iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: VI Month of publication: June 2025

DOI: https://doi.org/10.22214/ijraset.2025.71935

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue VI June 2025- Available at www.ijraset.com

Face Recognition Using Deep Learning

Vivek Kumar[1], Ritik Raj[2], Dr.Suveg Moudgil[3]
[1]Scholar, [2]Professor, School of Computer Science & Engineering, Galgotias University, Greater Noida

Abstract: Facial emotion recognition, an essential aspect of artificial intelligence (AI) and computer vision (CV), focuses on detecting and interpreting human emotions through facial expressions. This paper presents an approach utilizing deep learning (DL) models, including convolutional neural networks (CNN), binary-temporal scale convolutional neural networks (DTSCNN), recurrent neural networks (RNN), and residual networks (ResNet-50), to achieve real-time and accurate emotion recognition. The main goals of the study are real-time recognition,highaccuracy,minimallatency, and resilience to diverse conditions. Experiments conducted on benchmark datasets evaluate each model based on accuracy, processing speed, and facial orientations. This paper compares the performance of these models and highlights that the CNN model outperforms others, offering superior precision and robustness. Thisresearchcontributestoadvancingfacial emotion recognition with implications for human-computer interaction, psychology, marketing, and healthcare. The CNN ensemblemodelrepresentsasignificantstep forward in facial emotion recognition, providing a comprehensive solution with broad applications in various fields. Its effectiveness underscores the importance of continuous research and refinement of deep learning techniques for solving complex tasks in AI and CV.
Keywords: Facialemotionrecognition,deep neural networks, automatic recognition, database

I. INTRODUCTION

Affective computing is a multidisciplinary field aimed at developing systems that can recognize and interpret human emotions (Banafa,2016). It appearsthatyouhavenot provided any text to paraphrase. Please provide the text you'd like me to work on. The primary motivation for research in this area is the ambition to replicate empathy, allowing machines to identify and react to users' emotional conditions. Facial expressions play a vital role in communication, expressing intricate emotionalinformationduringexchanges. In nonverbal communication, the face acts as an essential medium for conveying feelings (Darwin and Prodger, 1996).

Using machine learning methods like facial recognition, facial expressions can be examined to gather emotional information and assess a person's emotional condition. Affectivecomputingfocusesonrecognizing andinterpretingemotionalstatestoenhance the quality of interactions between humans and machines. Systems possessing this ability might produce responses that are more pertinent and flexible, considering the emotional context of users (Banafa, 2016).

The applications of affective computing are wide-ranging. In marketing, understanding emotions is necessary for assessing the impact of advertisements or products on consumers. Many organizations are investing in affective computing technologies to address issues like workplace stress and to create video games that adapt based on players' emotional responses.

Thispaperexaminesthecreationofsoftware thatidentifiesuseremotionsthroughtheuse of computer vision methods and artificial intelligence algorithms. It also examines multiple emotional theories and approaches for assessing emotions via various algorithms. The software emphasizes emotion recognition using computer vision, especially through the application of convolutional neural networks (CNNs). Facial recognition utilizes the Multitask Cascade Convolutional Networks (MTCNN) framework, combining emotional theories with various assessment techniquestoeffectivelyevaluateemotional states. Facial recognition technology has evolved over the past few decades from a conventional method reliant on outdated machine learning techniques to the deep learning methods that are prevalent today. Previous methods, like PCA-based Eigenfaces, established a groundwork for the existing state-of-the-art by employing statistical techniques to map facial images into a reduced dimensional space. As suggested by Turk and Pentland in 1991, Eigenfacesfacilitatedtheextractionoffacial traits but were highly vulnerable to changes in lighting and various facial expressions.

A. *Facial Databases Accessible*

The success of facial emotion recognition (FER) systems largely depends on the quality and variety of the datasets used for training and evaluation. Numerous publicly availabledatasetshavebeendeveloped,each offering unique features related to emotion categories, annotation quality, demographic diversity, and image acquisition techniques.

Thispartexploresseveralwell-knownfacial emotion databases that have significantly influenced FER research.

B. FER2013(Facial Expression Recognition 2013)

The FER2013 dataset was introduced at the ICML2013Challengeandconsistsof 35,887 grayscale images showing facial expressions, with each image having a resolution of 48×48 pixels. The dataset is divided into training, public testing, and private testing subsets, with the images organized into seven emotional categories: anger, disgust, fear, happiness, sadness,surprise,andneutral.Itisfrequently used due to its large size and accessibility. However, the method of collecting data— using the Google image search API—introduces noise and potential mislabeling due to the automated nature of retrieving images.

C. EnlargedCohn-Kanade(CK+)Dataset

The CK+ dataset acts as a standard in facial expressionrecognition(FER),consistingof 593 sequences from 123 subjects. Each sequence begins with a neutral expression and concludes with a peak face. Emotions are divided into seven categories: rage, scorn, aversion, dread, joy, sorrow, and astonishment. CK+ features detailed Action Unit (AU) annotations based on the Facial Action Coding System (FACS), enabling accurateanalysisofexpressions.WhileCK+ isknownforitsstructuredandtidyformat,it is limited by a relatively small number of participants and a lack of ethnic diversity.

D. JAFFE (Japanese Women's Facial Expression Dataset)

The JAFFE dataset contains 213 images posed by 10 Japanese female models, each exhibiting seven facial expressions: happiness, sadness, surprise, anger, disgust, fear, and neutrality. The images are evaluated by Japanese participants, offering personal emotion labels. Owing to its straightforwardness and transparency, JAFFE is often employed for preliminary evaluationsofFERmodels.Nonetheless,its constraintsinvolvealimiteddatasetsizeand absence of demographic diversity.

E. AffectNet

AffectNet is among the largest datasets of facial expressions, comprising more than 1 million images sourced from the web. It comprises both categorical labels (eight distinct emotions) and dimensional annotations (valence and arousal scores). The dataset presents considerable variety regarding age, ethnicity, gender, and surrounding circumstances. While some annotations were produced by automated systems, a significant portion has been manually checked, rendering it appropriate forbothclassificationandregressiontasksin FER.

F. EmotioNet

TheEmotioNetdatasetincludesmorethan1 million facial images tagged with action unitsandemotionalcategories.Itisintended to examine the identification of facial expressions in natural settings. A portion of the dataset has been annotated by hand, whereas most depends on automated AU detectors. This extensive database aids in creating FER systems that can identify both simple and complex facial expressions.

G. RAF-DB (Real-world Affective Facial Expressions Database)

RAF-DB comprises more than 29,000 authentic facial images annotated with fundamental and complex emotions. Labels are sourced from the crowd and cross- verifiedtoguaranteequality.Thedatabaseis especially remarkable for its incorporation ofnuancedandintricateemotions,rendering it perfect for detailed emotion recognition tasks.

H. Oulu-CASIA NIR&VIS Facial Expression Dataset

The Oulu-CASIA database includes videos and static images of individuals exhibiting six fundamental facial expressions under different lighting scenarios and across both visible and near-infrared wavelengths. It comprisesinformationfrom80subjects, obtainedincarefully regulatedsettings.This dataset is especially valuable for research that focuses on multimodal emotion recognitionandillumination-invariantfacial analysis.

I. Comparative Examination

Compact datasets like CK+ and JAFFE are advantageousforpreliminarybenchmarking because of their precise annotations and regulated environments.Incontrast,extensivedatasets suchasAffectNetandFER2013arevitalfor trainingdeeplearningmodels that generalize more effectively across various real-world situations.

Additionally, the presence of Action Unit annotations in datasets such as EmotioNet and CK+ allows for the investigation of hybrid models that merge AUrecognition withemotion classification.

II. COMPARISONAND RESEARCHGAP

Traditional machine learning methods were mostly used in the early years of facial recognition research. Principal Component Analysis(PCA)wasusedintechniquessuch as Eigenfaces, which were introduced by Turk and Pentland in the early 1990s, to lower the dimensionality of face images. This technique made it simpler to identify and categorize facial photos by projecting them into a feature space. Eigenfaces, however, frequently performed worse in uncontrolled settings due to their extreme sensitivity to changes in lighting and facial expressions. Based on Linear Discriminant Analysis (LDA), Fisherfaces then enhanced Eigenfacesbyoptimizingtheratioofwithin- classvariationtobetween-classvariance.As a result, Fisherfaces were more resilient to illumination variations than Eigenfaces, but theywerestillnotcompletelyresistanttoall types of variability in real-world situations.

The introduction of deep learning brought about a dramatic change in the face recognition field.

Facebook's2014launchofDeepFace,which used a deep neural network trained on a sizable dataset to recognize faces with previouslyunheard-ofprecision,wasoneof the first projects in this field. DeepFace significantly outperformed conventional techniques, achieving near-human performance on the Labeled Faces in the Wild (LFW) benchmark.

Later,Google'sFaceNetintroducedtheidea offaceembeddings,pushingthelimitseven further. FaceNet mapped faces into a compact Euclidean space using a deep convolutionalneuralnetwork(CNN),where facial similarities were directly correlated with distances. Face recognition was transformed by this method, which made identification and verification procedures scalable and effective.

Finally, the problem of cross-domain generalization is still open. Models often performwellonthespecificdatasetstheyare trainedonbutdonotgeneralizewelltonew, unseendatasetsorreal-worldenvironments. This limits the scalability and effectiveness offacerecognitionsystemsandpointstothe needforfurtherresearchintomoreuniversal and adaptable algorithms.

These research gaps underscore the challenges of developing face recognition technologythatisefficient,fair,secure,and ethically responsible, thereby paving the wayformoreadvancedandreliablesystems in the future.

III. IMPLEMENTATION

Face Emotion Recognition (FER) is a rapidly evolving area within AI and CV (computer vision) dedicated to determining humanemotionsthroughfacialexpressions. Recognizing emotions from facial cues is essential for various applications, including human-computer interaction, security, marketing, and healthcare. While early emotionrecognitionsystemsreliedon manual feature extraction techniques, advancements in deep learning have been crucial in significantly improving the accuracy and efficiency of these systems.

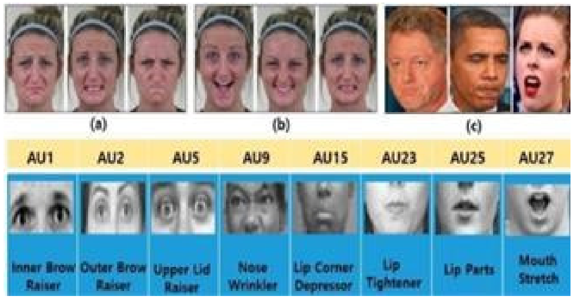
The emergence of deep learning, especially Convolutional Neural Networks (CNNs), hastransformedthefield ofFER.CNNsare capable of learning spatial hierarchies of features from images automatically, making them exceptionally effective for tasks like image classification. Unlike traditional methods, which necessitate handcrafted features such astheshapeoftheeyebrowsormouth,deep learning models can directly extract meaningful features from raw pixel data. This ability allows for the recognition of subtle facial expressions, which are crucial for detecting emotions like happiness, sadness, surprise, or anger.

Finally, emotion classification assigns an emotional label to the face based on the extractedfeatures.Theresultistypicallyone of several predefined emotions, such as happiness, sadness, surprise, anger, disgust, or fear.

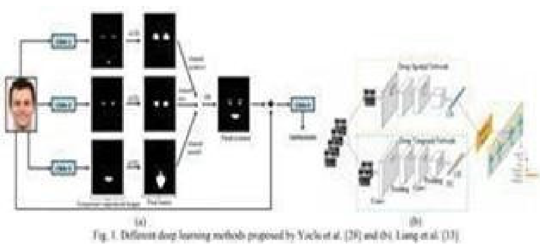
Training deep learning models for FER requires large, diverse, and accurately labeleddatasets.Widelyuseddatasetsinthis field include FER-2013, AffectNet, and EmotioNet, which contain millions of labeled images depicting various facial expressions. These datasets are essential for trainingmodelsthatcanrecognizeemotions across different lighting conditions, angles, and ethnic backgrounds, thus enhancing the model’s robustness and generalization. However,challengesremaininensuringthat models are not biased toward specific ethnicities or facial features.

A. CNN—convolutionalneural networks

Convolutional Neural Networks (CNNs) consistofseverallayersofneurons,with the convolutional layer being the key component. This layer handles an input vectorcomposedofpixelvalues,employing a range of filters that traverse the image to generate the layer's output (Simonyan and Zisserman,2014.)achievedgroundbreaking success by winning the 2012 ImageNet competition.

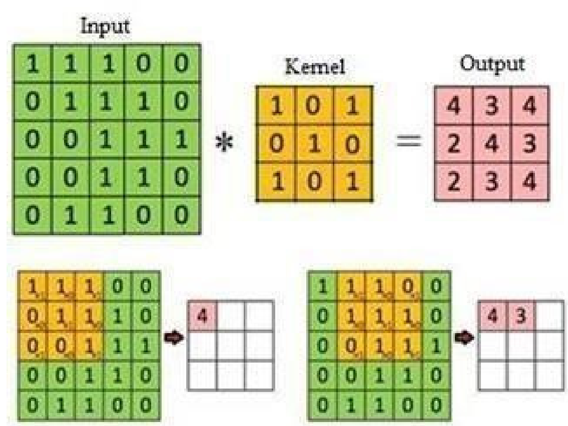


This success marked a pivotal moment in the evolution of the field. Following AlexNet, increasingly sophisticated CNN architectures have been introducedannually,consistentlysurpassing theperformancebenchmarkssetbyprevious models. Furthermore, AlexNet was the first architecture to implement ReLU activation functions, which have since become the standard in modern deep learning applications.



(Convolution algorithm functions through the application of a 3x3 kernel.)

Along with the filter size utilized in the convolutional layers, two important parameters influence their function: stride and padding, which entails surrounding the output generated by the filters with zeros (Mathworks, 2023). Scattered among the convolutional layers are intermediate layers that assist in managing non-linearity. By maintaining dimensions, these layers improvetheneuralnetwork'sresilienceand aid in minimizing the likelihood of overfitting. They frequently employ the ReLU activation function, which is more computationally efficient than conventional functionssuchasSigmoidorTanh(Thomas etal.,2005).Asthenetworkexploresfurther into additional convolutional layers, the activation maps increasingly identify more intricate features over broader areas of the image. The initial layers excel.at recognizing basic structures in specific regions, while deeper layers derive abstract representations by expanding on the previously identified features (Simonyan and Zisserman, 2014). A significant milestone in computer vision, the convolutional neural network known as AlexNet

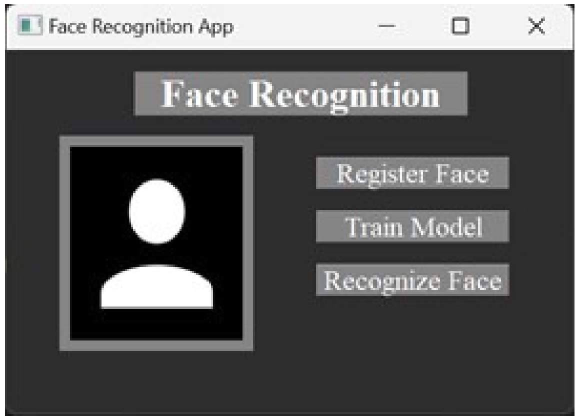


(Krizhevskyetal.,2017)

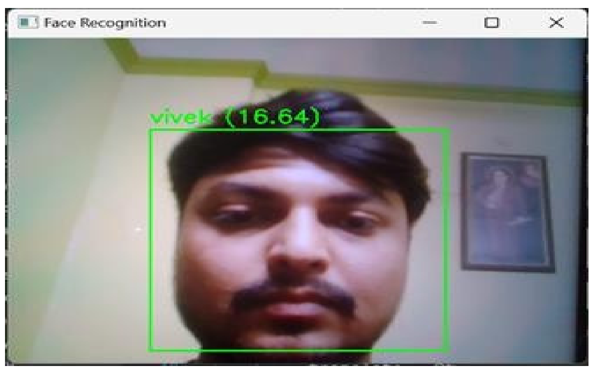
IV. RESULT

CNNs represent a type of neural network known for their remarkable effectiveness in taskslikefacerecognitionandclassification. CNNs are a form of multilayered, feed- forwardneuralnetworks.

CNNsaremadeup ofneuronsthathavelearningweights.Every filteracceptsspecificinputs,processesthem, and then applies a non-linearity [41]. As illustrated in Fig. 1, a standard CNN architecture is demonstrated. The architecture of CNN comprises layers such as convolution, pooling,rectified linear unit (ReLU), and fully connected.



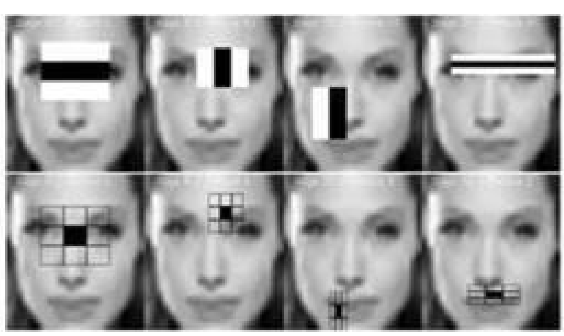
The convolution layer serves as the fundamental component of a convolutional network, handling the majority of intense computational tasks. The primary function of the convolution layer is to extract characteristicsfrominputdatathatisimage- based.By analyzingimagefeatures through small segments of the input image, convolution maintains the spatial relationship among pixels.



A collection of trainable neurons is employed to compress the input image.
This generates a feature map in the output imageandsubsequentlysuppliesthefeature maps to the next convolution layer as input data.

V. CONCLUSION AND FUTURE WORK

Inthisstudy,weexaminedtheapplicationof deep learning methods, particularly Convolutional Neural Networks (CNNs), along with the Haar Cascade algorithm for recognizingfacialemotions(FER).Through theintegrationofthesetechnologies,we demonstrated that emotion recognition systems can achieve high accuracy and efficiency, especially when combined with the robust face detection capabilities provided by Haar Cascade. The research highlights the growing potential of deep learning in enhancing FER systems and opensupavenuesforfutureadvancementsin the field.



(HairCharacteristics)

The CNN-based approach is particularly effective for FER due to its ability to learn hierarchical features from raw image data. Unlike traditional methods that relied on manually crafted features, CNNs automatically capture intricate patterns in facial expressions, leading to better performance in recognizing emotions. Moreover, CNNs are highly adaptable and can be trained on large, diverse datasets, improving their generalization capabilities. The Haar Cascade algorithm, while relatively older, remains a widely used techniqueforfacedetectioninFERsystems due to its speed and efficiency. This combination of CNNs for emotion classification and Haar Cascade for face detection proves to be a powerful tool in real-world applications.

Despite these advancements, numerous challenges continue to exist in the field of FER. One of the major challenges is the handling of variations in facial expressions across different demographics, including age, gender, and ethnicity. Additionally, real-timeemotionrecognitionremainsa difficult task due to the computational demandsofdeeplearningmodels,especially when applied to video streams or large datasets. Facial occlusions, such as glasses, masks, or hair, also pose difficulties for accurate emotion recognition. These issues can affect the overall performance of FER systems in real-world environments.

As we look ahead, there are many crucial aspects where enhancements can be implemented to improve the precision and resilience of FER systems. Initially, employingmoresophisticatedfacedetection algorithms like Multi-task Cascaded Convolutional Networks (MTCNN) could tackle some constraints of Haar Cascade, especially in identifying faces in difficult angles or with partial obstructions. Furthermore, incorporating attention mechanisms into CNN architectures might enhance the focus on key facial features, thereby boosting emotion recognition precision, particularly in instances of nuanced or blended expressions.

Regarding real-time processing, additional studiesarerequiredtoenhancedeeplearning models for quicker and more effective emotion recognition. Methods like model pruning, quantization, and hardware acceleration(forinstance,utilizingGPUsor dedicated AI hardware) may facilitate quicker processing rates without compromising accuracy. Moreover, investigating the possibilities of transfer learning, which involves adjusting pre- trained models for particular tasks, may lessen the time and data needed to develop effective FER systems.

Finally, the application of FER systems in real-world scenarios, such as healthcare, education, and human-robot interaction, presents exciting opportunities. As FER technologycontinuestoimprove,itcouldbe integrated into systems that provide emotional feedback, enhance user experiences, or assist in diagnosing emotionalandpsychologicalconditions.

However, ensuring the ethical use of FER technology, especially in sensitivecontexts, will be crucial in shaping the future of this field.

In conclusion, while deep learning techniques, particularly CNNs combined withtheHaarCascadealgorithm,havemade significant strides in FER, there is still considerable room for improvement. By addressingcurrentchallengesandexploring newmethodologies,futureFERsystemscan become more accurate, inclusive, and capable of real-time, real-worldapplications.

REFERENCES

[1] E.Sariyanidi,H.Gunes,etA.Cavallaro, « Automatic Analysis of Facial Affect: A Survey of Registration, Representation, and Recognition,IEEETrans.Pattern Anal. Mach.Intell., Oct.2014, doi: 10.1109/TPAMI.2014.2366127.

[2] C.-N.Anagnostopoulos,T.Iliou,etI. Giannoukos, « Features and classifiers for emotion recognition from speech: a survey from 2000 to 2011 », Artif. Intell. Rev., vol. 43,no. 2, pp. 155-177, févr. 2015, doi: 10.1007/s10462-012-9368-5.

[3] D. K. Jain, P. Shamsolmoali, et P. Sehdev, « Extended deep neural network for facial emotion recognition », Pattern Recognit. Lett., vol. 120, p.69-74,avr.2019,doi: 10.1016/j.patrec.2019.01.008.

[4] D.H.Kim,W.J.Baddar,J.Jang,et Y. M. Ro, « Multi-Objective Based Spatio-Temporal Feature Representation Learning Robust to Expression Intensity Variations for FacialExpressionRecognition», IEEETrans.Affect.Comput.,vol. 10,no.2,p.223-236,Apr.2019,doi:10.1109/TAFFC.2017.2695999.

[5] W. Zhao et al., "Face Recognition: LiteratureSurvey,"ACMComput. Surv.,vol.35,no.4,pp.399–458,2003.

[6] K.Delac,RecentAdvances in Face Recognition. 2008.

[7] A. S. Tolba, A. H. El-baz, and A. A. El-Harby, "Face Recognition: A Literature Review," Int. J. Signal Process.,vol.2,no.2,pp.88–103,2006.

[8] C. Geng and X. Jiang, "Face recognition using SIFT features," in Proceedings - International Conference on Image Processing, ICIP, pp. 3313–3316, 2009.

[9] S.J.Wang,J.Yang,N.Zhang,andC. G. Zhou, "Tensor Discriminant Color Space for Face Recognition," IEEE Trans.ImageProcess.,vol.20,no.9,pp. 2490–501, 2011.

[10] S.N.Borade,R.R.Deshmukh,andS. Ramu, "Face recognition using fusion of PCA and LDA: Borda count approach," in 24th Mediterranean Conference on Control and Automation, MED 2016, pp. 1164–1167, 2016.

[11] M.A.TurkandA.P.Pentland,"Face Recognition Using Eigenfaces," Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 72–86,

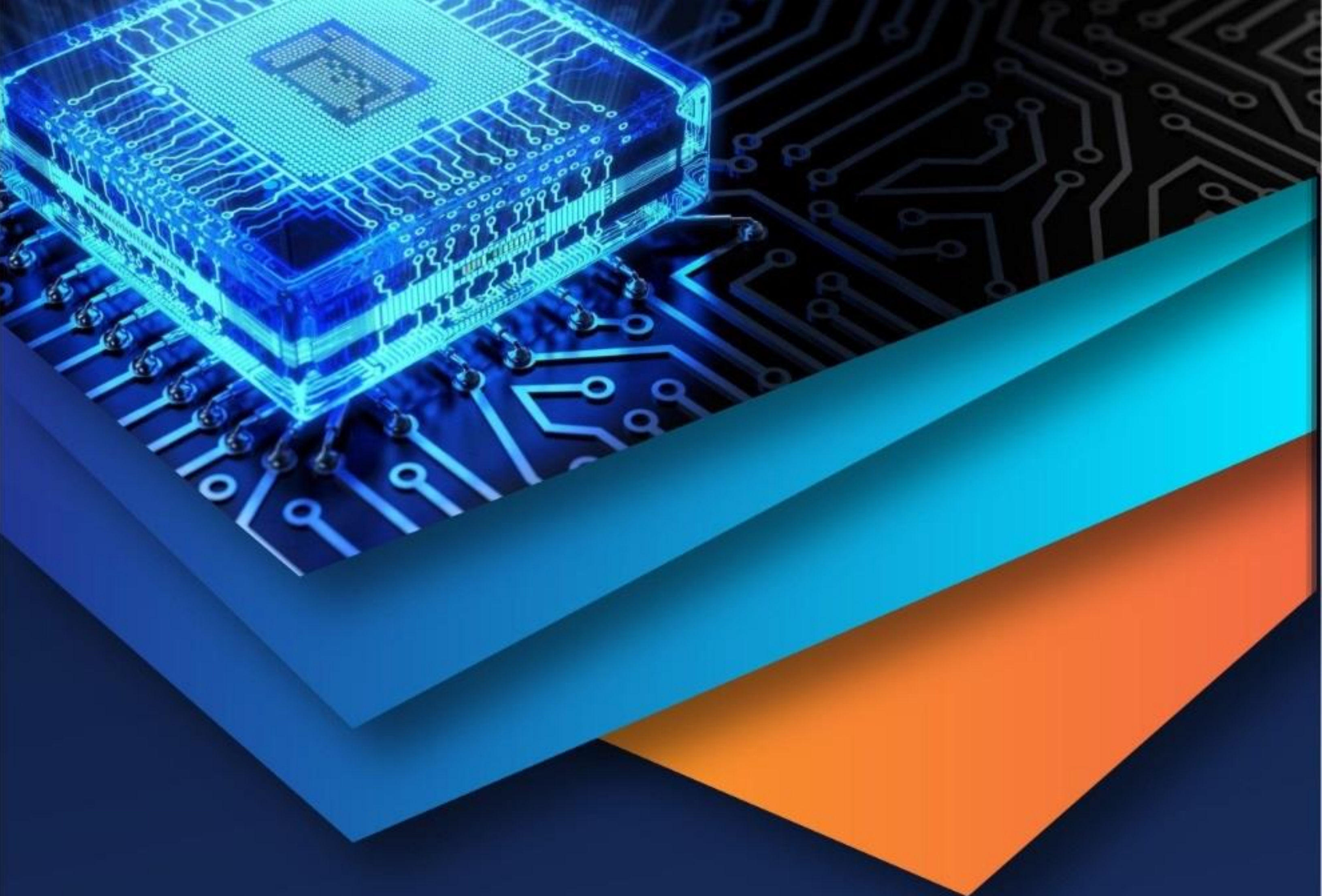
[12] 1991.

[13] M.O.Simón,"ImprovedRGB-D-T- based face recognition," IET Biometrics,vol.5,no.4,pp.297– 303,Dec.2016.

[14] O. Dniz, G. Bueno, J. Salido, and F. DeLaTorre,"Facerecognitionusing Histograms of Oriented Gradients," Pattern Recognit. Lett., vol. 32, no. 12, pp. 1598–1603, 2011.

[15] J.Wright,A.Y.Yang,A.Ganesh, S.S. Sastry, and Y. Ma, "Robust face recognition via sparse representation,"IEEETrans.Pattern Anal.Mach.Intell.,vol.31,no.2,pp. 210–227, 2009.

[16] C.Zhou,L.Wang,Q.Zhang,andX. Wei, “Face recognition based on PCA image reconstruction and LDA,” Opt. - Int. J. LightElectronOpt.,vol.124,no. 22, pp. 5599–5603, 2013.
[17] Z.Lei,D.Yi,andS.Z.Li,“Learning Stacked Image Descriptor for Face Recognition,” IEEE Trans. Circuits Syst. Video Technol., vol. 26, no. 9,pp.1685–1696, Sep. 2016.
[18] P. Sukhija, S. Behal, and P. Singh, “Face Recognition System Using Genetic Algorithm,” in Procedia Computer Science, vol. 85, 2016.
[19] S. Liao, A. K. Jain, and S. Z. Li, “Partial face recognition:t-free approach,” IEEE Trans. Pattern Anal.Mach.Intell.,vol.35,no.5,pp. 1193–1205, 2013.
[20] Z. Zhang, P. Luo, C. C. Loy, and X. Tang, “Learning Deep Representation for Face Alignment with Auxiliary Attributes,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 5, pp. 918–930, 2016.
[21] G. B. Huang, H. Lee, and E. Learned-Miller, “Learning hierarchicalrepresentationsforface verification with convolutional deep belief networks,” in Proceedings of the IEEE Computer Society ConferenceonComputerVisionand PatternRecognition,pp.2518–2525, 2012.
[22] S. Lawrence, C. L. Giles, Ah Chung Tsoi, and D. Back, “Face recognition: a convolutional neural- network approach,” IEEE Trans. NeuralNetworks,vol.8,no.1,pp.98–113, 1997.
[23] O. M. Parkhi, A. Vedaldi, and A. Zisserman, “Deep Face Recognition,”inProceedingsofthe British Machine Vision Conference 2015, pp. 41.1-41.12, 2015.
[24] Z.P.Fu,Y.N.Zhang,andH.Y. Hou, “Survey of deep learning in face recognition,” in IEEE International Conference on Orange Technologies,ICOT2014,pp.5–8,2014.
[25] X.Chen,B.Xiao,C.Wang,X.Cai,Z. Lv, and Y. Shi, “Modular hierarchical feature learning with deep neural networks for face verification,” Image Processing (ICIP),201320thIEEEInternational Conference on. pp. 3690–3694, 2013.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY