

Year of publication	Number of citations	Type of illness/	Pre-processing of	Feature extraction	Classification/	Database	Origin of the
---------------------	---------------------	------------------	-------------------	--------------------	-----------------	----------	---------------

Author(s)	Year of publication	Number of studies included	Number of disorders investigated	Number of speech signals	Machine Learning	Classification of the subjects	Features used for the classification	Performance detection		
Rikas R.K Sharma	Early Detection of Parkinson's disease through Voice	2014	15 Parkinson's	The signal is split into smaller frames of 15-40 ms duration. Pre-emphasis filter.	The F1T (Fast Fourier Transform) Mel-filter bank block and Discrete Cosine Transform (DCT) used to get the Mel-frequency Coefficients (MFCC). The pitch, formants, jitter, shimmer and glottal APRAT) were parameters also extracted from the samples.	Simple comparison of characteristics extracted from the samples.	21 people, 14 patients with Parkinson's and 7 healthy subjects.	Patients with Parkinson's have: average pitch, formants, jitter, shimmer variations, the pattern for glottal pulse in control subjects is similar and different in patients and there is greater Shimmer for patients achieving an Equal Error Rate (EER) of 20.30%.	No	
Evdaldis Vaciulytė, Arnasas Verikas, Adas Gelzinis, Marga Bacauskė	Detecting Parkinson's disease from sustained phonation and speech signals	2017	44 Parkinson's	Division of recording into short overlapping frames. Pre-emphasis filter and Hamming window function to extract some features.	Divided into global, long-term (low-pass, high-pass, pitch, statistical functions) and temporal and morphological descriptors (e.g. pitch maximum), spectral features, MFCC and PLPCC features, jitter and shimmer variations, pitch period entropy and more.	Random Forest (RF) built the classifier using various feature sets.	63 patients with Parkinson's Disease, 64 control subjects.	Kaunas University of Technology Lithuania	The best feature set was Essentia, with features achieving an Equal Error Rate (EER) of 20.30%.	Yes
Mehmet Bital ER, Emelike İSK, İbrahim İSK	Parkinson's Detection Based on Combined CNN and LSTM Using Enhanced Speech Signals with Variational Mode Decomposition	2021	28 Parkinson's	Framing of the signals of 25ms in length. Hamming window subsequently used. Denoising of the signal performed through Variational Mode Decomposition (VMD).	Mel-spectrogram and Fourier transforms applied (DFT)	Long short-term memory (LSTM) model and CNN are used in a combined form. The Softmax function is applied.	50 patients with Parkinson's and 50 control subjects.	Spanish-speaking PC-GITA dataset.	98.61 % accuracy rate	Yes
Jhon F. Moofarty, Patricia Angulo-Velez, Milton Sarria Paja	Automatic detection of Parkinson's disease from sustained phonation and speech signals	2020	51 Parkinson's	Downsampling to 16 kHz. Pre-emphasized using a first order finite impulse response filter. For MFCC computation specifically, triangular bandpass filters were used and for dynamic features Finite Impulse Response (FIR) filter were used.	MFCC, WIF and MHEC features extracted.	Hidden Markov Model (HMM), Gaussian Mixture Modelling (GMM), Support Vector Machine (SVM) and KNN. Random Cross Validation.	100 patients 50 control and 50 suffering from PD.	Not mentioned	Mean Hilbert Envelope Coefficients (MEHC), WIF and AAM. MFCC features provide a lot of information that can effectively discriminate patients suffering from Parkinson's from control patients.	No
N.P Narendran, Björn Schuller, Pavao Alku	The detection of Parkinson's disease using voice source information	2021	80 Parkinson's	Splitting of raw waveforms into segments of 250ms.	Features characterizing short-term features and long-term features (BBE), MFCC, glottal flow, phonation (jitter, pitch, perturbation, shimmer), prosody duration, fundamental frequency, energy). Use of Iterative Adaptive Shrinkage Filtering (IAF).	On the one hand SVM was used to study the classical pipeline approach, and on the other hand a combination of CNNs and MLP was developed for the end-to-end approach.	50 patients with Parkinson's and 50 control subjects.	Spanish-speaking PC-GITA dataset.	The best detection performance was achieved through the use of end-to-end systems with QCP-based glottal flow. 68% accuracy rate.	Yes
Rimmi Haavio, James Glass	Classifying Alzheimer's Disease Using Audio and Text-Based Representations of Speech	2021	60 Alzheimer's (AD)	Volume normalisation and splitting of signals with a duration of 10s each.	Text Features, CLAN features extracted by means of the Computerized Language Analysis (CLAN). BERT embeddings and Linguistic Inquiry and Word Count (LIWC) overall creating word vectors.	Linear Discriminant Analysis (LDA), Decision Tree (DT), K-nearest Neighbour (1) (ANN), SVM, RF, LSTM, CNN. These classifiers were then evaluated with Leave-one-subject-out (LOSO) cross-validation.	78 subjects with Alzheimer's and 78 control.	Not mentioned	SVM and RF classifiers trained on BERT embeddings and CLAN features achieved the highest accuracy of 85.4%. Using a PCA dimensionality reduction can also improve the classification tasks.	Yes
Evgenia Goldshteyn, Tieg Dang, Cecilia Mascolo	Automatic Detection of Obstructive Sleep Apnea Using Speech Signals	2011	35 Sleep Apnea (OSA)	Segmentation of speech signals.	Two different feature sets: short-term features and long-term features. All model coefficients and MFCCs, ARMA models, all through the Sequential Feature Selector (SFS) procedure.	8th order generalised method of moments (GMM) for males, and 4th order GMM for females. The classification error was later calculated with LOOCV.	67 subjects with sleep apnea and 26 control subjects.	Not mentioned	Specificity of 83% and 75% for male speakers and 86% and 94% for female speakers.	Yes
M.Kiran Reddy, Pury Heikkula, Y.Madhu, Keerthana, Kasimil Keki, Mikko Minkkinen, Heili Toipponen, Tarmo Nieminen, Pavao Alku	The automatic detection of heart using speech signals	2021	36 Heart failure (HF)	Linear phase FIR filter (cut-off frequency at 60Hz) was used to remove low-frequency noise, segmentation to non-overlapping files of 4s.	Glottal features extracted using glottal inverse filtering. MFCC features were extracted, and the Voicebox features, statistical measures.	A combination of the glottal and MFCC features were used to classify. The classifiers employed are : SVM, ET, RF (RF variant), AdaBoost, FNN	25 healthy patients and 20 patients suffering heart failure.	Finnish-speaking dataset.	The FFNN classifier trained using both glottal and MFCC features was the most successful at classifying patients suffering from heart failure with an accuracy of 81.5%.	Yes
Vasim Vakhichev, Elias Azarov, Aleksandra Petrosky, Yuliyia Rukhovich	Features extraction for the automatic detection of ALS disease from acoustic speech signal	todo	13 Amyotrophic Lateral Sclerosis (ALS)	Segments of speech signals containing vowels selected.	Different vowels are analysed to obtain the characteristics. There are harmonic features and format features such as the normal mutual formant location to form.	LDA	48 people, 26 patients suffering ALS, 22 control patients.	Russian-speaking dataset.	Accuracy of 88% in ALS detection.	Yes
Carlo Robotti, Giovanni Costantini, Giovanni Saggio, Valerio Cesami	Machine Learning-based Voice Assessment for the Detection of Positive and Recurrent COVID-19 patients	2021	104 COVID-19	Root Mean Square normalisation applied.	6373 unidimensional features extracted by using the configuration file of the INTERSPEECH2016.	The classification process was done using SVM, the Radial Basis SVM. The accuracy of the training set was calculated through a 10-fold cross-validation.	70 patients with COVID-19 70 recorded negative COVID-19 patients and 70 control subjects.	Conducted at three Italian COVID-19 units and healthy individuals recruited among hospital staff members.	90.24% accuracy in COVID-19 detection.	Yes
Tong Xia, Jing Han, Lorena Gendro, Ting Dang, Cecilia Mascolo	Uncertainty-Aware COVID-19 Detection from Imbalanced Sound Data	2021	35 COVID-19	Resampling of recordings to 16kHz, removal of silence at the start and end of recordings, audio normalisation establishing peak amplitude to 1.	The features extracted come from voice, cough and breathing and recordings and are normalisation into a vector. Mel spectrums are extracted from each recording.	10% and 20% of the database is used for validation and testing sets respectively. SVM was used to perform the classification.	469 positive COVID-19 patients and 1526 control subjects.	Crowdsourced data from a mobile app.	68% sensitivity and 69% specificity.	Yes
Gauri Deshpande, Björn W. Schuller	An Overview on Audio, Signal, Speech & Language Processing for COVID-19	2020	32 COVID-19	Highest accuracy achieved through this method in the report, although pre-processing techniques not mentioned.	STFT, MFCC and MFB acoustic features.	Convolutional Neural Network (CNN)	Google voice audio signals from 1.8 million Youtube videos.	Multiple	95% accuracy	Yes
Tusar Karil Dash, Soumya Mahna, Nishant Pandya, Suresh Chandra Satapathy	Detection of COVID-19 from speech signal using bio-inspired cepstral features	2021	39 COVID-19	Framing and windowing of the input signal. Downsampling to the signal to 16kHz.	Extraction of MFCCs, cepstral analysis carried out optimised by reducing feature space using Linear Discriminant Analysis Fisher's F-ratio to use maximum and minimum frequencies of MFCC and cepstral variance normalisation. The specific features are Triangular filter bank using mel-scale (TFBCC-M features), triangular filter bank using bark scale (TFBCC-B features), Triangular filter bank using human factor scale (TFBCC-H features) and Triangular filter bank using ERB scale (TFBCC-E features).	Support Vector Machine (SVM)	Crowdsourced from Cambridge University.	Volunteers from the United Kingdom.	96.85% Yes	
Mahmoud Al Iemali, Shadi Deshmukh, Rita Singh	Detection of Covid-19 through the analysis of vocal folds	2021	15 COVID-19	Sampled at 8kHz	Two studies performed, using either parameters derived from gamma, beta and alpha features, such as the horizontal displacements of the left and right vocal folds(x and y) or the features themselves	SVM, Decision Tree, Logistic Regression and AdaBoost.	512 patients, of which some where Covid-19 positive	Merlin Inc. private firm in Chile	Vocal fold oscillations are the most useful in positive patients, and its exclusiveness in relation to other voice pathologies is not studied in the paper.	No
Diadropovic, Muhammed İsmail, Maël Comel, noudiaa Mc Call, Guy Fagheraz	Detection of Covid-19 from voice, cough and breathing patterns: a novel floating point pre-training pattern	2021	50 COVID-19	Conversion of audio files to WAV format with 44.1 kHz sampling rate and 32-bit floating point. Cough recordings were segmented into individual cough signals and noise reduction was applied. Audacity v2.4.2 used to remove background noise using the spectral noise gating.	Use of standard acoustic feature sets: Geneva Minimalistic Acoustic Parameter Set (GeMAPS) selected according to their relevance to physiological and psychological voice production. Further use of GeMAPS and ComParE. Use of feature engineering transform to extract wavelet scattering features.	Boosting, Random Forests, Bagging, MLP.	Crowdsourced, with 84 positive patients and 1019 control subjects.	Multiple countries such as : Luxembourg, Serbia, France, Germany. Multiple genders and age ranges.	88.52% accuracy with 88.75% sensitivity and 90.87% specificity	Yes
Jing Han, Kun Qian, Meishu Song, Ziliang Yang, Zhao Ren, Shuo Liu, Huaiyuan Zheng, wei Ji, Tongmei Kokei, Xian Li, Zeng Zhang, Yoshiharu Yamamoto, Björn W.Schuller	An Early Study on Intelligent Analysis of Speech under COVID-19: Severity, Sleep Quality, Fatigue and Anxiety	2020	32 COVID-19	260 audio segments were converted to mono signals with a 16kHz sampling rate.	Computational Paralinguistics Challenge (COMPAR-E) set and the extended Geneva Minimalistic Acoustic Parameter Set (GeMAPS) extracted with the openSMILE kit.	SVM	52 Covid-19 patients from two Wuhan Hospitals.	Wuhan, China	Lack of control group implies that there cannot be enough data, further data is needed. Nevertheless, the severity of the disease may be estimated through the approach proposed in the paper.	No
Kawthar A. Al-Dhian	An adaptive speech signal processing for Covid-19 disease detection using deep learning approach	2020	25 COVID-19	Least-mean-Square (LMS) filtering method used for noise removal	Generative Adversarial Network (GAN).The filtered signal is directly analysed by the GAN classifier whereby the Covid threshold lies above 1.245 and a non-Covid patient lies below 0.60Hz.	Not mentioned, but both control subjects and patients suffering from Covid-19 are present.	Not mentioned	98.56% accuracy	Yes	
Ez-Eddine Hachani, Walid El-Shafai, Amged Sayed	CR19: a framework for automatic detection of COVID-19 in cough audio signals using machine learning algorithms for automated medical diagnosis applications.	2022	35 COVID-19	Pre-processing stage takes place but no specific form of it mentioned.	Fourier transform for cepstral coefficients (MFCC) for feature extraction. Multidimensional spectral and temporal characteristics are obtained.	Coswara data provided by the Indian Institute of Science (IISc).	India	90.78% accuracy	Yes	
Abdelfattah Hassan, Masoud, Danish Raza Rizvi, Mohamed Bader Alsabek	Covid-19 Detection System Using Recurrent Neural Networks	2022	29 COVID-19	PRAAT Pre-System used for noise reduction.	Features extracted include: kurtosis, Spectral Centroid (SC), Spectral Roll-off (SR), Zero Crossing Rate (ZCR), Mel-Frequency Cepstral Coefficients through logarithmic compression and Discrete Cosine Transform (DCT), from which the list and Second Derivatives of the MFCC were later computed.	LSTM	60 Control subjects and 19 Covid-19 patients.	United Arab Emirates Hospitals	98.2% accuracy in results found with the SVM algorithm with a 97% for coughing sounds and 88.2% for voices, with MFCCs features being the most precise in classification.	Yes
Madhuranada Pahan, Marisa Kopper, Robin Warren, Thomas Nesler	COVID-19 detection by cough, breath and speech using deep transfer learning and bottleneck features	2021	47 COVID-19	Downsampling to 16kHz, a filter applied to create a 20 ms Hamming window and a filter of pre-emphasis (0.95) is applied.	Extracted features include: kurtosis, mel-frequency cepstral coefficients (MFCCs) and linearly-spaced filter bank energies.	CNN, LSTM and Resnet50, then k-fold cross validation applied.	Overall 11 202 recorded cough sounds of patients suffering from Covid-19	6 datasets from multiple origins, one of which is specifically collected from a tuberculosis (TB) clinic near Cape Town, South Africa.	98.2% accuracy	Yes
Kotra Venkata Sai Ravi, Shareef Sibin Kalluri, Deepa Vijayanen	Covid-19 patient detection from telephone quality speech data	2020	17 COVID-19	No pre-processing activity specified	Fundamental frequency, short-term cepstrum, cepstral peak prominence, Harmonics to noise Ratio and short term mel-spectrum as low level features.	SVM	19 phone speakers, extracted from youtube videos, from which 10 are Covid-19 positive and 9 are control subjects.	Non-specific	92.7% Yes	
Chloé Brown, Jing Han, Tong Xia, Fátima Mangone Chauhan, Apinan Huathansamont, Pietra Cicuta, Andreas Grammenos, Dimitris Spathis, Cecilia Mascolo	Exploring Automatic Detection of COVID-19 from Crowdsourced Respiratory Sound Data	2021	30 COVID-19	Resampling of audio files to 22kHz, from which features are extracted at a segment and frame level.	Features extracted comprise: Temporal, Period, RMS, Spectral Centroid, Roll-off, Frequency, Zero-crossing, MFCC and derivatives from which further statistical features were extracted (mean, median, root-mean).	Logistic Regression (LR), Gradient Boosting (GB), Support Vector Machines (SVM) with a Radial Basis Function (RBF) kernel.	141 cough and breathing sounds, and 3 control groups: the first providing 296 samples, the second providing 32 and the third 20 samples.	Non-specific	80% accuracy distinguishing healthy users from Covid-19 patients	Yes
Laiba Verkh, Rusej de Pietro Alured Ghoneim, Mubarak Alirahouf, Haled N. Al-Kutbi, Giovanna Sannino	Exploring the Use of Artificial Intelligence Techniques to Detect the Presence of Coronavirus Covid-19 Through Speech and Voice Analysis	2021	61 COVID-19	Sampling at 44.1 kHz, a filter applied to create a 20 ms Hamming window and a filter of pre-emphasis (0.95) is applied.	Parameters extracted include: kurtosis, Fundamental Frequency, Jitter, shimmer and Harmonic to Noise Ratio (HNR) as well as Mel-frequency Cepstral Coefficients (MFCC) or Spectral Centroid or Roll-Off	Naive Bayes, SVM, K-Nearest Neighbour, AdaBoost, Decision Table and more.	The database comprises 229 control female subjects, and 721 male subjects as well as 23 female subjects with Covid-19 as well as 775 males with Covid-19.	Asia, Europe, America and Oceania	Best accuracy results found with the SVM algorithm with a 97%.	Yes
J.R Orozco-Arroyave, F. Hönig, J.D Arias-Londoño, J.F Vargas-Bonilla	Automatic detection of Parkinson's disease in running speech spoken in three different languages	2016	54 Parkinson's	Manual removal of silence in voice recordings at end and beginning. Windowing of speech frames of 40ms.	Feature extraction consisted of creating a set of 17 features including cepstral parameters and three measures of noise content: HNR, normalised noise energy, offset to noise excitation ratio.	SVM	The Spanish dataset included 50 Parkinson's and 50 control subjects, the German dataset consisted of 176 people, half control half patients, and finally 36 Czech native speakers of which 20 are diagnosed PD.	Spain, Germany, Czech Republic	Accuracy ranging from 85% to 99% depending on the language and speech task	Yes
Isra Nissar, Safar Al-Rasoud, Danish Raza Rizvi, Agni Mr. Agni Mr.	Voice-Based Detection of Parkinson's Disease through Ensemble Machine Learning Approach: A Performance Study	2018	28 Parkinson's	Data normalisation	MFCC using RFE and feature selection. Feature reduction is also performed.	9 Machine Learning Algorithms tested which include: Naive Bayes, SVM (linear and RBF kernel), logistic regression, k-nearest neighbour, Extreme Gradient Boost, decision tree.	188 Patients suffering from Parkinson's disease and 64 control subjects.	Non-specific	Highest accuracy achieved with the XGBoost classification technique with an 88.15%	Yes
T.Vita-Cañas, J.R Orozco-Arroyave, J.F Vargas-Bonilla, J.D Arias-Londoño	Modulation Spectra for Automatic Detection of Parkinson's Disease	2014	27 Parkinson's	None mentioned	Feature extraction for Automatic include PCA and LDA from the Signal Fast Fourier, through spectra analysis using MS and dynamic features (centroids and energy)	SVM	50 patients with Parkinson's and 50 control subjects.	Medellin Colombia	71.67% accuracy	Yes
Laetitia Jeancolas, habib Benali, Badr-Edine Berkefelat, Grazia Eddine Jean-Christophe Connot, Marie Vidalhet, Stéphane Lehenry, Djiana Patrovska Delacretaz	Automatic Detection of Early Stages of Parkinson's Disease through Quantitative Acoustic Voice Analysis with Mel-Frequency Cepstral Coefficients	2017	30 Parkinson's	Pre-amplification in the external sound card, and speech sampled at 96000 Hz, no subsequent pre-processing mentioned	MFCC feature extraction carried out with the Matlab software, followed by the Fast Fourier Transform and finally the Mel-Spaced filterbank.	Gaussian Mixture Model (GMM)	74 subjects of which 40 suffered from Parkinson's Disease.	Più-Salpêtrière Hospital Paris, France.	91% accuracy	Yes
E.O.A. Belacazar-Bolaños, J.R Orozco-Arroyave, J.D Arias-Londoño, J.F Vargas-Bonilla, E. Noh	Automatic Detection of Parkinson's Disease Using Means of Class-Specific Measures of Speech	2013	15 Parkinson's	Preprocessing by windowing of 40ms and 20ms overlap of voice frames	Harmonics to Noise Ratio, Cepstral Harmonics to Noise Ratio, Normalised Noise Energy and Glottal Excitation Ratio are extracted through Principal Component Analysis (PCA)	K nearest Neighbour (KNN)	50 patients with Parkinson's and 50 control subjects.	Medellin Colombia	62.29% Yes	
Laetitia Jeancolas, habib Benali, Badr-Edine Berkefelat, Grazia Eddine Jean-Christophe Connot, Marie Vidalhet, Stéphane Lehenry, Habib Benali	X-Vectors: New framework for automatic classification of Parkinson's Disease through Ensemble Machine Learning Approach: A Performance Study	2020	67 Parkinson's	Denoising through application of the Praat software in microphone recordings.	Kaldi Software used to extract overlapping windows of 20ms and 19 MFCCs. The first and second derivatives of the deltas were then calculated and added to the feature vectors named X-vectors.	GMM classification technique (VO) used for classification.	There were microphone and telephone recordings. For the microphone recordings there were 206 patients with Parkinson's and 91 control subjects. For the telephone recordings there were 10 Parkinson's patients and 61 control subjects.	Più-Salpêtrière Hospital Paris, France.	X-vectors are the optimal procedure for early Parkinson's detection with an increase in women of the 7.5%.	Yes
Kumara Shama, Anjana Krishna, Niranjani U.Chettyaya	Harmonic-to-Noise Ratio and Critical-Band Energy Spectrum of Speech and Acoustic Indicators of Laryngeal and Voice Pathology	2006	35 Laryngeal disorders	Segmentation into overlapping frames	Two sets of features are extracted: energy ratio of harmonics to noise components (HNR) and the energy spectrum at critical-band spacing.	K nearest Neighbour (KNN) classifier.	53 normal and 165 pathological voices.	Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Lab	Accuracy of 94.28% in HNR features and 92.38 in spectrum energy features.	Yes
J.R Orozco-Arroyave, E.A Belacazar-Bolaños, J.D Arias-Londoño, J.F Vargas-Bonilla, S. Rodol, J. Ruiz, K. Daegrov, F.Hönig, E. Noh	Characterisation Methods for the Detection of Multiple Voice Disorders: Neurological, Functional and Acoustic	2015	62 Laryngeal disorders	Windowing of 40ms and 20ms shift.	Features extracted include: noise measures (HNR, Voice Turbulence Index (VTI), Self Phonation Index (SPI), periodicity and stability of voice, spectral cepstral modeling and non-linear behavior. Largest Lyapunov Exponent (LLE), Lempel-Ziv Complexity (LZC), modelled Period Density Fluctuation (PDF) and the Determined Fluctuation Analysis (DFA).	Support Vector Machine (SVM).	Two databases were used in the investigation: 173 patients suffering from laryngeal disorders and the second database was composed of 200 patients and a control group of 199 subjects.	The first database pertained to the Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Lab and the second database pertained to the Universidad Politécnica de Madrid	Results with periodicity and stability features were the most suitable for discrimination between pathological patients and healthy ones. Accuracy values range around 96% for the multiple classified features.	Yes
Stefan Hadjilodorov, Boyan Boyanov, Bernard Testolin	Laryngeal Pathology by Means of Class-Specific Neural Maps	2000	20 Laryngeal disorders	No pre-processing of the signals mentioned	Features include: Pitch Period and its derivatives, stability of pitch period generation (STAB), the degree of dissimilarity of the shape (DISS), Noise-to-harmonics ratio (NHR), Low-to-high Harmonics ratio (LHER), and ratio-to-total cepstral energy, which are all included as components in a vector for classification.	KNN, Linear Discriminant Analysis (LDA), Self-organising map (SOM)	100 control subjects and 300 patients suffering from laryngeal disorders.	Phoniatric Department of the University Hospital in Sofia, Bulgaria	95.1% accuracy	Yes
Raissa Tavares, Raissa Correia, Silvana C. Costa, Benedito G. Aguiar Neto, Joseana Macedo Feschine	Optimising the Acoustic Signal for Voice Pathology Detection by using combined cepstral features	2010	21 Laryngeal disorders	Speech signals are processed by a 20 ms Hamming window with a 50% overlap and a filter of pre-emphasis (0.95) is applied.	The features extracted include: cepstral coefficients (CEP), weighted cepstral coefficients (WCEP), delta cepstral (DCEP), and weighted delta cepstral (WDCEP) parameters.	Vector quantisation technique (VQ) used for classification.	Out of 152 subjects, 44 patients with vocal fold edema, 55 cases of paralysis and the rest is the control group.	Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Lab.	98% accuracy rate between control subjects and edema -suffering patients and 96% between control group and paralysis.	Yes
Ali Akbar, Meisam Khalil Arjmandi	An efficient voice pathology classification scheme based on applying multi-layer linear discriminant analysis to wavelet packet-based features	2013	47 Laryngeal disorders	Only the sampling frequency of 44.1 kHz mentioned	Short-time Fourier transform (STFT) for feature extraction using energy features and non-linear behavior. Largest Lyapunov Exponent (LLE), Lempel-Ziv Complexity (LZC), modelled Period Density Fluctuation (PDF) and the Determined Fluctuation Analysis (DFA).	Linear discriminant analysis (LDA) used for classification optimisation followed by a multilayer neural network for classification.	53 control group and 205 patients.	Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Lab.	91% accuracy rate with SVM	Yes
Neel B. Pinto, Ingo R. Titz	Unification of perturbation measures in speech signals	1990	26 Laryngeal disorders	None mentioned	Perturbation measures (HNR, ZCR, rms, MEEI) through phonation parameters (amplitude and fundamental frequency)	Systematic but not automatic classification of patients done through sinusoid, Gaussian noise and linear trend modulations.	Database size not specified	Voice Acoustics and Speech Laboratory	Unified framework for perturbation analysis and measurement offered.	No
Marcelo de Oliveira Rios, José Carlos Pereira, Marcelo	Adaptive Estimation of Residual Signal for Voice Pathology Diagnosis	2000	43 Laryngeal disorders	Adaptive filtering is applied to create a residual signal from which features are extracted using the Square Residue (MSR), Excess Coefficient (EX), Spectral Slope of All Filter (SFA), JITTER, PEAK, peak of autocorrelation	Spectral and cepstral characteristics were calculated: mean Square Residue (MSR), Excess Coefficient (EX), Spectral Slope of All Filter (SFA), JITTER, PEAK, peak of autocorrelation	Statistical method of the Mann-Whitney test to evaluate the features.	73 speakers divided into 48 dysphonic and 25 controls.	Clinical Hospital of the Faculty of Medicine at Ribeirão Preto, Brazil	Greater accuracy of 94.78% achieved through Jitter feature.	Yes
Ghazaleh Yazini, Farhad Almazang, Roozbeh	Pathological Assessment of speech signals using nonlinear dynamical analysis	2010	40 Laryngeal disorders	Only mentions assessment of frequency of 44.1kHz	Phase space features and non-linear correlation measures (dimension (DD), largest Lyapunov Exponent (LLE), approximate entropy (ApEn), Ziv-Lempel complexity (ZL) and fractal dimension (FD).	SVM	1400 voice samples from 700 subjects	Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Lab.	Lyapunov exponent features provide a 94.44% accuracy rate.	No
Tadanao Araújo de Sousa, Vinícius J.D Vieira	Feature selection based on binary particle swarm optimisation and neural networks for pathological voice detection	2018	29 Laryngeal disorders	None mentioned	Two-dimensional and two-dimensional wavelet transform, Haralick texture features, PSD for feature selection.	MLP with k-fold cross validation	53 signals of a healthy larynx and 112 signals of a pathological larynx	Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Lab.	6% increase in accuracy through PSD optimisation to a 91%	Yes
Shih-Hao Fang, Yu Tao, Ji-Ying Chen, Ying-Hui Lai, Feng-Chuan Lin, Chi-Te Wang	Detection of Pathological Voice Using C-Expform Features: A Deep Learning Approach	2018	32 Laryngeal disorders	44,100 Hz sampling rate with sound amplification.	Mel frequency Cepstral coefficients derived from fast Fourier Transform, windowed by 40ms.	Deep Neural Network (DNN), SVM and Gaussian Model	402 pathological voices and 60 control subjects	Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech Lab.	DNN provides highest accuracy with 94.26%	Yes
R. Fraile, N. Sotelo Lechón, J.I Godino-Lorente, V.Omar-Ruiz	Automatic Detection of Laryngeal Pathologies in Records of Sustained Vowels by Means of Mel-Frequency Cepstral Coefficients Parameters and Differentiation of Parameters by Sex	2009	36 Laryngeal disorders	Sampling rate of 25kHz	Mel-frequency Cepstral coefficients derived from Fast Fourier Transform.	MLP classifier	53 control subjects and 173 pathological patients.	Speech record database commercialised by the American company Kay Elemetrics	Female voices have wider distributions so cepstral classification is more complicated to achieve.	No