European Journal of Interdisciplinary Research and Development
Volume-21 November 2023
Website: www.ejird.journalspark.org ISSN (E): 2720-5746

SUPPLYING CHORAL WITH ADVANCED TECHNIQUES FOR STUTTERED PEOPLE

Ali Bashar Hussein1,
Raid Rafi Omar Al-Nima1
Tingting Han 2
1 Department of Medical Instrumentation Technology Engineering,
Technical Engineering College of Mosul, Northern Technical University, Iraq
2 School of Computing and Mathematical Science, Birkbeck,
University of London, United Kingdom
Email: e.alibashar@gmail.com

Abstract

Stuttering can significantly impact the quality of a person's life. Stutterers may refrain from speaking and may waste opportunities to make friends, present own ideas, and opinions in public and to be disadvantaged in job interviews. In this paper, we suggest supplying choral or feedback speech with advanced techniques that can contribute to curing people who stutter. The advanced techniques are the Enhanced One-dimensional Local Binary Patterns (EOLBP) and Adapted Multi-Layer Perceptron for Regression (AMLPR). The EOLBP is a useful feature extraction and the AMLPR is a generative classifier. The Fluency Bank (FB) dataset which includes stuttered speech signals is utilized in this study. The best result of a very high accuracy (99.44%) is achieved in this study.

Keywords: Choral, Classifier, Feature Extraction, Stutter.

Introduction

Humans use speech to convey their emotions, thoughts, and ideas. Besides, articulation, voice and fluency make up speech [1]. On the other hand, there are many different types of speech disorders around the world including mutism, aphasia, spasmodic dysphonia, dysarthria, lisping, cluttering, apraxia of speech and stuttering/stammering. Furthermore, stuttering is one of the most occurring speech disorders which is considered here [2].

Stuttering is a kind of speech disorder in which speech fluency is disturbed by dysfluencies such as repetitions, prolongations and pauses [2]. *Repetition* is repeating the same small talking part more than once, *prolongation* is lengthening the duration of talking for a specific string and *pause* is important to be considered if a person is silent for a significantly long time, i.e., exceeding a particular amount of time [2]. Stuttering causes a problem for speakers because of emotional, psychological and shame-related issues. This problem affects a person who is speaking consistently or at the initiation time of speak. Moreover, stutterers' communication is generally not clear due to speech flow regarding time [3]. Figures 1 and 2 show examples of normal speech signal and abnormal speech signal of stuttered speech with prolongation, repetition and pause.

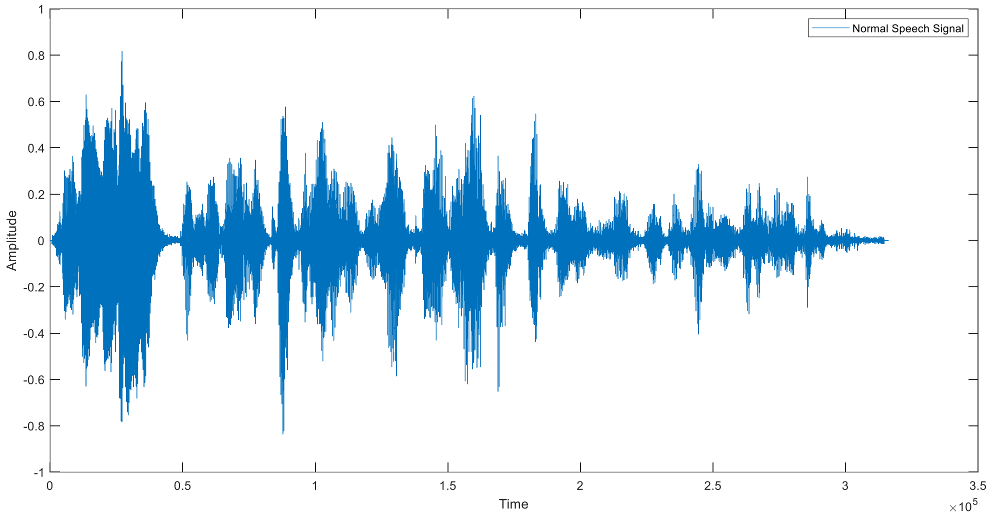


Figure 1: Example of normal speech signal

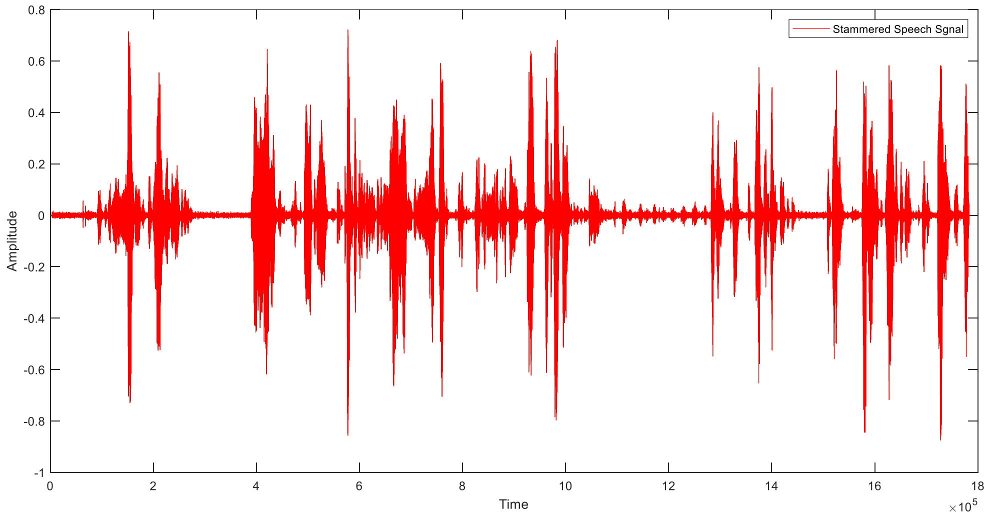


Figure 2: Example of stuttered speech signal

Choral speech is speech that is provided in unison with the stuttered person [4]. Choral speech, or speaking in unison, is an undeniable phenomenon that instantly causes fluent and natural speech in almost all people who stutter. This works well independently of linguistic content, situation or audience size [5]. Choral speech has long been known to rapidly generate fluent speech in even the most severe cases of stuttering [6]. The amazing effect of this choral phenomenon has led to provide simulation mechanism via Altered Auditory Feedback (AAF) as alternative solution for the choral effect, which assists in curing stuttered and stressed patients.
In studies of stuttering, the AAF refers to a collective term for cases in which speech signals are electronically changed such that speakers hear their voices differently than usual. Sometimes the altered speech signal is referred to as a second speech signal [7], [8]. Frequency-Altered Feedback (FAF), Delayed Auditory Feedback (DAF) and Masking Auditory Feedback (MAF) are the types of speech alteration that are most frequently employed to change speech signals [9].

DAF happens when the speaker's air-conducted voice signal is typically 50–100 ms late in reaching the listener. This causes speakers to hear what they just said after a delay [8]. FAF is the process of changing a speaker's voice frequency, often by 1/4 to 1 Octave. Speakers hear higher or lower pitches than typical for their own voice [9]. MAF means that the speaker's speech is simultaneously and continuously obscured by a different sound, usually noise (such as white noise) [10]. It has been investigated that DAF and FAF are more successful than MAF in reducing stuttered speech [11], [12].

Clinicians and stuttered people can now access increasingly sophisticated and miniaturized devices that provide DAF and FAF [9], [13]. In our work, we deliver choral speech of anti-stuttered signals instead of DAF and FAF signals. This can significantly remove and contribute in curing stutterers. The contributions of this paper are as follows:

1- Preparing and using the Fluency Bank (FB), a dataset for stuttering speeches.

2- Suggesting an effective anti-stammering algorithm.

3- Proposing the Enhanced One-dimensional Local Binary Patterns (EOLBP) feature extraction technique.

4- Presenting the Adapted Multi-Layer Perceptron for Regression (AMLPR) classifier technique.

5- Delivering choral or feedback speeches of anti-stuttered signals to achieve stuttering cures.

The rest of the paper is structured as follows: Section 2 provides review for relevant studies, Section 3 describes the utilized techniques, Section 4 demonstrates and discusses the results, and Section 5 declares the conclusion.

2. LITERATURE REVIEW

Stuttering can cause negative feelings of inadequacy, shame or embarrassment to stutterers. As a result, many studies were presented for this important topic.

In 2012, Hariharan *et al.* introduced the sample entropy and Least Square Support Vector Machine (LS-SVM) for speech stuttering evaluation. Before the feature extraction procedure, manual segmentation was used to identify the stuttering events. Bark scale, Mel scale and Erb scale filter banks were used to extract the sample entropy features after wavelet packet decomposition. The LS-SVM was used to evaluate the extracted features for repetition and prolongation. In this study, 39 speech samples from the University College London's Archive of Stuttered Speech (UCLASS) database were used. The wavelet packet filtering of Daubechies family-order 2 (db2) with the Erb scale provided the maximum accuracy of 96.84% [1].

In 2015, Kumar *et al.* designed and implemented a system for recognizing stuttered speech with silent pauses. The system will eliminate silent pauses from the speech and provide corrected and easily understandable speech style. The pre-emphasis, stutter removal, segmentation, feature extraction, Vector Quantization (VQ) codebook production and score-matching phases of the classification system were the main stages. This method eliminated stuttering by taking into accounts the fact that spoken speech has more energy than unspoken speech. The Mel Frequency Cepstral Coefficients (MFCCs) algorithm was used to extract the features. Training feature vectors of dysfluent speech were clustered by using K-means algorithm to create the VQ codebook, which were subsequently saved in the database. The database was matched with dysfluent speech using the Dynamic Time Warping (DTW) algorithm. Stutter-free speech was finally recognized after the correction of silent paused stutter speech [14].

In 2016, Bortoletto *et al.* designed anti-stuttering applications using the Digital Signal Processing (DSP) altered feedback system, which aimed to help people who stutter and have speech disorder by modifying the original voice and playing it back to the patient's ear. The system was consisted of three steps that carried out speech-processing algorithms. The initial step was to minimize or eliminate background noise. The second step was for the delaying procedure which causes delayed voice signal after noise removal, Delayed Auditory Feedback (DAF) function was implemented in this step. The third step was for modifying pitch frequency of input voice signal. Output voice was natural noise-free voice that has different pitch frequency from speaker's voice within same period of input signal. Output signals resulted with pitch frequencies that were slightly higher than input signals [15].

In 2017, Manjula *et al.* proposed using epoch features to identify and validate repetition/prolongation in stuttering speech by listening to glottal closure instants. Proposed technique was for identifying glottal closure instances in stuttering speech based on source signal phase harmonics. By using a Zero Frequency Filter (ZFF), effect of vocal tract resonances was removed from stuttering speech signal to extract source signal. Glottal activity region during excitation was highlighted by the suggested ZFF. Similarly, the normalized error helped in distinguishing between the areas of stuttering events and normal speech. Twenty participants between the ages of 15 and 35 were considered for the study by reading a unified text recorded at the All India Institute of Speech and Hearing (AIISH) rooms [16].

In 2018, Alharbi *et al.* presented a method to identify stuttering in children's speech that requires less supervision. By two-pronged strategies which focused on the most frequent stuttering occurrences in kids executing records of reading activity, it was aimed to identify stuttering events. First strategy was for the Automatic Speech Recognition (ASR) which enhances detections of specific stuttering events types (sound, word, part-word and phrase repetition, and revision) by using a task-oriented word (and sub-word) lattices. Enhanced results preserved 63% of stuttering occurrences. Second strategy was for the prolongation detector, which identifies segments corresponding to prolongation occurrences based on the correlation of subsequent voiced frames. The output here was added to the ASR system as a correction layer. The final result reported a raising percentage of preserved stuttering occurrences to 70% [17].

In 2019, Afroz and Koolagudi proposed that acoustic features could be used to identify and categorize pauses in stuttering speech. In this work, attempts were made to identify inaudible (silent or unfilled) pauses in stuttering speech. Using a dynamic threshold which was established based on energy and Zero Crossing Rate (ZCR), an automatic blind segmentation method was used to divide the speech signal into voiced and unvoiced regions. To find intra-morphic (unfilled) pauses in voiced regions, 4th formant frequencies were analyzed. For both stuttered speech and natural speech, the length of the intra-morphic pauses was analyzed. It was noted that segmenting of the intra-morphic pauses achieved an accuracy of 98% [18].

In 2020, Bhatia *et al.* developed a deep learning-based system for stutter diagnosis and treatment. The system aimed to assist stuttered people by determining the kind and intensity of stuttering, and also by recommending suitable therapies after establishing the relationship between the stutter descriptors and their success in speech therapies. The system focused on constructing a therapy suggestion agent using the Support Vector Machine (SVM) and stutter diagnosis agent utilizing the Gated Recurrent Convolutional Neural Network (GRCNN) on Mel Frequency Cepstral Coefficient (MFCC) audio features. Two GRCNN models were used in the study, they were separately trained to identify prolongation and repetition in speech audio. They attained the validation accuracies of 95% and 92%, respectively. The SVM model was trained to suggest best-suited therapy and it achieved 94% validation accuracy. The main problem with the models was that background noise corrupted audio speech signals, leading to a significant proportion of false positives [19].

In 2021, Vaidianathan *et al.* designed an enhanced Kaman filter and neural networks for stuttered speech recognition and classification. In this study, acoustic effects as pauses and noises made inside and outside were used to process speech signal depending on their caused disturbances. Here, an enhanced Kalman filter was introduced to reduce the impact of noise on speech signals. Performances of the filter and other filters were examined and compared using metrics like Mean Squared Error (MSE), Mean Absolute Error (MAE), Signal-to-Noise Ratio (SNR), Peak Signal to Noise Ratio (PSNR) and cross-correlation. CNN was then used to classify speech signals based on the extracted features [20].

In 2022, Sheikh *et al.* introduced adversarial learning and Multi-Task Learning (MTL) for robust stuttering detection, so that speech therapists could monitor stutterers' disfluencies and their progresses by automatically detecting and identifying the Persons Who Stutter (PWS). In this work, the impacts of MTL and adversarial learning were examined to learn the robust stutter features. The Stuttering Events in Podcasts (SEP-28k) stuttering dataset was used. It was recorded that the repetitions, blocks and interjections over the baseline were reached up to 10%, 6.78% and 2%, respectively [21].

From this literature, it can be investigated that many studies were presented based on the recognition and correction of stuttering/stammering. Otherwise, this study adds significant anti-stammering contribution by delivering choral (or feedback) speech to the people who stutter. To the best of our knowledge, no study considered such contributions with the suggested advanced techniques.

3. UTILIZED TECHNIQUES

3.1 Choral Effect on Stuttering

It has been proposed that choral speech effect is a type of direct imitation form, a basic and innate human ability which can be mediated by "mirror neurons" at the neuronal level [5]. First of all, choral speech has long been known as instantly producing spontaneous, reflexive and natural fluency sound in even the most severe cases of stuttering. In contrast, with typical post-therapeutic speech, the distinct characteristic of choral speech is the sensation of 'invulnerability' to stuttering. It is independent to phonetic context, audience size or situational environment. It also has been suggested that choral speech inhibits stuttering immediately by engaging mirror systems of neurons, innate primitive neural substrates which dominate the early stages of language development due to their capacity to fluently imitate gestural action sequences in reflexive way. In addition, mirror systems are primal in origin, thus, they take precedence over the much later-evolving stuttering disorder. Also, it has been suggested that the most effective method for stuttering treatment can involve re-engaging mirror neurons through choral speech or one of its derivatives (by using digital signal processing technology) to produce gestural mirrors, which are nature's way of instantly overriding the central stuttering block [6].

Choral effect produced by electronic devices allows the user to hear his/her own voice but with a time delay (DAF), tiny pitch shift (FAF) or a combination between them. Such modification creates an illusion of another person speaks as the stutterer (user) at the same time. Additionally, choral effect emulates nothing more than two persons speaking simultaneously. This effect has opened the door for effective stuttering therapy, as stutterers find it much easier to talk fluently when they are speaking with someone else. Finally, when the electronic device is used by someone who have stuttering, the engagement of the mirror neurons permits them to replicate the speech fluently [22].

3.2 Delivering Feedback (or Choral) Speech

Stuttering was actually eliminated (around 98%) across all utterances and all syllable locations in the true choral speech condition. On the contrary, stuttering for the AAF was reduced to an average of almost 68% as compared with the Nonaltered Auditory Feedback (NAF) condition [23].

Although stuttering reductions under AAF were significant, they weren't exactly the same as they were for choral speech. Stuttering reduction for choral speech was extremely effective even though the accompanist's voice was delayed from the adult who stuttered. It was also very robust when there was no chance for the adult stutterer and accompanist to interact in a dynamic way. Finally, it was again very efficient when the voice of the accompanist was replaced by the voice of an adult who stuttered. The aforementioned information approximates specific AAF features. Additionally, choral speech was highly influenced in reducing stuttering through changes in speech rate and through both unfamiliar and familiar passages [24]. The AAF requires speech to be initiated by the stutterer and the 'feedback' before it can directly inhibit stuttering, in contrast, choral signal is exogenously initiated and gives the most comprehensive fluency enhancement [23].

Due to the aforementioned advantages of choral speech comparing to AAF, we suggested to use the choral speech instead of the traditional ways of the DAF, FAF or combination of both. Since AAF is recommended to be a practical clinical choice for people who stutter and it is frequently used in conjunction with therapeutic techniques, particularly those help speech initiation [23]. So, our suggested method of the choral or feedback speech should be valuable and efficient.

We suggested that the anti-stammering output is the choral speech, since this speech has a different voice and different frequency from the stutterer's voice. It can be used as feedback speech which returns to the stutterer's ear to provide the choral effect, then, this contributes in providing the immediate fluent speech and stuttering cure.

3.3 Proposed Anti-stammering Algorithm and Feedback Speech Mechanism

In this paper, we have proposed an anti-stuttering algorithm to correct stuttering speech. Additionally, we have proposed a choral or feedback speech mechanism that can provide a significant curing by feeding back the choral speech to the stutterer's ear to acquire its effect. This is an alternative solution that may assist speech-language pathologists to consider instead of performing the routine methods to heal stuttered persons. The proposed algorithm and mechanism involve the following stages:

1- Preparing a dataset of stuttering and anti-stuttering signals.

2- Applying segmentations by partitioned the stuttered and the anti-stuttered signals into sentences.

3- Operating augmentations to segmented speech signals for the case of testing.

4- Extracting features by applying the suggested EOLBP approach.

5- Employing AMLPR in two phases: train and test.

6- Saving the final weights of training phase. Then, using them in the testing phase.

7- Providing anti-stuttering outputs for the tested stuttering signals and computing the performance.

8- Employing the anti-stuttering outputs which represent the feedback speech for a stuttered person's ear in order to provide the choral effect.

9- Getting fluent speech from the mouth of a stuttered person which can afford the stuttering cure.

Figure 3 illustrates the blockdiagram of the proposed anti-stuttering algorithm and Figure 4 illustrates choral or feedback speech mechanism. They show the flow of the stages.

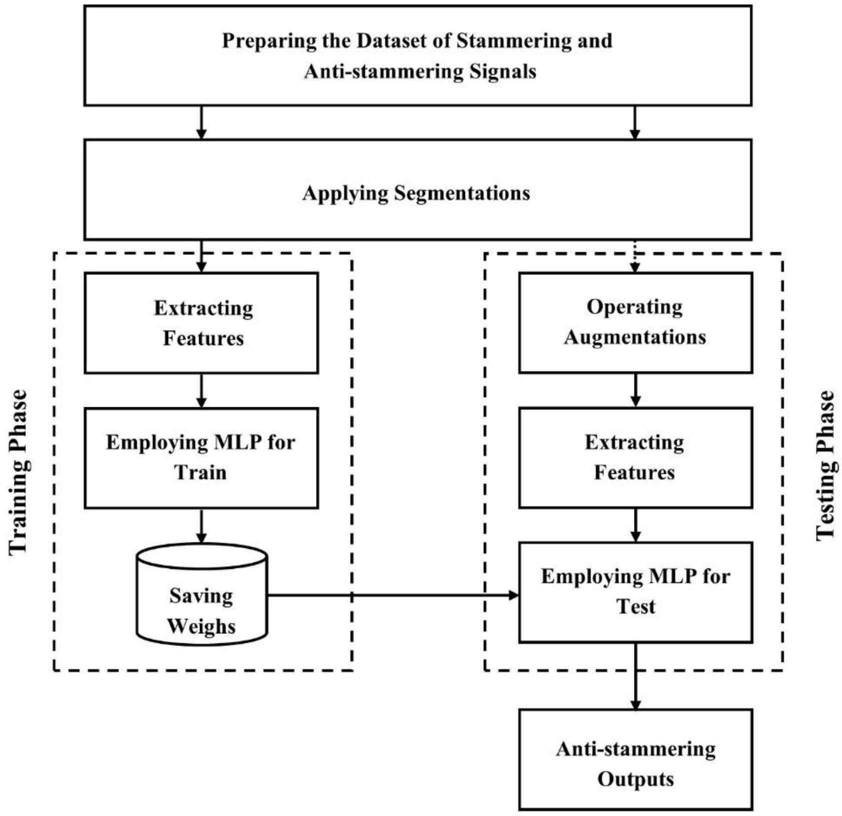


Figure 3: Blockdiagram of the proposed anti-stammering algorithm

European Journal of Interdisciplinary Research and Development
Volume-21 November 2023
Website: www.ejird.journalspark.org ISSN (E): 2720-5746

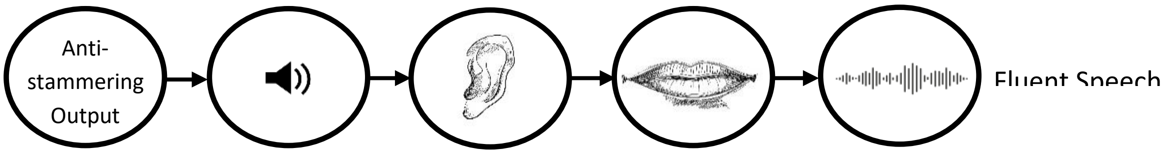


Figure 4: Choral or feedback speech mechanism

3.4 Preparing and Segmenting the Dataset of Stuttering Signals

FluencyBank (FB) dataset has been found so useful in this study. The FB is a shared dataset that are dedicated to fluency development. Also, it is special, valuable and reliable [25]. The FB is a new part of the larger TalkBank (TB) system, which has recently received funding from the National Science Foundation (NSF) and the National Institute on Deafness and other Communication Disorders (NIDCD) [26].

The preparation processes include:

1- Collecting the FB dataset. It includes 24 stuttered speech videos and 1 anti-stuttered speech video of reading unified long texts by different contributors of various ages.

2- Separating the stuttered speech voices (signals) from the stuttered speech videos for each contributor. Hence, 24 stuttered speech signals are obtained as (mp3) format.

3- Separating the anti-stuttered speech voice (signal) from the anti-stuttering speech video in order to be exploited as targets. They are also obtained as (mp3) format.

The segmentation processes include:

1- Partitioning each stuttered speech signal for each contributor into short sentences, resulting in 360 segmented signals for the whole contributors.

2- Partitioning each anti-stuttered speech signal into short sentences, resulting in 15 segmented signals (or targets) for each contributor.

3- Standardizing the sizes of segmented signals in order to have same size and be ready to use in the rest of algorithm.

3.5 Approached Feature Extraction

3.5.1 Conventional feature extraction

As a conventional feature extraction, the Local Binary Pattern (LBP) is efficiently employed in many works [27]. It is especially used for image recognition purposes. It is also adopted as a feature extraction for signal analysis [27].

The LBP is constructed by first comparing between a center value in a small part of a vector (window) and its neighbors [28]. Consequently, a binary number is produced from the comparison [29]. The produced binary number is then decoded to a decimal number. Such processes are carried out for the whole values of a vector. Then, histogram is exploited for the conventional LBP method [30]. Given the center value has the symbol of c and the neighbor value has the symbol of p the LBP is computed as follows [30]:

$$LBP = \sum_{p=0}^{7} s(g_p - g_c)\, 2^p \qquad (1)$$

where LBP is the produced LBP value, g_p is a value of p^{th} neighbor in a part of the vector, g_c is a value of center c in a part of the vector and $s(.)$ is expressed as:

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \quad (2)$$

In the conventional LBP, neighbor values in an image are arranged in a circle that covers 360°. Whereas, in one-dimensional (1D) vector, neighborhood values in a signal are arranged in a line of angle 180°. Therefore, we have investigated a 1D feature extraction based on the LBP which is suitable to be adopted for stammering signals.

3.5.2 Suggested EOLBP approach

Here, the EOLBP feature extraction is approached by enhancing the conventional LBP method. It considers 8 values of neighbors that are arranged in a horizontal direction around a center value for a part of the 1D stammering speech signal. This arrangement can afford reasonable performance as confirmed in [31].

The way of computing the EOLBP feature extraction for a 1D signal (vector) is explained in the following steps:

1- Considering a window of 9 values at the beginning of the 1D vector. It has a center value and neighbor values, which are arranged around the center value as 4 values on the left and 4 values on the right.

2- Taking into account the center value of a considered window is the threshold.

3- Comparing the threshold with each neighbor value in the same window. If the neighbor value is equal to or greater than the threshold, logic 1 will be assigned. Otherwise, logic 0 will be assigned.

4- Converting 8 logical values to a corresponding decimal value.

5- Moving the window to the right by only one value along the 1D vector and repeating the same processes from steps 2 to 4 until the whole 1D vector length is covered.

6- Dividing the corresponding original signal values by the computed decimal values.

It can be highlighted that the processes in steps 3 and 4 are equivalent to the calculations in equations (1) and (2) taking into account that all the values are arranged horizontally. Furthermore, the purpose of the process in step 6 is to maintain the variations in the feature extraction values according to the original signal values. Figure 5 shows a demonstration of essential processes in the suggested EOLBP approach.

3.6 MLP Overview

3.6.1 Fundamentals of MLP

First of all, the Artificial Neural Network (ANN) is one of the most widely used Artificial Intelligence (AI) methods. The fundamental idea of ANN is to simulate a mathematical neural network model as a biological neuron in a human's brain. Training and testing are the two phases to be carried out in this simulation. In the first phase, the ANN attempts to learn pairs of samples, which are the inputs and targets. In the second phase, new input samples are used to test the intelligence of the ANN model to see how well ANN will react to unique input samples that have never been seen before. Finally, the network's outputs can be compared with the desired targets and the performance is computed [32].

MLP is a type of ANN. It consists of an input layer (X units), hidden layer (Z units) and output layer (Y units). Both hidden and output layers have biases. The output units Y_k contain the biases

w_{0k}, while the hidden units Z_j contain the biases v_{0j} [33]. Figure 6 shows a standard form of MLP Network.

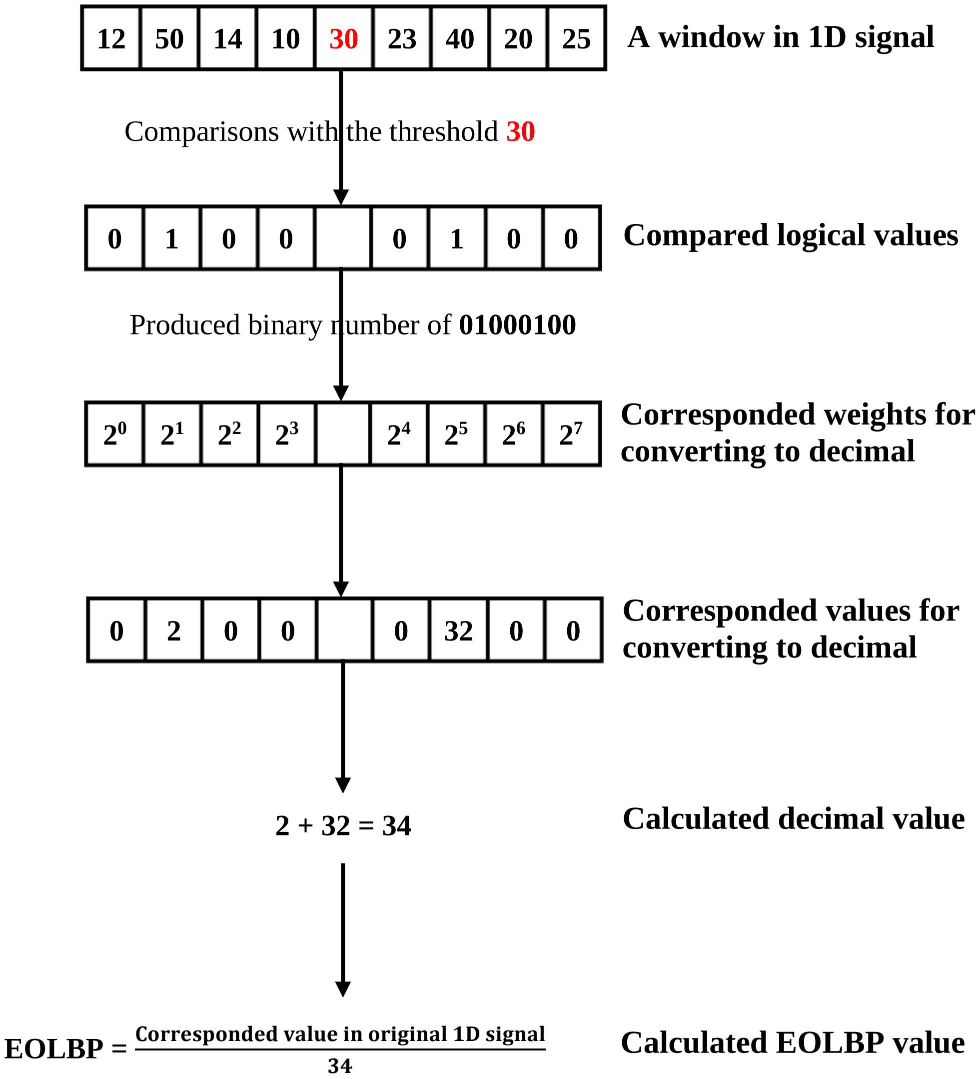


Figure 5: A demonstration of essential processes in the suggested EOLBP approach

During the feedforward phase, each input unit (X_i) $(i = 1,2, \dots, n)$ transmits an input value to each hidden unit via (**v**) weights. The hidden units (Z_j) $(j = 1,2, \dots, p)$ then compute their values of activations. The values are transmitted to the output units via (**w**) weights. Consequently, the output units (Y_k) $(k = 1,2, \dots, m)$ compute their values of activations. In the training phase, the output values Y_k and the target values T_k are compared. The error factor (δ_k) between the output and the target is calculated and distributed to the previous layer. It is also employed to update the weights and biases. Similarly, the error factor (δ_j) is computed in the hidden layer. The weights and biases between the input units and hidden units are updated using this error. All weights of v_{ij} and w_{jk} are accordingly modified, and all biases of v_{0j} and w_{0k} are accordingly updated too [33].

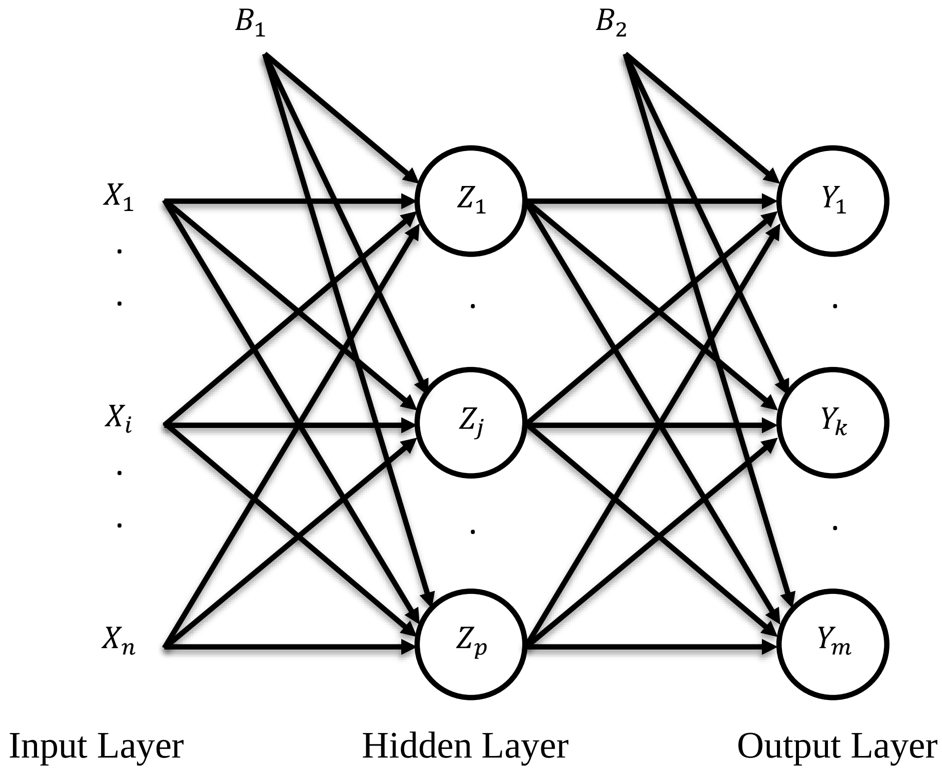


Figure 6: Standard form of an MLP Network

3.6.2 General training phase for the MLP

There are three training stages for the MLP: feed-forward of inputs to outputs, backpropagation of error and updating weights and biases [33]. All values of weights and biases here initialize to small random numbers [33]. The feedforward stage equations are expressed as follows [33]:

Feedforward:

$$z\_in_j = v_{0j} + \sum_{i=0}^{n} X_i\, v_{ij} \tag{3}$$

$$Z_j = f(z\_in_j) \tag{4}$$

$$y\_in_k = w_{0k} + \sum_{j=1}^{p} Z_j\, w_{jk} \tag{5}$$

$$Y_k = f(y\_in_k) \tag{6}$$

where z_in_j represents the input computations in each hidden unit and $(j = 1,2, \dots, p)$, v_{0j} represents a connection weight between a hidden and a bias units, X_i represents an input value in an input unit and $(i = 1,2, \dots, n)$, v_{ij} represents a connection weight between an input and a hidden units, Z_j represents a calculated value of a hidden unit, y_in_k represents an input computation in each output unit and $(k = 1,2, \dots, m)$, w_{0k} represents a connection weight between an output and a bias units, w_{jk} represents a connection weight between a hidden and an output unit, and Y_k represents the calculated output value of an output unit [33].

Backpropagation of error:

$$\delta_k = (T_k - Y_k)\, f'(y\_in_k) \tag{7}$$

$$\Delta w_{ik} = \alpha\, \delta_k\, Z_j \tag{8}$$

$$\Delta w_{0k} = \alpha\, \delta_k \tag{9}$$

$$\delta\_i\,n_j = w_{0k} + \sum_{k=1}^{m} \delta_k\, w_{jk} \tag{10}$$

$$\delta_j = \delta\_in_j\, f'(z\_i\,n_j) \tag{11}$$

$$\Delta v_{ij} = \alpha\, \delta_j\, X_i \tag{12}$$

$$\Delta v_{0j} = \alpha\, \delta_j \tag{13}$$

where δ_k represents a computed error at the output layer, T_k represents a provided target, Δ represents an updated value, α represents the learning rate, δ_in_j represents a calculated input error at the hidden layer and δ_j represents a computed error at the hidden layer [33].

Update biases and weights:

$$v_{0j}(\text{new}) = v_{0j}(\text{old}) + \Delta v_{0j} \tag{14}$$

$$w_{0k}(\text{new}) = w_{0k}(\text{old}) + \Delta w_{0k} \tag{15}$$

$$\mathrm{v}_{ij}(\text{new}) = v_{ij}(\text{old}) + \Delta v_{ij} \tag{16}$$

$$w_{jk}(\text{new}) = w_{jk}(\text{old}) + \Delta w_{jk} \tag{17}$$

In this phase, equations (3) to (17) are iterated until all trained MLP outputs are being so close to the provided targets.

3.6.3 General testing phase for the MLP

After the training phase, a testing procedure is performed to test the MLP performance. It has only one stage similar to the feed-forward stage of the training algorithm. However, the only difference is for the values of weights and biases, where in the testing phase they are initialized as the last values of the training phase. The equations of this phase are as follows [33]:

$$z\_in_j = v_{0j} + \sum_{i=0}^{n} X_i\, v_{ij} \tag{18}$$

$$Z_j = f(z\_in_j) \tag{19}$$

$$y\_in_k = w_{0k} + \sum_{j=1}^{p} Z_j\, w_{jk} \tag{20}$$

$$Y_k = f(y\_in_k) \tag{21}$$

In this phase, there are no analyses for the tested input values and no iterations to improve the calculated output values.

3.7 MLP for Regression

3.7.1 Fundamentals of Regression

Regression is an analysis technique used to establish a relationship between independent and dependent variables [34]. It has the ability to solve complex problems. models predict the results of dependent variables based on the independent variables. Fundamentally, regression relationship can be linear or non-linear [34]–[36]. Linear regression can be one of two types simple or multiple. Simple linear regression is a prediction process by employing a single independent variable or refers to the separate dependent and independent variables to explore the connection between two variables. Whereas, multiple linear regression is a prediction process with more than one independent variable or refers to a regression relationship between a single dependent variable and multiple independent variables [34]. Our work in this paper is related to the multiple linear regression type.

The proposed method can be considered as one of other developed artificial intelligence models as in [37], [38].

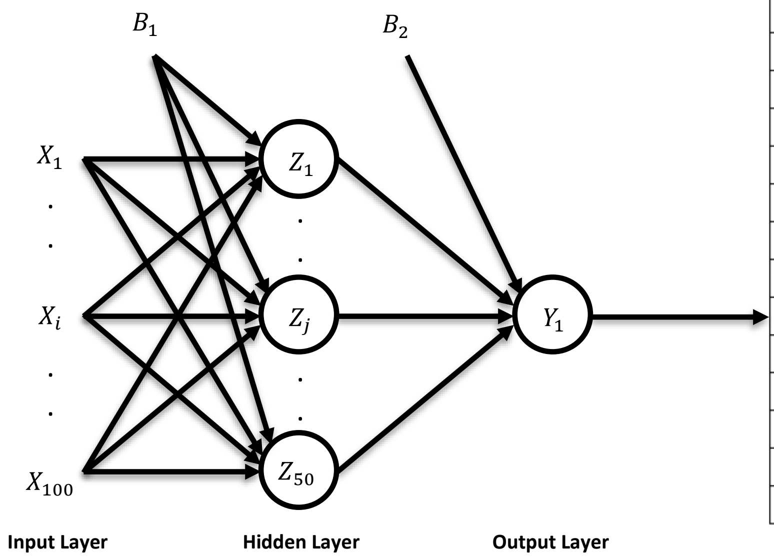
3.7.2 The AMLPR

With some changes in employed the MLP network, it has been adapted for regression. The AMLPR has the following specifications:

1. inputs of EOLBP feature extractions, where each input vector has 100 values and it enters parallel to the network.
2. targets of indexed values (from 1, 2, to 15), each refers to a certain anti-stuttering sentence.
3. input vectors each represents a feature extraction of a stuttering sentence.
4. one hidden layer consists of 50 nodes or neurons.
5. one AMLPR is assigned for a single stuttered person.

Figure 7 shows the architecture of the AMLPR. It works in two phases: training and testing. In the training phase, the AMLPR collects input training vectors of EOLBP feature extractions for stuttering sentences. It trains to produce output values that are closer to the provided corresponding targets. As mentioned, the target values are indexes, each refers to a determined anti-stuttering sentence. In the testing phase, augmentation process is used to evaluate stuttering sentences with expected changes. Also, augmentation has the capability to increase the number of stuttering samples. Augmentation type of adding noise to stuttering signals are utilized in this study as it is expected that signals of stuttered speech are not always pure (they usually mixed with noise).

We employed the first 100 values for each input vector which represent extracted features of starting stuttered signal. This offers advantages for the AMLPR to provide a complete fluent speech signal to the stutterers as reducing the time of delivering the full fluent sentence. This can help a stutterer in reducing negative feelings like embarrassment and stress to collect fluent speech by ear, then, produce anti-stuttered talk by mouth. This contributes in achieving anti-stutter cure.



Index	Anti-stammering Signals
1	1st anti-stammering signal
2	2nd anti-stammering signal
3	3rd anti-stammering signal
4	4th anti-stammering signal
5	5th anti-stammering signal
6	6th anti-stammering signal
7	7th anti-stammering signal
8	8th anti-stammering signal
9	9th anti-stammering signal
10	10th anti-stammering signal
11	11th anti-stammering signal
12	12th anti-stammering signal
13	13th anti-stammering signal
14	14th anti-stammering signal
15	15th anti-stammering signal

Lookup Table

Figure 7: The architecture of the AMLPR

4. RESULTS AND DISCUSSIONS

4.1 Dataset Description

The FB dataset has been used in this study. It is contributed by National Stuttering Association (NSA) members to help students learn more about actions and affective/cognitive characteristics of living with stuttering as adults [39].

FB dataset started establishing to its source in June 2017 with new contributors of adults who stutter. FB dataset has been collected from 25 participants of different stuttering degrees. Contributors or participants were aged from 24 to 62 years old. They formed what it can be called a clinical group, consisting of 16 males and 9 females who have been diagnosed with stuttering by the Institutional Review Board of the University of Maryland (Nan Bernstein Ratner, PI). It can also be investigated that the clinical group was instructed to read a Friuli passage from Stuttering Severity Instrument-4 (SSI4) which contains 369 syllables provided with the publishers' permissions. Furthermore, the speech dataset was found to be recorded in .mp4 form with the audio sampling rate of 48,000kHz for 22 participants and with the audio sampling rate of 44,100kHz for 3 participants [39].

4.2 Anti-Stammering Implementation

The proposed anti-stuttering algorithm is capable of correcting stuttered speeches and producing anti-stuttered speeches. It has been implemented in MATLAB, which is a high-performance programming language for technical computing. It contains powerful built-in routines that allow for a wide range of computations. All operations of the suggested EOLBP feature extraction and AMLPR network, as well as augmentations, have been established in MATLAB software in two phases: train and test.

4.3 Training Phase Results

The training phase has the following specifications:

- entering speeches with stuttering.
- applying the EOLBP feature extraction.
- employing the AMLP.
- using 100 nodes in the AMLPR input layer which is equivalent to the number of EOLBP feature extraction values for each stuttered speech of a sentence.
- utilizing 50 nodes in the AMLPR hidden layer.
- providing 1 node in the AMLPR output layer which is equivalent to the regression targets.
- applying the transfer function of tan sigmoid in the AMLPR hidden layer.
- employing the transfer function of pure linear in the AMLPR output layer.
- exploiting the training type of Scaled Conjugate Gradient (SCG) to train the AMLPR.
- utilizing 1000 epochs as the maximum number of AMLPR training epochs.
- using 1×10^{-14} as the minimum AMLPR training Mean Square Error (MSE).
- providing EOLBP feature extractions for sentences of stuttered speeches and a single AMLPR for each person.

Figure 8 shows examples of AMLPRs' training performance curves for speech signals with stuttering.

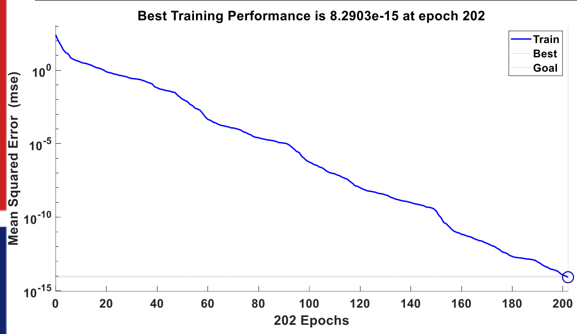
From the curves of figure 8(a-f) it can be noticed that at the beginning of each training process, the curves were far from the desired error. The curves are then heading down which means that the updating of weights has worked appropriately and trainings have done well. These indicate that training processes are successful.

Figure 9 demonstrates linear regression curves of AMLPRs in order to provide anti-stuttering speech signals at the end. They provide relationships between AMLPR outputs and desired targets after completed trainings.

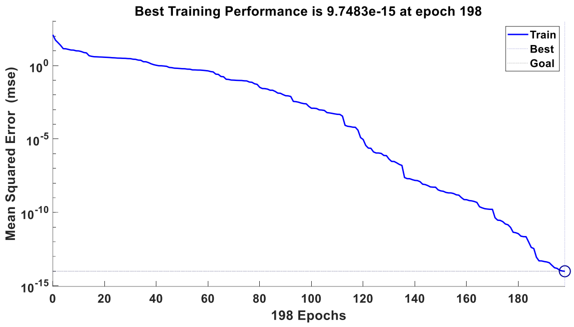
Figure 9(a-f) is another evidence to approve trainings success since all the data are exactly located on the best-fitted lines. They all achieve the best regression parameter (R) value, that is R=1 and this is very satisfactory result in our AMLPR training phase. This helps for obtaining suitable results in the testing phase.

Figure 10 provides gradient descent curves of AMLPRs during their training.

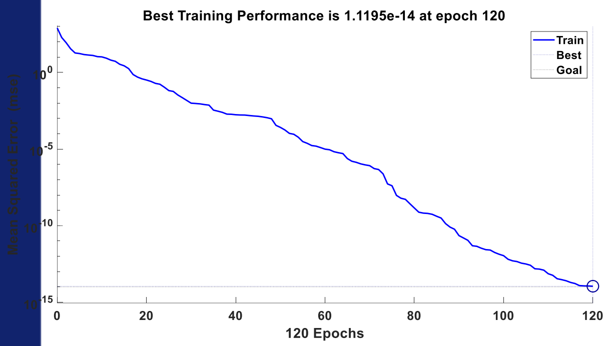
Gradient descent curves in figure 10(a-f) use an algorithm to reduce the error function in an MLP until acquiring best weights and producing nearest output values to desired targets. It can be seen how gradient descent curves are decreased during trainings to reduce the error functions. These are also referred to the successes of the training phase.



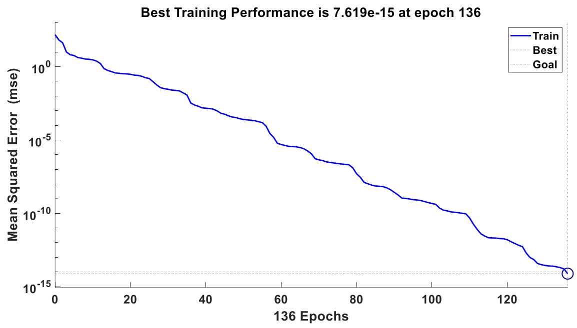
(a)



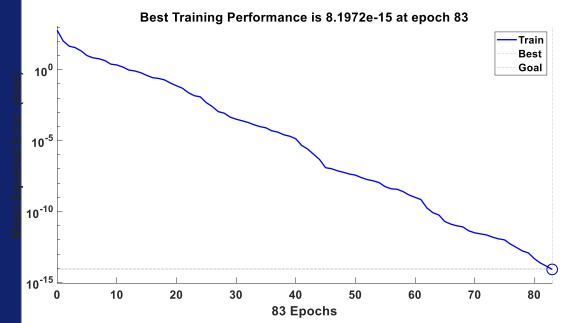
(b)



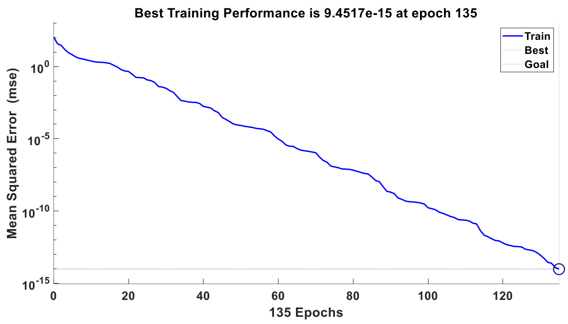
(c)



(d)



(e)



(f)

Figure 8: Examples of AMLPs' the training performance curves for speech signals with stuttering (a) for the first person, (b) for the second person, (c) for the third person, (d) for the fourth person, (e) for the fifth person and (f) for the sixth person

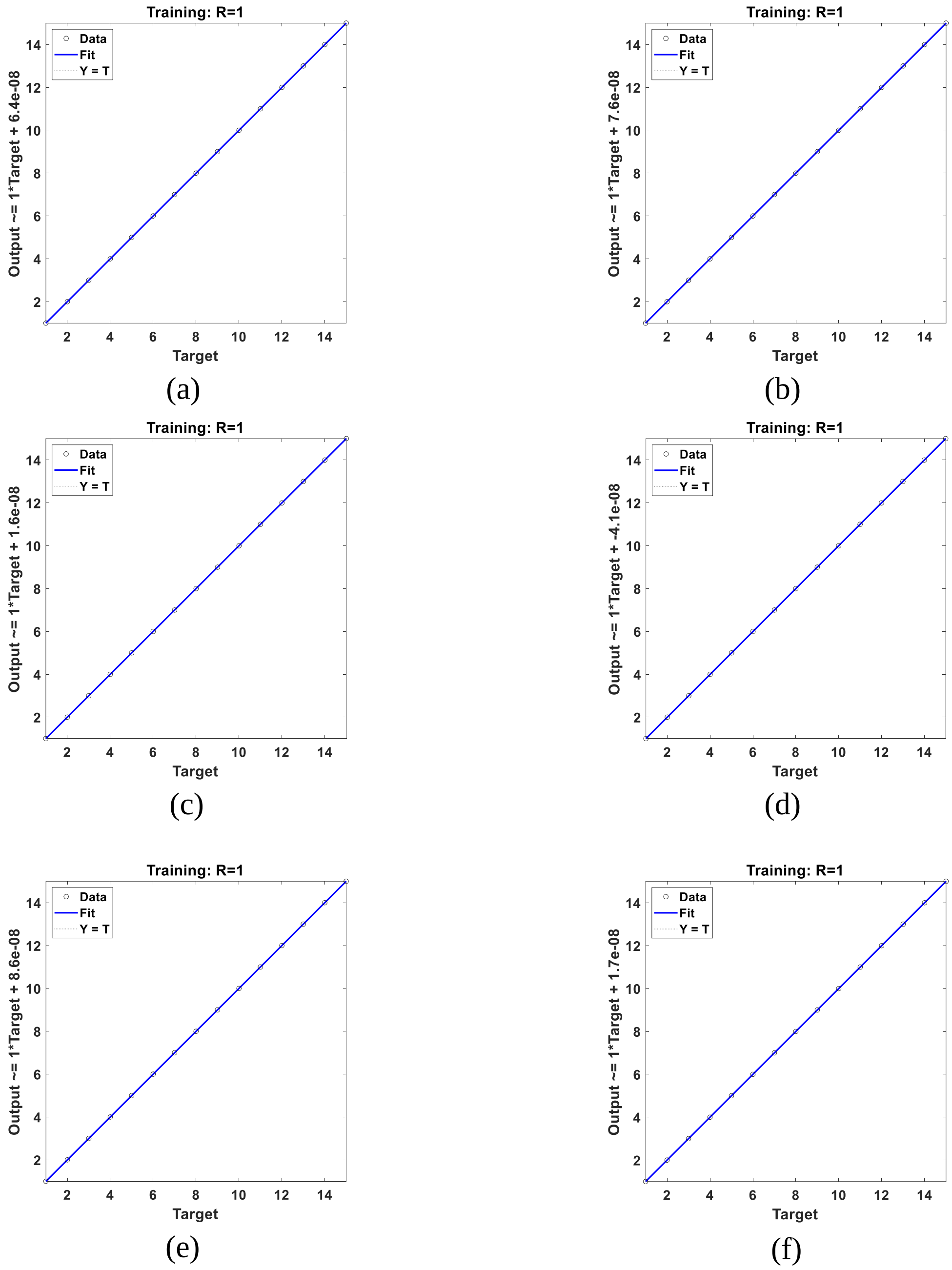


Figure 9: Linear regression curves of AMLPRs (a) for the first person, (b) for the second person, (c) for the third person, (d) for the fourth person, (e) for the fifth person and (f) for the sixth person

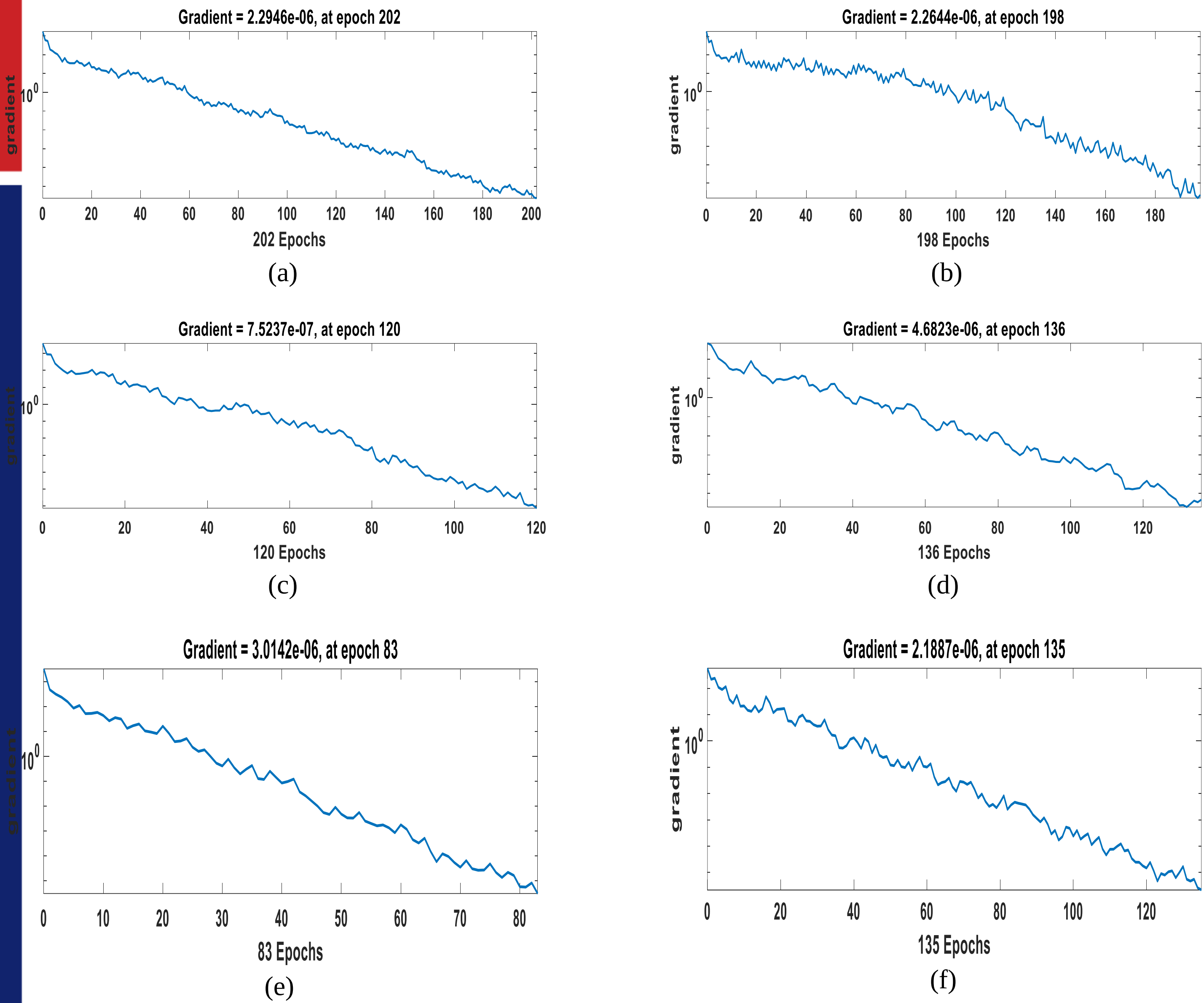


Figure 10: Gradient descent curves of AMLPRs (a) for the first person, (b) for the second person, (c) for the third person, (d) for the fourth person, (e) for the fifth person and (f) for the sixth person

4.4 Testing Phase Results

In the testing phase, augmentation is considered for signals of stuttered speeches. It is performed by adding random noises to the signals before implementing the EOLBP feature extraction and AMLPR. Here, two points can be highlighted. Firstly, using augmentation has the advantage of increasing data samples in the limited dataset. Secondly, it is expected that during testing phase speech signals with stuttering are handling noises, which come from environments and/or acquisition equipment. The augmentation operation to provide new stuttered signals for testing the performance of our anti-stuttering algorithm is applied by producing normally distributed random values along the sizes of speech signals with stuttering. Each augmented signal is processed by the suggested EOLBP feature extraction, then, the AMLPR is used to generate an output that refers to the anti-stuttering signal. Hence, the performance can be reached and calculated.

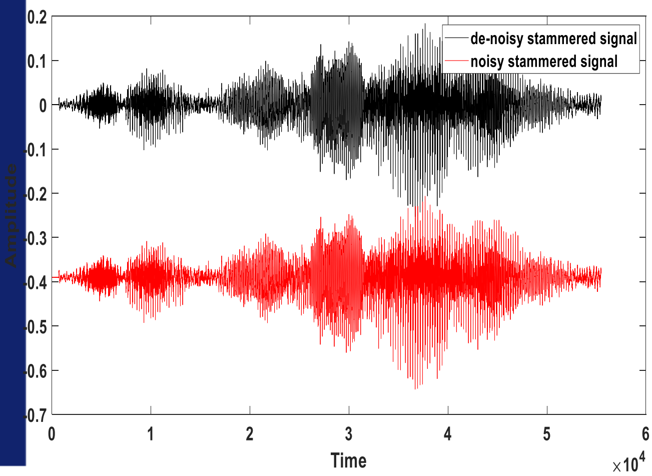
Figure 11 demonstrates examples of extracted features for 10% speech signals with stuttering, before and after augmentation.

For explaining the presentations of figure 14, the 100 values extracted features of each start stuttering sentence are not clear to show the differences between signals before and after augmentation process. So that a small percentage of 10% for any start of stuttering signal is used to present. Now, effects of augmentation are clearly shown.

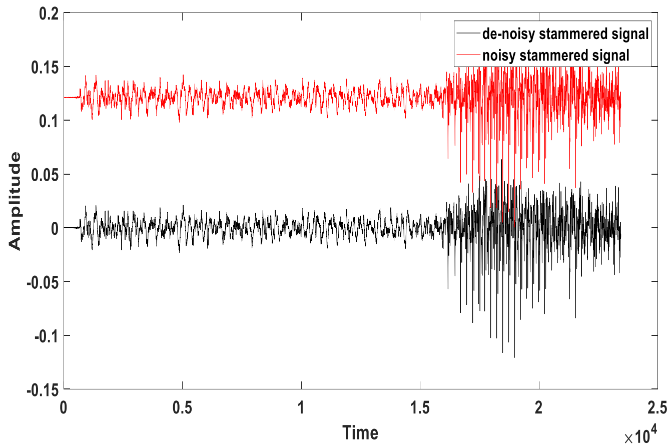
After testing the proposed anti-stuttering algorithm, completed pure speech signals can be obtained as given in figure 12.

For comparison purposes, the work in this paper can be compared with other techniques as in [40]–[50] after establishing fair basis. The performances of our AMLPR and other neural networks have been compared together by utilizing the same anti-stuttering procedure. Hence, the comparison would be fair. Table 1 shows the comparison results between our AMLPR and other methods of Radial Basis Neural Network (RBNN), Exact Radial Basis Neural Network (ERBNN), Cascade-Forward Neural Network (CFNN) and generalized Regression Neural Network (GRNN). A high regression accuracy of 99.44% have been attained by the AMLPR in the testing phase, with an error percentage equal to 0.56%.

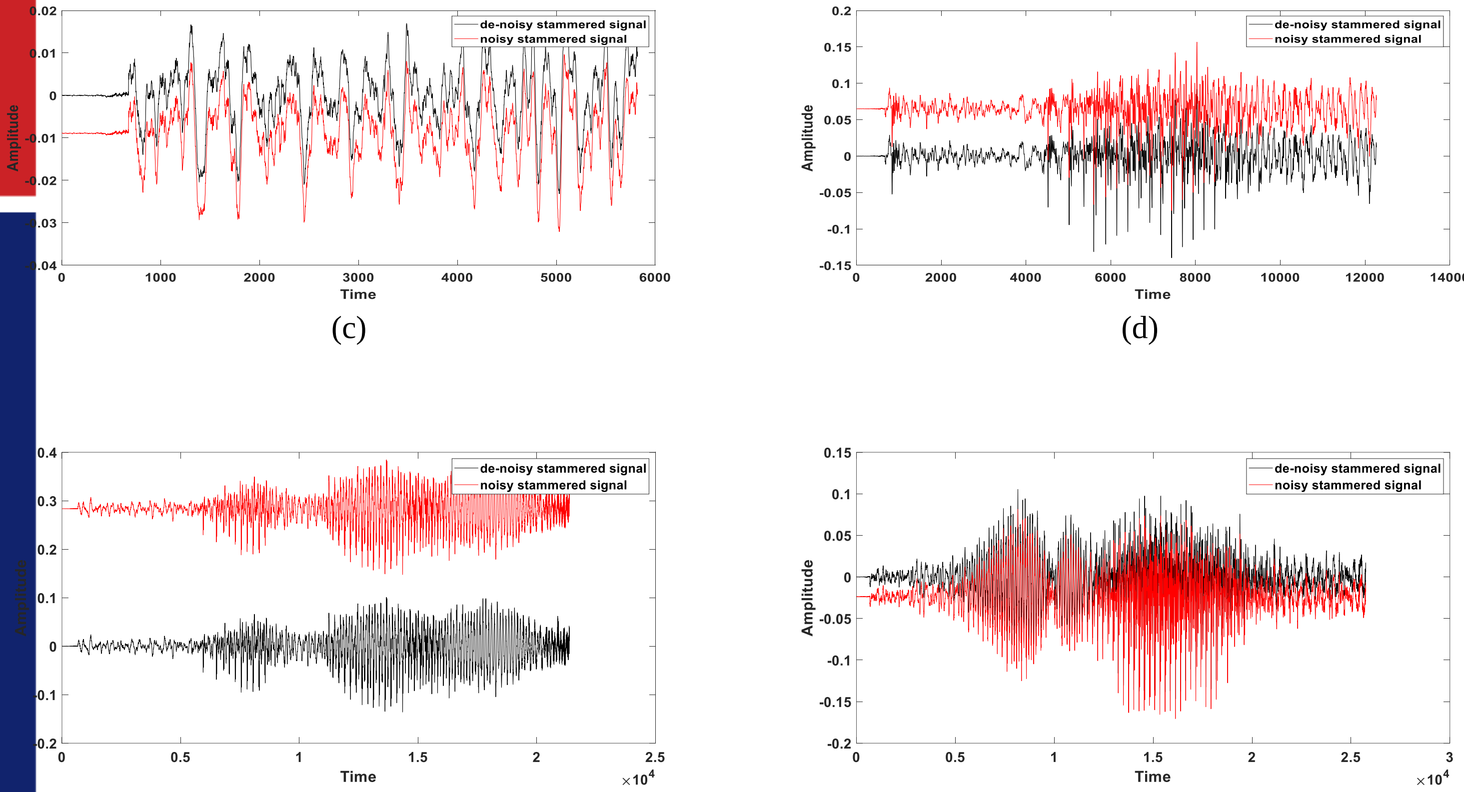
From this table, it can be observed that the GRNN is failed as it has achieved the accuracy of 6.66% with the error of 93.34%. The RBNN has obtained a moderate accuracy and error of 68.88% and 31.12%, respectively. The ERBNN and CFNN have attained comparable accuracies of 91.94% and 92.77% with comparable errors of 8.06% and 7.23%, respectively. Finally, it can be seen that the AMLPR has recorded the best accuracy of 99.44% with the best error of 0.56%. This yields that our adapted MLP method is so suitable for the anti-stuttering algorithm.



(a)



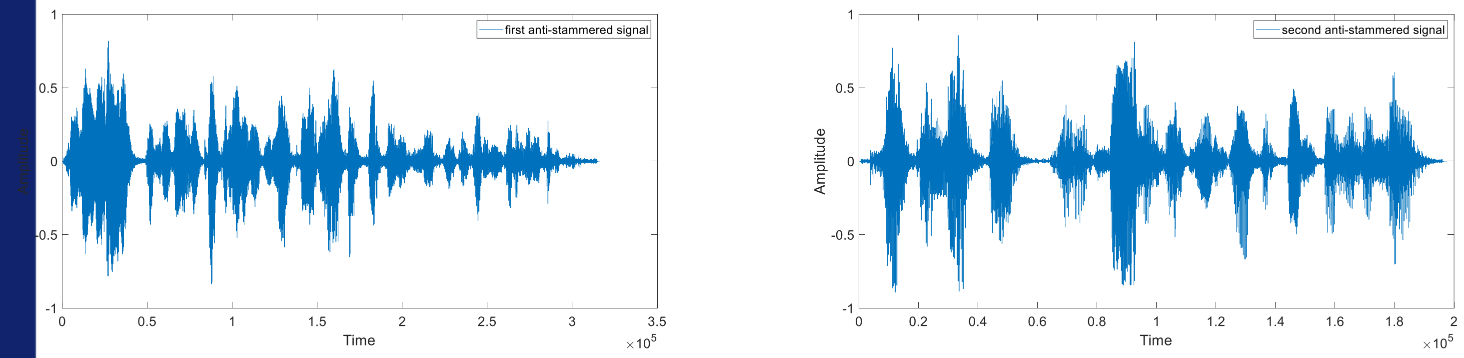
(b)



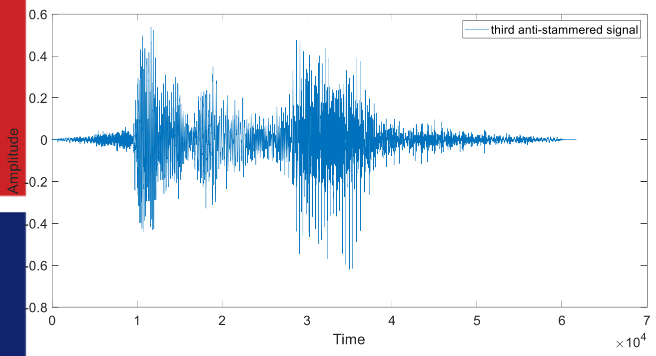
(c) (d)

(e) (f)

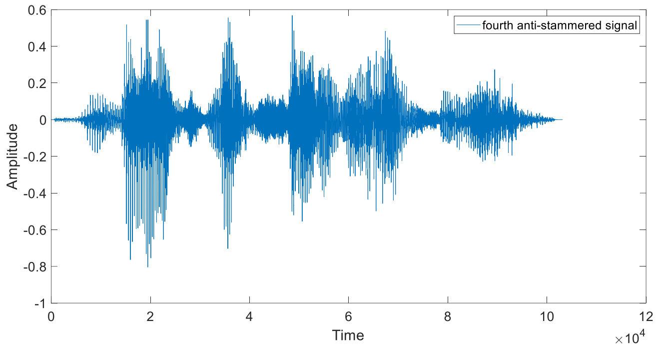
Figure 11: Examples of the first person's speech signals with stuttering before and after augmentation (a) for the first signal, (b) for the second signal, (c) for the third signal, (d) for the fourth signal, (e) for the fifth signal and (f) for the sixth signal



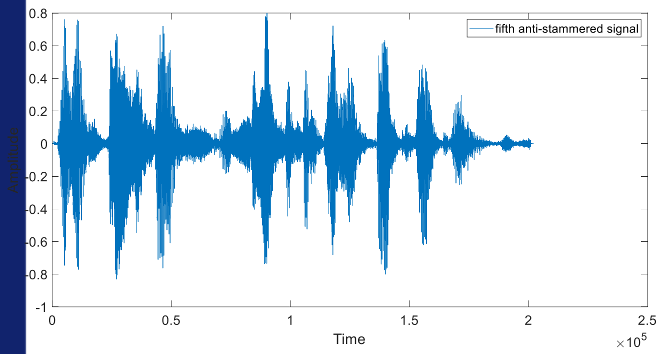
(a) (b)



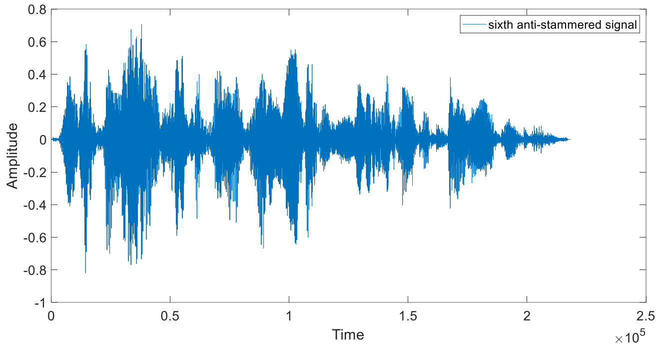
(c)



(d)



(e)



(f)

Figure 12: Examples of the first person's pure speech signals (without stuttering) (a) for the first signal, (b) for the second signal, (c) for the third signal, (d) for the fourth signal, (e) for the fifth signal and (f) for the sixth signal

Table 1: Comparison performances between our AMLPR and other neural networks by utilizing the same anti-stuttering procedure

Method	Accuracy	Error
GRNN	6.66%	93.34%
RBNN	68.88%	31.12%
ERBNN	91.94%	8.06%
CFNN	92.77%	7.23%
AMLPR	99.44%	0.56%

Fluent sentences that represent the feedback or choral speech can help in curing stutterers. Feedback speeches are delivered to a stuttered person's ear to get the choral effect as soon as possible. This is achieved by reducing the time of delivering the complete fluent sentence after providing just a small part of a stuttering sentence. Here, an advantage can be highlighted for the AMLPR in affording a complete fluent sentence to the stutterers as quickly as possible without needing to wait for the person who stutter to complete the entire sentence. This may contribute to decrease negative feelings of discomfort like stress and shame and let the person who stutter speak fluently.

5. CONCLUSION

In this study, we have investigated stuttering/stammering corrections for stuttered speeches. Efficient anti-stuttering algorithm was proposed. EOLBP was approached for extracting features from one-dimensional stuttered signals. AMLPR model was suggested to convert stuttering speech signals to anti-stuttering speech signals. In addition to employing the choral or feedback speech as a valuable mechanism for stuttering cure.

The choral speech mechanism can be used for stuttering cure by delivering anti-stammering signals to the stutterer's ear to get the effect of producing fluent speech signals by stutterer's mouth. By utilizing the starting of stuttered speech to produce a complete fluent sentence to a stutterer as quickly as possible without having to wait until finishing the whole sentence may contribute in decreasing negative feelings of stress and shame. This would let the stuttered person speak fluently and freely share his/her ideas with the other people.

FB dataset was utilized and prepared. It consisted of 720 stuttered speech signals of sentences, 360 signals were employed for the training phase and 360 augmented signals were applied for the testing phase. Furthermore, 15 anti-stuttered speech signals of sentences also from the FB dataset were exploited as targets. The results showed that the overall proposed anti-stuttering algorithm attained the best accuracy of 99.44% with the lowest error 0.56%.

REFERENCES:

[1] M. Hariharan, V. Vijean, C. Y. Fook, and S. Yaacob, "Speech stuttering assessment using sample entropy and Least Square Support Vector Machine," in 2012 IEEE 8th International Colloquium on Signal Processing and its Applications, IEEE, Mar. 2012, pp. 240–245. doi: 10.1109/CSPA.2012.6194726.

[2] R. Prabhu, P. P. Kumar, J. K. kumar, and D. E. L., "Speech Based Anti Stuttering Algorithm using Matlab," Int J Eng Adv Technol, vol. 9, no. 3, pp. 2152–2155, Feb. 2020, doi: 10.35940/ijeat.C5491.029320.

[3] S. Poornima and J. S. Kumar, "Feature Extraction and Signal Classification methods for Stuttering Speech Analysis," International Journal of Modern Computer Science and Applications, 2012, [Online]. Available: http://www.resindia.org

[4] Oliver Bloodstein Nan Bernstein Ratner Shelley B. Brundage, A Handbook on Stuttering, Fifth Edition., vol. 596 pages. San Diego, Plural Publishing, Inc., 1995.

[5] J. Kalinowski and T. Saltuklaroglu, "Choral speech: the amelioration of stuttering via imitation and the mirror neuronal system," Neurosci Biobehav Rev, vol. 27, no. 4, pp. 339–347, Jun. 2003, doi: 10.1016/S0149-7634(03)00063-0.

[6] J. Kalinowski and T. Saltuklaroglu, "Speaking with a mirror: engagement of mirror neurons via choral speech and its derivatives induces stuttering inhibition," Med Hypotheses, vol. 60, no. 4, pp. 538–543, Apr. 2003, doi: 10.1016/S0306-9877(03)00004-5.

[7] J. Kalinowski and V. N. Dayalu, "A common element in the immediate inducement of effortless, natural-sounding, fluent speech in people who stutter: 'the second speech signal,'" Med Hypotheses, vol. 58, no. 1, pp. 61–66, Jan. 2002, doi: 10.1054/mehy.2001.1451.

[8] V. K. Guntupalli, J. Kalinowski, T. Saltuklaroglu, and C. Nanjundeswaran, "The effects of temporal modification of second speech signals on stuttering inhibition at two speech rates in adults," Neurosci Lett, vol. 385, no. 1, pp. 7–12, Sep. 2005, doi: 10.1016/j.neulet.2005.05.010.

[9] M. Lincoln, A. Packman, and M. Onslow, "Altered auditory feedback and the treatment of stuttering: A review," J Fluency Disord, vol. 31, no. 2, pp. 71–89, Jan. 2006, doi: 10.1016/j.jfludis.2006.04.001.

[10] E. A. Antipova, S. C. Purdy, M. Blakeley, and S. Williams, "Effects of altered auditory feedback (AAF) on stuttering frequency during monologue speech production," J Fluency Disord, vol. 33, no. 4, pp. 274–290, Dec. 2008, doi: 10.1016/j.jfludis.2008.09.002.

[11] P. Howell, N. El-Yaniv, and D. J. Powell, "Factors Affecting Fluency in Stutterers when Speaking under Altered Auditory Feedback," in Speech Motor Dynamics in Stuttering, Vienna: Springer Vienna, 1987, pp. 361–369. doi: 10.1007/978-3-7091-6969-8_28.

[12] J. Kalinowski, J. Armson, A. Stuart, and V. L. Gracco, "Effects of Alterations in Auditory Feedback and Speech Rate on Stuttering Frequency," Lang Speech, vol. 36, no. 1, pp. 1–16, Jan. 1993, doi: 10.1177/002383099303600101.

[13] A. Stuart, S. Xia, Y. Jiang, T. Jiang, J. Kalinowski, and M. P. Rastatter, "Self-Contained In-the-Ear Device to Deliver Altered Auditory Feedback: Applications for Stuttering," Ann Biomed Eng, vol. 31, no. 2, pp. 233–237, Feb. 2003, doi: 10.1114/1.1541014.

[14] Y. P. S. C. O. P. V.Naveen Kumar, "Design and Implementation of Silent Pause Stuttered Speech Recognition System," International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 04, no. 03, pp. 1253–1260, Mar. 2015, doi: 10.15662/ijareeie.2015.0403012.

[15] A. C. Bortoletto, M. Minami, and C. S. Kurashima, "DSP altered feedback system for anti-stuttering applications," in 2016 IEEE International Symposium on Consumer Electronics (ISCE), IEEE, Sep. 2016, pp. 5–6. doi: 10.1109/ISCE.2016.7797363.

[16] G. Manjula, M. S. Kumar, and T. Kasar, "Identification and Validation of Repetitions/Prolongations in Stuttering Speech using Epoch Features," 2017. [Online]. Available: http://www.ripublication.com

[17] S. Alharbi, M. Hasan, A. J. H. Simons, S. Brumfitt, and P. Green, "A Lightly Supervised Approach to Detect Stuttering in Children's Speech," in Interspeech 2018, ISCA: ISCA, Sep. 2018, pp. 3433–3437. doi: 10.21437/Interspeech.2018-2155.

[18] F. Afroz and S. G. Koolagudi, "Recognition and Classification of Pauses in Stuttered Speech Using Acoustic Features," in 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN), IEEE, Mar. 2019, pp. 921–926. doi: 10.1109/SPIN.2019.8711569.

[19] G. Bhatia, B. Saha, M. Khamkar, A. Chandwani, and R. Khot, "Stutter Diagnosis and Therapy System Based on Deep Learning," arXiv preprint arXiv:2007.08003, Jul. 2020.

[20] B. Vaidianathan, S. Arulselvi, and B. Karthik, "Stuttered Speech Recognition And Classification Using Enhanced Kamnan Filter And Neural Network," in Proceedings of the First International Conference on Computing, Communication and Control System, I3CAC 2021, 7-8 June 2021, Bharath University, Chennai, India, EAI, 2021. doi: 10.4108/eai.7-6-2021.2308659.

[21] S. A. Sheikh, M. Sahidullah, F. Hirsch, and S. Ouni, “Robust Stuttering Detection via Multi-task and Adversarial Learning,” in 2022 30th European Signal Processing Conference (EUSIPCO), IEEE, Aug. 2022, pp. 190–194. doi: 10.23919/EUSIPCO55093.2022.9909644.

[22] “Stuttering and the Choral Effect | SpeechEasy | Stuttering Device.” Accessed: Sep. 27, 2023. [Online]. Available: https://speecheasy.com/stuttering-and-the-choral-effect/

[23] T. Saltuklaroglu, J. Kalinowski, M. Robbins, S. Crawcour, and A. Bowers, “Comparisons of stuttering frequency during and after speech initiation in unaltered feedback, altered auditory feedback and choral speech conditions,” Int J Lang Commun Disord, vol. 44, no. 6, pp. 1000–1017, Nov. 2009, doi: 10.3109/13682820802546951.

[24] M. Kiefte and J. Armson, “Dissecting choral speech: Properties of the accompanist critical to stuttering reduction,” J Commun Disord, vol. 41, no. 1, pp. 33–48, Jan. 2008, doi: 10.1016/j.jcomdis.2007.03.002.

[25] “FluencyBank.” Accessed: Jun. 01, 2023. [Online]. Available: https://fluency.talkbank.org/

[26] N. Bernstein Ratner and B. MacWhinney, “Fluency Bank: A new resource for fluency research and practice,” J Fluency Disord, vol. 56, pp. 69–80, Jun. 2018, doi: 10.1016/j.jfludis.2018.03.002.

[27] O. K. Toffa and M. Mignotte, “Environmental Sound Classification Using Local Binary Pattern and Audio Features Collaboration,” IEEE Trans Multimedia, vol. 23, pp. 3978–3985, 2021, doi: 10.1109/TMM.2020.3035275.

[28] L. Wang and D. C. He, “A New Statistical Approach for Texture Analysis,” Photogramm Eng Remote Sensing, pp. 61–66, 1990.

[29] T. Ojala, M. Pietikäinen, and D. Harwood, “A comparative study of texture measures with classification based on featured distributions,” Pattern Recognit, vol. 29, no. 1, pp. 51–59, Jan. 1996, doi: 10.1016/0031-3203(95)00067-4.

[30] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” IEEE Trans Pattern Anal Mach Intell, vol. 24, no. 7, pp. 971–987, Jul. 2002, doi: 10.1109/TPAMI.2002.1017623.

[31] M. Liu, Y. Tian, and Y. Ma, “Inner-Knuckle-Print Recognition Based on Improved LBP,” 2013, pp. 623–630. doi: 10.1007/978-3-642-34531-9_66.

[32] R. R. O. Al-Nima, F. H. Abdulraheem, and M. Y. Al-Ridha, “Using Hand-Dorsal Images to Reproduce Face Images by Applying Back propagation and Cascade-Forward Neural Networks,” in 2019 2nd International Conference on Electrical, Communication, Computer, Power and Control Engineering (ICECCPCE), IEEE, Feb. 2019, pp. 97–102. doi: 10.1109/ICECCPCE46549.2019.203755.

[33] Laurene V. Fausett, Fausett L.-Fundamentals of Neural Networks_ Architectures, Algorithms, and Applications (1994), 1'st. Financial Times Prentice Hall, 1994.

[34] Kavitha S, Varuna S, and Ramya R, “A comparative analysis on linear regression and support vector regression,” in 2016 Online International Conference on Green Engineering and Technologies (IC-GET), IEEE, Nov. 2016, pp. 1–5. doi: 10.1109/GET.2016.7916627.

[35] G. A. F. Seber and A. J. Lee, Linear Regression Analysis. Wiley, 2003. doi: 10.1002/9780471722199.

[36] A. I. Khuri, "Introduction to Linear Regression Analysis, Fifth Edition by Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining," International Statistical Review, vol. 81, no. 2, pp. 318–319, Aug. 2013, doi: 10.1111/insr.12020_10.

[37] A. S. I. Al-Obaidi, R. R. O. Al-Nima, and T. Han, "Interpreting Arabic sign alphabet by utilizing a glove with sensors," Int J Health Sci (Qassim), pp. 7170–7184, Aug. 2022, doi: 10.53730/ijhs.v6nS6.12018.

[38] W. R. Fathel, A. S. I. Al-Obaidi, M. A. Qasim, and M. M. M. Al-Hatab, "Skin Cancer Detection Using K-Means Clustering-Based Color Segmentation," Texas Journal of Engineering and Technology, vol. 18, pp. 46–52, 2023.

[39] "Teaching with FluencyBank." Accessed: May 25, 2023. [Online]. Available: https://fluency.talkbank.org/teaching/AWS.html

[40] R. R. O. Al-Nima, T. Han, T. Chen, S. Dlay, and J. Chambers, "Finger texture biometric characteristic: a survey," arXiv preprint arXiv:2006.04193, 2020.

[41] M. T. S. Al-Kaltakchi, R. R. O. Al-Nima, M. Alfathe, and M. A. M. Abdullah, "Speaker verification using cosine distance scoring with i-vector approach," in 2020 International Conference on Computer Science and Software Engineering (CSASE), IEEE, 2020, pp. 157–161.

[42] A. S. Anaz, R. R. O. Al-Nima, and M. Y. Al-Ridha, "Multi-encryptions system based on autoencoder deep learning network," Solid State Technology, vol. 63, no. 6, pp. 3632–3645, 2020.

[43] S. M. M. Najeeb, R. R. O. Al-Nima, and M. L. Al-Dabag, "Reinforced Deep Learning for Verifying Finger Veins.," International Journal of Online & Biomedical Engineering, vol. 17, no. 7, 2021.

[44] R. R. O. Al-Nima, M. K. Jarjes, A. W. Kasim, and S. S. M. Sheet, "Human identification using local binary patterns for finger outer knuckle," in 2020 IEEE 8th Conference on Systems, Process and Control (ICSPC), IEEE, 2020, pp. 7–12.

[45] M. R. Khalil, M. S. Majeed, and R. R. Omar, "Personal identification with iris patterns," AL-Rafidain Journal of Computer Sciences and Mathematics, vol. 6, no. 1, pp. 13–26, 2009.

[46] R. R. O. Al-Nima, N. A. I. Al-Obaidy, and L. A. Y. Al-Hbeti, "Segmenting finger inner surface for the purpose of human recognition," in 2019 2nd International Conference on Engineering Technology and its Applications (IICETA), IEEE, 2019, pp. 105–110.

[47] M. T. S. Al-Kaltakchi, R. R. O. Al-Nima, M. A. M. Abdullah, and H. N. Abdullah, "Thorough evaluation of TIMIT database speaker identification performance under noise with and without the G. 712 type handset," Int J Speech Technol, vol. 22, pp. 851–863, 2019.

[48] R. R Al-nima, "Design a biometric identification system based on the fusion of hand geometry and backhand patterns," IRAQI Journal Of Statistical Sciences, vol. 10, no. 1, pp. 169–180, 2010.

[49] R. R. Al-Nima, "Human authentication with earprint for secure telephone system," Iraqi Journal of Computers, Communications, Control and Systems Engineering IJCCCE, vol. 12, no. 2, pp. 47–56, 2012.

[50] R. R. Al-Nima, M. Y. Al-Ridha, and F. H. Abdulraheem, "Regenerating face images from multi-spectral palm images using multiple fusion methods," Telkomnika (Telecommunication Computing Electronics and Control), vol. 17, no. 6, pp. 3110–3119, 2019.