**FUNDAMENTAL OF STATISTICS AND AI**

**ASSIGNMENT**

**On**

###### EMOTION DETECTION WITH AI

**(CSE IV Semester )**

**2022-2023**

Submitted to: Submitted by:

Mr Amit Gupta sir Sanchit Manwal

Univ Roll No: 2119107

Class Roll No:60

B. Tech CSE

Section-D

Sem-IV

Session: 2022-2023

###### Bachelor of Technology (B. Tech.)

###### GRAPHIC ERA HILL UNIVERSITY, DEHRADUN

**Facial emotion recognition using deep learning: review and insights**

**Wafa Mellouka\* , Wahida Handouzia**

**a Laboratoire d’automatique de Tlemcen (LAT), Tlemcen university, BP 320, Chetouane Tlemcen 1300, Algeria**

**ABSTRACT**

This paper is about using computers to automatically recognize emotions through facial expressions, which is a useful tool for various fields. Researchers are using deep learning algorithms to develop better techniques for interpreting and coding facial expressions to improve computer predictions. The paper provides a review of recent studies in this field, focusing on the architecture and databases used, and compares the proposed methods and results. The goal of this paper is to guide future researchers by providing insights and recommendations for improving automatic facial emotion recognition through deep learning.

**Introduction**

Automatic emotion recognition involves using various sensors to capture verbal and non-verbal information, such as changes in facial expression, tone of voice, and physiological signals. Facial expressions are particularly important because they convey 55% of emotional information. Extracting features from facial expressions is a challenging task, but in 1978, Ekman and Freisen developed the Facial Action Coding System (FACS) to describe facial movements. Traditional methods for facial feature extraction include geometric and texture features, but recent advances in deep learning, specifically convolutional neural networks (CNN) and recurrent neural networks (RNN), have shown promising results. In this paper, recent advances in recognizing facial expressions using deep learning architectures are reviewed, with a focus on the period between 2016 and 2019. The paper includes an introduction to available public databases, a discussion of the state of the art in facial emotion recognition using deep learning, and a comparison of different approaches. The paper concludes with a discussion of future directions in this field.

**Facial emotion recognition using deep learning**

The use of deep learning methods has become popular in facial expression recognition (FER) due to its high recognition capacity. Researchers have proposed various deep learning techniques to improve FER performance using different databases. Pre-processing techniques such as data augmentation, rotation correction, cropping, down sampling, and intensity normalization have been applied before using convolutional neural networks (CNNs). Some researchers have proposed novel CNN architectures to solve problems such as over-fitting, facial occlusion, and gradient explosion. Spatio-temporal CNN and LSTM combinations have been proposed to recognize the temporal dynamics of facial expressions. BiLSTM networks have been used for classification into basic emotions such as happiness, disgust, surprise, anger, fear, sadness, and neutral. The proposed architectures have shown promising results in FER across different databases.

**COMPARISON**

In this paper, the interest of researchers in Facial Emotion Recognition (FER) via deep learning has been discussed. The automatic FER task involves data processing, model architecture, and emotion recognition. Preprocessing techniques such as image resizing, normalization, and data augmentation have been used in all the papers cited in this review. Several methods and contributions have been presented, achieving high accuracy. For spatio-temporal feature extraction, researchers have proposed different deep learning structures, such as a combination of CNN-LSTM, 3DCNN, and Deep CNN. The Softmax function and Adam optimization algorithm are the most commonly used CNN parameters. To test the effectiveness of proposed neural network architecture, researchers trained and tested their models in different databases, and the recognition rate varied from one database to another with the same DL model. Table 2 summarizes all the articles cited above, including architecture, database, and recognition rate. It can be concluded that CNN is the basic network of deep learning for FER, and combining CNN with RNN, especially LSTM network, achieves high precision in FER.

**CONCLUSION**

This paper reviewed recent research on Facial Emotion Recognition (FER) using deep learning techniques. The paper presented different CNN and CNN-LSTM architectures proposed by researchers and discussed the importance of preprocessing steps such as data augmentation, normalization, and cropping in improving accuracy. The paper also highlighted the limitations of FER, which is limited to recognizing only six basic emotions plus neutral, and the need to develop larger databases and more complex deep learning architectures to recognize all basic and secondary emotions. Multimodal analysis is now a growing area of research in FER, and researchers are working on creating powerful multimodal deep learning architectures and databases to improve emotion recognition. Finally, the paper emphasized that FER has potential for natural human-machine interaction, and advances in FER could lead to more natural interactions with machines in the future.

**Emotion Recognition and Detection Methods: A Comprehensive Survey**

**Anvita Saxena1 , Ashish Khanna1 , Deepak Gupta1, \* 1 Computer Science and Engineering Department, Guru Gobind Singh Indraprastha University, New Delhi, India Email: anvanvita2207@gmail.com; ashishkhanna@mait.ac.in; deepakgupta@mait.ac.in \*Corresponding Author: Deepak Gupta, Email:** [**deepakgupta@mait.ac.in**](mailto:deepakgupta@mait.ac.in)

**Abstract**

The paper provides an overview of human emotion recognition through artificial intelligence, which is a popular research field. The study analyzes more than a hundred papers and outlines all the emotion recognition models developed in the last decade. The study shows that emotion detection is mainly carried out through facial expression recognition, physiological signals recognition, speech signals variation, and text semantics. Generally, seven basic emotions are recognized through these methods. The study compares different methods employed for emotion detection and reveals that the best results were obtained by using Stationary Wavelet Transform for Facial Emotion Recognition, Particle Swarm Optimization assisted Biogeography based optimization algorithms for emotion recognition through speech, Statistical features coupled with different methods for physiological signals, and Rough set theory coupled with SVM for text semantics with respective accuracies of 98.83%, 99.47%, 87.15%, and 87.02%. The method of Particle Swarm Optimization assisted Biogeography based optimization algorithms with an accuracy of 99.47% on BES dataset gave the best results.

**Introduction**

This paper focuses on the field of Artificial Intelligence and its interdisciplinary nature with other fields such as Human Computer Interaction and Affective Computing. Emotions are a crucial part of human life, and methods for emotion analysis have been developed over the past few decades. Emotion recognition through computers has many applications, such as in creating smart homes and smart offices. The paper aims to provide an extensive and comprehensive study of significant facial, audio, physiological, and textual emotion detection and recognition methods developed in the last decade. A comparison was carried out based on features, datasets, and methodologies employed for the detection of emotions. The presented paper is a novel approach with a detailed comparison of all the significant emotion detection and recognition methods in the mentioned four domains. It also discusses the limitations associated with these methods and briefly touches upon the future scope and new emerging fields in this area.

**FACIAL EMOTION RECOGNITION**

The passage discusses various methods and techniques for facial expression recognition using machine learning algorithms. It covers methods like neural networks (MLP, RBF, Elman, Convolution Recurrent, etc.), Boosted Deep Belief Network, 3D meshes, fiducial point-based models, Boosting, Adaboost, Active Appearances Model, and Support Vector Machines (SVMs). Each method employs a different set of features and classifiers for expression recognition. Some methods use wavelet and Karhunen-Loeve transforms for feature extraction, while others use 2D DCT and Fourier spectrum of differences of the flow matrices for feature reduction. Bayesian Belief Networks (BBN) and Kalman Filters are used for temporal dependencies and muscle movement modeling, respectively. These methods aim to recognize human expressions such as happy, neutral, disgust, sad, fear, surprise, and anger.

**RESULT**

The text provides a comparison and analysis of different methods and models used for emotion recognition in various fields such as facial emotion recognition, speech signal recognition, and physiological signal recognition through EEG and ECG signals. For facial emotion recognition, the Static Wavelet Transforms model has the highest efficiency of 98.83%, and for feature-based techniques, the maximum accuracy was obtained through HMM and N-D HMM models with 94.41% accuracy. In physiological signal recognition, EQ radio had the highest accuracy among the methods, while EEG signals require extensive preprocessing before feature extraction, and Hybrid Filtering and Higher Order Crossings is the most efficient method among the given methods. Overall, there has been a significant improvement in the accuracy of emotion recognition using speech signals over the years. However, physiological signal recognition is still not widely employed due to the complexity of preprocessing and feature extraction.

**AI Based Emotion Detection for Textual Big Data:    Techniques and Contribution**

**Sheetal Kusal 1, Shruti Patil 2,\*, Ketan Kotecha 2,\*, Rajanikanth Aluvalu 3 and Vijayakumar Varadarajan 4**

**1 Symbiosis Institute of Technology, Symbiosis International (Deemed University),    Pune 412115, India; sheetal.kusal.phd2020@sitpune.edu.in 2 Symbiosis Centre for Applied Artificial Intelligence (SCAAI), Symbiosis Institute of Technology,    Symbiosis International (Deemed University), Pune 412115, India 3 Department of CSE, Vardhaman College of Engineering, Hyderabad 500001, India;    Rajanikanth.aluvalu@gmail.com 4 School of Computer Science and Engineering, The University of New South Wales,    Sydney, NSW 2052, Australia; vijayakumar.varadarajan@gmail.com \* Correspondence: shruti.patil@sitpune.edu.in (S.P.), drketankotecha@gmail.com (K.K.)**

**ABSTARCT**

The study examines text-based emotion detection using AI in social media big data as an upcoming area of Natural Language Processing research, and its potential applications in various fields. The study analyzed 827 Scopus and 83 Web of Science research papers from the years 2005-2020, providing a qualitative overview of different emotion models, datasets, algorithms, and application domains, as well as a quantitative bibliometric review of contributions, including publications, co-authorship networks, citation analysis, and demographic research distribution. The study also presents challenges and probable solutions, which can guide future research directions in this area.

**INTRODUCTION**

This passage discusses the prevalence of social media use and the emotional content found within it. Text-based emotion detection using artificial intelligence is a growing area of research that can be used to analyze this content, which is useful in various fields such as education, psychology, and software engineering. The study of text-based emotion detection involves Natural Language Processing (NLP), which uses techniques in linguistics and computations to help computers comprehend human language. Emotion detection is different from sentiment analysis in that it aims to identify specific emotions like happiness, sadness, and anxiety, rather than assigning a positive, negative, or neutral polarity. Emotion detection from text can be done explicitly, where emotion-bearing words are used, or implicitly, where emotions are identified without the use of such words. Challenges in emotion detection include short or incomplete texts, emojis, grammatical errors, and special characters.

**Text‐Based Emotion Detection: Overview**

The article discusses the process flow of text-based emotion detection using artificial intelligence. It involves creating datasets, text preprocessing, applying machine or deep learning, and classifying and predicting labels of unseen text. Text-based emotion detection has applications in product reviews, service reviews, online social media, conversational agents, etc. The article aims to represent qualitative and quantitative analyses of relevant research work, emotion modeling approaches, and publicly available datasets for text-based emotion detection. It also presents a bibliometric analysis of text-based emotion detection using Scopus and web of science databases. The article concludes with future work directions and challenges identified.

**Emotion Models—Brief Overview**

The article discusses different modeling approaches used for emotion detection, including the categorical emotion model, dimensional emotion model, and componential emotion model. The categorical emotion model, also known as the discrete emotion model, categorizes emotions into distinctive categories or classes. The Robert Plutchik model, the Paul Ekman model, and the OCC Model are commonly used models in this category. The Paul Ekman model differentiates emotions based on six basic classes, while the Robert Plutchik model assumes that few prime emotions appear in contrary sets and their amalgam creates intricate emotions. The OCC model classified emotions into twenty-two classes, including sixteen emotions in addition to the emotions Paul Ekman suggested as basic, covering a much broader representation of emotions. The article suggests that any of the categorical emotional models can be employed to depict emotions, but the OCC model has a broader emotional representation scope due to its larger number of classes.

The article describes different models of emotion classification. The dimensional emotion model divides emotions into three dimensions: valence, arousal, and power. Russell's 2D circumplex model and Plutchik's 2D wheel of emotions model are examples of this model. The article also describes the three-dimensional emotion model proposed by Russell and Mehrabian, which includes arousal, valence, and dominance. The componential emotion model, also known as an appraisal-based model, is an extension of the dimensional emotion model and includes appraisal theory. The article provides a comparative analysis of different emotion models, including their advantages and disadvantages. The final step is to choose an emotion detection approach based on the selected model.

**CONCLUSION**

The survey paper focuses on existing approaches to text-based emotion detection using artificial intelligence and highlights available datasets in the research domain. The paper provides insights into the predominant authors, publications, and future directions and challenges, including the need for new datasets, domain adaptation techniques, and the use of deep learning and ensemble techniques. The study recommends that research on artificial intelligence for text-based emotion detection should remain a focus of interest to improve understanding and enhance worldwide applications.

**Implementation of Artificial Intelligence Image Emotion Detection Mechanism Based on Python Architecture for Industry 4.0**

**Zhu Jinnuo , 1,2 S. B. Goyal , 2 Miretab Tesfayohanis , 3 and Yahye Omar 4 1 Nanchang Institute of Science & Technology, Nanchang, China 2 Faculty of Information Technology, City University, Petaling Jaya 46100, Malaysia 3 Department of Information Technology, Dambi Dollo University, Ethiopia 4 College of Computing & IT, Pepsi Area, University of Hargeisa, Hargeisa, Somaliland**

**ABSTARCT**

The paper discusses the use of artificial intelligence technology for emotion and behavior research, specifically using facial feature algorithms combined with machine learning for output detection and analysis. The paper analyzes the advantages of Python programs and conducts preliminary tests using the σE value algorithm. The paper proposes a new algorithm called σx, which takes the weighted sum, takes the logarithm, and then takes the square root. According to statistical analysis, the overall detection rate of the σx algorithm has been improved to around 80%, with less frequency fluctuation compared to existing algorithms. The paper's next direction is to use the Python main program to perform AI automatic facial emotion detection work by combining the σx algorithm with facial recognition feature algorithms through machine learning.

**Introduction**

Emotion detection technology has become a hot topic for research in the 21st century, as emotions play an important role in human life. Positive emotions can bring joy while negative emotions can lead to crises. The detection of emotions can be achieved through both artificial and machine modes. However, with the development of Industry 4.0, artificial intelligence technology has become more popular, and researchers are using machine learning to extract appearance and psychological data from the human body for emotion detection. This technology has been widely used in various fields, including intelligent policing, facial recognition, intelligent driving detection, EEG emotion, human neural network multimodal detection, and heart rate detection. Researchers have used different algorithms for emotion detection, such as AU algorithm, OpenVINO algorithm, multidimensional information fusion algorithm, and ECG emotion analysis and prediction algorithm. This paper analyzes existing algorithms and proposes an upgraded algorithm to improve emotion detection accuracy. The figure included in the paper illustrates the technical analysis of human behavior detection.

Human behavior analysis technology employs artificial intelligence techniques to extract relevant data such as the coordinates of limbs and in vivo bones in the human body structure and analyze behavior through changes in the coordinates of the bones and changes in the angles of the limbs. The main purpose of human behavior research is to understand the transformation methods and interaction behaviors between humans and artificial intelligence. Emotions stimulate behavior, and special methods can be used to obtain concurrent emotional data when analyzing human behavior. Artificial intelligence technology plays a decisive role in the analysis and prediction of human behavior.

The paper proposes a comprehensive analysis of human emotions and behavior through the use of effective machine learning methods. The research aims to provide a reliable technical foundation for future behavior and emotion integration studies. The algorithm implementation process in the human emotion detection phase is used to find the most effective basis for human behavior analysis. The model diagram in Figure 2 illustrates the proposed approach.

**Analysis of Results**

Based on the sample selection and classification principles adopted in the previous stage, a small amount of sample data was selected, and the preliminary detection process of human emotion and behavior data was completed using the initial training set. However, the detection accuracy is still far from the ideal state. The emotional sample data includes 8 categories, and the detection and recognition rate of the current mechanism is shown in Table 9. After improving the original Algorithms 1 and 2, the most advantageous core architecture of the original algorithm is selected to create the final machine learning mechanism, which shows a significant improvement in the overall detection rate compared to the original algorithm. However, the overall recognition rate fluctuates around 70% for each type of target data, which is relatively average. This is the conception stage of the paper, and the specific artificial intelligence technology and wearable technology to be used have not been determined yet.

**Preliminary Experimental Data Analysis**

The research work is divided into five stages, including machine learning, recognition ratio, sample testing, frequency fluctuation, and variance. The most significant improvement is observed in the machine learning and sample testing stages, resulting in a detection accuracy of over 70% after the optimization of the final algorithm. The recognition ratio, sample testing, and variance values also changed significantly during the experiment. Algorithm 3 shows a further reduction in the variance value in the detection process compared to the initial Algorithm 1.

**CONCLUSION**

This article presents a Python-based emotion detection system using modules such as OpenCV, NumPy, Idlib, and materlib5. After sample training data tests and optimization of existing algorithms, the logarithmic algorithm is proposed, which improves the detection rate and narrows the detection gap in target emotion detection. However, the research work has limitations, as the current detection range is expression recognition detection in images. Future research will need to expand the detection field to AI detection of facial emotions, requiring more familiarity with deep convolutional neural networks, facial location recognition, and discrete wavelet algorithms. The goal is to gradually shift to wearable technology with built-in emotion recognition. The future challenge is to use wearable devices to obtain human psychological data and integrate the mechanism of artificial intelligence behavior analysis and emotion detection under wearable technology. The research will continue to find more target sample data for the system to learn and improve the detection effect.

**Deploying Machine Learning Techniques for Human Emotion Detection**

**Ali I. Siam , 1 Naglaa F. Soliman , 2 Abeer D. Algarni,2 Fathi E. Abd El-Samie,2 and Ahmed Sedik 3 1 Department of Embedded Network Systems Technology, Faculty of Artificial Intelligence, Kafrelsheikh University, Kafr El-Sheikh, Egypt 2 Department of Information Technology, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh 84428, Saudi Arabia 3 Department of the Robotics and Intelligent Machines, Faculty of Artificial Intelligence, Kafrelsheikh University, Kafr El-Sheikh, Egypt**

**ABSTRACT**

The paper presents a real-time approach for implementing emotion detection using facial expressions in robotic vision applications. The proposed approach includes four phases: preprocessing, key point generation, key point selection and angular encoding, and classification. Key points are generated using the MediaPipe face mesh algorithm, and then encoded using a sequence of mesh generator and angular encoding modules. Principal Component Analysis (PCA) is used for feature decomposition, which is then fed into various machine learning techniques such as Support Vector Machines (SVM), k-Nearest Neighbor (KNN), Naïve Bayes (NB), Logistic Regression (LR), Random Forest (RF), and Multilayer Perceptron (MLP) classifiers. The techniques are evaluated on different datasets with different evaluation metrics, achieving a human emotion detection accuracy of 97%, which is superior among the efforts in this field.

**INTRODUCTION**

The recognition of human emotions is important for several applications such as augmented and virtual reality, human-computer interaction, and security systems. Facial expressions are a reliable way to detect human emotions for Human-Robot Interaction (HRI) applications as 55% of affective information is conveyed through facial expressions. This paper proposes a real-time emotion detection approach for robotic vision applications. The proposed approach consists of four phases: preprocessing, feature extraction and selection, feature decomposition, and classification. The feature extraction and selection are carried out by the MediaPipe face mesh algorithm based on real-time deep learning. The feature decomposition phase is performed by Principal Component Analysis (PCA) to enhance the accuracy of emotion detection. The extracted features are enrolled in a selected classifier such as Support Vector Machine (SVM), k-Nearest Neighbor (KNN), Naïve Bayes (NB), Logistic Regression (LR), or Random Forest (RF). An MLP deep neural network is also utilized. The introduced techniques are evaluated on different datasets using various evaluation metrics. The paper also presents a hardware implementation of the proposed models. The main contributions of this work include a novel fast and robust emotion detection framework, emotion face mesh based on automatic key point determination from face images, key point angular encoding for generating sensitive and distinguishable angular features, emotion classification using different machine learning techniques, and a comparison between the deployed techniques in terms of accuracy, scalability, and processing time. The paper is organized as follows: literature review, datasets utilized, proposed methodology, simulation results, and performance discussion.

**RESULTS**

The proposed framework is a facial expression classification model that is evaluated on two benchmark datasets, CK+ and JAFFE. The framework is developed using Python 3.9, OpenCV 4.5, SRGAN, MediaPipe 0.8.6, Scikit-learn 0.24.2, NumPy, Pandas, Math, OS, and Matplotlib libraries. The evaluation of the proposed model is based on five metrics, including accuracy, precision, recall, F1-score, and training time. To evaluate the performance of the proposed model, eight classifiers are employed, and the hyperparameters for each classifier are presented in Table 6. The classification is based on ten features extracted from images in each dataset using the procedure described in Section 4.

Learning curves are used to determine the cross-validation scores and behaviors for different training sizes for the adopted classifiers in the case of CK+. The confusion matrix for each classifier on the CK+ dataset using the proposed model is shown in Figure 14. It reveals that the per-class accuracies of Anger, Happy, and Surprise classes have higher values with all classifiers than those of other emotions, while the Contempt and Sadness classes have lower per-class accuracies.

Moreover, the performances of the proposed framework with eight classifiers on CK+, JAFFE, and RAF-DB datasets are presented in Tables 7–9. The classification report, including accuracy, precision, recall, and F1-score, as well as the training time taken for each classifier, is also shown in the tables. A visual comparison between the classifier accuracies across the used datasets is shown in Figure 16. Results indicate that the KNN classifier outperforms other classifiers in terms of accuracy, precision, recall, and F1-score. It achieved the best accuracies of 97% and 95% on CK+ and JAFFE datasets, respectively.

Additionally, the time required to train the KNN and Gaussian NB is 0.005 sec on CK+, which is the lowest time compared to those of other classifiers. The MLP and RF classifiers have the highest training times, which are 1.82 sec and 0.74 sec, respectively. The proposed models are also evaluated on the RAF-DB dataset, and the results reveal that the MLP and SVM models can be considered as good emotion detection models for this database, especially with an accuracy of 67% for both models.

Overall, the proposed framework provides a variety of models that are optimal for robust emotion detection environments. The framework uses state-of-the-art libraries and techniques to extract features from facial images and classify them into different emotions. The evaluation results demonstrate that the proposed framework achieves high accuracy, precision, recall, and F1-score with the KNN classifier, which is the best performing classifier among the eight classifiers used in this study.

**Discussion**

The proposed approach for human emotion detection demonstrates high performance based on simulation results. The encoding module has superior performance with deployed classifiers including KNN, SVM, and MLP. A comparison is presented between the proposed approach and works in the literature, indicating the proposed approach has superior performance compared to previous efforts in this field.

**CONCLUSION**

This paper proposes a novel approach for facial expression recognition as a solution for Human-Robot Interaction (HRI). The proposed approach consists of four phases: key point extraction, selection, mesh generator, and angular encoding modules. The generated feature maps are classified using several classification algorithms. The proposed approach is evaluated on CK+, JAFEE, and RAF-DB datasets and shows superior performance in terms of accuracy of detection and processing time evaluation metrics. The low dimensionality of extracted features enables the ML-based approaches to reach optimum performance in a short time with much lower computational cost than those of the DL-based approaches. Future work includes introducing a method for emotion detection from other modalities such as videos, spoken words, and written text, hardware implementation of the proposed approach, and further machine learning techniques such as dictionary learning and semi-supervised learning.

**Emotion AI, Real-Time Emotion Detection using CNN**

**Tanner Gilligan M.S. Computer Science Stanford University** [**tanner12@stanford.edu**](mailto:tanner12@stanford.edu)

**Tanner Gilligan M.S. Computer Science Stanford University** [**tanner12@stanford.edu**](mailto:tanner12@stanford.edu)

**ABSTRACT**

This paper proposes a CNN-based approach to real-time emotion detection using data from various datasets including the Extended Cohn-Kanade dataset, Japanese Female Facial Expression dataset, and custom images. The authors re-train a LeNet and AlexNet implementation and report accuracy above 97%. However, qualitative analysis of real-time images indicates that the models perform reasonably well but not as well as the quantitative results suggest. The authors also apply preprocessing steps to improve the model's performance.

**INTRODUCTION**

The paper focuses on the issue of emotion recognition in computer vision and presents a real-time emotion detection system using Convolutional Neural Networks (CNN). However, the authors note that the data sets for emotion detection are based on "acted" emotions, rather than "real" emotions, which can lead to problems in accurately detecting emotions. The authors mention the potential application of using emotion labels and prediction scores to predict emotion intensities based on social science research on emotions. However, the authors did not pursue this application due to the lack of existing data sets or research on the ground truth of emotion intensity prediction. The CNN model developed by the authors successfully classifies faces as one of the core seven emotions, including anger, contempt, disgust, fear, sadness, happiness, surprise, and neutral. The authors note that while the quantitative results of the model show high accuracy, the model may not perform as well in real-life scenarios where facial expressions may not be as exaggerated as those in acted emotions.

The authors of this paper developed a real-time emotion detection system, and their approach involved several steps. First, they built a dataset by collecting labeled facial expression data from multiple sources and processing it into a common format, including custom images of themselves and a friend. Next, they pre-processed the images by running facial detection software, cropping and scaling the images, manually eliminating poor quality images, applying a Gaussian filter, and subtracting the mean image of the training set. They also augmented the images to include reflections and rotations to improve robustness. Finally, they constructed a convolutional neural network (CNN) by utilizing pre-trained versions of AlexNet and LeNet in Caffe on AWS, re-training the first and last layers, and experimenting with various learning rate methods and parameters to generate a non-divergent model.

**Result Metrics**

The project team evaluated the performance of their real-time emotion detection model using two metrics: quantitative results and classification of livestreamed images. Quantitative results were measured using metrics such as precision, recall, and accuracy, based on the labeled dataset used to train and test the model. The team also examined how well the model performed in classifying livestreamed images, which were not labeled, making it difficult to quantify the results. In some cases, the team hand-labeled a second dataset to evaluate the performance of the model. However, since neither of the team members were actors, their facial expressions may not have been the best portrayals of the target emotions, which is a more realistic scenario for the application of the system.

**Conclusion**

The paper proposes a CNN-based emotion detection model that uses facial detection software and cloud computing to achieve high accuracy in classifying emotions. The model's accuracy on a custom data set reached 98.5%, and on the CK+ data set it was 97.2%. The model also showed more balanced accuracy results across the emotion spectrum and worked well with non-actor subjects. The authors suggest future work in creating a user interface for iteratively training the model, incorporating class imbalance handling in the network, and predicting the intensity of emotions on a continuous scale. They also propose the need for better data sets to understand real emotions. The code base for the project is available on Github.

**Emotion Recognition – A review**

**Sukhpreet Kaur1 , Nilima Kulkarni2**

**1 CSE,MIT SOE, MIT ADT University, Pune, India.**

**2 CSE,MIT SOE, MIT ADT University, Pune, India.**

**ABSTRACT**

This paper explores the importance of emotion recognition in various fields such as human-computer interaction, robotics, healthcare, biometric security, and behavioral modeling. The paper identifies different techniques for feature extraction and emotion classification using supervised and unsupervised machine learning algorithms. It also presents a comparative analysis of various machine learning algorithms used in referenced papers. The paper discusses the scope and applications of automatic emotion recognition systems and the parameters that can increase their accuracy, security, and efficiency. Overall, the paper provides insights into the current state of emotion recognition technology and its potential impact on different fields.

**INTRODUCTION**

This paper provides an overview of emotion recognition and its various applications. Emotion recognition can be achieved through soft biometrics, which include physical and behavioral traits, as well as human adhered characteristics. The paper explains that recognizing basic emotions is only the first step, and that factors such as valence, polarity, and arousal also play important roles in identifying one's state of mind.The paper also discusses the use of sentiment analysis to understand a person's opinion and attitude towards a particular topic. Machine learning techniques are used to extract features and patterns from collected data, including facial expressions, body movements, speech, text writing, and brain or heart signals.The applications of emotion recognition include human-computer interaction, biometric security, and affective computing. The paper explains how these applications can benefit from emotion recognition technology.The paper concludes by highlighting the need for more robust, secure, and efficient emotion recognition systems. It emphasizes the importance of addressing privacy and ethical concerns associated with emotion recognition technology. Overall, the paper provides a comprehensive overview of emotion recognition and its various applications, highlighting the potential benefits and challenges of this emerging technology.

**LITERATURE REVIEW**

In the first study [1], the recognition of basic emotions such as happiness, sadness, fear, anger, and neutral was performed using multiple body movements including head region, joints, upper and lower body movements, and arm bound space. Video datasets were used to extract motion or kinetic features from speed, space, and symmetry of various body parts. The authors applied ANOVA and MANOVA to compute the relevance of extracted features and normalization of features. Score and rank level fusion techniques were used to fuse the features, and the results showed an accuracy of 90% in walking, 96.6% in sitting, and 86.66% in action independent cases. The future scope suggested by authors was to include voice and facial expressions for better performance, remote sensing of emotions in case of emergency, and to recognize emotions from body movements more extensively.

In the second study [2], the authors used online handwriting (text dependent and text independent) and signatures to recognize emotional status such as happy, sad, and stress. The database used was collected by CIU based on physiological scenarios, including watching positive and negative videos, and with corresponding ground truth information such as identity, age, gender, ethnicity, and emotion. Extracted features included average pen velocity, max velocity, pressure in both x and y directions, altitude, and the number of times the pen passes through the midline. The authors used the K-Nearest Neighbor algorithm on the WEKA tool to classify emotions from preprocessed normalized features. Stress prediction achieved the highest accuracy from handwriting, and happiness was best recognized from signatures. The future scope suggested by the authors was to improve the database by making it a hybrid of both online and offline features to improve accuracy in all states of emotions.

The third paper [3] presents a feature extraction framework for emotion recognition from body movements. The framework uses video datasets to extract motion or kinetic features from various body parts under three scenarios: walking, sitting, and action independent cases. Geometric and temporal features are extracted, and ANOVA and MANOVA are applied to compute the relevance of the extracted features and normalize them. The system achieves high accuracy, and future work is suggested, including adding voice and facial expressions to improve performance and improving communication between humans and robots.

The fourth paper [4] uses online handwriting and signatures to recognize emotional status, such as happiness, sadness, and stress. Extracted features from the signature and handwriting include average pen velocity, max velocity, pressure in both x and y directions, altitude, and the number of times the pen passes through the midline. The features are normalized using z-score normalization, and K-Nearest Neighbour algorithm is used to classify the emotions from the preprocessed normalized features. The system achieves high accuracy, and future work is suggested, including improving the database by making it hybrid, combining online and offline features, to improve accuracy in all emotional states.

The fifth paper [5] presents an emotion recognition system that works under uncontrolled or natural conditions. The system focuses on facial expressions, body gestures, speech, and other sensors, and can recognize universal and non-universal expressions, facial muscle movements, and sentimental analysis parameters such as valence, arousal, and dominance. The authors suggest future work in challenging human intelligence to create machines that can understand and interpret human emotions under uncontrolled behaviors.

The sixth paper explains advanced datasets and feature extraction techniques to increase the efficiency of automatic facial emotion recognition (AFER) systems. The paper discusses the effects of facial expressions, including categorical, dimensional, physiological specificity, and microexpressions. The authors suggest future work in improving registration algorithms to work under an uncontrolled or natural environment, using 3D, RGB, and multimodal datasets to recognize microexpressions, and training machines using new datasets to form neural networks.

The seventh paper discussed a multi-modal system for enhancing facial expression recognition using deep neural network techniques. The paper used an extended database with added label preserving data to evaluate the results and addressed issues of data overfitting and data imbalance. The proposed model showed the highest accuracy of 94.41%. Future scope includes exploring multi-modal feature strategies applicable to poor quality facial images.

The eighth paper discussed the application of sentiment analysis to compare the popularity of McDonald's and KFC based on tweets. Sentiment analysis of unstructured Twitter data was performed using supervised and unsupervised machine learning algorithms. The paper found that McDonald's was more popular than KFC with the highest accuracy of 78% obtained through the maxtent algorithm. Future scope includes designing an algorithm that can automatically classify tweets.

The ninth paper proposed a supervised approach that built a topic-adaptive sentiment lexicon model (TaSL) that worked on the topic rather than the word to analyze sentiments. The TaSL model captured sentiment opinions or polarities of words under different topics and showed better performance of sentimental analysis compared to other approaches. Future scope includes using semi-supervised machine learning algorithms to remove data imbalance issues on unlabeled data.

The tenth paper introduced a dynamic database captured using RGB-Depth cameras for human body movement detection in 3D form. It could be used in applications such as human-computer interaction, robotics, and virtual reality. The paper discussed the process of data unification, body parts tracking, noise removal, and body parts' orientation estimation. The technique needs to be improved as it has shown less accuracy while hands and feet poses make variations frequently. The system has been built for detecting one person only at a time.

The Eleventh paper discusses various works in the field of sentiment analysis and emotion recognition using different techniques. The first paper presents the concept of affective computing and a personalized dataset called AESDD. The authors used various classification algorithms and found that the Ensemble classifier performed the best. Future work includes making multiple databases personalized to form a more robust system.

The twelveth paper proposes advancements in machine learning algorithms to recognize emotions and pain intensity using cloud-based GPU hardware. The authors used RGB frames from video input and applied face detection and image classifiers to recognize emotions. The results showed that remote cloud-based GPU environments increase time efficiency and security. Future work includes creating a larger database and obtaining a deep convolutional neural network by removing multilayers of neural networks.

The thirteenth paper presents a self-serving device called a psychometric analyzer that works on the history of the patient and recorded voice as an input. The system detected the intensity level, emotions, polarity, and subjectivity from the given input. The authors used supervised, unsupervised, and self-learning algorithms for data acquisition and classification. The results showed that the system was time-efficient. Future work includes building a model with the best possible time-cost-accuracy tradeoff.

The forteenth paper introduces two new trends in computational linguistics: affective computing and real-time brain signal machines for emotion recognition, and EEG brain signals. The authors used sentimental analysis architecture to find out positive/negative attitude, valence, and arousal. They used three approaches for recognizing text from speech and found that real-time brain signals (EEG) were needed for analyzing sentiments through polarity. Future work includes finding emotion dataset resources in languages other than English and improving accuracy by applying different feature extraction techniques.

The paper [15] presents an application framework for rehabilitation training centers to help stroke patients overcome negative mental states using virtual reality (VR). The authors simulated a ball-catching game using VR and recognized facial expressions, body movement, and voice modality to make the game interactive and enjoyable for patients. The system analyzed patients' color and 3D depth images and voice to understand their body skeleton structure through extracted features. By sensing the action of the limbs of the patient, it increased the difficulty level of the game and changed the position of the ball itself. The paper shows that this approach improves the mental state of patients and makes them excited for the next training sessions.

In [16], the authors explain the concept of soft biometrics, which are behavioral characteristics of a body that can provide some information about one's identity without precisely recognizing their identity. Soft biometrics include physical traits like height, weight, skin color, and behavioral traits such as voice, gait, and keystroke. The authors suggest that a multimodal approach, combining more than one modality, can improve the accuracy and efficiency of biometric recognition systems, but this would increase the cost of the system. To overcome the cost issue, soft biometrics can also be identified through the same sensing equipment as the primary biometric. Future work includes designing optimal fusion schemes for primary biometric and soft biometric and extracting soft biometrics without creating inconvenience to users.

The paper [17] presents a system to recognize non-linguistic facial expressions and head movements for finding the severity of depression. The authors establish a relationship between non-verbal expressions and depression because depressed patients avoid giving expressions and maintain distance from others. The system detected facial expressions and head movements from recorded videos and extracted features using SVM classifiers and 3D CSIRO trackers. The extracted features showed the difference between depressed patients and non-depressed patients. The measures for facial expressions and head movements were AUs (Action Units) and head amplitude, velocity. Results showed that affiliative expressions were high in depression abated patients and non-affiliative expressions were high in depressed patients. Future works include analyzing other emotions like fear, anger, and sadness in addition to valence and behavior, and designing completely independent automatic reliable systems for clinical research.

The paper [18] proposes a model based on the functioning of amygdala, responsible for producing emotional instincts in the human brain, using artificial intelligence (AI) affective computing framework to recognize both instantaneous and real-time emotions. The system recognizes instantaneous emotions from facial expressions and voice using an 8-layer convolutional neural network (CNN) and extracts real-time emotions using memory layer neural networks (MLNN). The system fuses instantaneous and real-time emotions using a hidden markov model (HMM) to produce internal or intracranial emotions. These emotions are also called personalized emotions that describe the dominance, influence, steadiness, and compliance of human behavior. The authors suggest future work on building an interactive AI framework to control human emotions.

Overall, these papers show the progress and challenges in automatic emotion recognition and suggest future work in improving the accuracy and efficiency of these systems.

**DISCUSSION**

The paper discusses the importance and demand for emotion recognition systems in the field of artificial intelligence and the internet of things. The paper also notes that to improve the accuracy of these systems, body movements need to be considered in addition to facial expressions and textual information. The paper cites several naturalistic databases that have been introduced to aid in the development of capable systems that can recognize compound and subtle emotions. Cloud storage and GPU are required to efficiently handle these large databases.The paper highlights that Convolution Neural Network (CNN) gives higher accuracy for feature extraction and classification compared to other approaches, as discussed in Table 2. Multimodal automatic emotion recognition systems have great potential for future use in various fields.The paper mentions that automatic emotion recognition systems can be used in healthcare for remote areas, by military personnel, and by youth and teenagers. Therefore, emotion recognition systems have a significant role in the tech-savvy world and need to be embedded powerfully in current technology. The paper concludes that the large number of applications of emotion recognition systems is a motivating factor behind the analysis of emotion recognition and sentiment analysis study.

**CONCLUSION**

The paper provides a survey of various research papers related to emotion recognition, sentimental analysis, and the applications of emotion recognition systems. The paper notes that convolution neural networks (CNNs) are self-learning algorithms that produce good results for naturalistic databases and are best suited for reducing data overfitting and data imbalance. The paper highlights that there is a need to increase the efficiency of emotion recognition systems in terms of accuracy, and work needs to be done using RGB datasets formed under uncontrolled conditions, deep neural networks as emotion classifiers, and multi-modal behavioral systems such as body movements, facial expressions, voice, etc., to form robust automatic recognition systems.The paper also highlights the need for improved security in emotion recognition systems by using cloud storage resources and cancelable biometrics. The paper notes that emotion recognition systems have various applications in fields like healthcare, virtual reality, and robotics. The paper concludes that the future scope of work in emotion recognition systems lies in developing more efficient and accurate systems using advanced technologies like deep neural networks and multi-modal behavioral systems.