99-107, May-2023**
https://doi.org/10.47001/IRJIET/2023.705011

[27] Saishu, Yuki, Amir Hossein Poorjam, and Mads Græsbøll Christensen. "A CNN-based approach to identification of degradations in speech signals." EURASIP Journal on Audio, Speech, and Music Processing 2021.1 (2021): 1-10.

[28] Submitter, I. C. S. M. D. I., et al. "Gender Recognition and Classification of Speech Signal." Proceedings of the International Conference on Smart Data Intelligence (ICSMDI 2021). 2021.

[29] Majkowski, Andrzej, et al. "Identification of gender based on speech signal." 2019 IEEE 20th International conference on computational problems of electrical engineering (CPEE). IEEE, 2019.

[30] Yıldız, Abdurrahman, and Umut Arıöz. "Analysing human voice and classification of voice frequencies according to smoking effect." 2016 24th Signal Processing and Communication Application Conference (SIU). IEEE, 2016.

[31] Shrivastava, Shilpa. "Biometric: types and its applications." International Journal of Science and Research (IJSR) (2015): 204-207.

**Citation of this Article:**

Ibrahim Khudhur Zaal, Yusra Faisal Mohammad, "Review "Smoker/Non-Smoker Classification of People Using a Speech Signal"" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 5, pp 99-107 May 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.705011

*******