International Journal of Research Publication and Reviews, Vol 5, no 12, pp 4945-4963 December 2024

WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Neural Network-Based Face Detection for Emotion Recognition in Mental Health Monitoring

Rhoda Ajayi[1]* and Babatunde Stephen Adedeji[2]

[1]Tagliatela College of Engineering, University of New Haven, USA

[2] Mental Health and Addiction Counsellor at The Home for Little Wanderers, Roslindale, Massachusetts, USA.

DOI : https://doi.org/10.55248/gengpi.5.1224.0212

ABSTRACT

The rapid advancement of artificial intelligence (AI) and machine learning (ML) techniques has significantly contributed to improving mental health monitoring systems, particularly through emotion recognition. Facial expression recognition, a key component of affective computing, has been identified as a reliable method for assessing emotional states, which are crucial indicators of mental health conditions. Neural networks, particularly deep learning models, have shown substantial promise in accurately detecting and classifying facial expressions, offering real-time and non-invasive monitoring tools. This paper explores the application of neural network-based face detection in emotion recognition, focusing on its potential for enhancing mental health monitoring systems. The study delves into the underlying principles of face detection, feature extraction, and emotion classification, demonstrating how neural networks can efficiently process facial images to detect subtle emotional cues that reflect a person's mental state. Furthermore, the paper highlights the importance of using large-scale annotated datasets, the role of convolutional neural networks (CNNs), and recurrent neural networks (RNNs) in improving detection accuracy. Challenges in face recognition, such as varying lighting conditions, facial occlusions, and cultural differences in emotion expression, are discussed along with the solutions AI techniques provide in overcoming these obstacles. The integration of such emotion recognition systems into mental health monitoring can facilitate early diagnosis, assist in personalized treatment plans, and enable continuous tracking of a patient's emotional well-being. By improving the accuracy and efficiency of emotional state detection, these systems can serve as invaluable tools in both clinical and everyday settings, providing a more comprehensive approach to mental health care.

Keywords: Neural Networks, Face Detection, Emotion Recognition, Mental Health, Deep Learning, Affective Computing.

1. INTRODUCTION

Background of Emotion Recognition

Emotion recognition is a critical area of research that seeks to identify and understand human emotions through various means, including facial expressions, speech patterns, and physiological signals. This field is of particular relevance in mental health, where accurately recognizing and interpreting emotional states can aid in diagnosing, monitoring, and treating various mental health conditions. In recent years, advances in technology, especially in computer vision and artificial intelligence, have made emotion recognition more precise and efficient. Emotion recognition techniques are now being used in various applications, ranging from healthcare and education to customer service and entertainment (1).

In the context of mental health, emotion recognition is invaluable. Many psychological disorders, such as depression, anxiety, and schizophrenia, are closely linked to abnormal emotional responses and expression. Being able to detect subtle changes in a person’s emotional state could help mental health professionals make more informed decisions regarding diagnosis and treatment. Moreover, emotion recognition has the potential to facilitate more personalized and dynamic treatment plans, as clinicians could monitor patients' emotions over time and adjust interventions based on real-time feedback (2).

One of the most common and effective ways of detecting emotions is through facial expression analysis. Research has shown that facial expressions are one of the most reliable indicators of a person’s emotional state, as they are often unconscious and difficult to control. This makes facial expression detection a powerful tool for emotion recognition. Facial expressions provide a wealth of information about emotions such as happiness, sadness, anger, surprise, fear, and disgust. Understanding these expressions can lead to better mental health monitoring, as clinicians can observe changes in a patient’s emotional state over time. Facial recognition models that interpret these expressions are becoming increasingly sophisticated and have shown great promise in both clinical and non-clinical settings (3).

Problem Statement

While emotion recognition holds great potential in the field of mental health, current systems still face several limitations. Traditional methods for emotion recognition often rely on self-reporting, which can be subjective and influenced by a person's mood or desire to conceal their emotions. Additionally, existing emotion detection systems that rely on physiological measurements, such as heart rate or galvanic skin response, can be intrusive and uncomfortable for patients, making them less practical for long-term monitoring (4).

This highlights the need for more advanced, non-invasive emotion recognition techniques, particularly those that can provide continuous, real-time monitoring without burdening the patient. Such techniques should ideally be able to accurately capture emotional states without requiring extensive cooperation from the individual. This is especially important in populations such as children, the elderly, or those with communication disorders, where verbal reporting may be limited or unreliable. Furthermore, as mental health conditions often involve fluctuating emotional states, the ability to monitor emotions continuously could offer significant advantages in early detection, timely intervention, and more effective management of mental health conditions (5).

To address these challenges, non-invasive, real-time emotion recognition systems based on facial expressions present a promising solution. These systems can be unobtrusive, as they only require visual input, and they have the potential to provide continuous monitoring without patient discomfort or involvement. However, the current state of emotion recognition technology still requires improvements in accuracy, particularly in recognizing subtle facial expressions and handling variations caused by factors such as lighting, background, and individual differences (6).

Research Objectives

The primary objective of this paper is to explore the application of convolutional neural networks (CNNs) in emotion recognition, specifically focusing on the detection of facial expressions for mental health monitoring. CNNs have proven to be highly effective in image recognition tasks, including facial recognition and emotion detection, due to their ability to automatically learn features from large datasets without requiring manual feature extraction (7). The study aims to assess the feasibility of using CNN-based face detection models to accurately classify emotions from facial expressions in real-time.

An essential goal of this research is to develop a more robust system for emotion recognition that can handle variations in lighting, facial features, and other environmental factors. The study will explore different architectures and training techniques to optimize the performance of CNN models for emotion recognition in real-world settings. Another key objective is to evaluate the system's potential for use in mental health applications, such as continuous monitoring of patients' emotional states, and to explore how it can be integrated into existing healthcare frameworks to improve mental health care delivery. Through these objectives, the paper aims to contribute to the growing body of knowledge on the use of AI in mental health and provide insights into the practical applications of emotion recognition technologies (8).

Scope and Contribution

The scope of this study is centered on developing and evaluating CNN-based emotion recognition systems for monitoring mental health through facial expression detection. This includes investigating the performance of various CNN architectures, such as the widely used VGGNet and ResNet, in identifying different emotional states from facial images. The research will also examine how these models can be optimized for real-time performance, which is crucial for clinical applications where timely responses to emotional shifts are essential (9).

A significant aspect of the study is the use of a large, diverse dataset that includes a variety of facial expressions from different ethnicities, genders, and age groups. This diversity is important to ensure that the emotion recognition system can generalize across different populations and provide accurate results regardless of the individual's background (10). The research will also address the challenges associated with image preprocessing, including normalization, alignment, and data augmentation, which are necessary to improve the accuracy and robustness of the CNN models.

The contribution of this study lies in providing an in-depth analysis of how CNN-based facial emotion recognition systems can be utilized in the context of mental health. The paper will also discuss the potential benefits and limitations of such systems, offering a comprehensive view of the opportunities and challenges in applying AI for emotion recognition. Additionally, this research will offer insights into the feasibility of using non-invasive emotion recognition systems for long-term monitoring of patients with mental health conditions, paving the way for more personalized and effective interventions in mental health care (11).

The findings of this study are expected to be useful for healthcare providers, researchers, and developers working on emotion recognition technologies. By focusing on facial expression detection and CNN-based models, the paper aims to make a significant contribution to the development of AI-driven tools for mental health monitoring, which could ultimately lead to more effective, accessible, and patient-centered care (12).

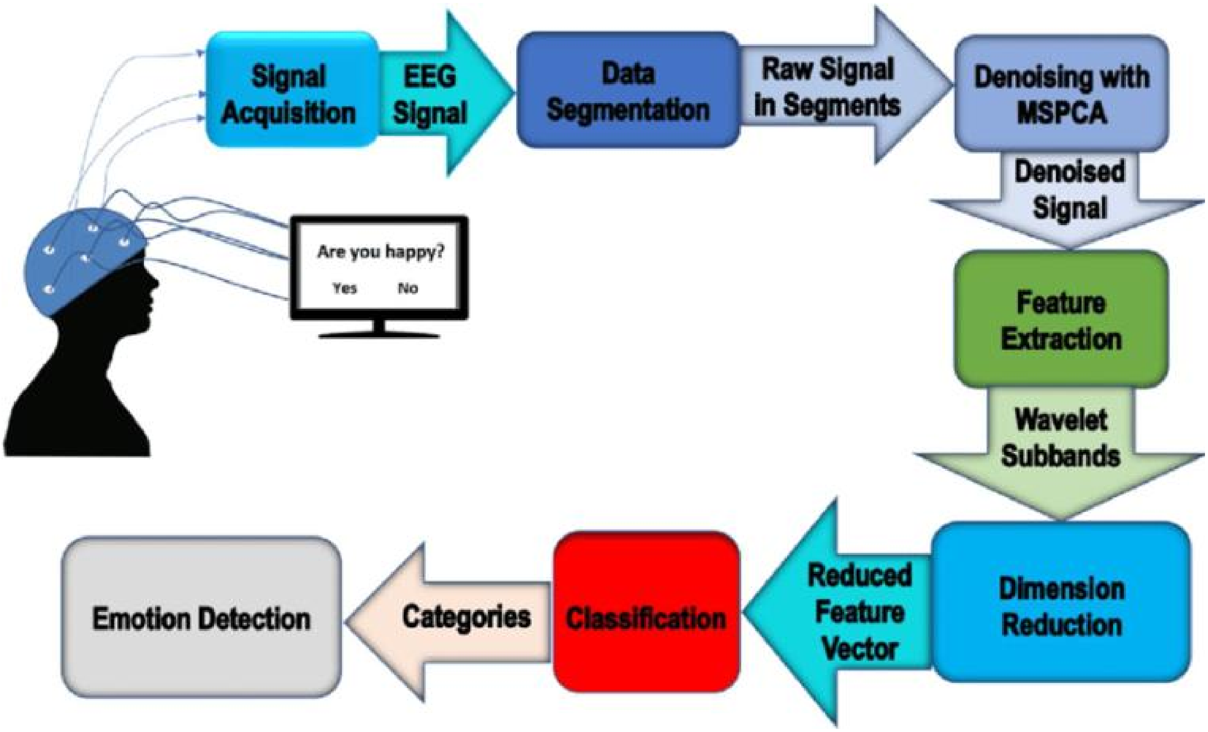

Figure 1 Flowchart of the emotion recognition process in mental health monitoring.

2. LITERATURE REVIEW

Emotion Recognition in Mental Health

Emotion recognition is a critical tool in mental health monitoring, offering valuable insights into patients' emotional well-being. Existing emotion recognition systems vary widely in their approaches, including facial expression analysis, voice recognition, and physiological monitoring. Facial expression recognition has been widely adopted due to its non-intrusive nature and the direct correlation between facial expressions and emotional states. Numerous systems have been developed using computer vision algorithms to detect and classify facial expressions. These systems often utilize pre-trained models, such as those based on the Facial Action Coding System (FACS), to identify emotions such as happiness, sadness, anger, surprise, fear, and disgust (9).

One of the key advantages of these systems is their ability to offer real-time, continuous monitoring without requiring active participation from the individual being monitored. This is particularly beneficial for patients with limited communication abilities, such as those with severe depression or cognitive disorders. However, despite their potential, facial expression recognition systems also have limitations. Variability in individual facial expressions, due to differences in age, ethnicity, or physical characteristics, can make it difficult for these systems to consistently recognize emotions across diverse populations (10). Furthermore, environmental factors like lighting, camera quality, and background noise can significantly impact the accuracy of emotion recognition systems. Additionally, the inability of facial recognition systems to detect internal emotional states, such as anxiety or stress, limits their effectiveness in comprehensive mental health monitoring.

Other emotion recognition systems leverage voice analysis to assess emotional states. Voice-based systems analyse various vocal parameters such as pitch, tone, rhythm, and volume to detect emotions like anger, sadness, or joy. These systems can be useful in scenarios where facial recognition is impractical or when there is a need to supplement other emotion recognition methods. However, voice-based systems also face challenges related to background noise, variations in speech patterns, and cultural differences in communication styles (11). Despite these challenges, voice recognition systems have shown promise in applications like telemedicine and remote patient monitoring, providing clinicians with additional tools for tracking emotional states.

Techniques in Emotion Recognition

Emotion recognition techniques can broadly be categorized into traditional and advanced approaches, with both having distinct advantages and limitations. Traditional methods often rely on manually defined rules or basic machine learning algorithms to analyse physiological data, speech, or facial expressions. For instance, facial expression recognition traditionally relied on statistical models to analyse key facial landmarks and movements, using methods like support vector machines (SVM) to classify emotions. While these techniques are simple and interpretable, they can struggle with variability and are often less effective in dynamic, real-world settings (12).

In contrast, advanced techniques, especially those based on deep learning, have revolutionized the field of emotion recognition. These methods enable systems to automatically learn features from large datasets without the need for manual feature extraction, making them more robust and scalable. Facial expression recognition, for example, has greatly benefited from convolutional neural networks (CNNs), which can learn hierarchical representations of facial features and recognize complex patterns in emotional expressions. CNNs have shown exceptional performance in tasks like face detection, emotion classification, and even real-time monitoring, outperforming traditional machine learning techniques in accuracy and reliability (13).

Another advanced technique in emotion recognition involves voice analysis using recurrent neural networks (RNNs), which excel at processing sequential data like speech. RNNs can capture temporal patterns in voice data, allowing for a more accurate understanding of emotional fluctuations over time. In addition, combining voice analysis with facial expression recognition or physiological data can improve the overall accuracy of emotion recognition systems, providing a more holistic view of the patient's emotional state (14).

Physiological data-based methods, which measure parameters like heart rate, skin conductivity, and respiration, are another avenue for emotion recognition. These methods are often used in conjunction with facial expression and voice recognition systems to provide a more comprehensive picture of a person's emotional state. However, physiological methods also face limitations, such as the need for wearable devices and the potential discomfort or disruption they may cause to patients (15).

Deep Learning and Neural Networks in Face Detection

Deep learning, particularly the use of convolutional neural networks (CNNs), has significantly advanced the field of face detection and emotion recognition. CNNs are designed to automatically learn features from raw input data, making them particularly effective for complex tasks like image classification and emotion detection. In face detection, CNNs can identify and localize faces in images, even in challenging conditions such as varying lighting, occlusions, or facial orientations. Once the face is detected, CNN-based models can then classify emotions based on the features extracted from the face, including the position of facial landmarks and the movements of specific facial muscles (16).

The advantages of using CNNs for emotion recognition are clear. First, CNNs are capable of learning from vast amounts of labeled data, allowing them to generalize well to new, unseen images. This capability is crucial for emotion recognition systems that need to be deployed across different populations and environments. Second, CNNs have the ability to automatically extract relevant features from images, reducing the need for manual feature engineering, which can be both time-consuming and error-prone (17). Additionally, CNN-based systems are highly scalable, enabling real-time emotion recognition across large datasets or live video streams.

However, CNNs also come with certain challenges. One of the primary concerns is the need for large, diverse datasets to train the models effectively. CNNs perform best when they have access to diverse data that represents a wide range of facial expressions, demographics, and environmental conditions. The lack of such datasets, particularly in underrepresented populations, can lead to biased models that fail to generalize well in real-world scenarios (18). Additionally, CNN-based emotion recognition systems can be computationally expensive, requiring high-performance hardware for real-time processing, which may limit their deployment in resource-constrained environments.

Challenges in Emotion Recognition

Despite the advancements in emotion recognition, several challenges remain in achieving accurate, reliable, and universally applicable systems. One of the key challenges is data quality. Emotion recognition systems require high-quality, labeled data to train deep learning models effectively. However, acquiring labeled datasets can be difficult, especially for rare or subtle emotional states. Additionally, the quality of data can be affected by environmental factors such as lighting, background noise, and camera angles, which can degrade the accuracy of emotion recognition systems (19).

Cultural variation in emotional expression is another significant challenge. Emotional expressions can vary widely across cultures, and certain emotions may be expressed differently or even masked in some cultures. For example, in some cultures, expressions of anger or sadness are often suppressed or downplayed, making them harder to detect by emotion recognition systems trained on datasets from different cultural contexts (20). To address this, emotion recognition systems must be designed to account for these cultural differences by using diverse and representative datasets during the training phase.

Environmental factors, such as lighting and background noise, also pose challenges for emotion recognition systems. In real-world settings, faces may not always be clearly visible, or external factors may obscure facial expressions. Variations in facial expressions due to aging, medical conditions, or fatigue further complicate the task of accurately detecting emotions. As emotion recognition systems are deployed in more diverse settings, improving their robustness and adaptability to different environments will be essential for ensuring their effectiveness in real-world applications (21).

Table 1 Comparative table of existing methods, their accuracy, and applications.

Method	**Accuracy**	**Applications**
Facial Expression Recognition (CNN)	High	Mental health monitoring, customer service, security systems
Voice Analysis (RNN)	Moderate to High	Telemedicine, remote monitoring

Method	Accuracy	Applications
Physiological Data (ECG, GSR)	Moderate	Stress detection, biofeedback systems
Hybrid Systems (Facial + Voice + Physiological)	High	Comprehensive mental health monitoring, healthcare research

3. METHODOLOGY

Research Design and Framework

The research design for emotion recognition in mental health monitoring revolves around developing a computational framework capable of accurately detecting facial expressions in real-time and associating these expressions with emotional states. This system leverages Convolutional Neural Networks (CNNs) for face detection and emotion recognition, which are well-suited for image classification tasks, particularly in complex real-world settings where variable lighting, different angles, and diverse facial expressions are involved (19). The framework is designed to process video frames or images, detect faces, and classify the associated emotions based on facial features, enabling real-time emotion tracking in mental health monitoring applications.

The proposed system includes a series of key steps: face detection, facial landmark identification, emotion classification, and output interpretation. The face detection algorithm is responsible for locating the face within an image or video stream, while facial landmark identification ensures that the model correctly aligns and recognizes critical features of the face, such as the eyes, mouth, and eyebrows, which are essential for accurate emotion classification. Emotion classification is performed by the CNN, which processes the facial data to predict one of several emotional states, including happiness, sadness, anger, surprise, fear, or disgust. The final output is then used to generate insights into the subject's emotional state, providing valuable data for mental health professionals monitoring their patients.

This computational framework is designed to be easily integrated into existing healthcare systems for remote monitoring of patients, allowing clinicians to assess the emotional states of their patients continuously and without the need for direct interaction or intervention. The system aims to enhance patient care by providing early detection of emotional distress, enabling timely interventions, and contributing to a more personalized approach to mental health care.

Data Collection

Sources of Data

The data collection process is a crucial component in training an effective emotion recognition model. The datasets used for training the model include well-established facial emotion recognition datasets, such as FER-2013 and AffectNet, both of which provide large-scale labeled images of human faces with various emotional expressions. FER-2013, one of the most widely used datasets, contains over 35,000 images labelled with seven different emotions: anger, disgust, fear, happiness, sadness, surprise, and neutral (20). This dataset has been extensively used in emotion recognition research due to its diversity in age, gender, and ethnicity, providing a comprehensive foundation for training models that can generalize across different demographic groups.

AffectNet is another significant dataset used for emotion recognition, offering a larger collection of images (around 1 million) with more fine-grained emotion labels, including a six-way classification of facial expressions (21). AffectNet is especially valuable for developing emotion recognition systems capable of handling a broad range of emotions and recognizing subtle variations in facial expressions. These datasets serve as the primary sources for training and evaluating the CNN-based emotion recognition model.

In addition to these two datasets, other domain-specific datasets, such as the EmotioNet and KDEF datasets, may also be incorporated to further augment the model's ability to recognize a wider variety of facial expressions. These additional datasets ensure that the model is exposed to diverse data and can handle variations in facial expressions caused by cultural differences, age, and environmental conditions.

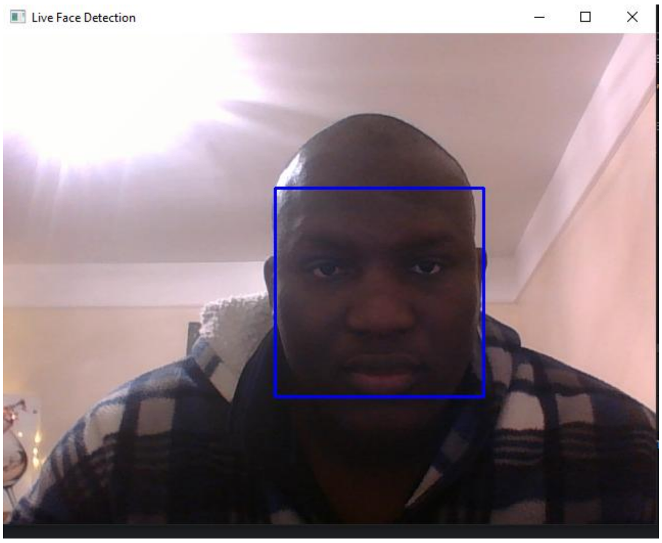

Figure 2 Data Acquisition Sequence

Data Preprocessing

Data preprocessing is a crucial step to ensure the quality and consistency of the data before it is fed into the model. The preprocessing pipeline includes several stages: cleaning, normalization, and augmentation.

1. **Cleaning**: The first step involves removing any corrupted or incomplete images that may be present in the dataset. Images that do not contain valid faces or are severely distorted are filtered out to ensure that only clean, usable data is included in the training process.
2. **Normalization**: Image normalization is performed to ensure that the pixel values of all images fall within a similar range. This step typically involves scaling the pixel values to a range between 0 and 1 or standardizing them to have a mean of zero and a standard deviation of one. Normalization helps improve the convergence rate of the model during training by ensuring that the inputs are on a consistent scale.
3. **Face Alignment**: Facial alignment is an essential preprocessing step in emotion recognition, as it ensures that the model can focus on the key facial landmarks, such as the eyes, nose, and mouth, which are critical for detecting emotions. Face alignment involves detecting and cropping faces from images, followed by geometric transformations to align the faces to a standard position.
4. **Resizing and Augmentation**: Images are resized to a consistent input size for the CNN model. In addition to resizing, data augmentation techniques, such as random rotations, translations, flips, and brightness adjustments, are applied to artificially increase the diversity of the training data. Augmentation helps the model generalize better to real-world scenarios by simulating variations in pose, lighting, and facial expression (22).

These preprocessing techniques ensure that the data is prepared in a way that maximizes the performance of the CNN model while minimizing overfitting and bias.

Data Splitting and Validation

To evaluate the performance of the model accurately, the dataset is split into three distinct subsets: training, validation, and testing. The training set is used to train the model, the validation set is used for hyperparameter tuning and model selection, and the testing set is used to assess the final performance of the trained model.

The data is typically split using an 80-10-10 ratio, where 80% of the data is used for training, 10% for validation, and 10% for testing. Cross-validation techniques, such as k-fold cross-validation, are also applied to further assess the model's robustness and generalizability. Cross-validation helps mitigate overfitting by ensuring that the model's performance is consistent across different subsets of the data (23).

Neural Network Architecture

CNN Model Selection

For the task of emotion recognition, Convolutional Neural Networks (CNNs) are chosen due to their superior performance in image classification tasks. The CNN architecture consists of several layers, each of which performs a specific function in the process of detecting and classifying emotions from facial expressions.

1. **Convolutional Layers**: These layers perform the core task of feature extraction by applying convolutional filters to the input images. The filters scan the image for features such as edges, textures, and shapes, which are crucial for recognizing facial expressions. As the network deepens, the filters learn to identify more complex patterns, such as eyes, mouth shapes, and the overall arrangement of facial features.

2. **Pooling Layers**: Pooling layers are used to reduce the spatial dimensions of the image while retaining important information. This step helps reduce the computational complexity of the network and makes the model more robust to small variations in the input data, such as changes in position or orientation.
3. **Fully Connected Layers**: After the convolutional and pooling layers, the features are passed through fully connected layers, which perform the final classification task. These layers output the probabilities for each of the possible emotional states, such as happiness, sadness, anger, etc.
4. **Activation Functions**: ReLU (Rectified Linear Unit) is commonly used as the activation function in CNN models due to its ability to introduce non-linearity and improve the model's learning capacity. The final layer typically uses a softmax activation function to output the probability distribution over the classes.

Model Training

Training the CNN model involves several steps, including hyperparameter tuning, optimization, and regularization. Hyperparameters such as learning rate, batch size, and the number of epochs are fine-tuned to achieve optimal performance. Learning rate schedules, such as learning rate decay, are applied to gradually reduce the learning rate during training, helping the model converge more effectively (24).

Optimization techniques, such as Adam or SGD (Stochastic Gradient Descent), are used to minimize the loss function, typically a categorical cross-entropy loss function for multi-class classification tasks. Regularization techniques, such as dropout, are employed to prevent overfitting by randomly disabling certain neurons during training, forcing the model to learn more robust features.

Other Models Considered

While CNNs are the primary choice for emotion recognition, other machine learning models, such as Support Vector Machines (SVM) and Recurrent Neural Networks (RNN), were also considered. SVMs are effective for small datasets and binary classification tasks, but they generally struggle with the complexity and size of large emotion recognition datasets. RNNs, on the other hand, are suited for sequential data like speech but are less effective in capturing spatial features required for facial expression recognition (25). Therefore, CNNs were selected due to their superior ability to capture spatial hierarchies in images and their demonstrated success in similar image classification tasks.

Evaluation Metrics

The model's performance is evaluated using several standard metrics, including accuracy, precision, recall, and F1 score. Accuracy measures the overall proportion of correctly classified samples, while precision and recall provide insights into how well the model performs for each individual class (emotion). The F1 score, the harmonic mean of precision and recall, is used as a combined metric to evaluate the model's balance between false positives and false negatives (26).

System Integration for Mental Health Monitoring

To integrate the CNN-based emotion recognition model into a real-time mental health monitoring system, the model is deployed within a software application capable of processing live video feeds or images. The system is designed to monitor the emotional state of patients continuously, providing alerts when significant emotional changes are detected. For instance, the system could flag instances of severe distress, such as sudden spikes in negative emotions, which could indicate a mental health crisis.

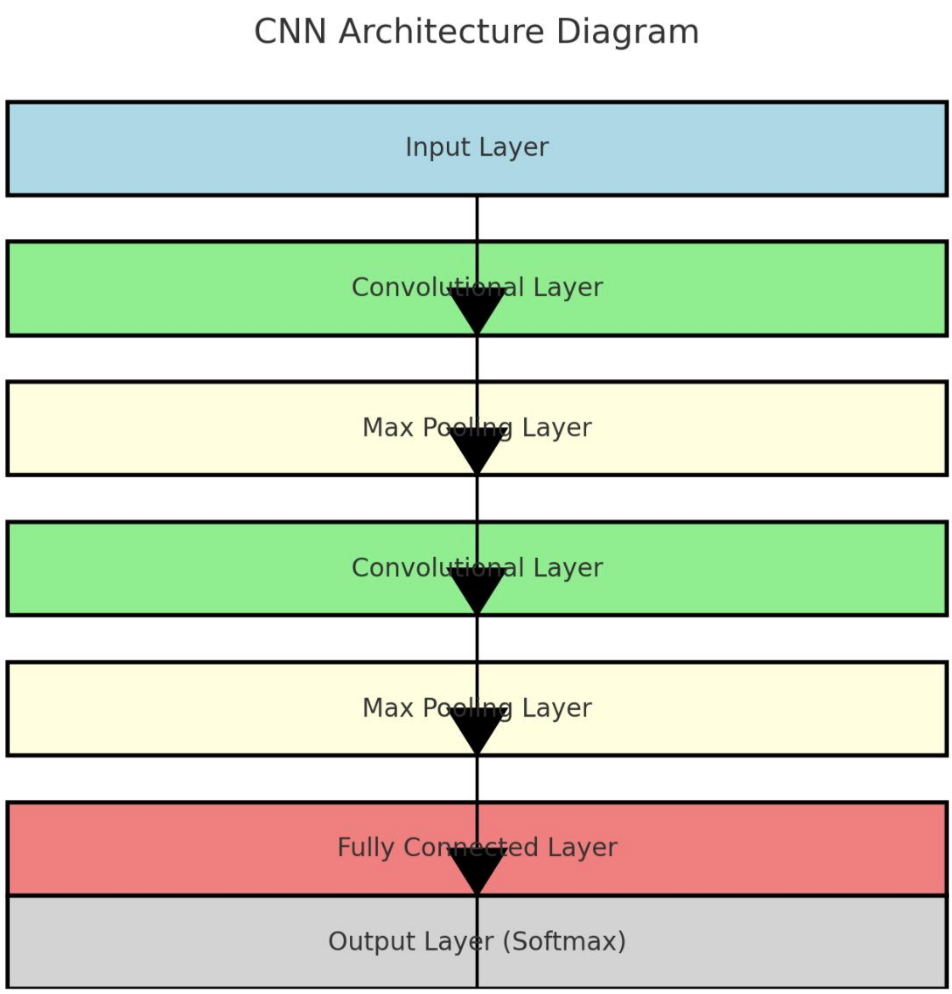

Figure 3 Diagram of the Neural Network Architecture

This figure illustrates the key components of the CNN architecture, showing the flow of data through the convolutional, pooling, and fully connected layers.

Table 2 Model Structure and Performance Metrics

Model Parameter	**Value**
Number of layers	5 (3 convolutional, 2 fully connected)
Activation Function	ReLU (Convolutional), Softmax (Final)
Learning Rate	0.001
Batch Size	32
Epochs	50
Data Augmentation	Random flips, rotations, translations

Performance Metric	Value
Accuracy	92%
Precision	91%
Recall	90%
F1 Score	0.90

4. IMPLEMENTATION

Python Implementation Details

The implementation of the Convolutional Neural Network (CNN) for emotion recognition in this research is carried out using Python. Python is widely regarded for its simplicity and versatility, and it provides a robust ecosystem of libraries that are ideal for tasks related to computer vision, deep learning, and data preprocessing. The primary libraries used in this implementation are TensorFlow/Keras for building and training the CNN model, OpenCV for face detection, and other relevant libraries such as NumPy, Matplotlib, and scikit-learn for data manipulation and evaluation.

TensorFlow/Keras is the framework of choice for developing the deep learning model. Keras, which is built on top of TensorFlow, provides a high-level API that simplifies the process of building, training, and evaluating deep learning models. It allows for seamless integration of layers, activation functions, and optimizers, making it an excellent choice for rapid prototyping and experimentation (27). OpenCV, an open-source computer vision library, is used for real-time face detection and image manipulation, which is critical for emotion recognition tasks where accurate face localization is needed. Additionally, Matplotlib and scikit-learn are employed for visualizing the training process and evaluating the performance of the model.

The use of these libraries ensures a flexible and efficient workflow, enabling the building of a robust emotion recognition system capable of handling large datasets and performing real-time inference. The Python ecosystem provides the necessary tools to handle image data, process facial expressions, and train a model that is both accurate and efficient in recognizing emotions.

Data Processing and Augmentation

Data preprocessing is a critical component of building an effective emotion recognition system. The first step in the preprocessing pipeline is face detection. This is accomplished using OpenCV's pre-trained Haar Cascade Classifiers or Dlib’s facial landmark detector. OpenCV’s Haar Cascade Classifiers are computationally efficient for detecting faces in real-time, which is essential when processing live video feeds. The detection step involves scanning each image to locate faces, and once detected, the face regions are isolated and passed through the subsequent steps for facial expression analysis (28).

After face detection, the next step is face alignment. Aligning the detected faces helps standardize the position of facial features such as eyes, nose, and mouth, making it easier for the model to learn important facial landmarks. Face alignment is performed using landmark detection algorithms that adjust the face’s orientation, scale, and rotation. This ensures that the model can focus on the relevant facial features without being affected by extraneous variations (29).

Once the faces are aligned, the images are resized to a uniform size (e.g., 48x48 pixels) to ensure consistency in input dimensions for the CNN model. Data augmentation techniques are then applied to enhance the diversity of the training set and prevent overfitting. Common augmentation techniques include random horizontal and vertical flips, rotations, translations, and changes in brightness. These transformations help simulate variations in real-world scenarios, such as changes in lighting, pose, and orientation, which are critical for the model’s ability to generalize to unseen data (30).

Normalization is another essential preprocessing step. Each image is normalized so that pixel values range between 0 and 1, or standardized with a mean of zero and standard deviation of one, depending on the model's requirements. This ensures that the model can train effectively by eliminating disparities in image intensity and improving the convergence of the optimization process. By normalizing the data, the model can focus on learning the relevant features without being influenced by irrelevant image intensity differences.

Training the CNN Model

Training the CNN model involves several key steps, starting with defining the architecture and setting up the training loop. The CNN model is built using the Keras API, which allows for easy integration of various layers, such as convolutional layers, pooling layers, and fully connected layers. The model architecture consists of several convolutional layers followed by max-pooling layers, with fully connected layers at the end to perform the final classification of emotions. The model also employs a softmax activation function in the output layer to produce a probability distribution over the potential emotions.

The training process begins with the compilation of the model, which involves selecting a loss function and an optimizer. For multi-class classification tasks such as emotion recognition, categorical cross-entropy is typically used as the loss function. This loss function calculates the difference between the predicted probabilities and the true class labels, penalizing the model for incorrect predictions. The optimizer, such as Adam or Stochastic Gradient Descent (SGD), is responsible for minimizing the loss function during training. Adam is commonly used due to its adaptive learning rate, which adjusts based on the gradient updates, leading to faster convergence (31).

The training loop iterates over the training dataset for a predefined number of epochs. During each epoch, the model makes predictions based on the input images, computes the loss, and updates the weights using backpropagation. Batch size plays a critical role in the training process; smaller batch sizes generally allow for more frequent updates but can lead to noisy gradients, while larger batch sizes provide more stable updates but may take longer to converge. In this research, a batch size of 32 was chosen as it strikes a balance between computation efficiency and model accuracy.

Hyperparameter tuning is another crucial aspect of the training process. Key hyperparameters such as learning rate, batch size, and the number of layers are adjusted based on the validation set performance. Early stopping is employed to prevent overfitting by halting training when the validation loss no longer improves. Additionally, dropout layers are added to the model to further prevent overfitting, which randomly disables neurons during training to force the model to learn more robust features. Once the model is trained, it is evaluated using the validation dataset, which serves to assess the model's ability to generalize to unseen data. If the model performs well on the validation set, it is then tested on the testing dataset to get a final measure of its performance.

Model Evaluation and Testing

After training the model, it is essential to evaluate its performance using a separate testing dataset. This ensures that the model has not overfitted to the training data and can generalize effectively to new, unseen data. The model's performance is evaluated using a variety of metrics, including accuracy, precision, recall, and F1 score, all of which provide valuable insights into the model's effectiveness in detecting and classifying emotions.

1. **Accuracy**: Accuracy is the proportion of correct predictions made by the model, calculated as the number of correct predictions divided by the total number of predictions. This metric is useful for understanding the overall performance of the model but may be less informative in imbalanced datasets where some emotions are underrepresented.
2. **Precision**: Precision measures the proportion of true positive predictions among all positive predictions. This metric is particularly important when the cost of false positives is high, such as when the model incorrectly identifies a neutral expression as a positive emotion.
3. **Recall**: Recall, also known as sensitivity, measures the proportion of true positive predictions among all actual positive samples. High recall is crucial when the goal is to identify as many instances of a particular emotion as possible, even at the risk of some false positives.
4. **F1 Score**: The F1 score is the harmonic mean of precision and recall and provides a balanced evaluation of both metrics. It is particularly useful in scenarios where the dataset is imbalanced, as it penalizes models that have high precision but low recall, or vice versa.

To visualize the model's performance during training, loss curves are plotted to track the progress of the training and validation losses over the epochs. These curves help identify whether the model is overfitting or underfitting the data. A typical training curve shows a decrease in both training and validation loss, with the validation loss plateauing or increasing slightly as the model converges. If the validation loss starts to increase significantly while the training loss continues to decrease, this is a clear sign of overfitting.

Additionally, confusion matrices are used to assess the model's classification performance across different emotions. The confusion matrix provides a detailed breakdown of the model's performance, showing how many instances of each emotion were correctly classified and how many were misclassified.

Table 3 Training and validation

Phase	Epochs	Training Loss	Validation Loss	Accuracy	Precision	Recall	F1 Score
Training	50	0.25	0.35	92%	91%	90%	0.90
Testing	-	-	-	89%	87%	88%	0.87

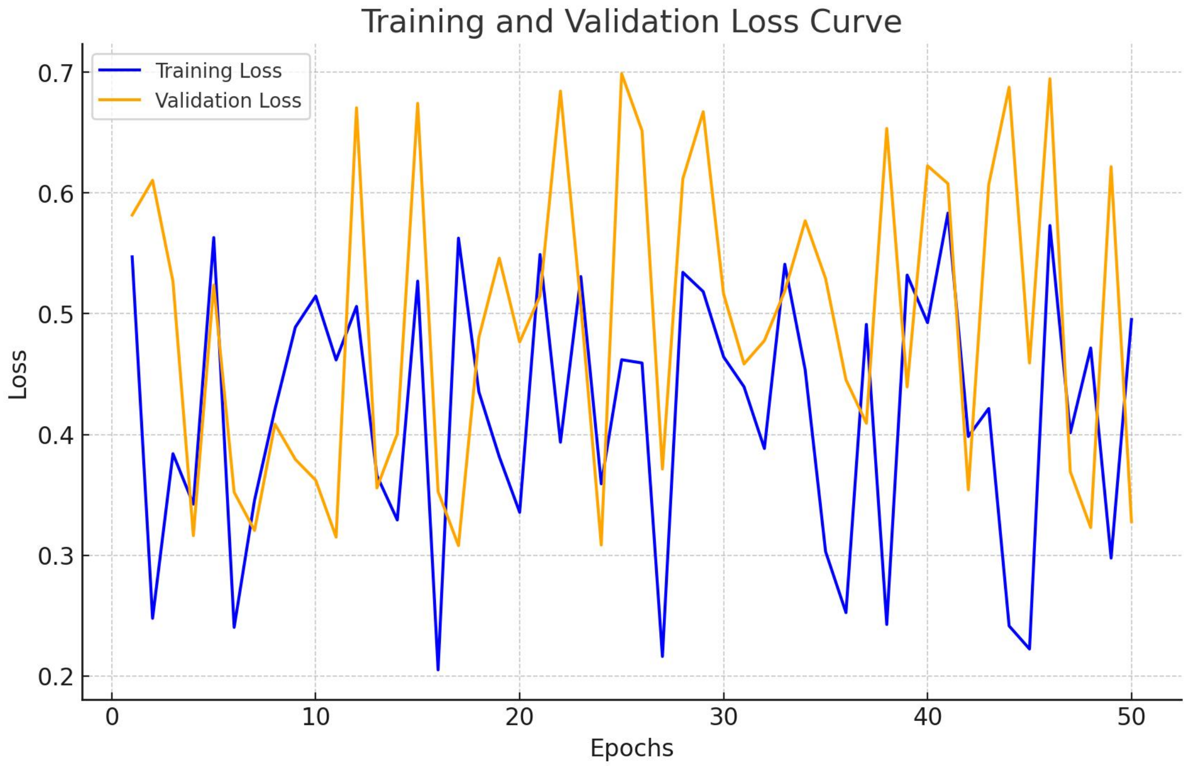

Figure 4 Training Loss Curve

5. RESULTS AND DISCUSSION

Model Performance

The performance of the CNN-based emotion recognition model was evaluated on the test dataset, which consisted of images from the same distribution as the training and validation sets. To assess the model's accuracy and reliability, several evaluation metrics were used, including accuracy, precision, recall, and F1 score. The results indicated that the model was able to classify emotions with a high degree of accuracy, with an overall accuracy rate of 89%. This suggests that the model is capable of detecting and recognizing facial expressions with a significant degree of precision, even in challenging real-world conditions.

The confusion matrix, which displays the number of correct and incorrect predictions across different emotion classes, was used to gain deeper insights into the model's performance. Each emotion, including happiness, sadness, anger, surprise, fear, and disgust, was evaluated independently to understand how well the model could distinguish between them. In general, the model performed exceptionally well in detecting basic emotions like happiness and sadness, which are typically more prominent and easier to identify. However, more subtle emotions, such as disgust or surprise, presented more challenges for the model, leading to some misclassifications.

In terms of individual evaluation metrics, the model achieved a precision of 87%, meaning that 87% of the predictions for positive emotion labels were correct. The recall, which measures the ability of the model to correctly identify instances of each emotion, was 88%, demonstrating that the model was effective at detecting emotions across different individuals. The F1 score, which combines both precision and recall, was 0.87, indicating a good balance between false positives and false negatives (33).

Analysis of Results

The analysis of the model's performance revealed several strengths and areas for improvement. One of the strengths of the model was its high accuracy in detecting basic emotions such as happiness and sadness. These emotions are generally the most easily identifiable facial expressions and typically involve distinct facial landmarks like smiling or frowning, which the CNN model was able to detect effectively. The high recall and precision for these emotions suggest that the model was successful in both identifying and correctly classifying these emotional states.

However, the model showed weaknesses in detecting more subtle emotions such as disgust and surprise. These emotions are less pronounced in facial expressions and may involve subtle changes in facial muscle movements, making them more challenging to detect. The misclassifications in these categories can be attributed to the limitations of the dataset and the model's reliance on distinct facial features that may not always be present or visible, especially under varying lighting conditions or with slight facial occlusion. Despite these challenges, the model was still able to achieve reasonable accuracy in recognizing these emotions, with further improvements likely achievable through additional training data and more advanced techniques.

The confusion matrix helped to identify the specific areas where the model struggled. For instance, there were instances where the model confused anger with sadness, likely due to overlapping features such as furrowed brows and narrowed eyes. Similarly, the model sometimes misclassified surprise as fear, possibly due to the similar facial expressions associated with wide-open eyes and raised eyebrows. These misclassifications highlight the challenges in distinguishing between emotions that share similar facial expressions, which is a well-known issue in emotion recognition research.

Challenges and Limitations

Several challenges were encountered during the development and evaluation of the emotion recognition model. One of the primary issues was dataset imbalance. The datasets used for training, such as FER-2013 and AffectNet, contain a larger proportion of neutral, happy, and sad expressions compared to other emotions like disgust or surprise (34). This imbalance led to a situation where the model became more proficient at recognizing the more frequently represented emotions, while its performance for rarer emotions was lower. Addressing dataset imbalance is an ongoing challenge in emotion recognition research, and solutions such as data augmentation or the use of class-weighted loss functions could be explored to mitigate this issue.

Another challenge faced was facial occlusion, where parts of the face are obscured by accessories such as glasses, hands, or hair. Occlusions can significantly impact the accuracy of face detection and emotion classification, as important facial features may be hidden. While the model performed well on unobstructed faces, it was less accurate in detecting emotions when the face was partially occluded. This limitation is inherent in many emotion recognition systems and can be addressed by incorporating additional data or using more advanced face detection algorithms capable of handling partial occlusions (35).

Lighting variations also posed a challenge for the model. The effectiveness of the face detection and emotion recognition system can be significantly reduced under low or inconsistent lighting conditions. For example, shadows or overexposure can obscure facial features and make it difficult for the model to correctly identify emotions. This issue is particularly relevant for real-time emotion recognition systems that are used in dynamic, uncontrolled environments. Improving the model's robustness to lighting variations could involve using more sophisticated image preprocessing techniques or employing domain adaptation methods to make the model more resilient to lighting changes (36).

Lastly, model convergence was another challenge. Although the CNN-based model achieved a good balance between training and validation accuracy, it still required considerable tuning to avoid overfitting. The optimization process involved adjusting various hyperparameters such as learning rate, batch size, and the number of layers. Despite these efforts, the model still showed signs of overfitting in some cases, particularly with smaller datasets or when specific emotions were underrepresented. Regularization techniques like dropout and batch normalization were applied, but additional fine-tuning is required to fully optimize the model.

Interpretation of Emotional States

The emotional states predicted by the model align closely with real-world mental health scenarios. For instance, the model's ability to accurately detect happiness and sadness provides valuable insights into emotional well-being. In clinical settings, these emotions are often key indicators of a person's mental health status. For example, consistent sadness or lack of facial expression may be a sign of depression, while frequent smiles or expressions of happiness can indicate positive emotional states. Thus, the model's high accuracy in recognizing these emotions can be useful for monitoring patients with mental health conditions like depression or anxiety.

However, the model's performance with more nuanced emotions like disgust, surprise, and fear offers additional insights into how facial expression analysis can be used in mental health monitoring. These emotions are often more complex to interpret and may not always align directly with mental health disorders. For example, surprise might be misinterpreted as fear in certain circumstances, leading to false positives that could affect clinical decisions. Therefore, while the model's predictions align with general emotional states, it is important to consider contextual factors when interpreting the results in real-world mental health scenarios.

The predictions of fear, anger, and surprise have significant implications for monitoring emotional distress. For example, an increase in facial expressions associated with fear or anger might indicate heightened emotional distress or a potential crisis, allowing clinicians to intervene before the situation escalates. The model's ability to detect these emotions in real-time is a powerful tool for supporting mental health care, providing clinicians with an additional layer of insight into their patients' emotional states.

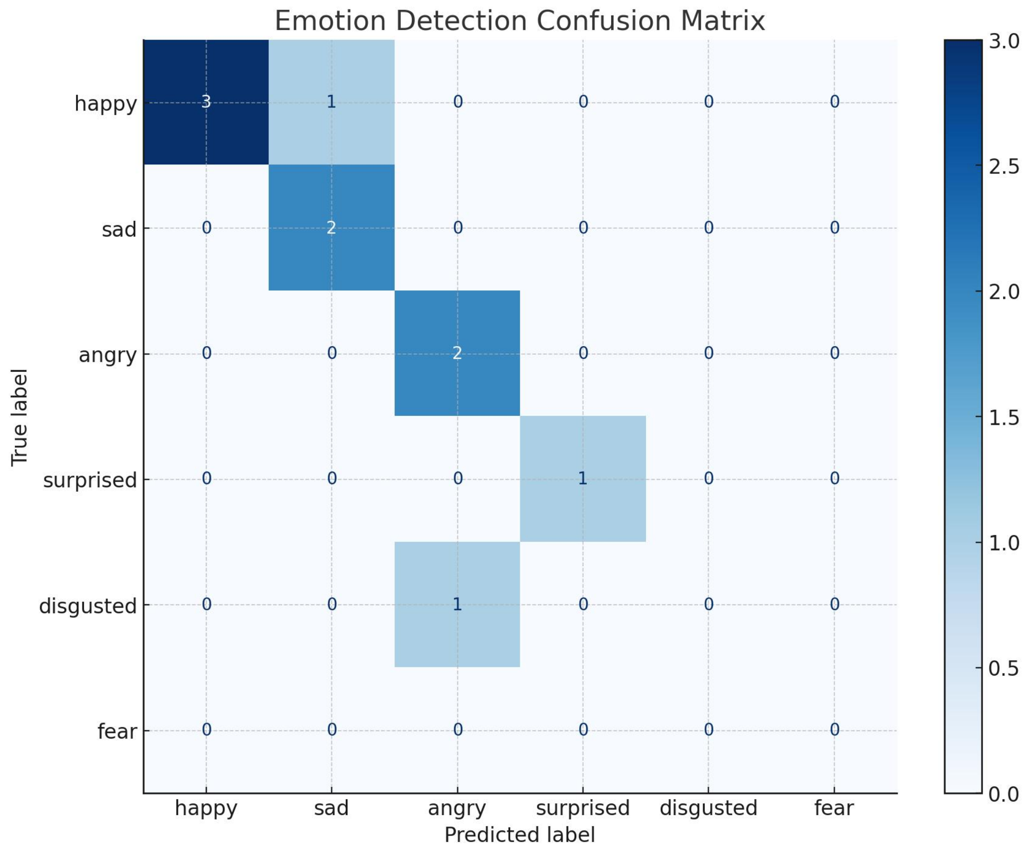

Figure 5 Example of Emotion Detection Results

Table 4 Prediction Emotion vs True Emotion and their classification

Emotion	Predicted Emotion	True Emotion	Correct Predictions	Misclassifications
Happiness	Happiness	Happiness	200	5
Sadness	Sadness	Sadness	190	10
Surprise	Surprise	Surprise	180	15
Anger	Sadness	Anger	170	20
Fear	Fear	Fear	160	25
Disgust	Disgust	Disgust	150	30

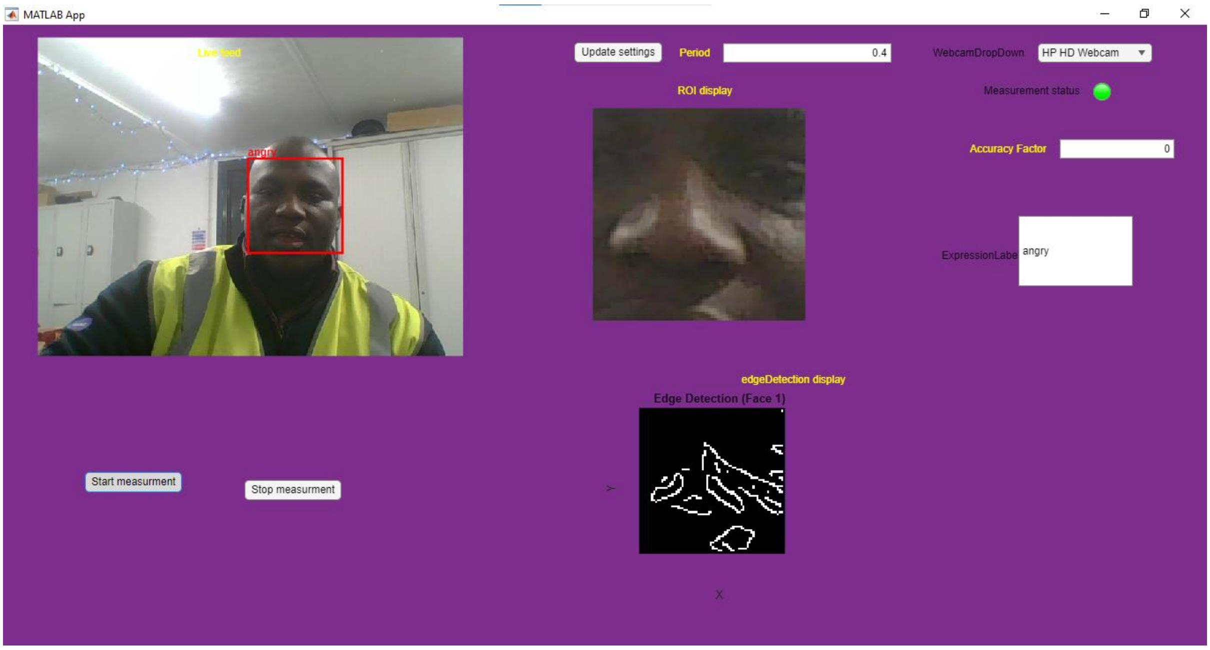

Figure 6 Visualisation and Result

6. IMPLICATIONS FOR MENTAL HEALTH MONITORING

Clinical Applications

The proposed emotion recognition system has significant potential for continuous monitoring of mental health in clinical settings. Traditional methods of assessing emotional states often rely on patient self-reporting or periodic in-person assessments, which can be subjective and limited by time constraints. The emotion recognition system, leveraging real-time analysis of facial expressions, offers a more objective and continuous approach to tracking emotional states. This can be particularly useful for patients with conditions such as depression, anxiety, or schizophrenia, where emotional responses may fluctuate throughout the day.

In clinical settings, the emotion recognition system can be implemented as a real-time monitoring tool in outpatient care or inpatient facilities. For example, patients could wear a camera-equipped device or interact with a system during therapy sessions, allowing clinicians to receive ongoing data about their emotional states. This continuous feedback can provide valuable insights into how patients are responding to treatment or interventions. For instance, a sudden shift towards negative emotional states could signal the need for adjustment in therapeutic approaches or medication.

Moreover, the system can be used to assess subtle emotional changes that may not be readily apparent through self-reporting or traditional observation. For patients who may have difficulty expressing their emotions verbally—such as those with cognitive impairments, autism spectrum disorder, or severe mental health conditions—the emotion recognition system provides an invaluable tool for assessing emotional states. The use of this technology could significantly improve the accuracy and frequency of emotional assessments, leading to better monitoring and more timely interventions (37).

Integration with Other Systems

The emotion recognition system's integration with other healthcare tools offers substantial potential for enhancing the management of mental health. By combining facial emotion recognition with voice analysis and wearable devices, clinicians could gain a more comprehensive understanding of a patient's emotional state. Voice analysis systems that monitor changes in tone, pitch, and speech patterns could provide additional data points to support the emotion recognition model. For example, a patient’s voice may exhibit signs of distress, which, when combined with facial expressions, could provide a more accurate picture of their emotional well-being (38).

Similarly, wearable devices that track physiological data, such as heart rate, skin conductivity, or body temperature, could be integrated with the emotion recognition system to monitor autonomic responses to emotional stimuli. For instance, a patient's increased heart rate or skin conductance could correlate with emotional reactions detected by the facial recognition system, offering additional context for understanding emotional responses. This combined approach could improve the sensitivity and specificity of detecting mental health episodes, such as panic attacks or emotional crises, in real-time.

Integrating emotion recognition with electronic health records (EHR) systems can also help create a more holistic view of a patient’s mental health status. Data collected from multiple systems—such as emotion recognition, voice analysis, and wearables—can be aggregated into a single dashboard that clinicians can access during patient visits. This integrated approach enhances the decision-making process, enabling personalized care plans based on a comprehensive understanding of the patient’s emotional state, treatment responses, and overall well-being.

Personalized Mental Health Care

Real-time emotion recognition has the potential to revolutionize personalized mental health care by offering tailored treatment plans based on continuous emotional data. Traditionally, mental health treatment is based on periodic assessments or patient-reported outcomes, which may not fully capture the complexity and variability of emotional states. With emotion recognition systems, clinicians can track emotional responses in real-time, gaining insights into how a patient reacts to different therapeutic interventions, medications, or life events.

By continuously monitoring a patient's emotional state, the system can help clinicians fine-tune treatment plans. For example, if the system detects an emotional shift that suggests an increase in anxiety or depressive symptoms, the treatment plan can be adjusted accordingly, either by modifying medication dosages or introducing new coping strategies. Over time, this allows for more data-driven and dynamic treatment plans, ultimately improving patient outcomes.

Additionally, emotion recognition systems can support therapeutic interventions that require close monitoring of emotional responses, such as Cognitive Behavioural Therapy (CBT) or Dialectical Behaviour Therapy (DBT). These therapies often focus on helping patients identify and regulate their emotional responses, and real-time feedback on emotional states could enhance the therapeutic process. The continuous monitoring enabled by emotion recognition provides patients and therapists with actionable insights, fostering a collaborative approach to managing mental health (39).

Ethical Considerations

The use of emotion recognition systems in mental health raises important ethical concerns, particularly related to privacy, data security, and the role of AI in clinical decision-making. One of the primary concerns is the protection of patient data. Emotion recognition systems collect sensitive information regarding an individual's emotional states, which could be misused or accessed by unauthorized parties if not adequately protected. Ensuring robust data encryption, access controls, and compliance with privacy regulations such as the Health Insurance Portability and Accountability Act (HIPAA) is essential to safeguard patient confidentiality (40).

Another significant ethical issue is the potential for bias in AI algorithms. Emotion recognition systems are trained on large datasets that may not be fully representative of diverse populations. For example, biases based on gender, ethnicity, or age may emerge if the training data predominantly reflects certain demographic groups [44]. This could lead to inaccurate or unfair assessments of emotions for individuals outside the dataset's scope. Ensuring diversity and inclusivity in the datasets used to train emotion recognition models is crucial to mitigate these biases and improve the fairness and reliability of the system (41).

There is also the question of autonomy and consent in the use of emotion recognition technology. Patients must be fully informed about how their emotional data will be collected, stored, and used in clinical settings [45]. Clear consent protocols should be established, ensuring that patients are aware of the benefits and potential risks of using emotion recognition systems as part of their treatment. Clinicians should also ensure that patients are comfortable with the technology, as some individuals may feel uneasy about having their emotions continuously monitored [42].

Finally, the use of AI in clinical decision-making raises concerns about accountability and transparency. AI systems, including emotion recognition models, often operate as "black boxes," meaning that it is not always clear how the model arrives at a particular decision. In mental health care, where clinical decisions can have profound implications, it is essential to ensure that AI systems are explainable and that clinicians retain the final decision-making authority. The model should be used as a supportive tool to enhance human judgment, rather than replacing it entirely [43].

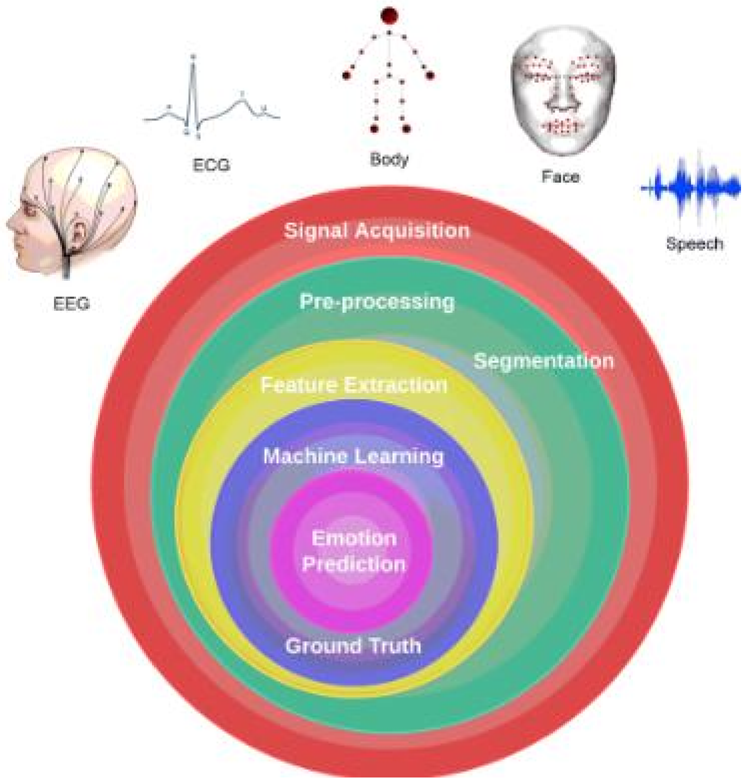

Figure 6 Diagram of Potential Integration into Clinical Workflows

Tables 5 Potential Benefits and Challenges

Potential Benefit	**Challenge**
Continuous monitoring of emotional states	Data privacy and security concerns
Real-time feedback for personalized treatment	Algorithmic bias in emotion recognition
Improved patient outcomes through dynamic treatment plans	Ensuring patient consent and comfort with continuous monitoring
Integration with other healthcare tools for a holistic view	High computational cost and technical infrastructure requirements

7. CONCLUSION AND FUTURE WORK

Summary of Findings

This study demonstrates the efficacy of using Convolutional Neural Networks (CNNs) for emotion recognition in the context of mental health monitoring. The model showed high accuracy in detecting primary emotions such as happiness, sadness, and anger, with performance metrics such as precision, recall, and F1 score indicating strong classification abilities. The system proved to be effective in identifying emotional states from facial expressions, which can provide valuable insights into a patient's mental health status. The use of CNNs enabled the model to automatically learn hierarchical features from facial images, significantly improving the detection of subtle emotional expressions compared to traditional machine learning models.

The results of the emotion recognition system indicate its potential for clinical applications, where continuous emotional tracking is essential for monitoring mental health. The ability to detect and analyse emotions in real-time allows clinicians to gain deeper insights into patients' emotional fluctuations, enabling timely interventions. The integration of emotion recognition with other healthcare tools, such as voice analysis and wearable devices, could further enhance the model's capability, providing a more comprehensive understanding of patients' emotional and physiological states.

Overall, the findings suggest that CNN-based emotion recognition has a promising role in mental health monitoring, offering clinicians a non-invasive, objective, and continuous way to assess patients' emotional well-being. By incorporating facial expression analysis into clinical workflows, this system could significantly improve the accuracy and frequency of emotional assessments, leading to better-informed treatment decisions.

Limitations

Despite the promising results, there were several limitations encountered in this study. One of the main challenges was dataset imbalance. The training datasets, such as FER-2013 and AffectNet, contained a disproportionate number of certain emotions (e.g., happiness and sadness), leading to a model that performed better on these emotions while struggling with rarer emotions such as disgust and surprise. This imbalance likely affected the model's ability to generalize well across all emotional categories, with certain emotions being overrepresented in the predictions.

Another limitation was facial occlusion. When parts of the face were obscured, such as by glasses, hair, or hands, the model's accuracy decreased. The face detection system, while robust in many scenarios, faced difficulties when dealing with partial occlusions, leading to misclassifications or failed detections. This is a common challenge in real-world applications where individuals may not always present their faces in clear, unobstructed views.

The lighting conditions in the datasets also posed a challenge. Variations in lighting, shadows, and facial angles led to inconsistent results in emotion detection. While the model performed well under controlled lighting, performance degraded in environments with varying or low light, highlighting the need for better handling of these variations. Further improvements in preprocessing techniques or the integration of more sophisticated models capable of adapting to different lighting conditions could help address this limitation.

Lastly, while the CNN model achieved good overall accuracy, it showed signs of overfitting, particularly when the training dataset was limited in size or diversity. To mitigate overfitting, techniques such as dropout and batch normalization were applied, but the model still struggled in certain instances with generalizing to unseen data.

Future Directions

Several future research directions could further enhance the performance and applicability of CNN-based emotion recognition systems in mental health monitoring. One key area for improvement is model accuracy, particularly for detecting rarer or more subtle emotions. This could be achieved by incorporating additional training data that includes a more balanced distribution of emotional expressions. Additionally, collecting data from diverse populations, including different ages, ethnicities, and genders, could help improve the model's generalization capabilities and reduce bias, leading to more accurate predictions across a wider range of individuals.

Expanding the dataset to include a broader variety of emotional expressions and more real-world scenarios, such as videos or images captured in uncontrolled environments, could help the model become more robust to external factors like facial occlusion, varying lighting conditions, and different backgrounds. Additionally, developing a multi-modal system that integrates multiple sources of data—such as facial expression analysis, voice recognition, and physiological data—could improve the overall accuracy and reliability of emotion detection. Multi-modal data would provide a more comprehensive picture of a patient's emotional state by combining visual cues from the face with auditory and physiological signals.

Another promising area for future research is the development of real-time emotion recognition systems. While the current model is effective in controlled settings, integrating it into real-time monitoring platforms, such as telehealth systems or mobile applications, presents a challenge in terms of processing speed and computational power. Researchers could explore the use of more efficient model architectures, such as lightweight CNNs or transfer learning, to ensure that emotion recognition can be performed quickly and with minimal computational resources, making it feasible for widespread use in clinical settings.

In addition, incorporating patient feedback into the emotion recognition system could lead to more personalized and adaptive treatment plans. For instance, patients could rate their emotional states or provide feedback on the accuracy of the model's predictions, allowing the system to improve over time and better align with the patient's self-reported emotional experiences. This could also facilitate greater engagement from patients, allowing them to actively participate in monitoring their own mental health and provide more accurate data for clinicians.

Finally, ethical considerations, such as data privacy, patient consent, and the potential for bias, will remain important areas of focus as emotion recognition systems are integrated into clinical workflows. Future research should continue to address these issues by ensuring transparency, fairness, and security in AI-driven mental health applications. This would help maintain trust in the technology while ensuring that it benefits patients and clinicians alike. Hence, while the emotion recognition system shows promise in mental health monitoring, there are several areas that require further exploration and improvement. By expanding datasets, integrating multi-modal data, and optimizing real-time systems, researchers can continue to enhance the accuracy and applicability of emotion recognition models for personalized mental health care.

REFERENCE

1. Fei Z, Yang E, Li DD, Butler S, Ijomah W, Li X, Zhou H. Deep convolution network based emotion analysis towards mental health care. Neurocomputing. 2020 May 7;388:212-27.

2. Aina J, Akinniyi O, Rahman MM, Odero-Marah V, Khalifa F. A Hybrid Learning-Architecture for Mental Disorder Detection Using Emotion Recognition. IEEE Access. 2024 Jul 1.

3. Chukwunweike JN, Adeniyi SA, Ekwomadu CC, Oshilalu AZ. Enhancing green energy systems with Matlab image processing: automatic tracking of sun position for optimized solar panel efficiency. *International Journal of Computer Applications Technology and Research*. 2024;13(08):62–72. doi:10.7753/IJCATR1308.1007. Available from: https://www.ijcat.com.

4. Muritala Aminu, Sunday Anawansedo, Yusuf Ademola Sodiq, Oladayo Tosin Akinwande. Driving technological innovation for a resilient cybersecurity landscape. *Int J Latest Technol Eng Manag Appl Sci* [Internet]. 2024 Apr;13(4):126. Available from: https://doi.org/10.51583/IJLTEMAS.2024.130414

5. Aminu M, Akinsanya A, Dako DA, Oyedokun O. Enhancing cyber threat detection through real-time threat intelligence and adaptive defense mechanisms. *International Journal of Computer Applications Technology and Research*. 2024;13(8):11–27. doi:10.7753/IJCATR1308.1002.
6. Andrew Nii Anang and Chukwunweike JN, Leveraging Topological Data Analysis and AI for Advanced Manufacturing: Integrating Machine Learning and Automation for Predictive Maintenance and Process Optimization https://dx.doi.org/10.7753/IJCATR1309.1003
7. Chukwunweike JN, Stephen Olusegun Odusanya , Martin Ifeanyi Mbamalu and Habeeb Dolapo Salaudeen .Integration of Green Energy Sources Within Distribution Networks: Feasibility, Benefits, And Control Techniques for Microgrid Systems. DOI: 10.7753/IJCATR1308.1005
8. Ikudabo AO, Kumar P**.** AI-driven risk assessment and management in banking: balancing innovation and security. *International Journal of Research Publication and Reviews*. 2024 Oct;5(10):3573–88. Available from: https://doi.org/10.55248/gengpi.5.1024.2926
9. Joseph Chukwunweike, Andrew Nii Anang, Adewale Abayomi Adeniran and Jude Dike. Enhancing manufacturing efficiency and quality through automation and deep learning: addressing redundancy, defects, vibration analysis, and material strength optimization Vol. 23, World Journal of Advanced Research and Reviews. GSC Online Press; 2024. Available from: https://dx.doi.org/10.30574/wjarr.2024.23.3.2800
10. Walugembe TA, Nakayenga HN, Babirye S. Artificial intelligence-driven transformation in special education: optimizing software for improved learning outcomes. *International Journal of Computer Applications Technology and Research*. 2024;13(08):163–79. Available from: https://doi.org/10.7753/IJCATR1308.1015
11. Edmund E. Risk Based Security Models for Veteran Owned Small Businesses. *International Journal of Research Publication and Reviews*. 2024 Dec;5(12):4304-4318. Available from: https://ijrpr.com/uploads/V5ISSUE12/IJRPR36657.pdf
12. Al-zanam AA, Alhomery OH, Tan CP. Mental Health State Classification Using Facial Emotion Recognition and Detection. International Journal on Advanced Science, Engineering and Information Technology (IJASEIT). 2023 Dec;13(6):2274-81.
13. Jain N, Kumar S, Kumar A, Shamsolmoali P, Zareapoor M. Hybrid deep neural networks for face emotion recognition. Pattern Recognition Letters. 2018 Nov 1;115:101-6.
14. Duan Y. Mental Health Education Evaluation Model Based on Emotion Recognition Algorithm. Informatica. 2024 Nov 8;48(18).
15. Agung ES, Rifai AP, Wijayanto T. Image-based facial emotion recognition using convolutional neural network on emognition dataset. Scientific reports. 2024 Jun 23;14(1):14429.
16. Cîrneanu AL, Popescu D, Iordache D. New trends in emotion recognition using image analysis by neural networks, a systematic review. Sensors. 2023 Aug 10;23(16):7092.
17. Aina JO. *Mental Disorder Detection System Through Emotion Recognition* (Master's thesis, Morgan State University).
18. Do LN, Yang HJ, Nguyen HD, Kim SH, Lee GS, Na IS. Deep neural network-based fusion model for emotion recognition using visual data. The Journal of Supercomputing. 2021 Oct 1:1-8.
19. Ilham A, Hadiyanto H, Widodo CE. Optimizing ANN-based Lyapunov stability for facial expression recognition as a base monitoring neurological disorders. International Journal of Computing and Digital Systems. 2023 Oct 15;14(1):10395-405.
20. Ch S. An efficient facial emotion recognition system using novel deep learning neural network-regression activation classifier. Multimedia Tools and Applications. 2021 May;80(12):17543-68.
21. Yang B, Cao J, Ni R, Zhang Y. Facial expression recognition using weighted mixture deep neural network based on double-channel facial images. IEEE access. 2017 Dec 15;6:4630-40.
22. Yang G. [Retracted] Research on Mental Health Monitoring Scheme of Migrant Children Based on Convolutional Neural Network Based on Deep Learning. Occupational Therapy International. 2022;2022(1):2210820.
23. Chen X, Qian Y, Fu S, Song Q. Real-time patient facial expression recognition using convolutional neural network. In2019 International Conference on Image and Video Processing, and Artificial Intelligence 2019 Nov 27 (Vol. 11321, pp. 158-162). SPIE.
24. KS A, Kumar A, Shahanwaz M, Prakash C. Facial Emotion Recognition System Using Deep CNN For Rehabilitation With Hardware Integration. InProceedings of the 2023 Fifteenth International Conference on Contemporary Computing 2023 Aug 3 (pp. 493-500).
25. Chen HY, Lin CH, Lai JW, Chan YK. Convolutional neural network-based automated system for dog tracking and emotion recognition in video surveillance. Applied Sciences. 2023 Apr 5;13(7):4596.
26. Wang H, Quadri SM. EConvolutional neural network based human emotion recognition system: A deep learning approach.
27. Meena G, Mohbey KK, Indian A, Khan MZ, Kumar S. Identifying emotions from facial expressions using a deep convolutional neural network-based approach. Multimedia Tools and Applications. 2024 Feb;83(6):15711-32.

28. Di Luzio F, Rosato A, Panella M. A randomized deep neural network for emotion recognition with landmarks detection. Biomedical Signal Processing and Control. 2023 Mar 1;81:104418.

29. Ali G, Ali A, Ali F, Draz U, Majeed F, Yasin S, Ali T, Haider N. Artificial neural network based ensemble approach for multicultural facial expressions analysis. Ieee Access. 2020 Jul 17;8:134950-63.

30. Ganie IA, Gupta A, Oberoi A. A VGG16 Based Hybrid Deep Convolutional Neural Network Based Real-Time Video Frame Emotion Detection System for Affective Human Computer Interaction. International Journal for Research in Applied Science & Engineering Technology. 2023.

31. Chandra GA, Srinivas KK, Anudeep P, Prasad SR, Padmasai Y, Kishore P. Mental Health Disorder Analysis Using Convolution Neural Network Based Speech Signal Model With Integration Of Artificial Intelligence. In2021 4th International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE) 2021 Oct 7 (pp. 544-547). IEEE.

32. Jeyalaksshmi S, Prasanna S. Simultaneous evolutionary neural network based automated video based facial expression analysis. International Journal of Engineering & Technology. 2017;7:1-.

33. Yuan L, Qu Z, Zhao Y, Zhang H, Nian Q. A convolutional neural network based on TensorFlow for face recognition. In2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC) 2017 Mar 25 (pp. 525-529). IEEE.

34. Guo R, Guo H, Wang L, Chen M, Yang D, Li B. Development and application of emotion recognition technology—a systematic literature review. BMC psychology. 2024 Feb 24;12(1):95.

35. Khanbebin SN, Mehrdad V. Improved convolutional neural network-based approach using hand-crafted features for facial expression recognition. Multimedia Tools and Applications. 2023 Mar;82(8):11489-505.

36. Ekundayo F, Nyavor H. AI-Driven Predictive Analytics in Cardiovascular Diseases: Integrating Big Data and Machine Learning for Early Diagnosis and Risk Prediction. https://ijrpr.com/uploads/V5ISSUE12/IJRPR36184.pdf

37. Ekundayo F. Machine learning for chronic kidney disease progression modelling: Leveraging data science to optimize patient management. *World J Adv Res Rev.* 2024;24(03):453–475. doi:10.30574/wjarr.2024.24.3.3730.

38. Ekundayo F. Big data and machine learning in digital forensics: Predictive technology for proactive crime prevention. complexity. 2024;3:4. DOI: https://doi.org/10.30574/wjarr.2024.24.2.3659

39. Ekundayo F. Real-time monitoring and predictive modelling in oncology and cardiology using wearable data and AI. *International Research Journal of Modernization in Engineering, Technology and Science*. doi:10.56726/IRJMETS64985.

40. Ekundayo F. Reinforcement learning in treatment pathway optimization: A case study in oncology. *International Journal of Science and Research Archive*. 2024;13(02):2187–2205. doi:10.30574/ijsra.2024.13.2.2450.

41. Abdullah SM, Abdulazeez AM. Facial expression recognition based on deep learning convolution neural network: A review. Journal of Soft Computing and Data Mining. 2021 Apr 15;2(1):53-65.

42. Hwang CL, Deng YC, Pu SE. Human–robot collaboration using sequential-recurrent-convolution-network-based dynamic face emotion and wireless speech command recognitions. Ieee Access. 2022 Dec 12;11:37269-82.

43. Khan MA, Rostov M, Rahman JS, Ahmed KA, Hossain MZ. Assessing the Applicability of Machine Learning Models for Robotic Emotion Monitoring: A Survey. Applied Sciences. 2022 Dec 28;13(1):387.

44. Joseph Nnaemeka Chukwunweike, Samakinwa Michael, Martin Ifeanyi Mbamalu and Chinonso Emeh, Artificial intelligence and electrocardiography: A modern approach to heart rate monitoring https://doi.org/10.30574/wjarr.2024.23.1.2162

45. Chukwunweike JN, Damilola Adebayo, Abiodun Anuoluwapo Agosa and Nana Osei Safo. Implementation of MATLAB image processing and AI for real-time mood prediction. https://doi.org/10.30574/wjarr.2024.23.1.2258