Facial Emotion Recognition for Therapy Session

Bhagyashree Sunil Kadam; Hitakshi Prakash
Chandan Yerusha Prashant Ingale; Sahil Ramdas
Jadhav Sheetal Vijay Kulkarni
Instrumentation Engineering, AISSMS
Institute of Information Technology, India

Abstract
The ability to accurately recognize another person's emotional state is a fundamental skill for psychotherapists, as it enables empathy and effective therapeutic interventions.[1] This study explores the application of facial emotion recognition (FER) technology in the context of therapy and training sessions. FER systems, utilizing advanced computer vision and machine learning techniques, can analyze facial expressions to identify and classify emotions such as happiness, sadness, anger, fear, surprise, and disgust.[1] By integrating FER into therapeutic settings, clinicians can gain objective insights into patients' emotional responses, leading to more tailored and effective treatments.[1] Furthermore, FER can assist in research by providing reliable and quantitative measures of psychotherapy process ratings, enhancing the validity and generalizability of findings. This study aims to demonstrate the potential benefits of FER in improving the quality and effectiveness of psychotherapy and training, ultimately contributing to advancements in mental health care.

Keywords:
Psychotherapy, Therapy session, FER, Computer vision and Mental health care.

1. Introduction

The Engaging with patients' emotions is a fundamental component of nearly all types of psychotherapy. However, it can be particularly challenging for psychotherapists, especially those in training, to assist patients in identifying, reflecting on, and experiencing their emotions. This difficulty may stem from various factors, including the individual characteristics and abilities of the patient, as well as those of the psychotherapist, such as mentalizing and reflective functioning, alexithymia, perceptual challenges, emotional and empathic skills, and intrapsychic or interpersonal biases. The dynamics of psychotherapeutic interactions are intricate, characterized by both verbal and non-verbal communication elements that influence one another and the therapeutic process. This study aims to explore the non-verbal dimensions of emotional communication and perception, specifically focusing on helping the trainee psychotherapists to identify non-verbal emotional cues in others and how this skill can be developed within psychotherapy training. Beyond the overt, verbal discussions regarding

DOI: https://doi.org/10.5281/zenodo.15589677

emotions, the ability of a psychotherapist to perceive and engage with patients' non- verbal emotional signals is regarded as a crucial competency in psychotherapy.[2] Non-verbal emotional expressions can manifest through various channels, such as facial expressions, body language, or vocal tone[2] and may sometimes appear only fleetingly before being concealed or altered by the individual (micro expressions).

2. Literature Survey-I

There are lot researches that have worked on Face Emotion Detection for Therapy Session.

2.1. Wafa Mellouka, Wahida Handouzi, (2009)“

Facial emotion recognition using deep learning: review and insights”, IEEE International Conference with Information and Automation, pp. 1040-1043. Expression-based emotion detection is an interesting field of research, the subject of research, and has applications in a variety of fields, including safety, health, and machine interfaces. Researchers in this field are looking for development techniques to code facial expressions, codes and extract visual features to perform better predictions through computers. The different types of architectures this offers and this was used to wrap up the methodology well. It highlights a wide range of articles in architecture and data records, showing progress achieved by comparing the methods used with the results achieved. This paper seeks to be both a study of the latest works and many suggestions for improving the field.

2.2. Amit Pandey, Aman Gupta, Radhey Shyam,

“FACIAL EMOTION DETECTION AND RECOGNITION,” (2022) Facial emotional expression lies within classification of face recognition, the recognition of facial expression has always been an easy task for humans but replicating the same in a computer algorithm is a challenge. Using recent and consistent breakthroughs in computer vision and machine learning, emotions could be detected in images, videos, etc. We propose a face expression recognition approach through the Deep Neural Networks which is a kind of deep learning modelling that is especially used for CNN and the image edge detection. After the facial expression image is normalized, the edge of each layer of the image is retrieved in convolution. The retrieved edge information is on each feature image to keep the texture picture's edge structure information. In this work multiple datasets are reviewed and analyzed for expression recognition training. This paper is about studying face emotion detection and recognition though learning algorithms and deep learning. This research article will provide deeper insight into Face emotion detection and Recognition It will also touch on the factors that affect its effectiveness.

2.3. Byoung Chul Ko - Due to its great

academic and commercial value, facial emotion (FER) detection has become one of the most fascinating topics in computer vision and artificial intelligence. Although FER can be performed if it relies on several sensors, this overview is directed to studies that are used only in view images, as visual representations are one of the main channels of information channels of interpersonal communication. This article provides an overview of the works related to FER that have been implemented over the years. This overview also focuses on current hybrid detection learning methods using folding networks (CNNs) for spatial attributes of a single frame and long short-term memory (LSTM) with fleeting features of consecutive frames.

DOI: https://doi.org/10.5281/zenodo.15589677

2.4. Jesús A, Ballesteros Gabriel, Ramírez V.

Fernando Moreira, Andrés Solano, Carlos A. Pelaez, Identificadores de Emociones Faciales a Partir del Cráneo con Redes Convolucionales Artificiales, Front. Computer. Sci., 31 January 2024" Sec.

Volume 6, 2024 Human-Media Interaction In this paper, we present a study in which we use artificial intelligence (AI) with computer vision algorithms to identify human emotions within video content while interacting with various types of visual stimuli Using the AI algorithms and customized image processing pipelines to discover users' facial expressions, the research is attempting to reveal the development of software for emotion detection

3. SCOPE of PROJECT -II

Face detection system is a computer program designed to automatically recognize or validate a person from a digital photo, or a video frame taken from a television source. One way to accomplish this is selecting facial characteristics from a facial database as well as from the image. It is most commonly used in security systems and is comparable to biometrics, fingerprints, iris recognition, and other similar technologies. For example, the algorithm may be considering the relative distance, size, and/or shape of the eyes, nose, check bones, and the jaw or some other features of the subject. the next step is to get similar matches from these attributes. Different algorithms rearrange the image ordered in the directory and then reduce the face data, keeping only the information that is valuable for face recognition. A probe image is then compared to the face data. One of the most successful prototypes has the advantage of combining both geometric and photometric methods. It includes applying of template matching methods to few facial features that creates somewhat a compressed face. Recognition algorithms are typically categorized into two groups: geometric ones that utilize specific facial traits and thus detect faces, or photometric ones that are based on a statistical method and hence distil images into values and compare the values with templates to remove any deviations.

3.3. Methodology

The entire manuscript should be in Times New Roman. Other font types may be used if required only for special purposes. Refer to Table 1 for font sizes.

3.3.1. Face Detection:

- All Gathering and preprocessing data: Compile a dataset with images labeled with the appropriate feelings (such as "angry," "disgust," "fear," "happy," "neutral," "sad," and "surprise"). You can use publicly available datasets like FER2013, CK+, or JAFFE. Preprocess the images:
  Resize images to a standard size (e.g., 48x48 pixel). Convert images to grayscale or RGB format depending on your preference. Pixel values should be normalized to fall between 0 and 1. Convert emotion labels into numerical format (e.g., one-hot encoding).
- Data Augmentation: Perform data augmentation to increase the diversity of your training dataset. Common augmentation techniques include rotation, flipping, zooming, and shifting.
- Model Building Build a CNN model using TensorFlow and Keras. Define the architecture of your CNN, including convolutional layers, pooling layers, and fully connected layers. Experiment with different architectures, activation functions, and regularization techniques (e.g., dropout) to improve the model's performance.
- Model Training: Make training, validation, and test puts from your dataset. Your CNN model should be trained on the training set using the Adam optimizer and categorical the crosses loss function. This component will assist in tracking the model's performance on the validation set and modifying the hyperparameters step in the event that overfitting is detected.

DOI: https://doi.org/10.5281/zenodo.15589677

- Model Evaluation: Finally, you can test the performance of the trained model using the test data. It is possible to compute metrics like F1-score, recall, accuracy, and precision for each feeling class. Visualize the model's performance using confusion matrices and ROC curves.
- Fine-tuning and Optimization: Fine-tune your model by experimenting with different hyperparameters and architectural modifications. Optimize the model's performance by adjusting learning rates, batch sizes, and training epochs.

3.3.1 Project Planning

- Fine-tuning and Optimization: Fine-tune your model by experimenting with different hyperparameters and architectural modifications. Optimize the model's performance by adjusting learning rates, batch sizes, and training epochs.

- Analysis of escape situations and exact types of problems through discussion with project guidelines.
- Research into the processes of various technologies used in systems.
- With the help of the guide, program specifications were determined and implemented in the project.
- It uses accelerometer sensors to interface the computer and a system built in for process and control.
- Testing, developing and troubleshooting to improve your user interface.

3.4. Details Of Designs, Working And Processes

Here's a detailed explanation of each step:

3.3.1.First Order Heading:

The first step is to detect faces in images or video frames. This can be achieved using pre-trained deep learning models such as Haar cascades, HOG detectors, or more advanced methods like convolutional neural networks (CNNs)[3]. Once faces are detected, their bounding boxes are identified.

3.3.2 Feature Extraction:

- After detecting faces, features need to be extracted from the facial regions.[4] These features can include geometric features (e.g., distance between eyes, mouth shape) and appearance-based features (e.g., texture, color)
- One common approach is to use facial landmarks, such as the positions of the eyes, nose, mouth, etc., which can be detected using pre-trained models like landmark estimation algorithms.
- Additionally, you can utilize deep mastering-based totally function extraction strategies, which include facial embeddings generated by means of networks like FaceNet or VGG-Face.

3.3.3 Emotion Classification:

- As soon as facial features are extracted, they're fed right into a machine mastering model for emotion type. This model could be a traditional machine studying set of rules like guide Vector
  - Machines (SVMs), Random Forests, or greater generally, deep mastering fashions which include Convolutional Neural Networks (CNNs).
- CNNs have shown massive success in photo- based tasks, together with facial expression reputation. They research hierarchical representations of capabilities immediately from the pictures. TensorFlow and Keras may be used to build and educate CNN fashions for emotion category.

3.3.4 Real-time Feedback/Intervention:

- In a realistic software, actual-time comments or intervention can be furnished primarily based on the detected emotions.[5] For anger control, this can involve imparting calming messages, suggesting coping strategies, or notifying a human supervisor if intervention is needed.

- Actual-time processing may be carried out using video streams from cameras.[6] OpenCV presents capabilities for taking pictures video frames and processing them in actual-

DOI: https://doi.org/10.5281/zenodo.15589677

time

- The emotion category model can be applied to every frame to constantly display the consumer's emotional nation.

3.3.5 Integration and User Interface:

- Finally, the facial recognition system wishes to be integrated into a user-pleasant interface. this may be a standalone utility or included into current platforms or gadgets.
- Python libraries inclusive of Tkinter, PyQt, or web frameworks like Flask or Django may be used to increase the person interface

3.3.6 Checking out and assessment:

- It's vital to very well check the system to make sure its accuracy, robustness, and usefulness. This entails checking out with numerous datasets, which includes unique demographics and environmental situations.

- Evaluation metrics inclusive of accuracy, precision, don't forget, and F1-rating may be calculated to evaluate the performance of the emotion reputation system.

3.4 Simulation

"Fig. 1", even at the beginning of a sentence. developing a simulation of face detection and popularity for anger control the use of Python and machine gaining knowledge of entails several steps. here's a excessive- degree overview of the way you can method this:

3.4.1 **Putting in surroundings:**

- set up Python and vital libraries like OpenCV, TensorFlow, or PyTorch for face detection and reputation.
- Optionally, installation a virtual environment to control dependencies. Each word must be capitalized except prepositions and conjunctions.

3.4.2 Statistics series:

- Annotate the dataset with labels indicating the emotional nation of each photograph (e.g., indignant, happy, sad).

3.4.3 **Records Preprocessing:**

- Convert the pics right into a format appropriate for education machine getting to know fashions.

3.4.4 **Face Detection:**

- Extract the detected faces from the snap shots for in addition processing.

3.4.5 **Feature Extraction:**

- Emotion popularity. -popular strategies include the usage of Histogram of oriented studying-based function extraction techniques.

3.4.6 **Model training:**

- Utilize techniques like cross-validation and hyperparameter tuning to optimize

3.4.7 **Version evaluation**:

- Compare the trained model the use of a separate validation dataset to evaluate its accuracy and generalization potential.
- Metrics which include accuracy, precision, take into account, and F1-score may be used to evaluate model performance.

3.4.8 **Integration:**

- Integrate the skilled version into a Python application or script that performs real-time face detection and emotion popularity.
- Make use of libraries like OpenCV to capture video from a webcam and procedure frames for emotion popularity.

DOI: https://doi.org/10.5281/zenodo.15589677

3.5. Figures

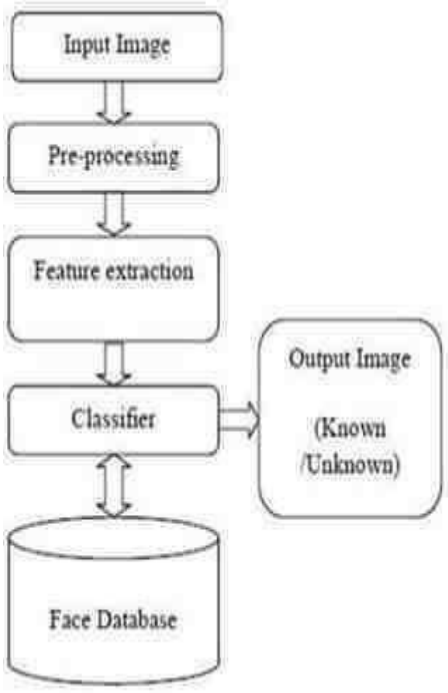

Fig.1. Block diagram

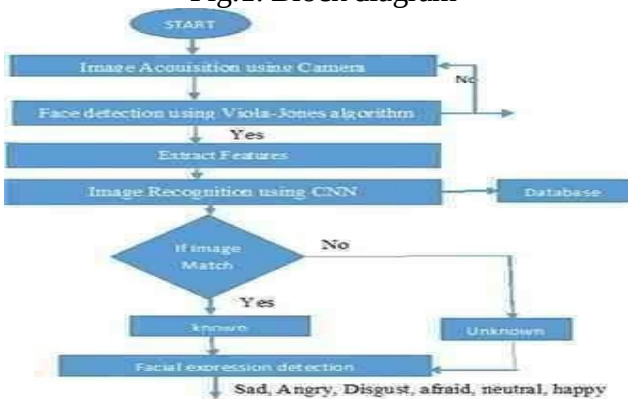

Fig.2. Flow Chart

3.4.Result

Facial recognition technology applied to anger management has yielded promising results, offering real-time detection and analysis of facial expressions associated with anger. By providing individuals with immediate feedback on their emotional states, this technology enhances self-awareness, empowering individuals to intervene before their emotions escalate. Through repeated use, users can learn to recognize their triggers and develop personalized strategies for emotional regulation. Therapists can utilize facial recognition data to tailor interventions to individual needs, leading to more effective treatment outcomes. Moreover, the accessibility of this technology expands support for anger management to a wider audience. In high-stress environments, such as workplaces or prisons, facial recognition enables proactive identification of individuals at risk of anger outbursts, facilitating timely interventions to prevent conflicts. While these advancements hold promise, ongoing research and ethical considerations are paramount to ensure responsible and effective implementation.

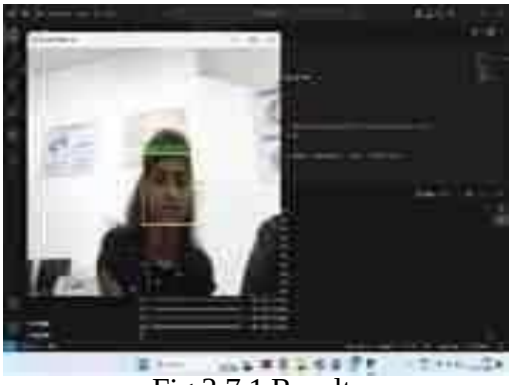

Fig.3.7.1 Results

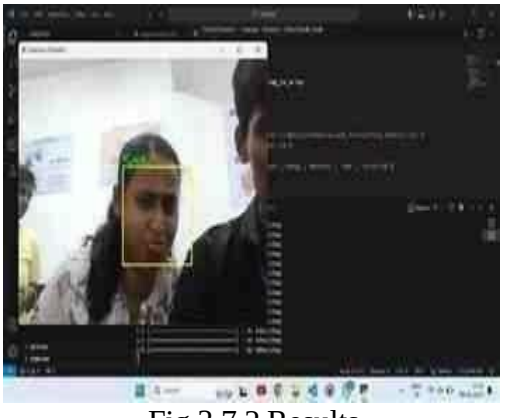

Fig.3.7.2 Results

3.5.Application

Applications are as follows:

- Detection of Facial Cues: Facial recognition algorithms can detect

DOI: https://doi.org/10.5281/zenodo.15589677

subtle changes in facial expressions associated with anger[7], such as furrowed brows, narrowed eyes, or tightened lips. By analyzing these facial cues, the technology can identify signs of anger in real-time

- Real-time Feedback: Facial recognition systems can provide real-time feedback to individuals, alerting them when signs of anger are detected.[7] This feedback can help individuals become more aware of their emotional states and take steps to manage their anger effectively.
- Training and Therapy: Facial recognition technology can be integrated into anger management training programs or therapy sessions. Therapists can use the data collected by the technology to track progress over time and tailor interventions to individual needs.
- Biofeedback Devices: Facial recognition technology can be integrated into wearable biofeedback devices that monitor physiological responses to anger, such as changes in heart rate or skin conductance. By combining facial recognition with other biometric data, these devices can provide a comprehensive picture of an individual's emotional state.
- Preventive Measures: In settings where anger management is crucial, such as prisons or high- stress work environments, facial recognition technology can be used as a preventive measure to identify individuals who are at risk of losing control and intervene before conflicts escalate.[8]
- Accessibility: Facial recognition technology can also be used to make anger management tools more accessible to a wider range of individuals, including those who may have difficulty expressing or recognizing emotions verbally.

4.Conclusion

This study has potential of Facial Emotion Recognition (FER) in enhancing psychotherapy and therapy sessions. The integration of FER provides objective, real-time insights into patients' emotional states, which help therapists to refine their skills through data-driven feedback. The integration of FER also offers quantifiable, reliable metrics for analyzing psychotherapy processes. Future research should focus on further refining the accuracy and effectiveness of these systems, as well as addressing ethical concerns to ensure responsible and equitable implementation.

Appendix

1. Techniques For Emotion Detection:

- Facial Expression Recognition (FER): This technique analyzes facial features like eyebrows, eyes, mouth, and wrinkles to infer emotional state. FER is a non-invasive method but can be influenced by cultural background and individual expression variations.
- Physiological Signal Analysis: This approach monitors physiological signals like heart rate, skin conductance, and muscle tension, which can change with emotional arousal. However, it requires specialized equipment and may not be practical for everyday use.
- Speech Analysis: Speech characteristics like pitch, volume, and prosody can provide clues about emotional state. However, speech analysis can be influenced by factors like background noise and speaking style.

2. Strengths Of Fer For Anger Management

- Increased Self-Awareness: Real-time feedback on facial expressions associated with anger can help individuals become more aware of their emotional state, especially if they tend to suppress anger cues.
- De-escalation Prompts: FER systems can trigger prompts or reminders for calming techniques (deep breathing, relaxation exercises) or suggest removing oneself from a situation when anger is detected.
- Data for Personalized Strategies: FER data collected over time can be used with a therapist to identify triggers and patterns related to anger outbursts. This data can inform the development of personalized coping mechanisms.

3. Accuracy Concerns:

DOI: https://doi.org/10.5281/zenodo.15589677

- FER technology is still under development, and accuracy can vary depending on the system and individual. Facial expressions alone are not always a reliable indicator of emotion.
- Cultural Differences: Facial expressions associated with anger can vary across cultures. FER systems may need to be trained on culturally diverse datasets to improve accuracy.
- Masking and Individual Variations: People can deliberately control their facial expressions (masking) or have unique ways of expressing emotions. FER systems may struggle with these variations.
- Privacy Concerns: The use of FER, especially in public spaces, raises privacy concerns. Individuals may not want their emotions constantly monitored.
- Over-reliance on Technology: FER should be a tool to support anger management, not a replacement for therapy or self-management techniques.

Acknowledgement

Our profound appreciation goes out to everyone who helped us finish this study. We want to start by sincerely thanking our mentors and advisors, whose wise counsel and perceptive criticism have greatly influenced the direction of this research. We are also incredibly grateful to the individuals who gave their time to participate in this study, since they gave us vital information and viewpoints. Their readiness to help with the investigation of emotional nonverbal communication in psychotherapy training has been really beneficial. We would especially want to thank our peers and coworkers, whose conversations and encouragement have deepened our comprehension of this intricate topic. Their support and helpful criticism have been invaluable in helping us improve our work.

References

R. D. Ward and P. H. Marsden, “Affective computing: problems, reactions and intentions,” vol. 16, 2004.

H. Zhang, Z. Qu, L. Yuan, and G. Li, “A face recognition method based on LBP feature for CNN,” 2017

M. Z. Uddin, M. M. Hassan, A. Almogren, M. Zuair, G. Fortino, and J. Torresen, “A facial expression recognition system using robust face features from depth videos and deep learning,” Comput. Electr. Eng., vol. 63, 2017.

J. Hakura, R. Domon, and H. Fujita, “Emotion recognition method using facial expressions and situation,” 2013.

A. Kumar and A. Agarwal, “Emotion recognition using anatomical information in facial expressions”, 2014.

P. SIGCHI (Group: U.S.), S. SIGART, J. G. IEEE Robotics and Automation Society, and Institute of Electrical and Electronics Engineers, Christchurch, NZ. IEEE Press, 2016.

B. Ryu, A. R. Rivera, J. Kim, and O. Chae, Local Directional Ternary Pattern for Facial Expression Recognition,” IEEE Trans. Image Process., vol. 26, 2017.

Q. Mao, Q. Rao, Y. Yu, and M. Dong, “Hierarchical Bayesian Theme Models for Multipose Facial,” vol.5,2014.

L. Zhang and D. Tjondronegoro, “Facial Expression Recognition Using Facial Movement Features,”, vol.2, 2011.

DOI: https://doi.org/10.5281/zenodo.15589677