**SPEECH EMOTION RECOGNITION**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Author/date | Topic | Concept/Theoretic-al Model | Dataset used | Findings | Limitations/Research gap |
| Kunxia Wang, Ning An, and Lian Li  IEEE 2014 | Speech Emotion Recognition based on wavelet packet coefficient model using wavelet pcket techniques | The concept revolves around utilizing wavelet packet techniques to enhance speech emotion recognition. | Dataset : EMODB | The highest classification accuracy was found up to 73.74% using SVM in EMODB dataset. | Limitations due to their short term analysis, which may not fully capture the non-stationary properties of voice signals |
| Roy Roy Tanmoy, Chakraverty Snehashish, Marwala Marwala Tshilidzi, Satyakama Satyakama Paul  IEEE,2018 | It introduces a novel feature set that incorporates Discrete Wavelet Transform (DWT) to breakdown speech signals and analyze dissimilarities from a neural emotional state. | The research paper employs three primary classification techniques.   * SVC with RBF Kernel * Gaussian Naive Bayes (GNB) * K-Nearest Neighbor (KNN) | Dataset :RAVDESS | SVM: 73.67%  GNB: 77.71%  KNN: 69.41% | The study conducted experiments on a single dataset (RAVDESS), which may limit the generalizability of the results to different datasets or real world applications. |
| Bjorn E Van Zwol, Mathijs A Langezaal, Lukas P A Arts, Albert Albert Gatt, and Egon L Van Den Broek  IEEE,2023 | It utilizes the Fast Continuous Wavelet Transform (fCWT) to enhance Deep Convolutional Neural Networks (DCNN) for Speech Emotion Recogniton (SER) | The DCNN architecture used in the study is based on the AlexNet model and final layer is modified to align with emotion labels. | Dataset: EMODB | The highest classification accuracy was found up to 71% in EMODB dataset. | Due to scarcity of data, the deep learning models induced by the fCWT may exhibit instability, impacting the reliability and consistency of the results. |
| Styliani Danai  IEEE,2020 | It focuses on developing an affective service based on Multi-Modal Emotion Recognition using Speech Emotion Recognition. | The model used in the research paper is an Affective service based on Multi-Modal Emotion Recognition. | Dataset: IEMOCAP (Interactive Emotional Dyadic Motion Capture Database) | The result of the study shows the utilization of datasets like IEMOCAP to train and evaluate models in emotion recognition tasks. | The research paper does not delve into specific limitations associated with the implementation of multi-modal emotion using speech emotion recognition. |