iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** III **Month of publication:** March 2024

DOI: https://doi.org/10.22214/ijraset.2024.58777

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue III Mar 2024- Available at www.ijraset.com

A Comprehensive Survey on Predicting Human Psychological State through Facial Feature Analysis

Ayush Bhalerao[1], Suraj Chamate[2], Shivprasad Bodke[3], Shantanu Desai[4], Prof. Somesh Kalaskar[5]
Marathwada Mitra Mandal's College Of Engineering, Pune, India

Abstract: ***This study uses convolutional neural networks (CNNs) and image edge computing techniques to analyze facial features and forecast human psychological states in a novel way. Real-time or almost real-time face expression identification is the goal of the suggested approach, which will advance the developing field of affective computing and have potential uses in mental health evaluation and human-computer interaction. In this study, a variety of datasets with a broad range of facial expressions are used to train CNN models, guaranteeing consistent performance across a range of users and cultural backgrounds. Furthermore, the effective detection of human faces in photos is achieved by the use of the Haar Cascade Classifier, which improves the overall accuracy and dependability of the emotion recognition system. The algorithms' efficiency is further increased by the addition of picture edge computing techniques, which makes them appropriate for deployment in contexts with limited resources. The suggested method's accuracy in identifying and categorizing facial emotions is demonstrated by the experimental findings, indicating its potential practical applications. This research has implications for building mental health monitoring systems and enhancing user experience through technology, which goes beyond affective computing. This research fills important gaps in mental health screening and assistance while also enhancing the capabilities of facial expression recognition systems and making human-computer interaction interfaces more responsive and intuitive.***
Keywords: Emotion detection, Convolutional Neural Network, edge computing, facial feature extraction, expression classification, deep learning.

I. INTRODUCTION

The amount of research on identifying emotions has skyrocketed in the last few years. Facial expres- sion recognition has gained popularity as a hot research issue in recent years for recognizing a wide range of fundamental emotions. In many applications, emotion-based facial expression detection is an essential component. Because there are countless applications for this emerging topic, the corporate sector is finding great profit from the development of emotion recognition systems. It has been demon- strated that the best method for the brain to identify human emotions in a range of contexts is through facial expressions. Perceiving and comprehending human emotions is critical for a variety of tasks, including computer interaction and mental health assessments. There has been a great deal of interest in figuring out a person's feelings from their facial features. Our goal is to develop intelligent algorithms that can swiftly and accurately recognize the emotions on people's faces in real-time. We employ some ingenious imaging techniques along with a unique kind of computer system known as a convolutional neural network (CNN) to do this. The primary objective is to develop a robust system that can identify and categorize facial emotions rapidly. This has the potential to be very helpful in a variety of contexts, such as the development of emotionally responsive technology or the provision of real-time emotional intelligence support.

This study is significant because of its prospective uses, which include mental health monitoring systems that can help medical practitioners and human-computer interfaces that can adapt to the emotional moods of their users. As technology As face expression detection technology develops, it may be possible to incorporate it into commonplace electronics to enhance user satisfaction and further the field of affective computing. In computer vision, convolutional neural networks (CNNs) provide revolutionary breakthroughs, particularly in face expression detection applications. CNNs have shown to be incredibly successful at identifying intricate patterns and characteristics in images; they were initially inspired by the visual processing of the human brain.

Our study involves utilizing CNN for real-time facial emotion identification. Applications in robotics, affective comput- ing, and human-computer interaction will find this to be a crucial task. The investigation of real-time facial expression detection relies heavily on CNN's flexibility and learning capacity, which offers both high levels of accuracy and efficient computing to capture the a wide range of human emotions.

©IJRASET: All Rights are Reserved

II. LITERATURE SURVEY

A. *Lightweight Convolutional Neural Network for Real-Time Facial Expression Detection*

There has been a movement in real-time facial expression identification towards lightweight convolutional neural networks (CNNs) in recent times. The suggested lightweight CNN intended for effective real-time facial emotion categorization is the main topic of this literature study. In order to recognize faces, this system incorporates a cascaded multitask convolutional network (MTCNN), which then outputs coordinates for the subsequent categorization of emotions. The model lowers parameters and increases portability by replacing fully connected layers with global average pooling, including residual modules, and using depth-separable convolutions. Impressive results are obtained with experimental validation on the FER-2013 dataset, attaining state-of-the-art accuracy (67%) with low memory consumption (3.1% of 16GB). The suggested lightweight CNN's successful design and future possibilities are highlighted in the conclusion, which also offers suggestions for potential enhancements to better tackle practical issues. The approach In this paper, we suggest fusing the strength of MTCNN with an interesting model we created for face emotion identification. Rather than using OpenCV's standard facial recognition, I utilize his MTCNN. due to MTCNN's superior facial recog- nition abilities and more precise emotion recognition. This adjustment is particularly helpful in images with a large number of faces. We drew inspiration for our face emotion detection model from Xception, an excellent tool for object recognition and learning. Our model should function well in a range of situations. In order to better comprehend someone's specific emotions, we are experimenting with a few different settings. Global Average Pooling streamlines the model's structure, removes completely connected layers, and enhances interpretability. This paper's algorithms go beyond the drawbacks of conventional techniques like OpenCV and present effective tests carried out on a reliable hardware configuration yielding outstanding face recognition outcomes. The literature study emphasizes how the suggested algorithm advances the field of real-time facial expression recognition while maintain- ing its uniqueness and efficacy.

B. *Context-Aware Emotion Recognition Based on Visual Relationship Detection*

This paper is all about understanding emotions in a smarter way, especially by looking at the sit-uation around a person. Rather than relying on facial expressions, language, and gestures as usual, explore what role the environment plays. In simple terms, we're suggesting a new idea where we payattention to how the main person in a scene is related to the things around them. This helps us under-stand emotions more deeply by considering both the space and the meaning of what is happening. Inour study, we added an extra attention mechanism to see how different things in a person's environ- ment affect emotions. When tested on his two datasets, CAER-S and EMOTIC, our method performed very well, even outperforming other methods on CAER-S. We admit that there are situations where it doesn't work perfectly, but overall, our approach is successful in understanding how the context of a scene influences people's emotions. The researchers plan to do more in the future, including testing the method on video and looking for different ways to express emotions. They are very interested in exploring how following people's gaze can help us understand emotions more deeply. That is, they talk about what they want to do next, summarize what they have added to the field, and emphasize theimportance of considering and understanding the context of the scene and the connections between the protagonist and everything around them. Emotions. The method of understanding emotions is very useful when examining how things visually relate to each other. It's like learning from what's around us. It also uses special attention mechanisms to further improve the model by focusing on important details in the context. Although our method works well, we recognize that there are challenges, suchas when emotions are misjudged or when the environment does not match a person's emotions. In the future, we will test this method on videos, experiment with different ways people express emotions, and explore how tracking eye movements can help understand emotions in context. It's a schedule. This research represents a major step forward in our deeper understanding of emotions. They have created a powerful system that pays attention to the context of a scene and its connections to the peo-ple and things around it. This is a useful guide for other researchers who want to make their emotion recognition systems more accurate and detailed. The way they did this could be very useful in the future of understanding emotions in different situations.

C. *Real-Time Implementation of Face Recognition and Emotion Recognition in a Humanoid Robot Using a Convolutional Neural Network*

This research aims to integrate two things that are important for humanoid robots: recognizing faces in real time and understanding emotions. Although these are usually considered different chal-lenges, this study proposes to tackle both challenges simultaneously using an intelligent computing system called a convolutional neural network (CNN). It's like teaching a robot to quickly recognize faces and understand people's feelings in the same way. This is a step that helps robots become moreaware of the people around them. We have seen how well-known computer models such as AlexNetand VGG16 can be used with humanoid robots.

We used data from his 30 electrical engineering stu-dents to teach the model to recognize faces and emotions. The results showed that VGG16 performed better, achieving 100They achieve great success in making intelligent computing systems (VGG16 and modified versions of AlexNet) work well on humanoid robots to recognize faces and emotions. Use specific numbers and percentages to indicate how accurate they are. This study also highlights that our method works quickly and in real time on humanoid robots, bringing more practical value toour research by introducing the possibility of distance measurement.

D. Happy Emotion Recognition from Unconstrained Videos Using 3D Hybrid Deep Features

In our study, we introduce a model called HappyER- DDF. This model aims to better detect whenpeople are feeling happy, especially in videos where things are more natural and less staged. We focuson facial expressions because facial expressions play a major role in understanding emotions, espe- cially happiness, and many researchers have recently become interested in this area. Current methods cannot handle large head movements and are not very accurate, especially when detecting happy emo- tions. Our solution, the HappyER-DDF model, addresses this problem. It leverages a smart combina- tion of 3D depth and distance features and uses a combination of 3D Inception-ResNet and a neural network called LSTM. This allows the model to understand not only the appearance of a face, but how it changes over time, and capture details associated with smiles and laughter. Our new model reflects the latest advances in facial expression recognition. This feature is unique in that it uses both3D depth and distance functions and combines them in an intelligent way. We use a combination of his two powerful neural networks: 3D Inception-ResNet and LSTM. These are kind of the latest andgreatest tools for understanding how facial expressions change over time in videos that are not staged or controlled.

III. IMPLEMENTATION OF EMOTION DETECTION USING CNN AND HAARCASCADE

A. Covolutional Neural Networks (CNN)

Many computer vision tasks, such as object identifica-tion, facial recognition, and image classification, have made extensive use of Convolutional Neural Networks (CNNs). CNNs have demonstrated promising results in emotion classification from facialphotos because of their capacity to autonomously learn discriminative features straight from raw pixeldata.

1) *Convolutional Layers (Conv2D):* CNNs' fundamental building components are convolutionallayers. In order to extract features, they apply a series of learnable filters to the input image and effectively convolve them across the whole image. To provide non-linearity, the design employs a number of convolutional layers, each followed by an activation function known as a Rectified LinearUnit (ReLU). With a kernel size of (3, 3), the first convolutional layer contains 32 filters, each of which is a 3x3 grid. In order to capture more intricate details, subsequent layers add more filters.

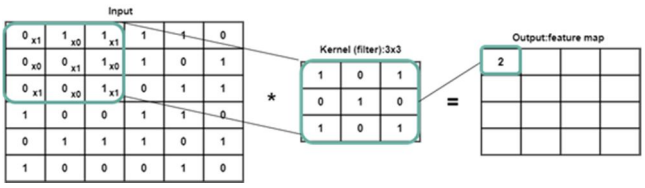


Fig. 1: Convolutional Layer

2) *MaxPooling Layers (MaxPooling2D):* By reducing the spatial dimensions of the feature maps, MaxPooling layers preserve the most crucial data. They do this by choosing the highest value that falls inside a given window. In order to downsample the feature maps and save computation, MaxPooling layers with a pool size of (2, 2) are implemented in the architecture after a few convolutional layers.

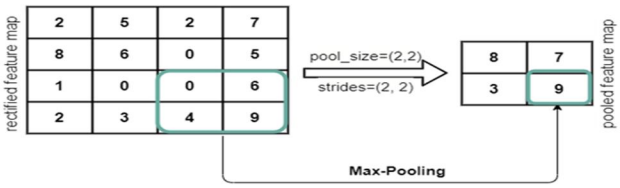


Fig. 2: Max Pooling Layer

3) *Dropout Layers:* By randomly removing a portion of the input units during training, dropout layers are utilized for regularization and to avoid overfitting. To regularize the model, dropout layerswith a dropout rate of 0.25 or 0.5 are incorporated into the architecture after a set of convolutional andfully connected layers.

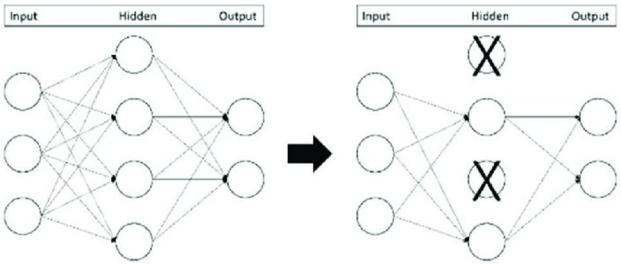


Fig. 3: Dropout Layer

4) *Flatten Layer:* The Flatten layer is used to convert the 2D feature maps into a 1D vector, whichcan be fed into a fully connected layer for classification. In the architecture, the Flatten layer is placedafter the last convolutional layer to prepare the data for the fully connected layers.

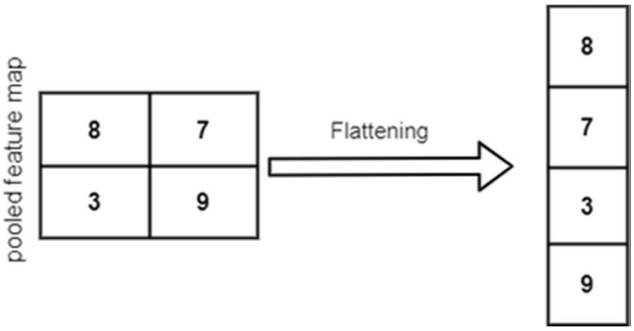


Fig. 4: Flattening Layer

5) *Fully Connected Layers (Dense):* Fully connected layers are conventional neural network lay-ers in which every neuron in a layer's predecessor and subsequent levels is connected to every otherneuron. Classification is carried out in the architecture using Dense layers with ReLU activation func-tions. Seven neurons with a softmax activation function that produce probability for each of the sevenemotion classes make up the final Dense layer.

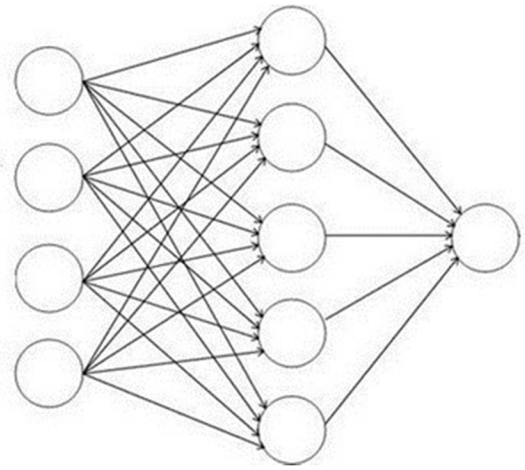


Fig. 5: Fully Connected Layer

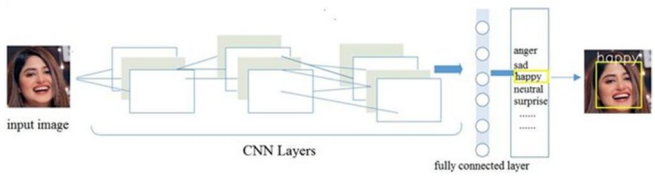


Fig. 6: Proposed Model Architecture

B. Haarcascade Classifier

The following procedures are involved in employing the Haar Cas-cade Classifier for face detection:

1) *Training the Classifier:* A sizable dataset of both positive and negative images is used to train the Haar Cascade Classifier. Negative photos feature backdrops or irrelevant items, while positive photographs feature the object of interest—in this example, human faces. utilizing a sequence of classifiers, usually based on the Ad- aboost algorithm, and feature extraction utilizing Haar-like features, the classifier learns to distinguish between the patterns of positive and negative images during training.
2) *Generating Cascade Classifiers:* After training, the Haar Cascade Classifier generates a cascade of classifiers, each of which is in charge of figuring out whether or not a certain area of the image has a face. Because the cascade architecture reduces the computational load by rapidly rejecting areas of the image that are unlikely to contain faces, it enables efficient calculation.
3) *Using the Classifier to Identify Faces:* The Haar Cascade Classifier is used to identify faces in an input image by swiping a window of various sizes at various scales. The classifier uses the learnt features to determine if the region inside the window resembles a human face at each position and scale. In the event that the classifier finds a face inside a region, it gives back the bounding box coordinates that show where the face was found.
4) *Post-processing and Validation:* After face detection, overlapping bounding boxes can be merged and the final collection of de-tected faces can be improved by using post-processing techniques such non-maximum suppression. Furthermore, validation procedures, like looking for facial landmarks or applying confidence thresholds, can be carried out to guarantee the precision and dependability of the recognized faces.

IV. METHODOLOGY

1) Step 1: Data Collection

The first and most important step in our study project is to obtain a large dataset of facial expres-sions that have been labeled with different emotions, such as anger, sadness, and happiness. In orderto achieve this, we make use of the FER 2013 dataset, a well-known benchmark dataset in the facialexpression recognition space. A wide variety of facial expressions, gathered from multiple platformsincluding Google Images, are included in the FER 2013 dataset. Images in the FER 2013 dataset aretagged with one of seven emotion categories: neutral, surprise, happiness, sorrow, disgust, and rage. Our model is trained on a wide range of facial expressions thanks to the dataset's broad spectrum ofemotions, which improves its capacity to identify emotions properly in real-world situations.

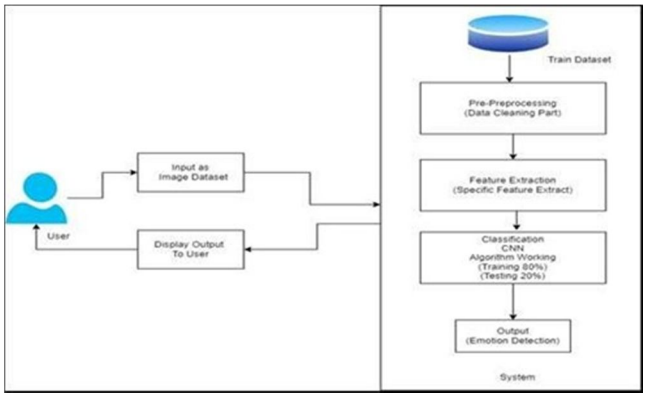


Fig. 7: System Architecture

2) Step 2: Data preprocessing

Preprocessing the photos after dataset collection is crucial for ensuring uniformity and improving model generalization. In this preprocessing stage, all photos are usually resized to a uniform resolution, pixel values are normalized to a common scale (e.g., [0, 1]), and data augmentation techniques (e.g., rotation, flipping, and scaling) are applied to enhance the dataset. Through this kind of preprocessing, we hope to reduce over fitting, increase model robustness, and make it easier for the model toacquire discriminative features.

3) Step 3: Encoding Labels

The emotion labels attached to each image must be encoded into numerical representations prior to model training. This procedure, called label encoding, gives every emotion category a distinct numerical identity. For example, we may assign the numbers 0 to 1, 2 to angry, sad, and joyful, respectively. We allow the machine to interpret and learn from the categorical data during training by encoding the emotion labels in this way.

4) Step 4: Split the Dataset

We divided the dataset into three separate subsets: training, validation, and test sets in order to precisely evaluate our model's performance. The validation set is used to check model performance and adjust hyper parameters while the training set is used to train the model's parameters. Lastly, the test set functions as a separate standard by which to measure the trained model's capacity to generalize to new data. To guarantee a balanced distribution of emotion classes across the training, validation, and test sets, appropriate stratification techniques are used.

5) Step 5: CNN Architecture

The Convolutional Neural Network (CNN) architecture that we have specifically designed for assessing facial features and forecasting human emotions forms the basis of our emotion classificationsystem. Typically, the CNN architecture consists of several layers: fully connected layers for classification, pooling layers for spatial down sampling, and convolutional layers for feature extraction. To effectively train and discriminate emotion-related data from facial photos, the CNN architecture carefully considers a number of aspects, including activation functions, kernel sizes, network depth, and regularization approaches.

6) Step 6: Model Training

After defining the CNN architecture, we use the labeled training data to train the model. Using an optimizer technique like Adam or Stochastic Gradient Descent (SGD), the model learns to minimize a predetermined loss function such category cross-entropy during training. This allows the model to map input photos to their corresponding emotion labels. Iteratively feeding batches of images through the network, calculating the loss, and backpropagating the updated model parameters are all part of the training process. Optimizing the network weights and biases to reliably predict emotions from facial expressions is the aim of model training.

7) Step 7: Validation

To avoid overfitting and adjust hyperparameters, it is crucial to keep an eye on the model's performance on a different validation set during the training process. We can decide on changes to learning rates, batch sizes, and architectural alterations to enhance generalization and convergence qualities byregularly assessing the model's performance on the validation set.

8) Step 8: Testing

After the model has been trained, we test the trained model's performance on an independent testset to see if it can generalize to new data. The efficacy of the model in reliably classifying emotions from facial expressions is determined by evaluating parameters including accuracy, precision, recall, and F1-score. The testing stage offers a vital reference point for assessing the trained model's depend-ability and practical applicability.

9) Step 9: Model Deployment

The trained model can be used for real-world applications like sentiment analysis on social net- working platforms, emotion identification in HCI systems, or tailored content recommendations after receiving appropriate performance validation. During the deployment phase, the trained model is integrated into production environments, memory footprint and inference speed are optimized, and robustness and dependability across a range of scenarios are ensured.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

10) Step 10: Monitoring and Updating

We stress the significance of ongoing monitoring and updating of the deployed model in real-world circumstances in the last step of our process. We can spot possible flaws or drifts in the model's behavior by keeping an eye on environmental changes, user feedback, and model performance metrics. Then, we can take proactive steps to maintain optimal performance and relevance over time by updating the model's parameters, retraining it on new data, or putting corrective measures in place. The lifespan and efficacy of our emotion classification system in identifying and deciphering human psychological states from facial expressions are guaranteed by this iterative process of observation and modification.

V. EXPECTED OUTCOME

Our research aims to predict people's psychological state by analyzing facial features using an intel-ligent computing system called a convolutional neural network (CNN). We expect our models to be able to look at a person's face and provide insight into their emotional well-being and psychologicalstate. By training a CNN with a diverse dataset of facial expressions, we hope to create a reliable system that can understand and predict a wide range of human emotions. This research could have important applications in areas such as mental health assessment, human-computer interaction, and other fields where understanding people's emotions is important. Our expected outcome is a powerful model that can accurately predict a person's psychological state based on facial features, contributingto the advancement of emotional intelligence technology.

VI. CONCLUSION

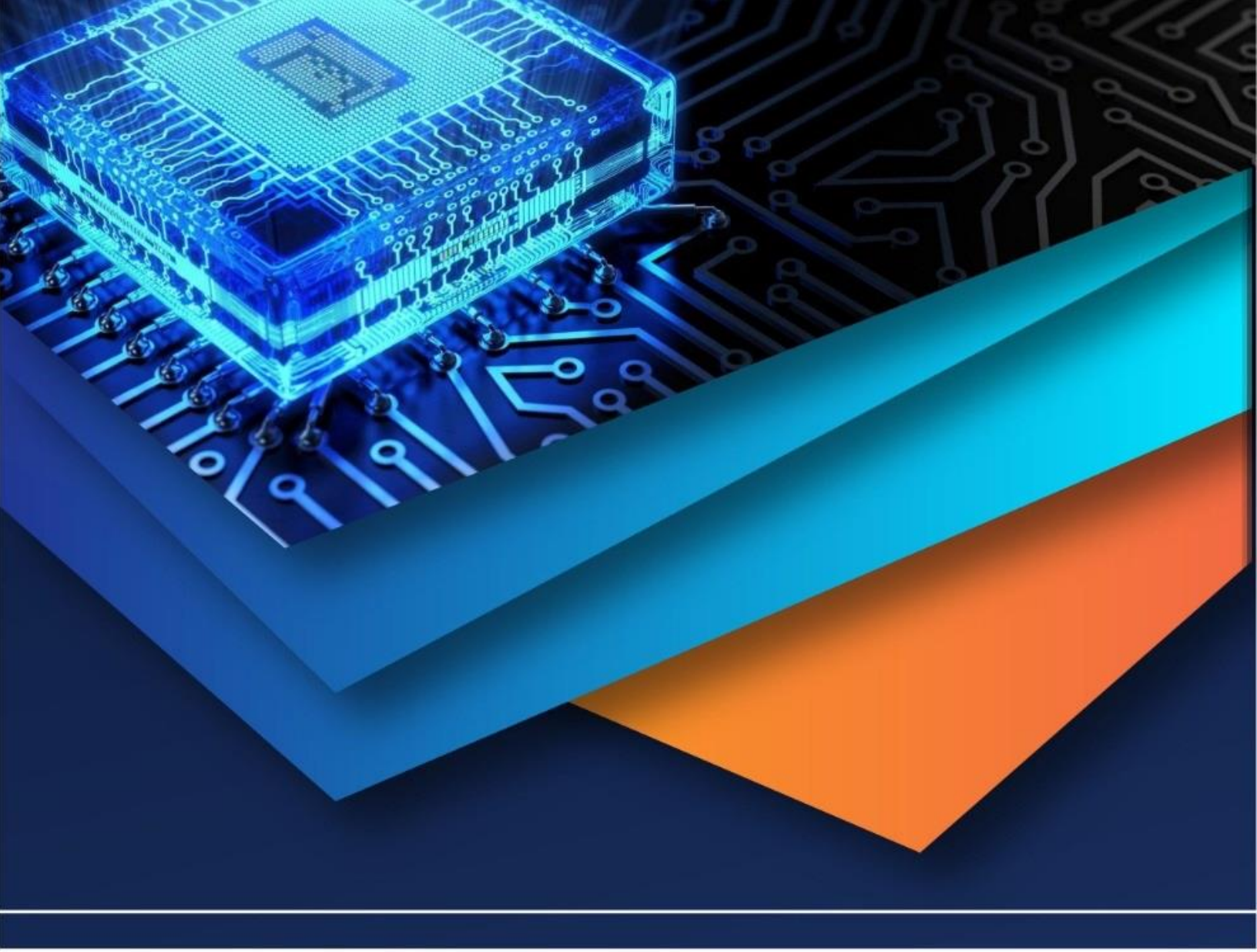
We've developed a smart way to look at people's faces and understand their emotions using convolutional neural networks (CNNs). We used a carefully selected dataset that included different types of people to ensure the system worked well for everyone. Before training the computer, we preparedthe data by resizing, adjusting, and adding diversity. This helps the system process all the details of people's facial expressions. Our special CNN architecture is like a recipe designed just for facial expression recognition. It has layers that find important features and others that make the final decisionabout how someone is feeling. We also added things like activation functions and dropout to make sure the system doesn't get too confident and stays accurate. Overall, our method is strong, adaptable,and good at figuring out emotions from facial expressions.

REFERENCES

[1] N. Zhou, R. Liang and W. Shi, "A Lightweight Convolutional Neural Network for Real-Time Facial Expression Detection," in IEEE Access, vol. 9, pp. 5573-5584, 2021, doi: 10.1109/AC- CESS.2020.3046715.
[2] M.-H. Hoang, S.-H. Kim, H.-J. Yang and G.-S. Lee, "Context-Aware Emotion Recognition Based on Visual Relationship Detection," in IEEE Access, vol. 9, pp. 90465-90474, 2021, doi: 10.1109/ACCESS.2021.3091169.
[3] S. Dwijayanti, M. Iqbal and B. Y. Suprapto, "Real-Time Implementation of Face Recognition and Emotion Recognition in a Humanoid Robot Using a Convolutional Neural Network," in IEEE Access, vol. 10, pp. 89876-89886, 2022, doi: 10.1109/ACCESS.2022.3200762.
[4] N. Samadiani, G. Huang, Y. Hu and X. Li, "Happy Emotion Recognition from Unconstrained Videos Using 3D Hybrid Deep Features," in IEEE Access, vol. 9, pp. 35524-35538, 2021, doi: 10.1109/ACCESS.2021.3061744.
[5] H. Zhang, A. Jolfaei and M. Alazab, "A Face Emotion Recognition Method Using Convolutional Neural Network and Image Edge Computing," in IEEE Access, vol. 7, pp. 159081-159089, 2019, doi: 10.1109/ACCESS.2019.2949741.
[6] S. K. Jarraya, M. Masmoudi and M. Hammami, "Compound Emotion Recognition of Autistic Children During Meltdown Crisis Based on Deep Spatio-Temporal Analysis of Facial Geometric Features," in IEEE Access, vol. 8, pp. 69311-69326, 2020, doi: 10.1109/ACCESS.2020.2986654.
[7] A. V. Savchenko, L. V. Savchenko and I. Makarov, "Classifying Emotions and Engagement in Online Learning Based on a Single Facial Expression Recognition Neural Network," in IEEE
[8] Transactions on Affective Computing, vol. 13, no. 4, pp. 2132-2143, 1 Oct.-Dec. 2022, doi: 10.1109/TAFFC.2022.3188390.
[9] F. Alrowais et al., "Modified Earthworm Optimization with Deep Learning Assisted Emotion Recognition for Human Computer Interface," in IEEE Access, vol. 11, pp. 35089-35096, 2023, doi: 10.1109/ACCESS.2023.3264260.
[10] J. Heredia et al., "Adaptive Multimodal Emotion Detection Architecture for Social Robots," in
[11] IEEE Access, vol. 10, pp. 20727-20744, 2022, doi: 10.1109/ACCESS.2022.3149214.
[12] T. Dar, A. Javed, S. Bourouis, H. S. Hussein and H. Alshazly, "Efficient-SwishNet Based Sys- tem for Facial Emotion Recognition," in IEEE Access, vol. 10, pp. 71311-71328, 2022, doi: 10.1109/ACCESS.2022.3188730.
[13] A. V. Savchenko, L. V. Savchenko and I. Makarov, "Classifying Emotions and Engagement in Online Learning Based on a Single Facial Expression Recognition Neural Network," in IEEE Transactions on Affective Computing, vol. 13, no. 4, pp. 2132-2143, 1 Oct.-Dec. 2022, doi: 10.1109/TAFFC.2022.3188390.
[14] S. K. Vishnumolakala, V. S. Vallamkonda, S. C. C, N. P. Subheesh and J. Ali, "In-class Stu- dent Emotion and Engagement Detection System (iSEEDS): An AI-based Approach for Respon- sive Teaching," in 2023 IEEE Global Engineering Education Conference (EDUCON), Kuwait, Kuwait, 2023, pp. 1-doi: 10.1109/EDUCON54358.2023.10125254.

[15] S. Haddad, O. Daassi and S. Belghith, ”Emotion Recognition from Audio-Visual Informa- tion based on Convolutional Neural Network,” in 2023 International Conference on Con- trol, Automation and Diagnosis (ICCAD), Rome, Italy, 2023, pp. 1-5, doi: 10.1109/IC- CAD57653.2023.10152451.
[16] A. S. Kumar and R. Rekha, ”Improving Smart Home Safety with Face Recognition using Machine Learning,” in 2023 International Conference on Intelligent Systems for Communica- tion, IoT and Security (ICISCoIS), Coimbatore, India, 2023, pp. 478-482, doi: 10.1109/ICIS- CoIS56541.2023.10100592.
[17] A. Sharma, V. Bajaj and J. Arora, ”Machine Learning Techniques for Real-Time Emotion Detec- tion from Facial Expressions,” in 2023 2nd Edition of IEEE Delhi Section Flagship Conference (DELCON), Rajpura, India, 2023, pp.1-6, doi: 10.1109/DELCON57910.2023.10127369.
[18] Y. S. Deshmukh, N. S. Patankar, R. Chintamani and N. Shelke, ”Analysis of Emotion Detection of Images using Sentiment Analysis and Machine Learning Algorithm,” in 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2023, pp. 1071-1076, doi: 10.1109/ICSSIT55814

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)