A Survey Of Machine Learning Techniques OnSpeech Based Emotion Recognition And Post Traumatic Stress DisorderDetection

Chappidi Suneetha[1], Dr.Raju Anitha[2]

Abstract:

This paper reviews the literature on a wide array of methodologies utilizedfor emotion recognition from stressedspeech. These techniques include models for the detection of post-traumaticstress disorder (PTSD), neural networks, and long and short-term memory networks (LSTM). The relevance of selecting alternative classification models has also been emphasized. The crucial issues to consider for future emotion recognition research in general,specifically in theindian context, have been highlighted where applicable. Finally, possible future trends in stressed emotion recognition are discussed, including the use of region based convolutional neural network (RCNN), you look only once (YOLO),recurrent neural networks (RNN), and LSTM techniques to solve new challenges in this domain.

Keywords:stressed emotion recognition, CNN, YOLO, RNN, LSTM

DOI Number: 10.4704/nq.2022.20.14.NQ88010 **Neuro Quantology 2022; 20(14):69-79**

1. Introduction

Emotional stress detection computers strive to make human-computer interfaces more effective and natural. Stress, receives little attention,in spite ofcertaintythat onesuffers stress at work or in everyday settings. These scenarios, for example, while driving during rush hour [2], might be perilous. In these cases, an intelligent car management system could be extremely useful, relaxing the driver by playing new music or alerting him to his present emotional state. This example shows how valuable an automatic stress recognizer may be [3][4]. However, such applications are currently challenging to deploy because data is difficult to come by, and the available datasets are largely military recordings [5].

Speech is createdas a sequence of complexsynchronized articulator movements, respiratory system airflow, and vocal system in physiological time [6]. While variations in articulator posture produce speech, not all utterances produced by a speaker will be identical in every way. This reason is that, in many circumstances, the person is under emotional stress, which influences the utterance and causes articulator movements to deviate[7][8]. Listeners can handle or process tiny variations in human communications far better compared to automatic human-machine interaction [9]. Stress and its impacts on human speech production, perception, and automated speech systems are currently poorly understood [10]. The complexity of the human speech signal is that comprises data about the speaker, his or her state, the acoustic surroundings, the person's aim, their language history, accent and dialectical features, and other paralinguistic elements [11].One of the major goals of this survey is to discuss stressed emotion recognition methodsin addition to voice and speaker recognition. In voice

***Corresponding Author**:- Chappidi Suneetha
Address: [1]Research Scholar, Department of Computer Science and Engineering. Koneru Lakshmaiah Education FoundationVaddeswaram, India. ,E-mail: 183030075@kluniversity.in
[2]Associate Professor, Department of Computer Science and Engineering. Koneru Lakshmaiah Education FoundationVaddeswaram, India., Email:anitharaju@kluniversity.in
Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest
Received: **Accepted:**

recognition systems, the inequalityat intervals in training and test utterance has a wide reflection on performance. Observing thatreaction of stress on the speech production system of humans will undoubtedly enhancethe performance of speech recognition technologies and aid in the synthesizing of speech to replicate emotions [12][13].

It is necessary to have the knowledge on the qualities of "stress" before moving further towards analysis and development of such aforementioned systems, as the stress qualities will define the need for a speech based system [14]. Stress is a challenging concept to define because it exists on a spectrum and is not always a binary choice [15]. A single definition cannot, in general, include all instances. For practical purposes, most definitions may be regarded as a little hazy. Despite this, we will use principles of linguistics to illustrate the importance of syllable in our elucidation. As a consequence, we use the term "speech under stress" to describe a scenario in which a person is speaking under duress and the production process of speech is altered [16]. "Speech under neutral conditions" refers to speaking that takes place in stress-free, pressure-free conditions. Usually as outcome, stress is a psychological state that arises in acknowledgementto expanding threat or task demand and is often associated with specific emotions. Even for cases in whichahumandoes not desire it, these changes also might have effect on human speech. Any deviation from the neutral manner of speech, including speaking style, word selection, and body language, is classified as communication under stress. These changes maycast an effect onthe person's speech, even if it's undesired. Speaking during stress is defined as any departure from the neutral mode of speech, encompassing speaking style, word selection, word usage, including sentence length [17]. As a result, communication under stressful settings refers to speech produced outside of a natural, conversational framework due to some environmental cause or emotional state [18]. The physical and emotional position of a speakers maydiffer across different stressfulsettings. Some examples include emergency responders,essentially police, fire, ambulance, and military soldiers engaged in peacekeeping or military activities [19].

High cognitive load, sleep deprivation, irritation over misleading evidence, emotion such as anxiety, illness, emotional exhaustion, and other modern-day multitasking situations all contribute to stress. The following areas have a higher concentration of emotions:

1. Diagnostics–involvesthe actions, conversation analysis including potential threats [20].
2. Protection as well as security — undersea explorers, spacecraft astronauts, power distribution technicians, [21], army forces confronting a panel of assessors, but also criminal justice trainees [22][23].
3. Psychology - patients' emotional states [24].

Many investigations have been fetched on workload, task training stress, in noisy environments [25].

Stress from cognitive or physical strain, the Lombard effect from a loud environment, accent shift, and linguistic variability are all elements that influence speech output and the speaker, as seen in figure 1. These numerous variables or settings, impair the enactment of automatic speech organisms & human speech perception. In noisy, stressful circumstances requiring voice search along with speech recognition, speaker verification, and dialogue systems, we discovered that decreasing noise alone is insufficient to counteract performance losses. Even if loud stressful events could be avoided entirely, stress-induced production variability, particularly the Lombard phenomenon, also has a considerably greater impact upon speech system stability.

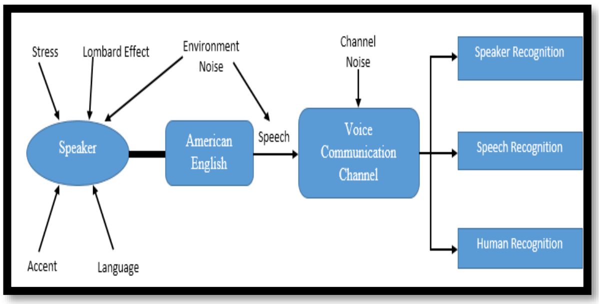

Figure.1 Stress and other factors have an impact on the speech system

Understanding the multitude of ways that stress besides emotion affects speaking output in real-world settings is more important than speech technology progression. There are three key areas of speech important from the standpoint of communication:

a) The physical creation of speech
b) For voice and speaker recognition, speech systems, and technology use variations from the acoustic signal, and
c) Hearing and human perception, especially resolve whether a person is stressed or not stressed.

2.Literature Review

There are various approaches are present for Stress emotion recognition. Some of the important literature that covers the more important emotion recognition is discussed below:

*Lalitha S, Gupta D, et al.[26]*describe a novel speech emotion recognition (SER) system that is multilingual, that takes into account, five distinct languages and dialect variation. Standard speech emotion databases areusually generated by mixed-lingual corpora. A compact feature set is also created, which includesdiscrete combinations of increased perceptive traits and modified H-coefficients along with a diverse set of speech features. It also provides evidence for detecting emotions in SER technologies that are linguistic or mixed-language using samples from a variety of languages and dialects.

*Passardi S, Peyk P, et al. [27]*Post traumatic stress disorder (PTSD) detection based on facial attributesisyet to be researched. People with PTSD reported higher levels of thought or emotion stopping feelings and alexithymia without all these, especially emotional deregulation indicated a poorer level of recognition of unpleasant facial expressions. Although face recognition and reactions to perceived facial emotion are unaffected by PTSD, alexithymia is more significant than depressive symptoms for recognizing unpleasant emotions in the face.

*A. Bérubé, C. Blais, et al. [28]*Mothers were shown photographs of their children's faces with changing boundaries of the six fundamental emotional expressions. The sensitive behaviors of partnerships between children and their mothers were then coded. Each mental instability questionnaire assessed the mothers' records of childhood maltreatment. The goal of this research is to identify the abusive history and maltreatment during childhood which impacts the bondage with the mother.

*Haider F, Pollak S, et, al.[29]*Reducing the number of qualities utilized in inductive inference is one technique to attain this goal. We compare three approaches in this study, we apply state-of-the-art features extraction approaches such as infinite latent feature selection (ILFS), reliefand fisher (generalizable fisher scores) to something like a new deep learning strategy called active feature selection (AFS). Three emotion recognition data sets are evaluated by combining two pairs of conventional audible phonetic elements Berlin Database of Emotional Speech(EmoDB), Surrey Audio-Visual Expressed Emotion (SAVEE)and Italian Emotional Speech Database (EMOVO).

*Egger M, Ley M, et, al.[30]*This study should aid practitioners, researchers, and engineers in determining the exact elucidation for a prearranged situation. The use of a video camera to examine facial features is a non-contact method. Whether utilizing computers, cell phones, even tablets with built-in cameras, this would be very beneficial. Smart wearable's make coming into contact with a person, performing stochastic activities, subsequent clinicalstandards with electro-dermal activities as well as heart-related signals maybe entered in a discreet manner. Heterogeneous sentimental supercomputers outperform unimodal solutions in terms of categorization accuracy.

*Yin Z, Liu L, et, al. [31]*Locally robust electroencephalogramfeature selection (LRFS) is a novel technique for finding generalizable characteristics of electroencephalography (EEG) in a variety of patient subgroups. These obtained EEG features were initially characterized through probability density functions in the LRFS paradigm. The inter-individual stability of EEG parameters arealleged for the resemblance from all probability density functions in between two subjects. The resulting stability is being used to identify locally robust EEG properties, using emotional margin loss indicating the relevance of each feature.

Pandey P, Seeja KR, et, al. [32] Emotionrecognition electroencephalography waves were used to demonstrate sentiment identification process. The intrinsic mode functionsare used to extract two features for emotion categorization: the power spectral density's (PSD's) optimal values and also the first discrepancy.Empirical Mode

eISSN: 1303-5150

Decomposition (EMD) andvariational mode decomposition(VMD) are two methods for extracting intrinsic mode functions(IMFs) from EEG recordings. On the database for emotion analysis using physiological signals(DEAP) benchmark dataset, the methodology is examined, and the top three IMFs' VMD-based characteristics outperform EMD-based traits, according to the findings. The deep neural network(DNN)'s classifier outperforms support vector machine (SVM) classifier when it comes to subject-independent emotion identification.

*Imani M, Montazer GA, et, al. [33]*Emotions are extremely important in the learning process. In electronic learning (e-learning) platforms, emotions must always be put into consideration. Various emotionswere identified from the algorithmslisted along with their pros and cons which limit their usage inonline-instruction systems. The methodology of emotion recognition in multi-modal, dependsupon theseinvestigation conclusions.Byadding information from facial expressions, user messages, and bodily motions, multimodal systems surpass single-modal systems.

*Kranefeld I, Blickle G, et, al. [34]*When General Mental Ability (GMA) and conscientiousness were accounted for, the researchattested the relation between emotion recognition ability for faces (ERA-F) and occupational position. When age had been taken into an account, ERA-F predicted socioeconomic status. The link vanished when GMA and conscientiousness were eliminated from the equation. Although ERA-F besides conscientiousness had no association, GMA had a stronger effect than conscientiousness. It is anticipated that it will encourage further organizations to discover the relationship between employment segregation and ERA using different approaches and from different perspectives to get a fuller picture of the ERA-career success relationship.

Visser LN, Tollenaar MS, et, al. [35] Overall effects of these two types of dermatologists' emotion-oriented conversation upon patients' medical training recollection, including the feasibility of mediated by such a lowering throughout stressful situations, were explored. Furthermore, the impact of personal characteristics on moderation was looked into. An oncologist's response in words to theemotional feelings of patients were altered during a recorded, prearranged bad-news session. The three situations that were developed were emotion-oriented speech, continuousconversation and silence based on emotions. A total of 217 people were assigned to all three circumstances at random.

*Hartling C, Fan Y, et, al. [36]*Effects of early childhood distress in emotion recognition on the face is moderated by a profile of genetic variability linked to variation in the hypothalamic-pituitary-adrenal(HPA)axis activity. If the findings are replicated in a different cohort, they could point to a biological foundation for vulnerability to the negative effects of ELS on social and emotional functioning. The entire discovery adds to either a growing body of evidence emphasizing the importance of this process in determining ELS impacts or to a growing body of evidence emphasizing the importance of this mechanism in determining ELS impacts.

*Zeier P, Sandner M, et, al, [37]*Divide 80 healthy people into two groups, one working on a stress 'two groups either working on a stress'(TSST)target and remainingon control task. Both groups then pay attention onupdatedemotion management paradigm. In this exercise, participants are asked to generate distinct reappraisals of diverse emotional scripts as possible to reduce unpleasant emotions. We can analyze individual fluency (number of reappraisals) and flexibility (diversity of reappraisals) scores by recording the participants' reappraisals.

Wirkner J, Ventura-Bort C, et, al, [38] To measure sympathetic and neutrality scenario interpretation and afterward recognition memory in female volunteers, high cortisol quantities have been used as a potential biomarker for psychological stress. Event-related potentials demonstrated that emotional material elicited higher late positive potentials (LPPs) than neutral content during picture viewing. Total LPP volumes are associated positively with hair cortisol levels, and these brain potentials changed in response to long-term hair cortisol levels.

Smith KE, Leitzke BT, et, al, [39] The researchers investigated how child psychological evaluations are influenced by the interaction of stress exposure. A total of 89 young people ages (11–15years) weredisclosed to one-of-a social stressor via video. The children took an emotion detection exam before and participants witnessed combined representations including facial movements as well as circumstances suggesting consistent versus irreconcilable facial expressions following the stressor. Observation of the eyes was used to track alterations affecting attentiveness towards stimuli. To label theemotions for children, expose them to acute stress and acquire face relevant data is most recommended.

*Martin D, Croft J, et, al, [40]*conducted a comprehensive assessment of literature upon this correlation among probability feature for illusion or psychosis and emotion recognition processing deficiencies. Since there was enough data, meta-analyses were performed; alternatively, descriptive descriptions have been used. In both overall and specific facial emotion recognition deficiencies, meta-analyses were conducted.

Table.1. Comparative study of stress emotion recognition techniques.

Author	Year	Methodology	Advantages	Accuracy
Lalitha S, Gupta D	2020	Multilingual and mixed-lingual	_	High accuracy
Nawaz R, Cheah KH	2020	EEG	Utilizing more temporal domain data.	Excellent accuracy
Passardi S, Peyk P	2019	Post-traumatic stress disorder	Detect even the tiniest facial muscle activations	Medium accuracy
Bérubé A, Blais C	2020	Emotion recognition and maternal sensitive behaviors	Provides a consistent foundation for all mothers.	_
Haider F, Pollak S	2020	Evaluation of automatic feature selection methods	_	High accuracy
Egger M, Ley M	2019	Emotion Recognition from Physiological Signal Analysis	Increased precision.	Higher accuracy
Yin Z, Liu L	2020	The EEG feature selection of Locally robust	Involving medical care, driving, and safety-critical operations.	Average accuracy
Pandey P, Seeja KR	2019	VMD and Deep Learning	_	Obtained an accuracy of 83.33%
Imani M, Montazer GA	2019	E-Learning environments	It benefits from both the consistency requirement and the nonlinear transformation.	*Highest accuracy*
Kranefeld I, Blickle G	2020	GMAand conscientiousness	_	Higher scores indicate higher accuracy
Visser LN, Tollenaar MS	2018	Oncologists' emotion-oriented communication	Such a design's ecological validity is inevitably limited.	_
Hartling C, Fan Y	2018	Interplay of HPA line genetics	Evaluating higher-order functions	_
Zeier P, Sandner M	2018	Potential resilience mechanism	_	_
Wirkner J, Ventura-Bort C	2019	Chronic stress and emotion	_	_
Smith KE, Leitzke BT	2020	Responses to chronic andvideo-based laboratory stress	Significant savings in both human resources and time spent conducting the process.	_
Martin D, Croft J	2019	Schizophrenia and facial emotion recognition	_	_

The above comparison table represents stress emotion recognition techniques and indicates that the case of emotion recognition by e-learning achieved the highest accuracy. Compared to the above techniques Approaches for recognition of emotions, with such a focus on e-learning settings achieves high accuracy as well as it takes the Stability requirement as well as exponential modification each have their merits.

3. Studies Related To Stress Emotion Recognition Based Rcnn&Yolo

Neural networks (NN) are made up of layers of processing units called neurons, each one has its own set of connections. Each neuron multiplies an initial value by weight, adds the results to additional values flowing into the same neuron, modifies the result by the bias of the neuron, and finally normalizes the output with an activation function.

A neural network (NN) is a system of neurons, or simply an artificialneural networks(NN) made up of artificial neurons or nodes, in today's terms [41].A neural network can therefore be physiological or synthetic, and it could be applied to solve artificial intelligence challenges. Weight values are used to depict the connectivity among neurons in the biological analogy. Relationship of excitability have a positive weight, while inhibitory links are negative bonds. The inputs are given weight before being totalled. The output's amplitude is finally controlled by an activation function. For illustration, a suitable frequency range is often 0 to 1, but it could alternatively be -1 to 1.

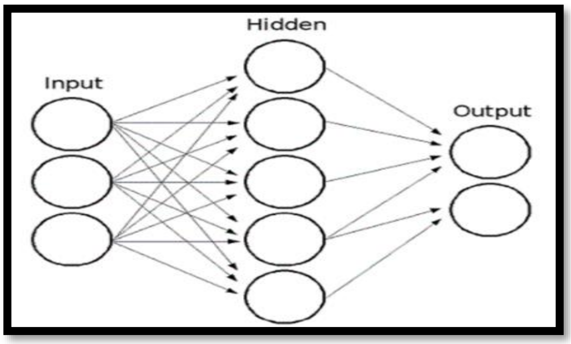

Figure.2 Structure of Neural Network

A. Neural Networks Of Various Types In Deep Learning

There arethree majorkinds of neural networks which provide the foundation to most pre-trained deep learning models, however current analysis specifically looks into CNN and RNN:

B.. Convolution Neural Network (Cnn)

Convolution neural network(CNN)and its robust architectural and functional variations is currently the vogue inside the community of deep learning. Those neural networks are employed in several activities as well as sectors, although image and video analytics is where they're most prevalent.

C. Advantages Of Convolution Neural Network (Acnn)

Convolutional neural network (CNN) identifies the limitations but it does not broadcast anything. The extraction about the suitable knowledge from the incoming data is aided by these filters. It grasps the spatial nature of an appropriate feature input.

In a target speech under stress, geographical features depend on the organization of levels of stress and also the emotions which emerge among them. These aid in the accurate identification of sentiments [42].

*Wang S, Li J, et, al. [43]*The Laban Movement Analysis (LMA) method is a dance emotion expression method that comprises the characteristic parameters. Authentic dance movement data is converted onto three unique interpretation factors to produce dance emotion data, using body structure, spatial orientation, as well as pressure effect. The dancing emotion collection is tested and trained using CNN and LSTM hybrid neural network models. Eventually, decision tree algorithms, random forests, CNNs, and LSTMs were established and checked for capability to illustrate the CNN-LSTM model's usefulness.

*Kwon S, et, al. [44]*demonstrated a one-dimensional adaptation to help neural network-based end-to-end genuine SER model. Then use a multi-learning approach in which the system acquired long-term situational associations whilst identifying geographically relevant psychological effects using voice signals. In an attempt to discover a connection, researchers employed the residual blocks with such a skip connection component, emotional cues, as well as the sequence learning subsystem to acquire the long-term contextual relationships throughout the input data.

Yao Z, Wang Z, et, al. [45] developed a confidence-based fusion technique in which several multi-task learning-based sub-classifiers were used for normative identification for emotional states: DNN plus expression high-level statistical functions(HSFs), convolution neural network (CNN) with multiple segment-level mel-spectrograms (MS), as well as RNN with multiple frame-level low-level descriptors(LLDs). To combine various CNN or RNN outcomes, researchers used a balanced pooling technique based on transfer learning. The brain network was willing to concentrate upon psychologically important regions, arguably as a result of the the activation function.

*Cimtay Y, Ekmekcioglu E, et, al. [46]*employed state-of-the-art CNNtopologies which were pre-trained to considerably improve specific topic recognition accuracy. In contrast to earlier studies that generate spectrum band power statistics using EEG readings, they employed unprocessed EEG data without demodulating, pre-adjustments, as well as acceptability. Through, eliminating minutiae extraction from the training technique, the chance of omitting unique information inside the original reducing power, as well as the deep neural network's ability to identify unknown features is increased.

*Khare SK, Bajaj V, et, al. [47]*multiple CNNs are proposed for automated feature extraction and categorization. To begin, the suggested approach using only a time-frequency notation, turns the sanitized EEG data sets into some kind of picture. To integrate frequency EEG signals onto pictures, a normalized pseudo-Wigner–Ville allocation is used. Pre-trained ‘alex net’, ‘ResNet50’,‘visual geometry group’(VGG16), and also a configurable CNN, were given those images. Overall accuracy, precision, Mathew's coefficient of correlation, F1-score, as well as the false-positive rate of four CNNs are all evaluated.

*Sarath S, et, al. [48]*focused on evaluating sentiments hidden throughout the face, such as stress, worry, or excitement, that can cause a pleasant diffusion around the features. Those infrared graphics were estimated by making use of the thermal value of the pixel than the intensity of every pixel. Image processing methods were used forpulling out the region of interest which also included segregating the object from its backdrop. Given the wealth of publicly available methodologies for recognizing emotional states, economy, resilience, as well as efficiency, could all be enhanced by combining several modalities, which are focused on sensory facial features, thermal face imaging, as well as structural data.

Laroca R, Severo E, et, al. [49] describe a reliable and effective automatic License Plate

Recognition(ALPR) system that is based on the cutting-edge YOLO detection system. CNNs are formed, then fine-tuned at every ALPR phase to guarantee that they have been adaptable in a variety of circumstances. Researchers design a two-stage strategy for graphic identification and classification utilizing modest data augmentation including such upside-down license plates and reversed scripts.

Chen RC, et, al. [50] offer a robust automatic recognition system suitable for swiftly detecting license plate numbers whilst also ensuring accurate results. Theclassic YOLO was altered to make a compact YOLO singular detector containing 36 models. To prevent issues regarding microscopic element categorization, the detector uses a sliding window with all categories from each frame. The value loss was 0.40, while average detection, as well as recognition latency, wasrecorded as 825.81 ms, according to the research. The approach essentially affects a specific development stage identification and tracking technique, which doesn't involve feature extraction or echo cancellation.

Table.2. Comparative Study of Stress emotion recognition using RCNN & YOLO technique

Author	*Year*	*Methodology*	*Advantage*	*Accuracy*
Wang S, Li J	2020	CNN and LSTM	Extracts deep features but also maintains the temporal relationship of the data, making it perfect for extracting dance emotion elements.	The accuracyis 92%
Kwon S	2020	CNN Based on Multi-Learning Trick Approach	IEMOCAP and EMO-DB datasets were used as benchmarks.	73% accuracy
Yao Z, Wang Z	2020	HSF-DNN, MS-CNN and LLD-RNN	–	Weighted accuracy of 57.1% and unweight accuracy of 58.3%
Cimtay Y, Ekmekcioglu E	2020	EEG Emotion Recognition	Adaptive categorization has the advantage of preserving the original distribution data.	It achieves 72% accuracy
Khare SK, Bajaj V	2020	CNN	Stability as well as simplicity	High accuracy
Sarath S	2020	ThermalImages using Yolo Algorithm	–	Accuracy can be enhanced
Laroca R, Severo E	2018	A Robust Real-Time Automatic License PlateRecognition	They use far larger private datasets, which is a considerable advantage, especially for deep learning algorithms.	Improved accuracy
Chen RC	2019	Sliding-Window Darknet-Yolo DL	It employs both CNN for feature learning and automated localization to get around the character segmentation technique.	High accuracy

Stress emotion recognition utilizing RCNN algorithms can be seen in comparison table 2. Dance emotion recognition depending on labanactivity recognition,employing CNN as well as LSTM, achieves a significant precision of 92 basis points. It also has an outstanding experience, besides feature extraction of dance emotions compared to the other methodologies[16]. Yolo states, in contrast to the processes discussed above, are used to recognize stress feelings. Darknet-Yolosliding-window automatic license recognition is combinedwith CNN for learning features and automated localization. Deep learning achieves great accuracy while avoiding the character segmentation step.

4. Studies Related To Speech Based Post Traumatic Stress Disorder Detection

Passardi S, Peyk P, et, al. [51] High mental capability such as emotion control are impaired in patients with PTSD. Facial mimicry (automated facial adaptations towards observed facial emotion) is not analyzedtill now in the literature on PTSD. Similarly, although abnormalities in recognizing facial emotions have also been documented, the underlying causes are unknown.

*Ragsdale KA, Jones KR, et, al. [52]*adopted a 2-week course of treatment which involved long term exposure, cognitive schema therapy, with psychological capacity, independent of the pathophysiology, particularly resolving PTSD plus intellectual difficulties brought on to by mild Traumatic Brain Injury(mTBI). Another of the service's main objectives should be to establish fully prepared expectancies to minimize the incidence of capability of sensing from treating suspected mTBI symptoms. The PTSD+TBI partial hospitalization was finalized by thirty patients with filled or idiopathic intracranial hypertensionPTSD and historiography of self-reported TBI.By exploring the required neurobiological rationale for 'eye movement desensitization and reprocessing'(EMDR), researchers propose a potential that a recently introduced neurophysiological focused nine construct model for experienced selfhood might assist address this difficulty.

*Patterson J, Martin CJ, et, al. [55]*Researchers have developed a subjective procedure called interpretative phenomenological analysis, which particularly incorporates a specific framework, to understand better, the lived experience of patient advocate engagement. To learn more about how women suffering from post-traumatic stress disorder (PTSD) engage with physicians throughout care delivery, individualized face-to-face tests were held with six women who met the entire treatment guidelines for Post-Traumatic Stress Disorder Post Childbirth (PTSD-PC) including six midwives who offered postpartum care. Particularly, where individuals sense throughout their relationships and what that implies to them.

*Gilmore AK, Lopez C, et, al. [56]*present additional predictive analytics with a current randomized controlled testing, estimatingthe worth of prolonged exposure(PE) provided in-person.Especially, PE administered through telehealth among female veterans having military sexual trauma(MST)related PTSD is presented. Military sexual trauma is pervasive amongst military personnel, and that can lead to depression like post-traumatic stress disorder and anxiety. Spite about substantiality there are a handful of scientific proof treatments for MST-related PTSD, include extended exposure (PE) psychotherapy, it is uncertain if there are any features that correlate towards treatment termination too early.

Table.4. Comparativestudy of detecting speech-based post-traumatic stress disorder technique

Author Name	*Year*	*Methodology*	*Advantage*	*Accuracy*
Passardi S, Peyk P	2020	Post-traumatic stressdisorder	Detecting even just the smallest facial muscle activations, but does not show whether the muscle activation was generated by facial mimicry or was an effective response to the stimuli.	Higher accuracy
Ragsdale KA, Jones KR	2020	Treatment of Gentle Traumatic injury of Brain and Post-traumatic stress disorder	—	Medium accuracy
Vujanovic AA, Lebeaut A	2019	Post-Traumatic Stress also Alcohol Use Disorders	—	—
Fingelkurts AA, Fingelkurts AA	2019	Three-dimensional model of the experientialselfhood	—	—
Patterson J, Martin CJ	2019	Post-Traumatic Stress Disorder (PTSD) post-childbirth	—	—
Gilmore AK, Lopez C	2020	Military Sexual Trauma-Related Posttraumatic Stress Disorder	—	—

An investigation involving stress emotion recognition as well as post-traumatic stress disorder approaches is mentioned in Table 3. Comparative analysis towards the components discussed earlier in this process, cranial imitation, cranial emotion, as well as symptomology in post-traumatic stress disorder possess a high level of faultlessnessand also the added benefit efficacious to determine for the gentle specific facial action potentials. Nevertheless, it doesn't provide any documentation upon whether muscle activation has been triggered by facial mimicry or if it was a right answer towards the stimuli.

4.Studies Related to PTSD Using RNN and LSTM

*Zhang T, Wu J, et, al. [57]*in their study state, that in order to represent the speech signal more abstractly, traditional prosodic acoustic components are introduced and contrasted with new i-vector features. For mapping the qualities to emotion labels, a RNNtechnique ispresented.

*Ma J, Tang H, et, al. [58]*in their work state that the multimodal residual (MMRes) LSTM connectivity is just a bidirectional residual LSTM network. To understand the connection between the EEG along with other physiological particulars, the MMRes-LSTMis given high input values at every LSTM layer across the modality. It incorporates the spatiotemporal shortening pathways of the residual network as well as the LSTM, enabling quick learning emotion-related high-level information. DEAP, a publicly accessible dataset with EEG-based feeling of love, hate, guilt etc, recognition, has been used to figure out the proposedstructure.

*Nakisa B, Rastgoo MN, et, al. [59]*proposed a new framework for leveraging differential evolution (DE) to automatically improve LSTM hyperparameters. Here the initial study focused on hyperparameter optimization within the framework for emotional identification in such a systematic manner. Utilizing a novel algorithm derived through embedded devices, they evaluate and compare the conceptual methodology versus various state-of-the-art hyperparameter optimization methodologies. These results indicate how tweaking LSTM hyperparameters improves four-quadrant dimensional emotional expressions by 14% whilst also increasing precision by 14%.

*Nath D, Singh M, et, al. [60]*in their study, employed the DEAP dataset, which is made up of pre-processed EEG and physiological data and is freely available. Their research is focused on analysing EEG signals to construct a headgear model that can track emotions in real-time. This article evaluates the recognition accuracy of several predictors for such valence versus arousal subdomains utilizing band power, a frequency-domain feature extracted from EEG data.

Table.5. A comparison of RNN and LSTM techniques for recognizing post-traumatic stress disorder

Author	Year	Methodology	Advantage	Accuracy
Zhang T, Wu J	2015	I-Vector Feature and RNN Model	—	—
Ma J, Tang H	2019	Multimodal Residual LSTM Network	Gaining a better grasp of the correlation by exchanging values among modalities is beneficial.	Accuracy of 92.87 % and 92.30 %
Nakisa B, Rastgoo MN	2018	Long Short-Term Memory HyperparameterOptimization	Swiftly converge by producing a large number of diverse answers.	Average accuracy 90.5%
Nath D, Singh M	2020	LSTM Network	—	High accuracy 98.3%

Table 5 shows findings of a study employing RNN and LSTM approaches, identifying stress emotions in PTSD. Whiledifferentiatingtheabove-noticedmethodologies, we deduce that emotion recognition depends on LSTM and EEG.The LSTM network is quite effective.

5.Conclusion

The drawing of speech into emotions,aids in the naturalness of current speech systems' performance. In the recent past, anexhaustive amount of work has been done in this area. This review study conducted on fairly recent work in stress emotion identification may pique the interest of the research community in filling some critical research gaps. This paper discussed most of the recent research undertaken in stressed-speech emotion recognition, with significanceof stress under speech emotion frameworks, speech features, and models of classification. The paper alsoaddressesmany researchquestions in theinvestigation of voice-based emotion identification. Following a thorough examination of previous work, seminal articles, beneficial in this topic work were picked for investigation.

References

Latif, S., Rana, R., Qadir, J. (2018) Adversarial machine learning and speech emotion recognition: Utilizing generative adversarial networks for robustness November 2018.

Liu, P., & Wang, X. (2018). A semieager classifier for open relation extraction. Hindawi Mathematical Problems inEngineering.https://doi.org/10.1155/2018/4929674, Vol. Article ID 4929674, pp. 9.

Lotfian, R., & Busso, C. (April 2019). Curriculum learning for speech emotion recognition from crowd sourced labels. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 27(4), 815.

Lubis, N., Sakti, S., Yoshino, K., Nakamura, S. (2019) Positive emotion elicitation in chat-based dialogue systems. IEEE/ACM Transactions on Audio, Speech, and Language Processing 27(4).

Schuller, B. W. (2018). Speech emotion recognition: Two decades in a Nutshell, Benchmarks, and Ongoing Trends. Communications of the ACM, 61(5), 90–99.

Shahin, I., Nassif, A. B., Hamsa, S. (2019) Emotion recognition using hybrid Gaussian mixture model and deep neural network. IEEE Access 7, pp. 26777–26787.

Marmar, C. R., Brown, A. D., Qian, M., Laska, E., Siegel, C., Li, M., Vergyri, D. (2019). Speech-based markers for post-traumatic stress disorder in US veterans. Research Article, Wiley 36(7), pp. 607–616, DOI: 10.1002/da.22890.

Song, P. (2019). Transfer linear subspace learning for cross-corpus speech emotion recognition. IEEE Transactions on Affective Computing, 10(2), 265–275.

Tarunika, K., Pradeeba, R. B., Aruna, P. (2019) Applying machine learning techniques for speech emotion recognition. In 9th International Conference on Computing, Communication and Networking Technologies, DOI: 10.1109/ICCCNT.2019.8494104.

Kerkeni, L., Serrestou, Y., Mbarki, M., Raoof, K., Mahjoub, M. A., Cleder, C. (2019) Automatic speech emotion recognition using machine learning. IntechOpen.

Ding, N., Ye, N., Huang, H., Wang, R., Malekian, R., 2018. Speech emotion features selection based on BBO-SVM. In: IEEE 10th International Conference on Advanced Computational Intelligence, ISBN: 978-1-5386-4362-4.

Kerkeni, L., Serrestou, Y., Mbarki, M., Raoof, K., Mahjoub, M. A., Cleder, C. (2019) Automatic speech emotion recognition using machine learning. IntechOpen.

Huang, Z., & Epps, J. (2018) An investigation of partition-based and phonetically-aware acoustic features for continuous emotion prediction from speech. IEEE Transactions on Affective Computing (Early Access), Electronic ISSN: 1949-3045.

Yin, Z., Zhao, M., Zhang, W., Wang, Y., Wang, Y., & Zhang, J. (2019).Physiological-signal-based mental workload estimation via transfer dynamical autoencodersin a deep learning framework. Neurocomputing, 347, 212–229.

Yang, S., Yin, Z., Wang, Y., Zhang, W., Wang Y., & Zhang J. (2019). Assessing cognitive mental workload via EEG signals and an ensemble deep learning classifier based on denoising autoencoders. Computers in Biology and Medicine, 109, 159-170.

Wang, X., Chen, X., Cao, C. (2020). Human emotion recognition by optimally fusing facial expression and speech feature. Signal Processing: Image Communication, 84, 115831.

Taran, S., & Varun B. (2019)., Emotion recognition from single-channel EEG signals using a two-stage correlation and instantaneous frequency-based filtering method. Computer Methods and Programs in Biomedicine, 173, 157–165.

Torres, E. N., Wei, W., Hua, N., & Chen, P. J. (2019). Customer emotions minute by minute: How guests experience different emotions within the same service environment. International Journal of Hospitality Management, 77, 128–138.

Richhariya, B., Tanveer, M., & Rashid, A.H. (2020). Diagnosis of Alzheimer's disease using universum support vector machine based recursive feature elimination (USVM-RFE). Biomedical Signal Processing and Control, 59, 101903.

Rao, K. P., Rao, M.V.P. C. S., & Chowdary, N. H. (2019). An integrated approach to emotion recognition and gender classification. Journal of Visual Communication and Image Representation, 60, 339-345.

Raza, H., Rathee, D., Zhou, S.-M., Cecotti, H., & Prasad, G. (2019). Covariate shift estimation based adaptive ensemble learning for handling non-stationarity in motor imagery related EEG-based brain-computer interface. Neurocomputing, 343, 154-166.

Pereira, E. T., Gomes, H. M., Veloso, L. R., &Mota, M. A. (2018). Empirical Evidence Relating EEG Signal Duration to Emotion Classification Performance. IEEE Transactions on Affective Computing, in press. doi: 10.1109/TAFFC.2018.2854168.

Jain, D. K., Shamsolmoali, P., & Sehdev, P. (2019). Extended deep neural network for facial emotion recognition. Pattern Recognition Letters, 120, 69–74.

King, R. B., & dela Rosa, E. D. (2019). Are your emotions under your control or not? Implicit theories of emotion predict well-being via cognitive reappraisal. Personality and Individual Differences, 138, 177–182.

Hung, J. C., Lin, K.-C., & Lai, N.-X. (2019). Recognizing learning emotion based on convolutional neural networks and transfer learning. Applied Soft Computing, 84,105724.

Lalitha S, Gupta D, Zakariah M, Alotaibi YA. Investigation of multilingual and mixed-lingual emotion recognition using enhanced cues with data augmentation. Applied Acoustics. 2020 Dec 15;170:107519.

Passardi S, Peyk P, Rufer M, Wingenbach TS, Pfaltz MC. Facial mimicry, facial emotion recognition and alexithymia in post-traumatic stress disorder. Behaviour research and therapy. 2019 Nov 1;122:103436.

Bérubé A, Blais C, Fournier A, Turgeon J, Forget H, Coutu S, Dubeau D. Childhood maltreatment moderates the relationship between emotion recognition and maternal sensitive behaviors. Child abuse & neglect. 2020 Apr 1;102:104432.

Haider F, Pollak S, Albert P, Luz S. Emotion recognition in low-resource settings: An evaluation of automatic feature selection methods. Computer Speech & Language. 2021 Jan 1;65:101119.

Egger M, Ley M, Hanke S. Emotion recognition from physiological signal analysis: a review. Electronic Notes in Theoretical Computer Science. 2019 May 4;343:35-55.

Yin Z, Liu L, Chen J, Zhao B, Wang Y. Locally robust EEG feature selection for individual-independent emotion recognition. Expert Systems with Applications. 2020 Dec 30;162:113768.

Pandey P, Seeja KR. Subject independent emotion recognition from EEG using VMD and deep learning. Journal of King Saud University-Computer and Information Sciences. 2019 Nov 14.

Imani M, Montazer GA. A survey of emotion recognition methods with emphasis on E-Learning environments. Journal of Network and Computer Applications. 2019 Dec 1;147:102423.

Kranefeld I, Blickle G. Emotion recognition ability and career success: Assessing the roles of GMA and conscientiousness. Personality and Individual Differences. 2021 Jan;168:110370.

Visser LN, Tollenaar MS, van Doornen LJ, de Haes HC, Smets EM. Does silence speak louder than words? The impact of oncologists' emotion-oriented communication on analogue patients' information recall and emotional stress. Patient education and counseling. 2019 Jan 1;102(1):43-52.

Hartling C, Fan Y, Weigand A, Trilla I, Gärtner M, Bajbouj M, Dziobek I, Grimm S. Interaction of HPA axis genetics and early life stress shapes emotion recognition in healthy adults. Psychoneuroendocrinology. 2019 Jan 1;99:28-37.

Zeier P, Sandner M, Wessa M. Emotion regulation in the face of stress-Investigating reappraisal inventiveness as a potential resilience mechanism. Psychoneuroendocrinology. 2019 Feb 1;100:S5.

Wirkner J, Ventura-Bort C, Schwabe L, Hamm AO, Weymar M. Chronic stress and emotion: Differential effects on attentional processing and recognition memory. Psychoneuroendocrinology. 2019 Sep 1;107:93-7.

Smith KE, Leitzke BT, Pollak SD. Youths' processing of emotion information: Responses to chronic and video-based laboratory stress. Psychoneuroendocrinology. 2020 Dec 1;122:104873.

Martin D, Croft J, Pitt A, Strelchuk D, Sullivan S, Zammit S. Systematic review and meta-analysis of the relationship between genetic risk for schizophrenia and facial emotion recognition. Schizophrenia research. 2020 Apr 1;218:7-13.

Hopfield, J. J. (1982). "Neural networks and physical systems with emergent collective computational abilities". Proc. Natl. Acad. Sci. U.S.A. 79 (8): 2554–2558. Bibcode:1982PNAS...79.2554H. doi:10.1073/pnas.79.8.2554. PMC 346238. PMID 6953413.

Zheng L, Li Q, Ban H, Liu S. Speech emotion recognition based on convolution neural network combined with random forest. In2018 Chinese Control And Decision Conference (CCDC) 2018 Jun 9 (pp. 4143-4147). IEEE.

Wang S, Li J, Cao T, Wang H, Tu P, Li Y. Dance Emotion Recognition Based on Laban Motion Analysis Using Convolutional Neural Network and Long Short-Term Memory. IEEE Access. 2020 Jul 8;8:124928-38.

Kwon S. MLT-DNet: Speech emotion recognition using 1D dilated CNN based on multi-learning trick approach. Expert Systems with Applications. 2021 Apr 1;167:114177.

Yao Z, Wang Z, Liu W, Liu Y, Pan J. Speech emotion recognition using fusion of three multi-task learning-based classifiers: HSF-DNN, MS-CNN and LLD-RNN. Speech Communication. 2020 Jun 1;120:11-9.

Cimtay Y, Ekmekcioglu E. Investigating the use of pretrained convolutional neural network on crosssubject and cross-dataset EEG emotion recognition. Sensors. 2020 Jan;20(7):2034.

Khare SK, Bajaj V. Time-frequency representation and convolutional neural network-based emotion recognition. IEEE transactions on neural networks and learning systems. 2020 Jul 31.

Sarath S. Human Emotions Recognition from Thermal Images using Yolo Algorithm. In2020 International Conference on Communication and Signal Processing (ICCSP) 2020 Jul 28 (pp. 1139-1142). IEEE.

Laroca R, Severo E, Zanlorensi LA, Oliveira LS, Gonçalves GR, Schwartz WR, Menotti D. A robust real-time automatic license plate recognition based on the YOLO detector. In2018 International Joint Conference on Neural Networks (IJCNN) 2018 Jul 8 (pp. 1-10). IEEE.

Chen RC. Automatic License Plate Recognition via sliding-window darknet-YOLO deep learning. Image and Vision Computing. 2019 Jul 1;87:47-56.

Passardi S, Peyk P, Rufer M, Wingenbach TS, Pfaltz MC. Facial mimicry, facial emotion recognition and alexithymia in post-traumatic stress disorder. Behaviour research and therapy. 2019 Nov 1;122:103436.

Ragsdale KA, Jones KR, Nichols A, Watkins LE, Penna S, Rauch SA, Rothbaum BO. Clinical Effectiveness of an Intensive Outpatient Program for Integrated Treatment of Posttraumatic Stress Disorder and Mild Traumatic Brain Injury. Cognitive and Behavioral Practice. 2020 Sep 19.

Vujanovic AA, Lebeaut A, Zegel M, Smit T, Berenz EC. Post-traumatic stress and alcohol use disorders: recent advances and future directions in cue reactivity. Current opinion in psychology. 2019 Dec 1;30:109-16.

Fingelkurts AA, Fingelkurts AA. Eye movement desensitization and reprocessing for post-traumatic stress disorder from the perspective of three-dimensional model of the experiential selfhood. Medical hypotheses. 2019 Oct 1;131:109304.

Patterson J, Martin CJ, Karatzias T. Disempowered midwives and traumatised women: Exploring the parallel processes of care provider interaction that contribute to women developing Post Traumatic Stress Disorder (PTSD) post childbirth. Midwifery. 2019 Sep 1;76:21-35.

Gilmore AK, Lopez C, Muzzy W, Brown WJ, Grubaugh A, Oesterle DW, Acierno R. Emotion Dysregulation Predicts Dropout from Prolonged Exposure Treatment among Women Veterans with Military Sexual Trauma-Related Posttraumatic Stress Disorder. Women's Health Issues. 2020 Nov 1;30(6):462-9.

Zhang T, Wu J. Speech emotion recognition with i-vector feature and RNN model. In2015 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP) 2015 Jul 12 (pp. 524-528). IEEE.

Ma J, Tang H, Zheng WL, Lu BL. Emotion recognition using multimodal residual LSTM network. InProceedings of the 27th ACM international conference on multimedia 2019 Oct 15 (pp. 176-183).

Nakisa B, Rastgoo MN, Rakotonirainy A, Maire F, Chandran V. Long short term memory hyperparameter optimization for a neural network based emotion recognition framework. IEEE Access. 2018 Sep 3;6:49325-38.

Nath D, Singh M, Sethia D, Kalra D, Indu S. An Efficient Approach to EEG-Based Emotion Recognition using LSTM Network. In2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA) 2020 Feb 28 (pp. 88-92). IEEE.