The American Journal of Engineering and Technology
ISSN 2689-0984 Volume 08 - 2026

Efficient Affective State Identification in Vocal Signals through Machine Learning-Based Neural Frameworks

Dr. Litia K. Amaru
School of Engineering and Technology University of the South Pacific (Marshall Islands Campus) Majuro, Marshall Islands

Received: 15 Oct 2025 | Received Revised Version: 01 Nov 2025 | Accepted: 15 Nov 2025 | Published: 30 Nov 2025

Volume 07 Issue 11 2025

Abstract

Efficient identification of affective states from vocal signals remains a critical challenge in biomedical signal processing and computational intelligence due to the inherent variability, noise sensitivity, and non-stationary characteristics of speech data. This study proposes a machine learning-based neural framework that integrates wavelet packet decomposition, adaptive feature optimization, and hybrid classification models for robust emotional and pathological voice state recognition. The system leverages multiresolution analysis to extract discriminative acoustic features while employing optimization-driven feature selection strategies inspired by evolutionary computation and neural adaptation principles.

The proposed framework is grounded in signal decomposition techniques such as lifting wavelet transforms and wavelet packet representations, which enable efficient localization of temporal and spectral speech characteristics. Feature engineering is further enhanced through statistical descriptors including Mel-frequency cepstral coefficients, jitter measures, and complexity-based acoustic parameters. These features are subsequently processed using machine learning classifiers such as support vector machines, linear discriminant analysis, and hybrid neural architectures to achieve high classification accuracy.

The methodology is informed by prior research on pathological voice classification and adaptive signal processing, where wavelet-based feature extraction and genetic algorithm optimization have demonstrated strong performance in distinguishing subtle variations in vocal patterns (Ariased-Londono et al.; Saidi & Almasganj). Additionally, the study incorporates biologically inspired computational principles that align with neural adaptation mechanisms described in computational neuroscience literature (Doya, 1999), enhancing interpretability and adaptive learning capacity.

Experimental design considerations highlight robustness across noisy and heterogeneous speech datasets, particularly using benchmark corpora such as the Disordered Voice Database. The proposed framework emphasizes computational efficiency while maintaining high sensitivity in affective state classification tasks.

The findings indicate that hybrid machine learning models combining wavelet-based feature extraction with neural optimization significantly outperform conventional statistical classifiers in terms of accuracy, robustness, and generalization capability. The study contributes to advancing affective computing systems by providing a scalable and interpretable framework for vocal emotion and pathology detection, with applications in healthcare diagnostics, human–computer interaction, and intelligent assistive systems (Anoop et al., 2018).

Keywords: Affective computing, voice signal processing, wavelet packet transform, support vector machines, neural networks, emotional speech classification, feature optimization, biomedical signal analysis

Cite This Article: Amaru, D. L. K. (2025). Efficient Affective State Identification in Vocal Signals through Machine Learning-Based Neural Frameworks. The American Journal of Engineering and Technology, 7(11), 256–264. Retrieved from https://theamericanjournals.com/index.php/tajet/article/view/8005

1. Introduction

The analysis of human vocal signals for affective state identification has become a central research area in computational intelligence, biomedical engineering, and human–computer interaction systems. Vocal signals carry rich paralinguistic information that reflects not only linguistic content but also emotional, physiological, and pathological states of the speaker. Efficient extraction and classification of these affective states are essential for applications such as clinical diagnosis of voice disorders, mental health monitoring, adaptive communication systems, and intelligent virtual assistants.

Traditional speech processing systems primarily focused on linguistic content recognition, often neglecting the expressive and emotional dimensions embedded in acoustic patterns. However, recent advancements in machine learning and neural computation have enabled the development of systems capable of interpreting subtle variations in pitch, intensity, spectral distribution, and temporal dynamics. These developments have significantly expanded the scope of speech processing from purely linguistic recognition to affective state analysis.

Despite progress, accurate classification of vocal affective states remains challenging due to several inherent limitations. Speech signals are highly non-stationary and susceptible to environmental noise, speaker variability, and recording conditions. Furthermore, emotional expressions in speech are often ambiguous and overlapping, making it difficult to define clear decision boundaries between affective categories. These challenges necessitate the development of robust feature extraction and classification frameworks capable of handling high-dimensional and noisy data.

Wavelet-based signal processing techniques have emerged as a powerful tool for addressing these challenges. Unlike Fourier-based methods, wavelet transforms provide localized time-frequency representations that are well-suited for analyzing transient and non-stationary speech signals. Adaptive wavelet frameworks, such as lifting schemes, further enhance computational efficiency and flexibility in feature extraction (Sweldens; Claypoole et al.). These methods have been widely applied in pathological voice classification, demonstrating strong capability in capturing subtle vocal irregularities.

In parallel, machine learning approaches such as support vector machines and neural networks have demonstrated significant success in pattern recognition tasks involving high-dimensional biomedical data. Support vector machines, in particular, are effective in handling nonlinear classification problems by constructing optimal hyperplanes in transformed feature spaces (Cortes & Vapnik). Neural network models, on the other hand, provide adaptive learning mechanisms capable of modeling complex nonlinear relationships between acoustic features and affective states.

The integration of wavelet-based feature extraction with machine learning classifiers offers a promising hybrid approach for affective speech analysis. This combination allows for efficient representation of speech signals while maintaining strong classification performance. Additionally, optimization techniques such as genetic algorithms further enhance system performance by selecting the most discriminative feature subsets, thereby reducing dimensionality and improving generalization capability.

The relevance of this research extends to biomedical applications, particularly in the detection of voice pathologies and neurological disorders. Studies utilizing datasets such as the Disordered Voice Database have demonstrated that acoustic variations can serve as reliable indicators of vocal fold abnormalities and neurological impairments. These findings highlight the importance of developing computational frameworks that can effectively distinguish between normal and pathological speech patterns.

The objective of this study is to develop a computationally efficient and robust framework for

affective state identification in vocal signals using machine learning-based neural architectures. The proposed system aims to integrate wavelet-based feature extraction, optimization-driven feature selection, and hybrid classification models to improve accuracy and robustness across diverse speech datasets.

The scope of this research encompasses both emotional and pathological voice classification, thereby bridging the gap between affective computing and biomedical signal analysis. The significance of this study lies in its potential to enhance diagnostic accuracy in clinical applications while also improving user experience in intelligent communication systems.

This study further aligns with prior research on neural computation and speech emotion recognition, which demonstrates the effectiveness of artificial neural networks in modeling complex acoustic-emotional relationships (Anoop et al., 2018). By building upon these foundational works, the present research contributes toward the development of scalable, adaptive, and interpretable affective computing systems.

2. Literature Review

Research on affective state identification in vocal signals has evolved through multiple interdisciplinary domains, including signal processing, machine learning, biomedical engineering, and computational neuroscience. Early approaches primarily relied on handcrafted acoustic features and statistical classification techniques, while recent studies emphasize hybrid frameworks combining wavelet transforms, evolutionary optimization, and neural networks.

Wavelet-based analysis forms a foundational pillar in modern speech signal processing. Sweldens introduced the lifting scheme as an efficient framework for constructing biorthogonal wavelets, significantly reducing computational complexity while maintaining perfect reconstruction properties. This advancement enabled adaptive wavelet design tailored to non-stationary biomedical signals (Sweldens). Similarly, Claypoole et al. extended adaptive wavelet transforms through lifting-based optimization, demonstrating improved signal representation in time-varying environments. These contributions are particularly relevant for speech signals, where transient acoustic variations carry critical affective information.

In biomedical voice analysis, wavelet packet decomposition has been extensively used for pathological classification tasks. Arias-Londoño et al. demonstrated that combining complexity measures with Mel-frequency cepstral coefficients significantly improves discrimination between healthy and disordered voices. Their study highlighted the importance of multi-domain feature integration, particularly in noisy clinical environments. Similarly, Akbari and Arjmandi applied multi-layer linear discriminant analysis to wavelet packet features, achieving improved classification accuracy in voice pathology detection tasks. These studies collectively emphasize the effectiveness of wavelet-based representations in capturing fine-grained vocal abnormalities.

Feature optimization has also played a crucial role in enhancing classification performance. Genetic algorithms have been widely used to select optimal feature subsets from high-dimensional acoustic representations. Behroozmand and Almasganj demonstrated that genetic algorithm-based feature selection significantly improves classification accuracy in pathological speech analysis by eliminating redundant and noisy features. Carr and Samanta further established genetic algorithms as robust optimization tools for feature selection in complex machine learning pipelines.

Support vector machines (SVMs) have been extensively applied to speech classification problems due to their strong theoretical foundation in statistical learning theory. Cortes and Vapnik introduced SVMs as maximum-margin classifiers capable of handling nonlinear separability through kernel functions. In voice pathology detection, Saidi and Almasganj showed that wavelet-adapted SVM frameworks outperform traditional classifiers, particularly in small-sample biomedical datasets. This demonstrates the suitability of SVMs for high-dimensional acoustic feature spaces.

Neural network-based approaches further enhance classification performance by enabling nonlinear mapping between acoustic features and affective states. Fezari et al. demonstrated that adaptive neural classifiers outperform traditional statistical methods in voice disorder detection. Majidnezhad further integrated genetic algorithms with artificial neural networks, resulting in improved diagnostic accuracy and feature adaptability. These studies highlight the importance of hybrid neural optimization frameworks in affective speech processing.

Additional research emphasizes feature robustness through jitter estimation and LPC spectrum analysis.

Silva et al. showed that jitter-based features provide valuable information for detecting pathological voices. Cordeiro et al. demonstrated that LPC spectrum peak analysis can effectively capture vocal fold irregularities. These features complement wavelet-based representations by providing complementary spectral and temporal information.

Biomedical datasets such as the Disordered Voice Database have played a significant role in benchmarking classification systems. These datasets provide annotated pathological and normal speech samples, enabling standardized evaluation of computational models. Studies using these datasets consistently demonstrate that hybrid models combining wavelet features, SVMs, and neural networks outperform isolated feature-classifier combinations.

Despite significant progress, several research gaps remain. First, many existing systems rely on static feature extraction methods that fail to capture dynamic temporal variations in speech emotion. Second, computational complexity remains a concern in real-time applications, particularly when using multi-layer wavelet decomposition combined with evolutionary optimization. Third, generalization across datasets remains limited due to variability in recording conditions, speaker demographics, and emotional expression styles.

The theoretical positioning of this study lies at the intersection of wavelet-based signal processing, machine learning classification, and neural optimization frameworks. It extends prior work by integrating adaptive feature extraction with biologically inspired neural computation principles, as suggested in computational neuroscience literature (Doya). Additionally, the inclusion of artificial neural network-based emotional modeling aligns with prior findings in speech emotion recognition systems (Anoop et al., 2018), reinforcing the validity of neural architectures in affective computing.

Overall, the literature indicates a clear transition from handcrafted feature-based systems to hybrid machine learning frameworks that integrate signal decomposition, feature optimization, and neural classification. However, there remains a need for unified architectures that balance computational efficiency, robustness, and interpretability in real-world affective speech applications.

3. Methodology

3.1 System Overview

The proposed methodology introduces a hybrid computational framework for affective state identification in vocal signals. The system integrates wavelet packet-based feature extraction, adaptive optimization, and machine learning classifiers to achieve robust emotional and pathological voice classification. The architecture consists of four primary stages: preprocessing, feature extraction, feature optimization, and classification.

Let the input speech signal be denoted as S(t)S(t)S(t). The objective is to learn a mapping:

f:S(t)→Cf: S(t) \rightarrow Cf:S(t)→C

where CCC represents affective or pathological class labels.

3.2 Signal Preprocessing

3.2.1 Noise Reduction

Speech signals are first denoised using wavelet lifting-based denoising techniques. The signal is decomposed into approximation and detail coefficients:

S(t)=Aj+∑j=1nDjS(t) = A_j + \sum_{j=1}^{n} D_jS(t)=Aj+j=1∑nDj

Thresholding is applied to detail coefficients to remove noise components while preserving relevant speech structures (Gavrovska et al.).

3.2.2 Normalization

Amplitude normalization ensures consistency across recordings:

Snorm(t)=S(t)max|S(t)|S_{norm}(t) = \frac{S(t)}{\max|S(t)|}Snorm(t)=max|S(t)|S(t)

This step reduces speaker-dependent variability.

3.3 Feature Extraction

3.3.1 Wavelet Packet Decomposition

Wavelet packet decomposition is used to represent speech signals across multiple frequency bands. Each node in the decomposition tree captures localized spectral information.

Wj,k(t)=∫S(t)ψj,k(t)dtW_{j,k}(t) = \int S(t)\psi_{j,k}(t)

dtWj,k(t)=∫S(t)ψj,k(t)dt

This representation enhances sensitivity to subtle vocal variations.

3.3.2 Acoustic Feature Set

The extracted feature vector includes:

- Mel-frequency cepstral coefficients (MFCCs)
- Linear predictive coding (LPC) coefficients
- Jitter and shimmer measures
- Complexity-based features

These features capture both physiological and acoustic characteristics of vocal signals.

3.3.3 Statistical Feature Representation

Mean, variance, skewness, and entropy are computed over wavelet coefficients:

E=−∑p(x)logp(x)E = - \sum p(x) \log p(x)E=−∑p(x)logp(x)

Entropy captures signal irregularity associated with pathological or emotional variation.

5.4 Feature Selection and Optimization

Genetic algorithms are employed to select optimal feature subsets.

3.4.1 Chromosome Encoding

Each chromosome represents a binary feature selection vector.

3.4.2 Fitness Function

F=αAccuracy−βComplexityF = \alpha Accuracy - \beta ComplexityF=αAccuracy−βComplexity

This balances classification performance and computational efficiency.

3.4.3 Evolutionary Operations

- Selection: roulette wheel
- Crossover: single-point
- Mutation: random bit flip

This ensures exploration of optimal feature subsets (Carr; Samanta).

3.5 Classification Models

3.5.1 Support Vector Machine

SVM constructs an optimal hyperplane:

f(x)=wTx+bf(x) = w^T x + bf(x)=wTx+b

Kernel functions enable nonlinear classification.

3.5.2 Artificial Neural Network

A feedforward neural network is used:

y=σ(Wx+b)y = \sigma(Wx + b)y=σ(Wx+b)

Backpropagation optimizes weights.

3.5.3 Hybrid Ensemble Model

Final classification is obtained through weighted fusion:

C=∑iωiCiC = \sum_i \omega_i C_iC=i∑ωiCi

This improves robustness and reduces variance.

3.6 System Training Strategy

Training is performed using cross-validation. Dataset is split into training and testing sets (70/30). Performance is evaluated using:

- Accuracy
- Precision
- Recall
- F1-score

3.7 Computational Complexity

- Wavelet decomposition: O(nlogn)O(n \log n)O(nlogn)
- Genetic optimization: O(g·p)O(g \cdot p)O(g·p)
- SVM training: O(n2)O(n^2)O(n2)

3.8 Limitations

- High computational cost in genetic optimization
- Sensitivity to parameter tuning
- Dependency on dataset quality

4. Results

The proposed machine learning-based neural framework for affective state identification demonstrates strong

performance across multiple evaluation dimensions, including classification accuracy, robustness to noise, and generalization across heterogeneous speech samples. The integration of wavelet packet-based feature extraction with genetic algorithm-driven optimization significantly enhances the discriminative capability of the feature space, particularly for complex emotional and pathological speech patterns.

Experimental evaluation indicates that wavelet packet decomposition provides a highly localized representation of vocal signals, enabling the model to capture subtle variations in frequency sub-bands associated with emotional arousal and vocal tract irregularities. Compared to traditional Fourier-based representations, the wavelet-based features exhibit improved sensitivity to transient speech dynamics, which is critical for distinguishing closely related affective states such as sadness and neutral speech.

The feature optimization stage using genetic algorithms reduces dimensionality while preserving classification-relevant attributes. This leads to improved computational efficiency and reduces overfitting, especially in high-dimensional acoustic feature spaces. The optimized feature subset consistently outperforms full feature sets, confirming that redundant acoustic descriptors negatively impact classification stability.

In classification performance, the support vector machine model demonstrates strong baseline accuracy due to its ability to construct optimal decision boundaries in nonlinear feature spaces. However, the hybrid ensemble model, which combines SVM and artificial neural network outputs, achieves superior results in terms of both accuracy and F1-score. This improvement is attributed to the complementary strengths of both classifiers: SVM provides structural margin maximization, while neural networks enhance nonlinear adaptive learning capacity.

Noise robustness evaluation shows that the proposed framework maintains stable performance under moderate environmental noise conditions. Wavelet-based denoising effectively suppresses high-frequency distortions while preserving essential speech characteristics. However, performance degradation is observed under extreme noise conditions, indicating limitations in preprocessing robustness when signal-to-noise ratios fall below critical thresholds.

Across multiple test scenarios, the system demonstrates consistent classification behavior for both emotional and pathological speech datasets. Emotional categories such as happiness and anger are classified with higher accuracy due to distinct spectral and prosodic features, whereas neutral and low-intensity emotions show relatively higher confusion rates.

Comparative analysis with conventional machine learning pipelines confirms that the proposed hybrid framework outperforms standalone classifiers and non-optimized feature extraction systems. The combination of wavelet packet features, genetic optimization, and neural classification yields a balanced trade-off between accuracy and computational cost.

Overall, the findings validate that integrating adaptive signal processing techniques with machine learning-based neural frameworks significantly enhances affective state identification in vocal signals, aligning with prior observations in neural speech emotion recognition research (Anoop et al., 2018).

5. Discussion

The experimental findings highlight the effectiveness of combining wavelet-based signal processing with machine learning classifiers for affective state identification. The results confirm that multi-resolution analysis plays a critical role in capturing emotionally relevant speech features that are not easily detectable using conventional time-domain or frequency-domain methods.

One of the key observations is that wavelet packet decomposition significantly improves feature representation by enabling localized analysis of speech signals across multiple frequency bands. This is particularly important for emotional speech, where variations in pitch, intensity, and spectral energy distribution occur in short time intervals. The ability of wavelet transforms to preserve temporal resolution while analyzing frequency components contributes directly to improved classification performance.

The genetic algorithm-based feature selection mechanism further enhances system efficiency by eliminating redundant and less informative features. This optimization process not only reduces computational complexity but also improves generalization by minimizing overfitting risks. However, the stochastic nature of genetic algorithms introduces variability in feature selection outcomes, which may affect reproducibility under different initialization conditions.

Support vector machines provide a strong theoretical foundation for classification due to their margin maximization principle. Nevertheless, their performance is highly dependent on kernel selection and parameter tuning. In contrast, artificial neural networks demonstrate greater adaptability in modeling nonlinear relationships but require larger datasets for optimal performance. The hybrid ensemble approach effectively mitigates these individual limitations by combining the strengths of both models.

Despite these advantages, certain limitations remain. The system's performance degradation under high-noise conditions indicates that wavelet-based denoising, while effective, is not sufficient for extreme acoustic distortions. Additionally, the computational overhead introduced by genetic optimization may limit real-time applicability in resource-constrained environments.

From a theoretical perspective, the integration of adaptive signal processing and neural computation aligns with biological information processing principles observed in neural systems (Doya, 1999). The hierarchical structure of feature extraction, optimization, and classification mirrors the layered organization of sensory and cognitive processing in the human brain.

In practical terms, the proposed framework has significant implications for biomedical diagnostics, particularly in detecting voice disorders and neurological conditions. It also holds potential for affect-aware systems in human–computer interaction, where understanding user emotional states is essential for adaptive responses.

Overall, the study demonstrates that hybrid machine learning frameworks offer a robust and scalable solution for affective speech analysis, although further improvements in noise resilience and computational efficiency are required for deployment in real-world applications.

6. Conclusion

This study presented a hybrid machine learning-based neural framework for efficient affective state identification in vocal signals. The proposed system integrates wavelet packet decomposition, genetic algorithm-based feature optimization, and hybrid classification models to achieve robust and accurate recognition of emotional and pathological speech patterns.

The results demonstrate that multi-resolution wavelet analysis significantly enhances feature representation by capturing localized spectral and temporal characteristics of speech signals. The inclusion of genetic algorithms improves system efficiency by selecting optimal feature subsets, thereby reducing dimensionality and enhancing classification stability. Furthermore, the hybrid combination of support vector machines and artificial neural networks provides improved generalization and classification accuracy compared to standalone models.

The study highlights the importance of integrating signal processing techniques with machine learning approaches to address the complexity of affective speech analysis. While the proposed framework shows strong performance under moderate noise conditions, challenges remain in handling extreme acoustic environments and reducing computational complexity for real-time applications.

Future research should focus on developing deep learning-based hybrid architectures and integrating adaptive noise-robust preprocessing techniques. Additionally, exploring transfer learning approaches may improve generalization across diverse speech datasets and speaker populations.

In conclusion, the proposed framework contributes to advancing intelligent affective computing systems by providing a structured, efficient, and biologically inspired approach to vocal emotion and pathology detection (Anoop et al., 2018).

References

1. Akbari, M. Khalil Arjmandi, "An Efficient Voice Pathology Classification Scheme Based on Applying Multi-Layer Linear Discriminant Analysis to Wavelet Packet-Based Features ", Biomedical Signal Processing and Control, vol. 10, pp. 209–223, March 2014.
2. I. R. Fontes, P. T. V. Souza, A. D. D. Neto, A. M. Martins, L. F. Q. Silveira, "Classification System of Pathological Voices Using Correntropy ", Hindawi Publishing Corporation, Mathematical Problems in Engineering, pp. 1–7, August 2014.
3. M. Gavrovska, M. P. Paskas, I. S. Reljin, "Wavelet Denoising within the Lifting Scheme Framework," Telfor Journal, Vol. 4, No. 2, pp. 101–106, 2012.
4. J. Arias-Londono, J. Godino-Llorente, N. Saenz-Lechon, V. Osma-Ruiz, G. Domınguez, "Automatic Detection of Pathological Voices Using Complexity

Measures, Noise Parameters, and Mel-Cepstral Coefficients ", IEEE Transactions on Biomedical Engineering, vol. 58, No. 2, February 2011.

5. R. L. Claypoole, R. G. Baraniuk, R. D. Nowak, "Adaptive Wavelet Transforms via Lifting ", Proceeding of Acoustic, Speech and Signal Processing, Vol. 3, pp. 1513–1516, May 1998.
6. Cortes, V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, Issue 3, pp. 273–297, September 1995.
7. H. Cordeiro, J. Fonseca, C. M. Ribeiro, "LPC Spectrum first Peak Analysis for Voice Pathology Detection," International Conference on Health and Social Care Information Systems and Technologies (HCIST), vol. 9, pp. 1104–1111, December 2013.
8. Disordered Voice Database, version 1.03, Massachusetts Eye and Ear Infirmary, KAY Elemetrics Corporation, Boston, MA, Voice and Speech Lab., 1994.
9. M. Fezari, F. Amara, I. M. M. El-Emary, "Acoustic Analysis for Detection of Voice Disorders Using Adaptive Features and Classifiers ", International Conference on Circuits, Systems and Control, pp. 112–117, February 2014.
10. V. Majidnezhad, "A novel hybrid of genetic algorithm and ANN for developing a high efficient method for vocal fold pathology diagnosis," EURASIP Journal on Audio, Speech, and Music Processing, vol. 3, January 2015.
11. M. Islam, I. Parvez, H. Deng, P. Goswami, "Performance Comparison of Heterogeneous Classifiers for Detection of Parkinsons Disease Using Voice Disorder (Dysphonia) ", International Conference on Informatics, Electronics and Vision (ICIEV), pp. 1–7, May 2014.
12. J. Carr, "An Introduction to Genetic Algorithms ", Senior Project, pp. 1–40, May 2014.
13. Y. Liu, Wensheng Cai, Xueguang Shao, "Intelligent background correction using an adaptive lifting wavelet ", Chemometrics and Intelligent Laboratory Systems, Vol. 125, pp. 11–17, June 2013.
14. V. Penfield, T. Rasmussen, The Cerebral Cortex of Man. A Clinical Study of Localization of Function. New York: Macmillan, 1950.
15. P. T. Hosseini, F. Almasganj, T. Emami, R. Behroozmand, S. Gharibzade, F. Torabinezhad, "Local Discriminant Wavelet Packet Basis for Voice Pathology Classification ", the 2nd International Conference on Bioinformatics and Biomedical Engineering (ICBBE), pp. 2052–2055, May 2008.
16. P. Saidi, F. Almasganj, "Voice Disorder Signal Classification Using M-Band Wavelets and Support Vector Machine ", Circuits System and Signal Processing (CSSP), vol. 34, pp. 1–12, January 2015.
17. P. Salehi, "The Separation of Multi-Class Pathological Speech Signals Related to Vocal Cords Disorders Using Adaptation Wavelet Transform Based on Lifting Scheme," Cumhuriyet University Faculty of Science Science Journal (CSJ), vol. 36, No. 3, pp. 2371–2382, May 2015.
18. K. Doya, What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? Neural Networks 12(7-8) 961-974, 1999.
19. R. Behroozmand, F. Almasganj, "Optimal Selection of Wavelet- Packet-Based Features Using Genetic Algorithm in Pathological Assessment of Patients Speech Signal with Unilateral Vocal Fold Paralysis " Journal of Computers in Biology and Medicine, Vol. 37, pp. 474–485, April 2007.
20. W. Sweldens, "The Lifting Scheme: A Custom-Design Construction of Biorthogonal Wavelets ", Applied and Computational Harmonic Analysis, Vol. 3, Issue 2, pp. 186–200, April 1996.
21. S. Samanta, "Genetic Algorithm: An Approach for Optimization (Using MATLAB) ", International Journal of Latest Trends in Engineering and Technology (IJLTET), vol. 3 Issue. 3, pp. 261–267, January 2014.
22. S. Tanaka, Theory of self-organization of cortical maps: mathematical framework. Neural Networks 3(6) 625-640, 1990.
23. G. Strang, T. Nguyen, "Wavelets and filter banks ", wellesley - cambridge press, ISBN: 0-9614088-7-1, 1996.
24. Silva, L. Oliveira, M. Andrea, "Jitter Estimation Algorithms for Detection of Pathological Voices," EURASIP Journal on Advances in Signal Processing, Hindawi Publishing Corporation, No. 9, pp. 1–9, January 2009.
25. J. C. Saldanha, T. Ananthakrishna, R. Pinto, "Vocal Fold Pathology Assessment using PCA and LDA ", International Conference on Intelligent Systems and Signal Processing, pp. 140–144, March 2013.
26. N. Erfanian Saeedi, F. Almasganj, F. Torabinejad, "Support vector wavelet adaptation for pathological voice assessment ", Computers in Biology and Medicine, vol. 41, pp. 822–828, June 2011.
27. N. Erfanian Saeedi, F. Almasganj, "Wavelet adaptation for automatic voice disorders sorting ", Computers in Biology and Medicine, vol. 43, pp.

699–704, March 2013.

28. Anoop, V., Rao, P.V., Aruna, S. (2018). An Effective Speech Emotion Recognition Using Artificial Neural Networks. In: Reddy, M., Viswanath, K., K.M., S. (eds) International Proceedings on Advances in Soft Computing, Intelligent Systems and Applications . Advances in Intelligent Systems and Applications, vol 628. Springer, Singapore. https://doi.org/10.1007/978-981-10-5272-9_36.