JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 1731 - 1740
https://publishoa.com
ISSN: 1309-3452

Audio Signal Based Stress Recognition System using AI and Machine Learning

Megha Gupta
Department of Computer Engineering NHITM, Mumbai University
Thane, India
meghagupta@nhitm.ac.in

Dr. Shubhangi Vaikole
Department of Computer Engineering
DMCE, Mumbai University
Airoli, India
slv.cm.dmce@gmail.com

ABSTRACT

The detection of stress from speech signals has recently received a lot of interest. Individual speech is a verbal means for people to communicate with one another. The human speech reflects the speaker's mental state. To ensure that the person is in a healthy state of mind, proper classification of these speech signals into stress categories is essential. Speech is frequently used to detect whether a person is in a stressful or routine scenario. These can lead to the reliable classification of speech signals into separate stress types, showing that the individual is in a healthy state of mind. In this paper, stress identification and classification algorithms are built using machine learning (ML), artificial intelligence (AI), and Mel frequency cepstral coefficient (MFCC) feature extraction methodologies. Because most current stress indicators are intrusive, requiring samples from patients' bodies, this study was done to identify methods to detect stress without introducing instruments into the body. This study illustrates how stress can be detected by using speech signal analysis approaches.

Keywords— AI, Stress, Speech, emotion, machine learning, MFCC, RNN, LSTM

I. INTRODUCTION

Stress faced by humans can be defined as an imbalanced mental condition [1] that occurs when an individual is not able to cope with the demands imposed on them [2]. Because of its negative consequences, detecting stress is a major topic among mental health professionals and researchers. [3]. It's been used to improve polygraph tests [4], emergency call recognition [5], and HCI, among other things[6]. As a consequence of the responsibilities and expectations placed on the people, they tend to become stressed. The stress may be exacerbated if they realize how difficult it will be to adapt to a new environment. [7]. Episodic stress is defined as a pattern of upsetting events that occur more frequently but then fade away. It has been connected to hurried and stressful lives [8]. The most damaging sort of stress is Chronic stress, which arises when stressors, such as family issues, job-related strain, or neediness, are present for an extended period [9]. It's crucial to assess and manage stress soon, when the stress level is either acute or episodic, to prevent it from reaching its peak level and to help mitigate the risks [10]. During the last 20 years, various domains such as forensics, human-machine interfaces, healthcare, emergency services, smart settings, teaching-learning, and real-time situations have shown a strong interest in stress detection. [11]. Stress is a multidimensional phenomenon that has a variety of physical and psychological repercussions. Digestion functions [12], blood pressure [13], skin issues [14], dietary patterns [15], effectiveness [16], judgement [17], and general wellness [6] are all affected by stress. Telomeres, which are structures on the ends of chromosomes, shrink in response to stress, limiting the creation of new cells [18]. According to "The Global Burden of Disease" [19], a World Health Organization (WHO) report issued in 1996, heart disease will be replaced as the second most common disability by depression, stress, and mental illnesses.

A. Problem Definition

Stress is a major concern of society in today's era. According to Cigna Corporation's worldwide well-being research (2019), 82% of Indians are coping with stress because of jobs, health-related issues, and economic difficulties. Suicidal behavior is frequently preceded by stressful conditions such as marital and family problems, legal concerns, and job loss. For decades, researchers have tried to recognize stress using a variety of methods to reduce the number of suicides and depression cases. Building a method to recognize stress before it turns long-lasting is crucial. The suggested research work presents a method that uses speech signals to detect human mental stress.

B. Proposed System Overview

For various applications and requirements, several stress detection approaches have been proposed. However, there is room for improvement in terms of the exactness of the outcomes. A great amount of research has been done on the fundamental emotions, there has been less information on how to detect and analyze them. Based on the literature review and consequent research questions revealed from the findings, the current effort aims to work towards the creation of a system for stress recognition. Deep learning algorithms like RNN and LSTM are used for enhancing the results of the proposed system.

C. *Organization of the Paper*

This paper is organized into 5 sections. In section 1, the introduction of this research is given along with the significance of stress detection. Then, the problem statement, and overview of the proposed method are presented. In section 2, a literature survey on stress detection using speech signal is presented and previous research on stress detection systems are studied. Section 3 comprises background, dataset analysis, and the proposed stress detection system along with the selection of features. This is followed by an experimental result, a performance analysis of the proposed method with the traditional classifiers together with a discussion of future work in section 4. The concluding remarks and summary of achievements are presented in the final section.

II. LITERATURE REVIEW

Conventional methods of assessing stress include interrogating the person and questioning about stress to gain better insights into their condition and perceiving the expressions on the person's face. People under stress might widen their pupils, modify the contour of their eyebrows, and increase the rate at which they blink their eyes. [20] are limited as there's a high likelihood that distressing situations may go unnoticed. The modern high-level technique comprises guiding life stress interviews using instruments such as the "Stress and Adversity Inventory", "UCLA Life Stress Interview", and "Life Events and Difficulties Schedule" [21]. In turn, the most popular method is to administer brief self-report questionnaires such as the Perceived Stress Scale [22]. Item specificity and validity are typically lacking in self-report questionnaires, and interview-based examinations can be laborious as well as expensive [23]. Furthermore, both methods are retrospective and are subject to (often unmeasured) levels of cognitive bias and social desirability, which could affect the final evaluations' authenticity, reliability, and validity [22]. Earlier stress detection methods are investigated in depth in this paper, with a focus on the ways by which earlier works tackled a few of the concerns. It discusses the spurs causing stress, the stress measuring approaches, the information collected, and the use of ML techniques in these studies. Researchers have created a range of approaches for stress detection and studying the factors that trigger it to reduce long-term mental stress [24] issues. H. Lin et al. proposed automatic stress detection techniques based on cross-media microblog data [25]. The data is acquired by looking at components of tweet texts and representations, such as tweeted photographs, comments, and favorites, in the form of overlap features. With the help of the built-in Neural Network, the authors were able to understand the types of stress types based on the attributes provided.

A. *Speech Signal for Stress detection*

Voice signals can be used to detect stress in a variety of ways. It is used in psychology to track the various degrees of stress in patients with various stress disorders and to deliver the relevant therapy. Aviators, marine divers, and armed captains confronting police administration may have their stress levels monitored to assist determine a system's safety and security. Stress detection is also useful for speaker recognition, uncovering deception, and also for identification of threat calls in a few criminal situations [26]. When preparing to speak, people should consider which order of words would successfully convey the anticipated message. Stress can alter these assessments, causing variations in language, syntax, and speech pace. In the future, these could be used as audio stress indicators. [27, 28]. Stress, furthermore, brings about extra changes. The body modulates the tension of various muscles to produce sound waves by pushing air out of the vocal tract and via the vocal folds [29]. By increasing tension of muscles and rate of breathing, stress modifies the method of generation of speech and hence the sound of speech [30,31]. A speech-based stress recognition application for android [24] detects stress levels from human speech. This program was created in a variety of environments, with a variety of speakers and scenarios.

Kevin Tomba et al. [32] performed stress detection using various datasets. SVM (Support Vector Machines) and ANN (Artificial Neural Network) algorithms were employed. Speech analysis features such as mean energy, mean intensity, and MFCCs have been proved to be advantageous.

RNN classifier was used by the N.P. Dhole et. al.. The researchers have worked on [11] BERLIN and HUMAINE Datasets. They also used the RNN on real datasets created by the Audacity software. The efficiency percentage was not calculated, even though it worked well for the audio signal.

Mansouri et al. [33] used Wavelet and Neural Network to create and implement an emotion recognition system using speech signals. The accuracy was determined to be acceptable. The procedure, however, is laborious, and recognition of stress was not considered.

1) Analysis

The detection of stress using speech signals promises to be a fascinating subject of study, and a review of existing studies would be beneficial to future research. Table I summarises several research that all focus on the same topic, the databases employed, the benefits of the approaches, and the scope of development.

TABLE I. REVIEW OF PAST WORK IN THE SAME FIELD OF RESEARCH

Title	**Dataset used**	**Technique**	**Pros of Technique**	**Scope for Improvement**

"Stress Detection Through Speech Analysis" Kevin Tomba et. al. (ICETE 2018) [32]	RAVDESS, the Berlin Emotional Database (EMO-DB), and the Keio University Japanese Emotional Speech Database (KeioESD)	SVMs and ANNs have been chosen.	For speech analysis, mean energy, mean intensity, and MFCCs have proven to be useful features.	Works only on audio.
"Study of Recurrent Neural Network Classification of Stress Types in Speech Identification" N.P. Dhole, S.N. Kale (IJCSE 2018) [11]	BERLIN and HUMAINE Datasets	Recurrent Neural Network	Also works on real datasets created using Audacity software	Efficiency percentage not calculated. Works only on audio.
"Designing and Implementing of Intelligent Emotional Speech Recognition with Wavelet and Neural Network" Mansouri et. al. (IJACSA 2016) [33]	EMO-DB and SAVEE	ANN classifier	Accuracy is good	Time-consuming method. Stress detection was not considered
"Stress Detection from Speech and Galvanic Skin Response Signals" H. Kurniawan et. al. (IEEE 2013) [34]	Own Dataset	SVM	Accuracy is Good	Count of Stress classes and Stress conditions were not reported
"Detecting stress and depression in adults with aphasia through speech analysis" S. Gillespie et. al. (IEEE 2017) [35]	APHASIA DATASET	SVR	Not Reported	Not Reported

III. Speech Signal-based Stress Recognition System

A. Background

Human emotions play an important role in everyday human interactions. This is required for sensible as well as knowledgeable decision-making. It assists us in matching and grasping the emotions of others, which enables us to communicate our feelings and make suggestions to others. According to research, human social connection is heavily influenced by emotion. Feelings and emotions indicate a great deal of information about one's psychological state. The non-invasive aspect of using speech signals for stress detection is critical for both users and the creation of a large database. A range of sentiments presented by the speech of humans changes depending on the circumstances. The seven archetypal emotions that comprise all of these sensations are calm, neutral, sadness, fear, disgust, surprise, happy, and anger [36], [37]. Recognition of the human emotional state by retrieving the information from the audio signal is an essential topic that must be handled by the proposed system. The audio signal comprises contextual information concerning space and time, and also frame dependencies.

To improve the efficacy and correctness of the feature extraction process, first, the pre-processing of the input speech signal is done. Pre-processing processes include filtering, framing, and windowing. Filtering is a method to decrease noise in an audio signal induced by external variables or during recording. The pre-emphasis filter's goal is to boost the speech signal's intensity in higher frequencies that are suppressed during speech signal synthesis in the vocal tract. After a pre-processing stage, the basic voice quantities such as pitch, energy, and formants are extracted from the spoken stream.

These raw voice/speech quantities are sent to the feature extraction module. The features F1, F2,... Fn are extracted at the feature extraction stage which are then passed on to the classifier. The classifier then processes the features. Eventually, this classifier will be able to discern a person's emotions[38][39].

The fully-connected layers take the output and use a SoftMax layer to help them make their ultimate decision. The conventional Speech Emotion Detection (SED) system is depicted in Figure 1.

B. Dataset Analysis

In the Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) dataset, we show that our LSTM-RNN-based model for Stress identification from the speech signal outperforms other approaches, obtaining state-of-the-art results to the best of our knowledge. We chose RAVDESS over other datasets because of this.

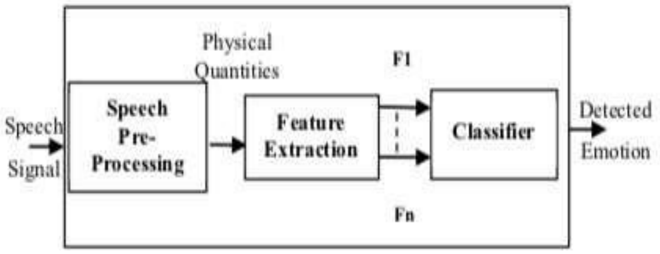

Fig. 1 Speech Emotion Detection system

a) Dataset Gathering

It's challenging to find a dataset that has stressed or unstressed versions of the same speaker. We used RAVDESS, a multimodal dataset that combines emotive audio and audiovisuals for stress-related research.

b) Dataset Specifics

There are 7356 files in the RAVDESS dataset totaling 24.8 GB in size. The 24 professional actors (12 females, 12 males) in this gender-balanced and validated dataset each vocalize two lexically matched lines in a neutral North American accent. Calm, happy, sad, angry, fearful, surprise, and disgust are some of the different expressions used in speech. Normal and Strong are the two levels of emotional intensity for each expression, as well as a neutral expression. Audio-Video, Audio-Only, and Video-Only are the three modality forms available in all situations. 247 untrained study volunteers from North America provided ratings. A total of 72 people took part in the test-retest investigation. Both inter-rater reliability and emotional validity were found to be quite good. To aid researchers in the selection of stimuli, corrected accuracy, and composite "goodness" scores are provided. 1440 files are included in the Speech file (Audio Speech Actors 01-24.zip, 215 MB). 60 trials per actor multiplied by 24 actors equal 1440. In total, there are 7356 files each with its name in the RAVDESS collection. A seven-part numerical filename has been used. These IDs, as illustrated in figure 2, define the stimulus features.

Identifier	Coding description of factor levels
Modality	01 = Audio-video, 02 = Video-only, 03 = Audio-only
Channel	01 = Speech, 02 = Song
Emotion	01 = Neutral, 02 = Calm, 03 = Happy, 04 = Sad, 05 = Angry, 06 = Fearful, 07 = Disgust, 08 = Surprised
Intensity	01 = Normal, 02 = Strong
Statement	01 = "Kids are talking by the door", 02 = "Dogs are sitting by the door"
Repetition	01 = First repetition, 02 = Second repetition
Actor	01 = First actor, . . ., 24 = Twenty-fourth actor

Fig. 2 Filename Identifiers

02-01-06-01-02-01-12.mp4 contains the following information: (02) indicates that Modality is Video-only (01) conveys that the vocal channel is Speech (06) tells us that the audio has Fearful emotion (01) signifies Normal Emotional intensity(02) communicates that the audio consists of vocals of "Dogs are sitting by the door" statement (01) indicates first Repetition (12) audio is of the 12th Actor who is a Female.

Figure 3 depicts the flowchart for RAVDESS development and validation

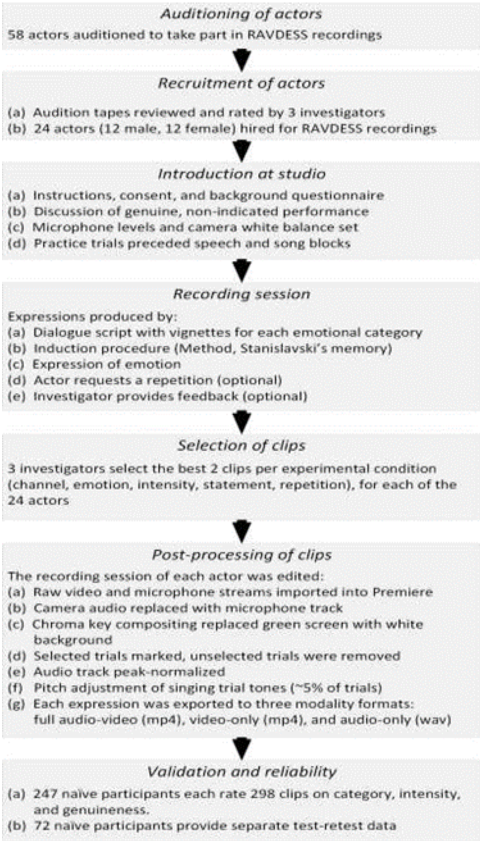

Fig. 3 RAVDESS construction and validation flowchart

Only the audio portion of the dataset will be considered, not the video piece. All actors (01-24) have audio-only files available as two separate zip files (200 MB each): The speech file (Audio_Speech_Actors_01-24.zip, 215 MB) contains 1440 files: 60 trials per actor x 24 actors = 1440. For the real-time dataset, 10 speech samples were recorded with one trial each to validate the classification accuracy.

C. MFCC (Mel-Frequency Cepstral Coefficients)

The cepstrum contains information about the rate of change in different spectrum bands. In sound processing, MFC (Mel Frequency Cepstrum) is a depiction of a sound's short-term power spectrum. It is built on a nonlinear Mel frequency scale and linear cosine transform of a log power spectrum. Mel-frequency cepstral coefficients (MFCCs) are the coefficients that combine to form an MFC [35]. Cepstral picture from the audio sample makes up MFCC (a nonlinear "spectrum-of-a-spectrum"). On a Mel scale, frequency bands of MFC are spaced uniformly. MFC's evenly spaced frequency bands are closer to the response of the auditory mechanism of humans as compared to the cepstrum's frequency bands which are spaced linearly. Figure 4 shows feature extraction using MFCC.

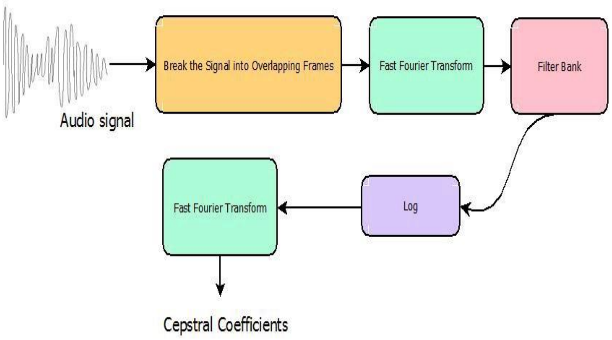

Fig. 4 MFCC Features Extraction

D. The Proposed System for Human Stress Recognition

The proposed system is a Human stress recognition system using deep learning. The method employs voice parameters like MFC as input. The system uses a novel algorithm by combining LSTM (Long Short term memory) with RNN. The deep learning approach is used to determine human mental stress. The flowgraph for the proposed methodology is depicted in figure 5.

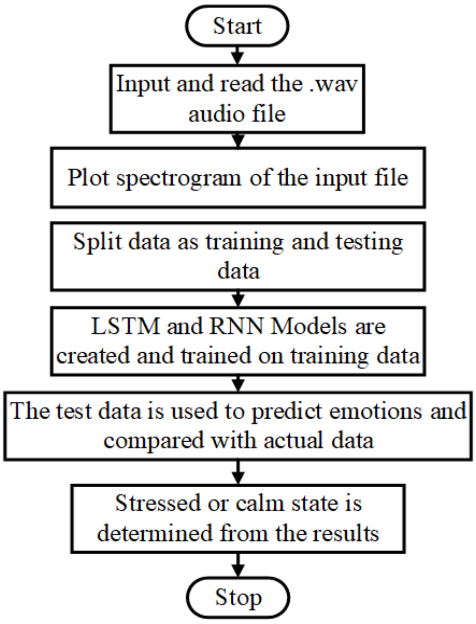

Fig. 5 Flowgraph for Stress recognition using Speech Signals.

In the proposed approach, two closely related ML algorithms viz., RNN and LSTM are cascaded together as shown in the flowchart (Figure 6). Cascading improves the convergence time of the combined algorithm. Here, speech signals are pre-processed and filtered using MFCC. The features of the input signal are extracted at this stage. These features are sent to 4 neurons RNN and 10 neuron LSTM working in parallel to each other. The RNN module which does not have a cell state generates the required labels for these features while the LSTM module is used for emotion prediction only. The LSTM module receives labels from the RNN module as its first input and the extracted features from MFCC as its second input. This combined approach reduces the size of the LSTM module by reducing the number of neurons required for emotion prediction. E.g., a 40-neuron LSTM module is replaced by a 10-neuron LSTM module with a 4-neuron RNN module to achieve the same result at a faster rate. Moreover, to prevent the model from overfitting, the LSTM module employs a dropout layer by randomly setting other edges of the hidden layer to zero. This reduces the convergence time to about 3/4 of the traditional approach.

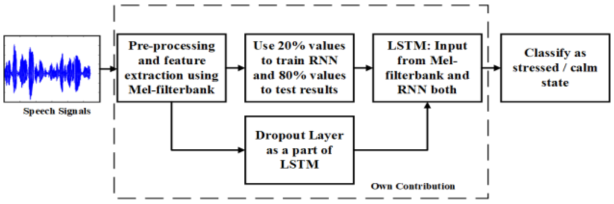

Fig. 6 Speech Signals based Human Stress detection

The flowchart for the LSTM with the dropout layer used in this method is illustrated in figure 7.

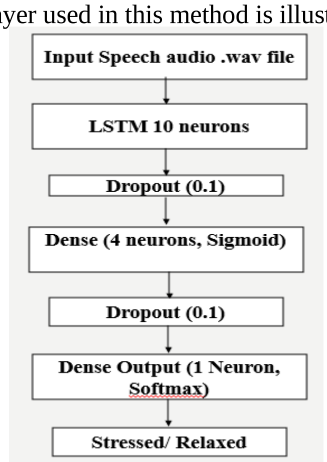

Fig 7. Flowchart for LSTM with dropout layer used in this method

D. Selection of Features

Humans use pitch, articulation rate, energy, and Mel frequency cepstral coefficients(MFCC) to discern between different moods. A complete graph depicting the quality of speech based on 6 different emotions is available (Banse and Scherer, 1996). Using a combination of mean energy and mean intensity, this table distinguishes the five emotions happy, disgust, sad, fear/anxiety (stress), and anger. Because they perform better in similar situations, MFCCs are employed for speech analysis. As a result, emotions are classified using the MFCCs, mean energy, and mean intensity. [40]

1) Analysis of Frequency Distribution concerning Time

A signal processing method called Time-frequency analysis examines a signal in both the time-frequency domains at the same period by employing different time-frequency representations. An appropriate threshold for separating the object signal from the background noise is chosen to evaluate the legitimacy of multiple noise removal algorithms. A scientific and engineering metric that compares a signal's strength to the quantity of noise in the surroundings is the Signal to Noise ratio (SNR). The normalized noise magnitude is depicted in Figure 8.

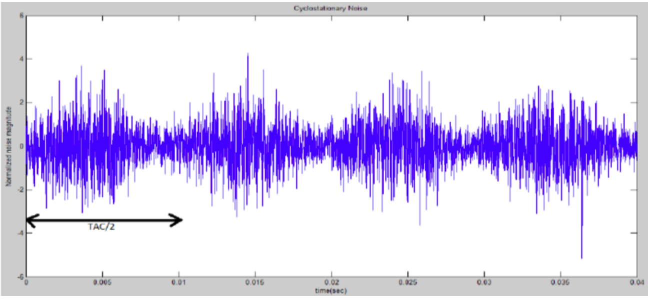

Fig. 8 Noise magnitude (normalized)

The SNR measures how strong a signal is in comparison to how loud background noise is. Decibels represent signal-to-noise ratios (dB). To express signals with a wide dynamic range, the logarithmic decibel scale is commonly utilized.

IV. RESULTS AND DISCUSSIONS

The proposed system is carried out using the RAVDESS dataset. For both training and testing, 1440 spoken utterances of twelve male and twelve female actors having 6 different emotions anger, anxiety, disgust, neutral, and sadness are considered. Also, 10 speech utterances were recorded at our end with happy and sad emotions to test the validity of the proposed algorithm. The proposed LSTM+RNN classification method with dropout layer was compared with the random forest method and support vector machines for analysis. The Hamming windows for actor 1 and actor 2 are shown in figure 9 and figure 10 respectively.

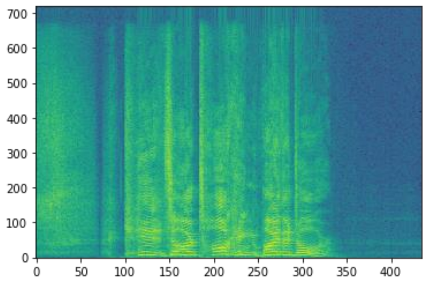

Fig. 9 Hamming Window for audio file from Actor 1

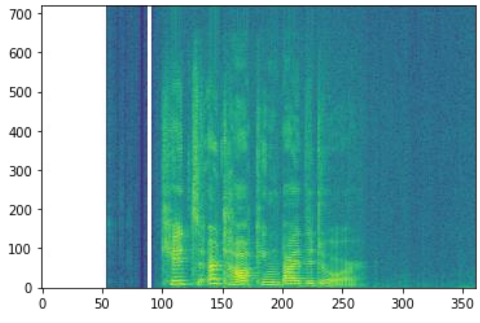

Fig. 10 Hamming Window for audio file from Actor 2

The confusion matrix for the RAVDESS dataset using the proposed novel RNN- LSTM algorithm is shown in figure 11.

Emotion Class	Anger	Calm	Fearful	Happy	Sad
Anger	**98.25**	0.00	0.00	0.00	0.9
Calm	0.1	**94**	0.08	0.5	0.00
Fearful	0.55	0.07	**90**	1.2	0.4
Happy	0.00	0.02	0.25	**94.25**	0.00
Sad	0.02	0.00	0.4	0.7	**91.5**
Overall Accuracy					93.6

Fig. 11 Confusion Matrix for RAVDESS dataset using the proposed algorithm

The binary classification for mental stress could be done by using a 10-neuron layer using the proposed LSTM + RNN algorithm using dropout. The SoftMax function is used to improve selectivity for binary state classification. The classification accuracy of the proposed system is compared with the other classifiers. The performance analysis of the proposed system on the RAVDESS dataset is shown in table II.

TABLE II. PERFORMANCE ANALYSIS OF THE PROPOSED SYSTEM ON RAVDESS

Classification Accuracy %	SVM	RNN	MFCC (LSTM)	MFCC(LSTM +RNN) Proposed Algorithm
Angry	**52.6**	**67.9**	**85.9**	**98.25**
Calm	**34.5**	**62**	**86.7**	**94**
Fearful	**44.3**	**74**	**83.6**	**90**
Happy	**57.9**	**70.3**	**88.4**	**94.25**
Sad	**68.3**	**60.3**	**82.3**	**91.5**
Overall Accuracy	**51.52**	**66.9**	**85.38**	**93.6**

Accuracy for prediction for the proposed algorithm for the RAVDESS dataset is 93.6%. Comparatively, obtained results are 15-50% better than other algorithms.

The performance analysis of the proposed system on the real-time dataset is shown in table III. The accuracy of prediction for the proposed algorithm for real-time data is 65%. This is a better result than the algorithms used for comparison. However, by recording the speech audio in a professional environment and proper selection of neurons and values for drop-out layers, the accuracy of the proposed algorithm can be improved for real-time recorded files.

TABLE III. PERFORMANCE ANALYSIS OF THE PROPOSED SYSTEM ON A REAL-TIME DATASET

Classification Accuracy %	RNN	SVM	MFCC (LSTM)	MFCC(LSTM+ RNN)
Happy	**30**	**20**	**30**	**60**
Sad	**50**	**50**	**50**	**70**

Overall Accuracy	**40**	**35**	**40**	**65**

A. Key Aspects

The proposed LSTM + RNN classifier method with a drop-out layer was implemented using Python programming language for speech signal processing-based mental state detection. The accuracy of the proposed algorithm is 94% and 65% for the RAVDESS dataset and real-time dataset respectively which is better than traditional methods used for classification by 15-50%.

V. CONCLUSION

Stress is a negative emotional condition. that alters biochemistry, physiology, and behavior. Stress is a huge problem in today's world, and issues at work e.g. cumbersome tasks and the necessity to adapt to constant change perpetuate the problem. People suffer from medical problems as a result of chronic stress, while businesses suffer significant financial losses. As a result, monitoring stress levels is critical for detecting stress early on and avoiding serious long-term consequences. The need to manage chronic stress in people gave rise to the concept of stress detection. Our method is an excellent starting point for detecting stress and improving the quality of life. For the RAVDESS dataset, our proposed novel LSTM-RNN algorithm could outperform other algorithms, However, robust tuning of the proposed method and using more data for training can further improve the result accuracy. In the case of a Real-time dataset, we got a better result than other classification algorithms. However, by recording the speech audio in a professional environment and proper selection of neurons and values for drop-out layers, the accuracy of the proposed algorithm can be improved for real-time recorded files.

REFERENCES

[1] G. P. Chrousos, “Stress and disorders of the stress system,” Nat.Rev. Endocrinology, vol. 5, pp. 374–381, 2009.
[2] J. M. Koolhaas, A. Bartolomucci, B. Buwalda, S. F. D. Boer,G. Flugge, S. M. Korte, P. Meerlo, R. Murison, B. Olivier, P. Palanza, G. Richer-Levin, A. Sgoifo, T. Steimer, O. Stiedl, G. V. Dijk,M. Wohr, and E. Fuchs, “Stress Revisited: A critical evaluation of the stress concept,” Neurosci. Biobehavioral Rev., vol. 35, pp. 1291–1301, 2011.
[3] R. Lazarus, Stress and Emotion: A New Syhthesis. New York, NY, USA: Springer, 2006.
[4] I. Pavlidis and J. Levine, “Thermal image analysis for polygraph testing,” IEEE Eng. Med. Biol. Mag., vol. 21, no. 6, pp. 56–64, Nov./ Dec. 2002.
[5] L. Lefter, L. J. M. Rothkrantz, D. A. V. Leeuwen, and P. Wiggers, “Automatic stress detection in emergency (telephone) calls,” Int. J. Intell. Defence Support Syst., vol. 4, no. 2, pp. 148–168, 2011.
[6] J. Zhai and A. B. Barreto, “Realization of stress detection using psychophysiological signals for improvement of human-computer interactions,” in Proc. IEEE Southeast Con, 2005, pp. 415–420.
[7] Hunt J, Eisenberg D. Mental health problems and help-seeking behavior among college students. J Adol H 2010; 46: 3-10.
[8] J. Bakker, M. Pechenizkiy, N. Sidorova, What’s your current stress level? Detection of stress patterns from GSR sensor data, in Proceedings – IEEE International Conference on Data Mining, ICDM, IEEE, 2011, pp. 573–580. vol. 1, doi: http://dx.doi.org/10.1109/ICDMW.2011.178. http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6137431>
[9] T.W. Colligan, E.M. Higgins, Workplace stress, J. Workplace Behav. Health 21 (2) (2006) 89–97, http://dx.doi.org/10.1300/J490v21n02_07. <http://www.tandfonline.com/doi/abs/10.1300/J490v21n02_07>.
[10] N. Sharma, A. Dhall, T. Gedeon, R. Goecke, Thermal Spatio-temporal data for stress recognition, EURASIP J. Image Video Process. 2014 (1) (2014) 28, http://dx.doi.org/10.1186/1687-5281-2014-28. <http://jivp.eurasipjournals.com/content/2014/1/28>.
[11] N.P. Dhole, S.N. Kale, “Study of Recurrent Neural Network Classification of Stress Types in Speech Identification” International Journal of Computer Sciences and Engineering, Paper Volume-6, Issue-4 E-ISSN: 2347-2693
[12] H. Monnikes, J. Tebbe, M. Hildebrandt, P. Arck, E. Osmanoglou, M. Rose, B. Klapp, B. Wiedenmann, and I. Heymann-Monnikes. Role of stress in functional gastrointestinal disorders. Digestive Diseases, 19(3):201-211, 2001.
[13] D. Carroll, G. D. Smith, M. J. Shipley, A. Steptoe, E. J. Brunner, and M. G. Marmot. Blood pressure reactions to acute psychological stress and future blood pressure status: a 10-year follow-up of men in the Whitehall ii study. Psychosomatic Medicine, 63(5):737-743, 2001.
[14] A. Garg, M.-M. Chren, L. P. Sands, M. S. Matsui, K. D. Marenus, K. R. Feingold, and P. M. Elias. Psychological stress perturbs epidermal permeability barrier homeostasis: implications for the pathogenesis of stress-associated skin disorders. Archives of dermatology, 137(1):53{59, 2001.
[15] T. C. Adam and E. S. Epel. Stress, eating, and the reward system. Physiology & behavior, 91(4):449{458, 2007.
[16] S. J. Lupien, F. Maheu, M. Tu, a. Fiocco, and T. E. Schramek. The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. Brain and cognition, 65(3):209{37, Dec. 2007.
[17] G. Keinan. Decision-making under stress: Scanning of alternatives under controllable and uncontrollable threats. Journal of personality and social psychology, 52(3):639, 1987.
[18] https://www.healthline.com/health-news/mental-eight-ways-stress-harms-your-health-082713#6 “Nine Ways Stress Is More Dangerous Than You Think”, Brian Krans August 4, 2016
[19] A. D. Lopez and C. J. Murray. The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Harvard School of Public Health, 1996.
[20] J A. Raij, et al., "mStress: Supporting Continuous Collection of Objective and Subjective Measures of Psychosocial Stress on Mobile Devices," In Proc. Of ACM Wireless Health 2010, San Diego, 2010.
[21] Slavich, G.M. (2019). Stressnology: The primitive (and problematic) study of life stress exposure and the pressing need for better measurement. Brain, Behavior, and Immunity, 75, 3–5. doi: 10.1016/j.bbi.2018.08.011
[22] Monroe, S.M. (2008). Modern approaches to conceptualizing and measuring human life stress. Annual Review of Clinical psychology, 4, 33–52. doi: 10.1146/annurev.clinpsy.4.022007.141207
[23] Shields, G.S., & Slavich, G.M. (2017). Lifetime stress exposure and health: A review of contemporary assessment methods and biological mechanisms. Social and Personality Psychology Compass, 11, e12335. doi: 10.1111/spc3.12335

[24] H. Lu, et al., "StressSense: Detecting Stress in Unconstrained Acoustic Environments Using Smartphones," In Proc. of UbiComp 2012, Pittsburgh, 2012.
[25] H. Lin et al., "Psychological stress detection from cross-media microblog data using Deep Sparse Neural Network," 2014 IEEE International Conference on Multimedia and Expo (ICME), Chengdu, 2014, pp. 1-6.
[26] Chandrasekar Vuppalapati, Mohamad S Khan, Nisha Raghu, Priyanka Veluru, Suma Khursheed "System to Detect Mental Stress Using Machine Learning And Mobile Development" Proceedings of the 2018 International Conference of Machine Learning and Cybernetics, Chengdu China IEEE 2018.
[27] G. Giannakakis, D. Grigoriadis, and M. Tsiknakis, "Detection of stress/anxiety state from EEG features during video watching," in Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE., 2015, pp. 6034–6037.
[28] Gaikwad, Paithane "Novel Approach for Stress Recognition using EEG Signal by SVM Classifier" Proceedings of the IEEE 2017 International Conference on Computing Methodologies and Communication IEEE 2017
[29] Titze, I.R. (2000). Principles of voice production. Salt Lake City, UT: National Center for Voice and Speech.
[30] Sondhi, S., Khan, M., Vijay, R., & Salhan, A.K. (2015). Vocal indicators of emotional stress. International Journal of Computer Applications, 122, 38–43. doi:10.5120/21780-5056
[31] Zhou, G., Hansen, J.H., & Kaiser, J.F. (2001). Nonlinear feature-based classification of speech under stress. IEEE Transactions on Speech and Audio Processing, 9, 201–216. doi:10.1109/89.905995
[32] Kevin Tomba, Joel Dumoulin, Elena Mugellini, Omar Abou Khaled and Salah Hawila "Stress Detection Through Speech Analysis" Proceedings of the 15th International Joint Conference on e-Business and Telecommunications (ICETE 2018) Volume 1: DCNET, ICE-B, OPTICS, SIGMAP and WINSYS, pages 394-398.
[33] Mansouri, Mirvaziri, Sadeghi "Designing and Implementing of Intelligent Emotional Speech Recognition with Wavelet and Neural Network" International Journal of Advanced Computer Science and Applications, Vol. 7, No. 9, 2016.
[34] H. Kurniawan, A. V.Maslov and M. Pechenizkiy, "Stress detection from speech and Galvanic Skin Response signals," Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems, Porto, 2013, pp. 209-214 https://doi.org/10.1109/CBMS.2013.6627790.
[35] S. Gillespie, E. Moore, J. Laures-Gore, M. Farina, S. Russell and Y. Y. Logan, "Detecting stress and depression in adults with aphasia through speech analysis," 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, 2017, pp. 5140-5144.https://doi.org/10.1109/ICASSP.2017.7953136.
[36] A. Baum. Stress, intrusive imagery, and chronic distress. Health psychology, 9(6): 653, 1990.
[37] N. Sharma and T. Gedeon. Objective measures, sensors and computational techniques for stress recognition and classification: A survey. Computer methods and programs in biomedicine, 108(3):1287–1301, 2012.
[38] J. Lee and I. Tashev. High-level feature representation using recurrent neural networks for speech emotion recognition. Interspeech 2015 September 6-10, 2015, Dresden, Germany 2015. Pp 1537-1540
[39] I. R. Murray and Arnott J. L. Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. The Journal of the Acoustical Society of America, 93(2):1097–1108, 1993.
[40] L. Woodrow. Anxiety and speaking English as a second language. RELC Journal, 37(3):308–328, 2006.