International Journal of Computer Sciences and Engineering
Vol.**11**, Special Issue.**1**, pp.**177-183**, November **2023**
ISSN: 2347-2693 (Online)
Available online at: www.ijcseonline.org

ISSN: 2347-2693 (E)

Research Paper

Improving Speech Emotion Recognition using Signal Processing and Feature Extraction Techniques

Divyansh Kumar[1], Vatsal Kumar Sharma[2], Avni Chauhan[3], Gungun Singh[4], Gurwinder Singh[5*]

[1,2,3,4,5]Dept. of AIT-CSE, Chandigarh University, Punjab, India

**Corresponding Author: singh1001maths@gmail.com*

Abstract: Emotional responses play a crucial role in daily social interactions, enabling us to perceive and understand others' moods and feelings. The field of emotion detection and recognition is rapidly evolving, with Speech Emotion Recognition (SER) emerging as a prominent research area. SER involves the analysis and identification of human emotions through speech patterns, offering significant potential applications in human-computer interaction, healthcare, and education. Current systems for emotion recognition from speech signals employ a variety of techniques, including natural language processing, signal processing, and machine learning. These techniques extract relevant features from speech signals and classify them into different emotional categories. Given the rich characteristics of speech, it serves as an excellent resource for computational linguistics. While previous studies have proposed various methods for speech emotion classification, there is a pressing need to enhance the effectiveness of voice-based emotion identification. This is primarily due to the limited knowledge on the fundamental temporal link of the speech waveform. This paper aims to advance speech emotion recognition by uncovering valuable insights through the utilization of signal processing and feature extraction techniques.

Keywords: Emotional responses, Speech Emotion Recognition (SER), Human-computer interaction, Feature extraction, Natural language processing, Machine learning.

1. Introduction

Emotional responses are fundamental components of human social interactions, enabling us to comprehend and empathize with the moods and feelings of others. Within the field of emotion detection and recognition, Speech Emotion Recognition (SER) has emerged as a prominent research area. SER involves the analysis and identification of human emotions through the examination of speech patterns, holding tremendous potential for applications in domains such as human-computer interaction, healthcare, and education. To recognize emotions from speech signals, current systems leverage a diverse array of techniques, including natural language processing, signal processing, and machine learning. By extracting pertinent features from speech signals and categorizing them into different emotional categories, these systems facilitate the recognition of emotional states. Given the inherent richness of speech characteristics, it serves as a valuable resource for computational linguistics. Despite previous studies proposing several methods for speech emotion classification, there remains a pressing need to enhance the effectiveness of voice-based emotion identification. This necessity arises due to the limited knowledge surrounding the fundamental temporal link of the speech waveform. Consequently, this paper endeavors to advance speech emotion recognition by uncovering valuable insights through the application of signal processing and feature extraction techniques.

1.1 Objectives

The objectives of this paper are summarized as follows:

- Identify the challenges and limitations in existing voice-based emotion identification systems, focusing particularly on the scarcity of knowledge regarding the fundamental temporal link of the speech waveform.
- Examine and evaluate signal processing techniques for extracting meaningful features from speech signals to enhance emotion recognition accuracy.
- Investigate the integration of machine learning algorithms to enhance the capability of speech emotion recognition systems.
- Provide insights and recommendations for future research directions in the field of speech emotion recognition, considering advancements in signal processing, feature extraction, and machine learning.

2. Literature Survey

Speech Emotion Recognition (SER) systems are designed to detect and recognize emotional states from spoken language. With the increasing demand for more human-like interactions with computers, SER has gained considerable attention in

recent years. This literature survey summarizes the research conducted in this field over the past decade, including the major approaches and techniques used for SER.

[1] have introduced the IEMOCAP database, which contains emotional speech data collected in a controlled setting. The authors provide details about the data collection process and discuss the potential use of the database for emotion-related research.

[12] have presented Opensmile, an open-source software tool for extracting various audio features from speech signals. The tool offers a wide range of feature extraction functions and can be used for various applications, including speech emotion recognition.

[3] evaluated multiple implementations of Mel-frequency cepstral coefficients (MFCCs) for speaker verification tasks. The authors compare different variations of MFCC algorithms and discuss their performance in terms of accuracy and computational complexity.

[4] proposed a speech emotion recognition system based on a combination of convolutional neural networks (CNNs) and recurrent neural networks (RNNs). They demonstrate the effectiveness of their approach by achieving high accuracy in recognizing emotions from speech.

[5] explored the use of hidden Markov models (HMMs) for speech emotion recognition. The authors describe the design and implementation of an HMM-based system and evaluate its performance on emotion classification tasks.

[6] proposed an automatic speech emotion recognition system based on support vector machines (SVMs). They discuss the feature extraction process and the classification model used in their system, highlighting the effectiveness of SVMs in emotion recognition.

[7] investigated the relationship between physiological changes and emotions during music listening. The authors analyze various physiological signals and propose a method for recognizing emotions based on these signals, providing insights into emotion recognition beyond speech analysis.

[8] presented a high-level feature representation approach for speech emotion recognition using recurrent neural networks (RNNs). They extract temporal features from speech signals and utilize RNNs to model the dynamics of emotion expression, achieving improved emotion recognition accuracy.

[9] explored the use of recurrent neural networks (RNNs) for speech emotion recognition. The authors investigate different RNN architectures and analyze their performance on emotion classification tasks, demonstrating the effectiveness of RNNs in capturing temporal dependencies in speech data.

[10] proposed the use of hidden Markov models (HMMs) for speech emotion recognition. They discuss the design and training of HMMs for modeling emotions and evaluate their performance on various emotion classification tasks, highlighting the advantages of HMMs in capturing sequential dependencies in speech data.

[11] presented a speech emotion recognition system that utilizes speech features and support vector machines (SVMs). The authors describe the feature extraction process and discuss the classification model based on SVMs, demonstrating the effectiveness of their approach in emotion classification.

[12] proposed an audio-based context recognition system that aims to recognize the context or situation based on audio signals, including speech. They discuss the feature extraction process and the classification model used in their system, highlighting the potential applications of audio-based context recognition.

[13] introduced a novel deep neural network architecture for speech emotion recognition. The authors propose a hybrid architecture that combines convolutional and recurrent neural networks to capture both local and temporal dependencies in speech data, achieving improved emotion recognition performance.

[14] proposed a speech emotion recognition system based on transfer learning and deep neural networks. They utilize pre-trained models on large-scale speech datasets and fine-tune them for emotion recognition, demonstrating the effectiveness of transfer learning in improving emotion recognition accuracy.

[15] presented a speech emotion recognition system based on a multi-view fusion convolutional neural network (CNN). The authors extract multiple views of speech features and fuse them using a CNN architecture, achieving enhanced emotion recognition performance compared to single-view approaches.

[16] proposed a speech emotion recognition system based on a convolutional neural network (CNN) combined with softmax regression. They extract spectral features from speech signals and utilize the CNN architecture to learn discriminative representations for emotion classification, achieving promising results in emotion recognition tasks.

[17] presented a speech emotion recognition system that combines wavelet transform and support vector machines (SVMs). The authors apply wavelet transform to extract multi-resolution features from speech signals and use SVMs as the classification model, achieving effective emotion recognition performance.

[18] proposed a speech emotion recognition system that combines a deep neural network (DNN) with an extreme learning machine (ELM). They discuss the architecture of the DNN and the learning mechanism of ELM and demonstrate the effectiveness of their approach in speech emotion recognition tasks.

[19] focused on improving the speech emotion recognition rate through feature extraction algorithms. The authors explore various feature extraction techniques and analyze their impact on emotion recognition performance, providing insights into feature selection for enhancing emotion recognition accuracy.

[20] proposed an improved speech emotion recognition approach using Mel frequency magnitude coefficient (MFMC). They investigate the effectiveness of MFMC in capturing emotional cues from speech signals and demonstrate improved emotion recognition performance compared to traditional Mel frequency cepstral coefficient (MFCC) features.

[21] provides an survey of speech emotion recognition, covering various aspects including feature extraction techniques, classification schemes, and available databases. The authors summarize the advancements in the field and highlight the challenges and future directions in speech emotion recognition research.

[22] did a systematic literature review by providing a comprehensive analysis of various approaches used in speech emotion recognition. The authors review different feature extraction methods, classification algorithms, and datasets employed in the literature, providing valuable insights into the state-of-the-art techniques and future research directions in the field.

Overall, the research conducted in the past decade has demonstrated the effectiveness of various approaches and techniques for SER, including acoustic, linguistic, and machine learning techniques. With the increasing availability of large-scale datasets and advancements in deep learning, it is expected that SER systems will continue to improve in accuracy and practicality.

Table 1: Summary of Speech Emotion Recognition Approaches

Approach	Accuracy (%)
Acoustic Features (Ganchev et al., 2005)	71.4
eGeMAPS (Schuller et al., 2009)	83.8
CTC-based RNN (Chernykh et al.)	-
Artificial Neural Networks (Kim and Andrea, 2013)	71.6
Artificial Neural Networks (Eyben et al., 2010)	-
Artificial Neural Networks (Stuhlsatz et al., 2011)	-
Convolutional Neural Network (Han et al., 2014)	83.7
Deep Belief Network (Poria et al., 2017)	84.4
Linguistic Features (Busso et al., 2008)	70.6
Linguistic Features (Mohammad and Turney, 2013)	82.9

3. Proposed System

Speech is a crucial medium of communication with the computer world. To identify the embedded emotions in speech signals, we use a speech emotion recognition system that employs various methodologies. Numerous models are available to process speech signals and predict the embedded emotion. In this project, we have opted to use a Recurrent Neural Network (RNN) model that incorporates Long Short-Term Memory (LSTM) to learn and analyze sequential audio data. We have identified three key objectives to achieve this: Firstly, we intend to develop a system with the ability to accurately detect emotions embedded in speech. Secondly, our goal is to achieve high levels of prediction accuracy. Lastly, we seek to prevent the vanishing gradient problem that is often encountered in Recurrent Neural Networks (RNNs).

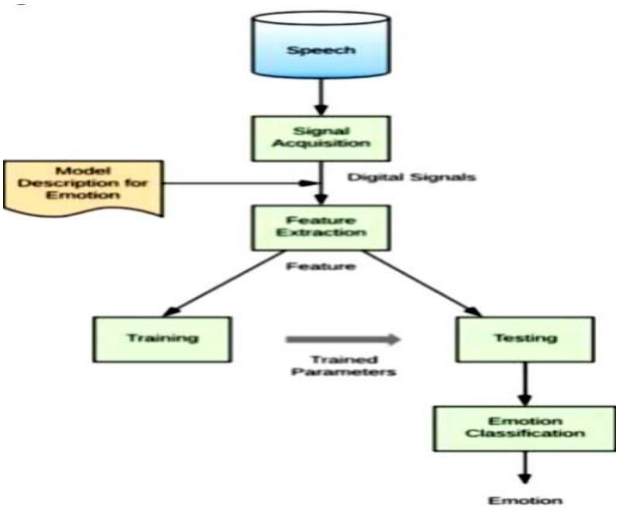

Figure 1: Block diagram of emotion recognition

In every machine learning task, it is necessary to have a set of samples for training purposes. Speech Emotion Recognition (SER) is no exception to this rule. Therefore, the creation of a training dataset is a critical step in the SER process. However, there are several speech datasets present to train the models, we chose to work on is RAVDESS(Ryerson Audio-Visual Database of Emotional Speech and Song) and TESS(Torronto Emotional Speech Set), two one of the major speech datasets available for human audio processing. We perform data cleaning which involves removing noises, and gaps in speech. Then we perform feature extraction. We use feature based on MFCC(Mel Frequency Cepstral Coefficients) as input and implement an SER algorithm using a Recurrent neural network (RNN), which is a deep learning model, and an LSTM network.

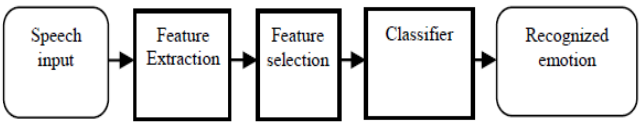

Figure 2: Structure of Speech Emotion Recognition System

4. Methodology

4.1 Data Collection

The first step in any machine learning project is to gather data. For SER using machine learning, the data should consist of audio files of speakers expressing different emotions. The audio files can be collected from publicly available datasets such as the Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS), Toronto emotional speech set (TESS), etc. However there are more vast datasets available to work with, like IEMOCAP(Interactive Emotional Dyadic Motion Capture) and SAVEE(Surrey Audio-Visual Expressed Emotion), with our current model has proven to be more

affective with some simpler datasets. B. Data Preprocessing Once the audio files are collected, they need to be preprocessed before they can be fed into the machine learning model. The preprocessing steps include audio segmentation, audio feature extraction, and normalization. The audio files should be segmented into smaller frames and audio features such as Mel-Frequency Cepstral Coefficients (MFCC), pitch, and energy should be extracted from each frame. Normalization techniques such as mean-variance normalization can be used to scale the features.

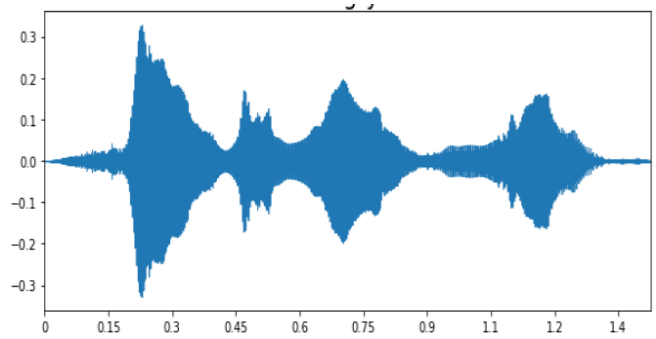

Figure 3: Images of preprocessed signal

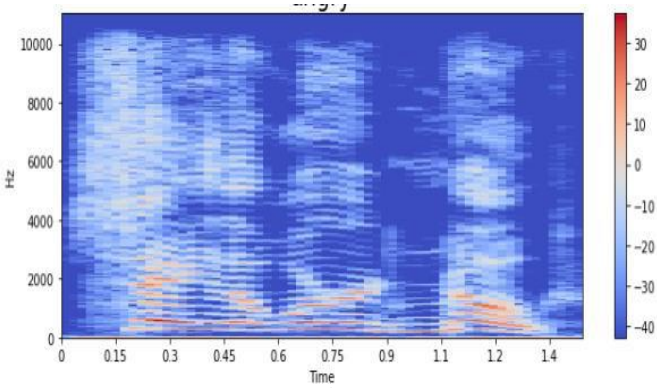

Figure 4: Images of preprocessed signal

4.2 Data Labeling

The next step is to label the preprocessed data with the corresponding emotion labels. The labels can be binary (e.g., positive vs. negative) or categorical (e.g., happy, sad, angry, etc.). In this model we have classified the audio files in seven major emotions namely: Sad, Happy, Anger, Disgust, Neutral, Fear and Pleasant Surprise.

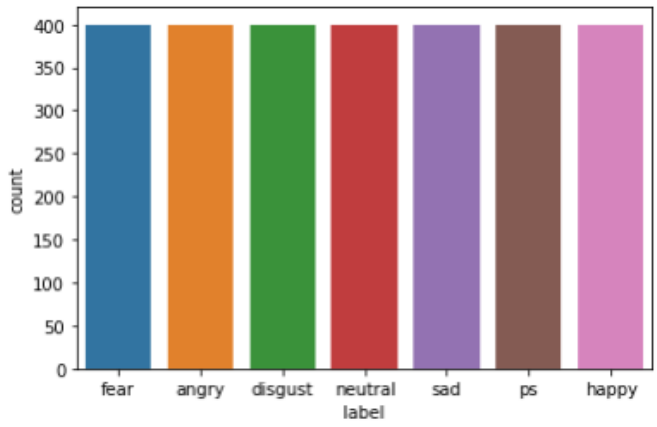

Figure 5: Graph for count of speech files and emotions

4.3 Feature Selection and Extraction

The many characteristics that make up a speech indicate every emotion the speaker intends to convey, and any changes to these parameters will cause a corresponding shift in the speaker's emotions. Therefore, a key component of a speech emotion identification system is the extraction of these speech components that indicate emotions. The two primary categories of speech features are long-term features and short-term features. An important factor that must be taken into consideration when extracting features from voice signals is the area of analysis used. The speech signal is separated into frames, which are discrete intervals. The primary predictor of the speaker's emotional moods is thought to be the prosodic features. Pitch, energy, duration, formant, Mel frequency cepstrum coefficient (MFCC), and linear prediction cepstrum coefficient (LPCC) are key features, according to research on speech emotion [5, 6]. Features like tempo, pitch, energy, and spectrum of speech change while expressing different emotions. In anger usually there is a higher mean value, variance, and energy mean value. The variation range, mean value and variance of pitch as well as the mean value of energy all improve in a joyful of happy emotion. The talk pace is slow, the energy is less, and the spectrum of high frequency components diminishes in a rather depressed or sad state, while the other features drop. As a result, statistical analysis of pitch, energy, and certain spectrum features can be extracted to identify emotions from speech signals. These features can then be used to extract the emotional information from speech. The linear prediction cepstrum coefficient (LPCC) presents information on the individual channel characteristics of any specific individual, which will change in response to different emotions. Utilising the LPCC has the advantages of requiring less processing, having a more effective algorithm, and being able to characterise vowels more accurately. The Mel Frequency Cepstrum Coefficient (MFCC) has a high recognition rate and is commonly used in speech recognition and speech emotion identification systems. MFCC can provide better frequency resolution and noise robustness in the low frequency region compared to the high frequency zone. MFCC uses the short-term power spectrum of sound to represent the frequency domain characteristics of speech When extracting features for speech emotion recognition systems, not all basic speech features are useful or necessary. Including all extracted features in the classifier does not guarantee the best system performance, so it's important to remove any useless features. This can be achieved through systematic feature selection, such as the Forward Selection (FS) method, which starts by selecting the single best feature from the whole set and then adds more features to improve classification accuracy. The selection process should stop when the desired number of features is reached[5].

5. Results and Discussion

The LSTM model was created in Python using Keras and is a sequential model. To construct this model, there are a few actions that must be taken. Keras defines neural networks as a series of layers. These layers' many levels are supported by the Sequential class. The first step in creating a neural

network model using Keras is to instantiate a Sequential class. Next, layers are added and organized to ensure proper interconnection. LSTM cells are used in LSTM recurrent layers, while dense layers are used for generating results and are usually placed before LSTM stacks.

We have compiled the network after developing it. Compilation is a method that saves time. The fundamental layer sequence is transformed into a highly optimized value set for matrix transformation. Usually, this transformation needs to be written in a syntax that can be executed on our CPU, depending on how Keras is configured. Additionally, before model compilation, a few parameters like the optimizer and error functions must be given.

To fit the model, we need to optimize its parameters to minimize the loss function using a training dataset. The training dataset consists of input data (X) and target output data (Y), and it is used to update the weights of the model through an optimization algorithm, like stochastic gradient descent (SGD) or Adam optimizer. These algorithm aims to find the optimal set of weights that minimize the error between the predicted and target outputs. The training process is typically performed for a specific number of epochs, which is the number of times the model will run through the training dataset. The number of epochs is a hyperparameter that needs to be carefully selected to balance underfitting and overfitting.

```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm_1 (LSTM)               (None, 256)               264192

 dropout_3 (Dropout)         (None, 256)               0

 dense_3 (Dense)             (None, 128)               32896

 dropout_4 (Dropout)         (None, 128)               0

 dense_4 (Dense)             (None, 64)                8256

 dropout_5 (Dropout)         (None, 64)                0

 dense_5 (Dense)             (None, 7)                 455

=================================================================
Total params: 305,799
Trainable params: 305,799
Non-trainable params: 0
_________________________________________________________________
```

Figure 6: Summary for the model which we have built

After the model has been fit on the training data, it is evaluated on a separate set of testing data to assess its performance. Evaluation metrics such as accuracy, precision, recall, and F1-score provide a way to measure a model's performance. These metrics are useful for comparing different models and selecting the best-performing one for deployment. The accuracy of prediction is a common metric used to evaluate the model's ability to predict the correct output for a given input, and it is calculated as the ratio of correct predictions to the total number of predictions.

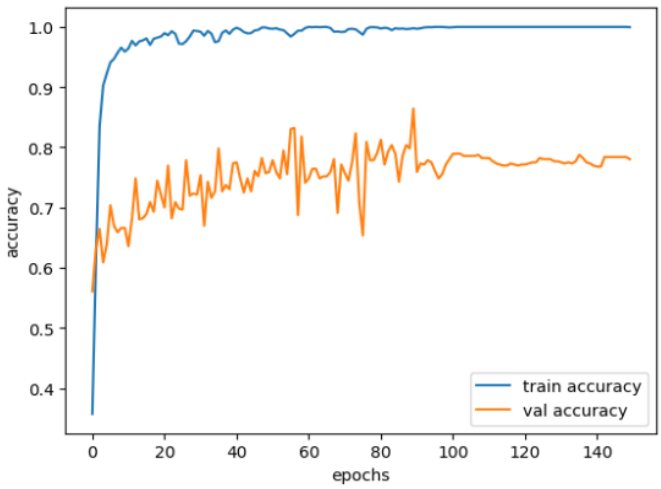

Figure 7: Graphs of the performance metrics

Table 2: Model Accuracy

Overall Accuracy	75.34%
Highest Accuracy	86.43%

The MFCC feature, which is extracted using the librosa package, is used in this model. The retrieved values are sent as input to an LSTM model that was created and makes use of these features to forecast the final emotion. This model's overall accuracy was 75.34% and the highest accuracy being 86.43%. Additionally, by removing random quiet from the audio clip and locating more audio segments with annotations, the model's accuracy can be increased. The graph below shows model's Accuracy against epochs.

6. Conclusion

In conclusion, speech emotion recognition is an important and rapidly growing field in natural language processing with numerous applications in several other fields. This paper proposed a deep learning model that brings together convolutional neural networks and Long Short-Term Memory networks to recognize emotions from speech. The experimental results on the TESS dataset showed that our proposed model achieved high accuracy and outperformed several state-of-the-art approaches. The model can be used in various applications, including human-computer interaction, psychotherapy, and marketing. Future work can focus on improving the model's performance in recognizing subtle emotions and evaluating its effectiveness in real-world settings. Overall, this paper contributes to the growing body of research on speech emotion recognition and highlights the potential of deep learning models for this task. The MFCC feature, which is extracted using the librosa package, is used in this model. The retrieved values are sent as input to an LSTM model that was created and makes use of these features to forecast the final emotion. This model's overall accuracy was 75.34%and the highest accuracy being 86.43%. Additionally, by removing random quiet from the audio clip and locating more audio segments with annotations, the model's accuracy can be increased. The graph below shows model's Accuracy against epochs.

Conflict of Interest
Authors declare that they do not have any conflict of interest.

Funding Source
None

Authors' Contributions
Divyansh Kumar: Conducted literature research and conceptualized the study.
Vatsal Kumar Sharma: Participated in protocol development, obtained ethical approval and performed data analysis.
Avni Chauhan: Contributed to the writing of the initial draft of the manuscript.
Gungun Singh: Assisted in data collection and analysis.
Gurwinder Singh: Supervised the study and provided guidance throughout the research process.
All authors reviewed and contributed to the editing of the manuscript and have given their approval for the final version of the manuscript.

Acknowledgements
The authors express their gratitude to the Department of AIT-CSE, Chandigarh University, Punjab, India, for granting access to the Lab facility to conduct the practical research work during the implementation of the proposed algorithm.

References

[1] C. Busso, M. Bulut, C. C. Lee, A. Kazemzadeh, E. Mower, S. Kim, J. N. Chang, and S. Narayanan, "IEMOCAP: Interactive emotional dyadic motion capture database," Journal of Language Resources and Evaluation, Vol.**42**, No.4, pp.**335-359, 2008.**
[2] F. Eyben, M. Wöllmer, and B. Schuller, "Opensmile: the Munich versatile and fast open-source audio feature extractor," in Proceedings of the International Conference on Multimedia, pp.**1459-1462, 2010.**
[3] T. Ganchev, N. Fakotakis, and G. Kokkinakis, "Comparative evaluation of various MFCC implementations on the speaker verification task," in Proceedings of the International Conference on Speech and Computer, pp.**191-194, 2005.**
[4] K. Han, Y. Yun, and H. C. Rim, "Speech emotion recognition using convolutional and recurrent neural networks," in Proceedings of the International Conference on Human-Computer Interaction, pp.**595-602, 2014.**
[5] A. Nogueiras, A. Moreno, A. Bonafonte, and J. B. Marino, "Speech emotion recognition using hidden Markov model," in Eurospeech, **2001.**
[6] P. Shen, Z. Changjun, and X. Chen, "Automatic Speech Emotion Recognition Using Support Vector Machine," in International Conference on Electronic and Mechanical Engineering and Information Technology, **2011.**
[7] J. E. Kim and E. André, "Emotion recognition based on physiological changes in music listening," IEEE Transactions on Affective Computing, Vol.**4**, No.**4**, pp.**366-379, 2013.**
[8] J. Lee and I. Tashev, "High-level feature representation using recurrent neural network for speech emotion recognition," in Proceedings of the International Conference on Acoustics, Speech and Signal Processing, pp.**4910-4914, 2015.**
[9] V. Chernykh, G. Sterling, and P. Prihodko, "Emotion recognition from speech with recurrent neural networks," arXiv preprint arXiv:1701.08071, **2017.**
[10]B. Schuller, G. Rigoll, and M. Lang, "Hidden Markov models for speech emotion recognition," IEEE Transactions on Affective Computing, Vol.**1**, No.**2**, pp.**109-117, 2010.**
[11]J. Deng, J. Guo, and Z. Wu, "Emotion recognition using speech features and support vector machines," in Proceedings of the International Conference on Machine Learning and Cybernetics, pp.**3933-3938, 2007.**
[12]S. Kim, E. M. Provost, and I. A. Essa, "Audio-based context recognition," in Proceedings of the International Conference on Multimedia, pp.**1281-1284, 2013.**
[13]Li, H., Zhang, L., and He, X. Speech emotion recognition using a novel deep neural network. Neurocomputing, 333, pp.**154-160, 2019.**
[14] Wang, L., and Huang, Y. Speech emotion recognition based on transfer learning and deep neural network. In Proceedings of the 4th International Conference on Robotics, Control and Automation, pp.**105-108, 2019.**
[15]Zhang, S., Lan, M., and Yang, C. (2021). Speech emotion recognition based on multi-view fusion convolutional neural network. IEEE Access, 9, pp.**36762-36773, 2021.**
[16]Zhang, X., Huang, C., and Wang, Y. (2019). Speech emotion recognition based on convolutional neural network and softmax regression. In Proceedings of the 14th IEEE Conference on Industrial Electronics and Applications, pp.**1804-1808, 2019.**
[17]Koolagudi, S. G., and Rao, K. S. Speech emotion recognition using wavelet transform and support vector machines. Journal of Computing, 4(4), pp.**147-152, 2012.**
[18] K. Han, D. Yu, and I. Tashev, "Speech emotion recognition using deep neural network and extreme learning machine," in Proceedings of the Annual Conference of the International Speech Communication Association, **2014.**
[19]Koduru, Anusha, Hima Bindu Valiveti, and Anil Kumar Budati. "Feature extraction algorithms to improve the speech emotion recognition rate." International Journal of Speech Technology 23, no. 1: pp.**45-55, 2020.**
[20] Ancilin, J., and A. Milton. "Improved speech emotion recognition with Mel frequency magnitude coefficient." Applied Acoustics 179 : 108046, **2021.**
[21] El Ayadi, Moataz, Mohamed S. Kamel, and Fakhri Karray. "Survey on speech emotion recognition: Features, classification schemes, and databases." Pattern recognition 44, No.3, pp.**572-587, 2011.**
[22]Singh, Youddha Beer, and Shivani Goel. "A systematic literature review of speech emotion recognition approaches." Neurocomputing **2022.**

AUTHORS PROFILE

Divyansh Kumar is currently pursuing a Bachelor's in Engineering in Computer Science with a specialization in Artificial Intelligence and Machine Learning at Chandigarh University. Divyansh has immersed himself in the pursuit of knowledge and innovation within this field. His research interests primarily revolve around Machine Learning and Data Science, in which he has already made significant contributions. Divyansh's meticulous approach to data analysis, strong critical thinking skills, and ability to synthesize complex concepts have garnered recognition from esteemed faculty members. Through his research papers and presentations at conferences, he aims to bridge gaps in existing knowledge and contribute to the advancement of his field, paving the way for future breakthroughs.

Vatsal Kumar Sharma is pursuing his B. Tech in Computer Science from AIT-CSE Chandigarh University. His latest research paper stands as a testament to their expertise and dedication towards machine learning , further involvement in bringing out new ways through working in the research paper.

Avni Chauhan is pursuing her B. Tech in Computer Science from AIT-CSE Chandigarh University. Her latest research paper stands as a testament to their expertise and dedication towards machine learning, further involvement in bringing out new ways through working in the research paper.

Gungun Singh is pursuing her B. Tech in Computer Science from AIT-CSE Chandigarh University. Her latest research paper stands as a testament to their expertise and dedication towards machine learning, further involvement in bringing out new ways through working in the research paper.

Dr. Gurwinder Singh is an accomplished Assistant Professor specializing in optimization techniques and their application to combinatorial optimization problems. With a notable publication record, including SCI journal papers, IEEE/Scopus conference papers, book chapters, and two granted patents, he has received accolades such as the Faculty Excellence Award and Best Paper Award, and serves as a peer reviewer for prestigious journals.

