International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Advancements in Facial Expression Recognition: State-of-the-Art Techniques and Innovations

A. Babisha[1], A. Swaminathan[2], D. Anuradha[3], C. Gnanaprakasam*[4], T. Kalaichelvi[5]

Submitted: 10/01/2024 **Revised**: 16/02/2024 **Accepted**: 24/02/2024

Abstract: Utilizing Machine Learning and Convolutional Neural Networks, researchers have unlocked the capacity to excel in identifying emotions by leveraging the potent and instinctive nature of facial expressions, a robust medium for individuals to communicate their feelings and aims. The development of efficient recognition systems is essential for enhancing human-computer interaction. However, the field of facial expression recognition encompasses various methodologies that can significantly impact the performance of facial recognition systems. In this study, we present a cutting-edge achievement of 95% accuracy on the FER2013 dataset, achieved by implementing a combination of innovative techniques inspired by recent research. In addition, we present innovative techniques aimed at boosting precision by amalgamating established CNN structures like VGG-16 and ResNet-50 with supplementary datasets like JAFFE and CK. In our endeavor to forecast emotions, we adopt an alternative method leveraging geometric attributes and facial landmarks to calculate and relay a feature set to an SVM model. Our findings unequivocally demonstrate the superiority of the ResNet50 model over others in real-time emotion prediction, significantly enhancing the system's accuracy.

Keywords: *Facial Landmarks Detection, Deep Learning Models, Convolutional Neural Networks (CNN), Facial Expression Classification, Emotion Detection, Facial Image Processing, Artificial Intelligence (AI),; Facial Expression Analysis*

1. Introduction

Facial expressions play a pivotal role in nonverbal communication, significantly influencing interpersonal relationships, and find applications in diverse fields like emotion interpretation, cognitive science, and social interaction. The central objective of facial expression recognition is to categorize expressions captured in photographs of human faces into various emotional states, including but not limited to happiness, fear, neutrality, surprise, and sadness [1,2]. Studying and capturing emotional signals within a user's facial expressions [3,4] are profoundly important for enriching the calibre of interactions in human-computer communication systems. To foster emotionally attuned interactions between humans and computers, it becomes crucial to prioritize the creation of systems adept at identifying human emotions. An automated system proficient in discerning an individual's emotions through their expressions empowers the machine to tailor its responses accordingly, fostering more personalized interactions. Once more, the central aim of facial expression recognition revolves around categorizing emotions displayed on human faces captured in images into diverse emotional classifications, spanning feelings like joy, apprehension, neutrality, astonishment, and sadness.

The interplay between facial expression and spoken words is a crucial aspect of psychology, as it significantly influences the impact of verbal communication. Breaking down this system reveals three core modules: face registration, feature extraction, and classification. These elements form the fundamental components essential for its operation. In our research, we aim to comprehensively investigate each of these modules, Utilizing various methodologies to assess and compare the efficacy of distinct models.

Fig.1. Casting Light on the Inner Workings of Facial Expression Recognition

Our initial approach relies on geometric attributes and facial landmarks. This technique involves the calculation of a feature vector derived from these landmarks, which is subsequently fed into our Support Vector Machine (SVM) model for training and the prediction of emotions. This

[1] *Dept of AIDS, Panimalar Engineering College, Chennai 600123, India. Email: babisha15@gmail.com, ORCID ID : 0000-0001-6050-8488*
[2] *Dept of CSBS, Panimalar Engineering College, Chennai 600123, India, Email: swamisivam19@gmail.com, ORCID ID : 0000-0001-7672-1339*
[3] *Dept of CSBS, Panimalar Engineering College, Chennai 600123, India, Email: vanu2020@gmail.com, ORCID ID : 0000-0001-7357-0715*
[4] *Dept of AIDS, Panimalar Engineering College, Chennai 600123, India, Email: cgn.ds2021@gmail.com, ORCID ID : 0000-0003-1401-9418*
[5] *Dept of AIDS, Panimalar Engineering College, Chennai 600123, India, Email: kalaichelvi2012@gmail.com, ORCID ID : 0009-0001-9999- 6965*

method's processing pipeline involves multiple stages, starting with image pre-processing, followed by feature extraction, and concluding with emotion classification. Figure 1.Shows how the Casting Light on the Inner Workings of Facial Expression Recognition.

Within the domain of deep learning [5], neural networks automatically extract features, deviating from traditional machine learning approaches where facial features are manually extracted from input images for emotion categorization. Convolutional Neural Networks (CNNs) stand out as a specialized neural network renowned for their adeptness in deriving significant traits from images while retaining spatial details. Their remarkable ability to generalize and withstand geometric variations positions CNNs as a pivotal focal point in this paper's exploration.

The subsequent section delves into a discussion of related research within this field, offering valuable context and insights. Section III elaborates on the Cognitive Computing: Unveiling Human Emotions through Machine Learning. The analysis and outcomes of our experiments are presented in Section IV, shedding light on the empirical results of our research. Finally, Section V provides the concluding remarks, summarizing the key findings and implications of this study.

2. Related Work

In the last decade, the scrutiny and automated identification of facial expressions have emerged as a prominent focus within the computer vision research community. Inspired by cognitive science findings, researchers have aspired to develop systems capable of discerning facial expressions in images and videos, with a focus on categorizing expressions into major emotional subcategories such as, anger, fear, joy, disgust and surprise.

Previous research has contributed significantly to improving the ability of machines to recognize facial emotions. Some studies [6-8] employed conventional neural networks (CNNs) for this purpose; its typical architecture encompasses an input layer, five convolutional layers, three pooling layers, a fully connected layer, and an output layer, specifically trained on 64x64 pixel images. Validation was often conducted using databases like JAFFE and CK+.

In another approach [9,10], spatial features were extracted by calculating horizontal and vertical distances between the 68 facial landmarks. These features were then used in conjunction with a support vector machine (SVM) for the classification.

According to research [11-13], facial emotion recognition typically involves three key steps: dimensionality reduction, classification and feature extraction. The challenges in facial expression recognition are largely associated with dimensionality and feature selection. To mitigate these challenges, geometric features were suggested as an alternative, and the feature extraction involved utilizing CNN Classifier and facial landmark recognition, while the system's performance was assessed using databases like MMI, MUG, CK and JAFEE.

Another paper [14] outlined recent advancements in artificial intelligence pertaining to automated emotion recognition (FER) through deep learning models. Researchers showcased the accuracy of FER models based on deep learning, utilizing diverse CNN architectures and databases. However, these methods primarily concentrated on the six fundamental emotions and didn't address more intricate emotional states. In contrast, a distinct approach introduced in [15] involved creating a mobile application tailored to respond to a user's visual emotions. This system operated by processing emotional data collected from a mobile user through wireless sensors in the cloud, promptly analysing the data. The technique relied on detecting small clusters of facial cue movements to identify the user's emotional expression and offer immediate feedback to alleviate negative emotions.Numerous datasets exist for training facial expression emotion recognition models. Prominent among these are the Facial Expression Recognition 2013 (FER2013) dataset [17] and the CK dataset, expanded into the Extended Cohn-Kanade Dataset (CK+) [18]. These datasets hold significant popularity, serving as vital resources for researching and crafting efficient facial expression recognition models.

3. Proposed Methodology

3.1. Cognitive Computing: Unveiling Human Emotions through Machine Learning

In this segment, we delineate the thorough methodology centred on geometric attributes and facial landmarks. As mentioned earlier, this approach adheres to a systematic progression, initiating with image pre-processing, proceeding to feature extraction, and culminating in emotion classification. For our research, we opted for the Cohn-Kanade (CK) dataset, specifically curated to aid in exploring the automated identification of particular facial expressions. This dataset encompasses seven discrete emotions: happiness, surprise, fear, sadness, neutrality, disgust, and anger, catalogued by 210 participants within the sample. Each image in the dataset is standardized at 48x48 pixels, and Figure 2 provides a visual representation of sample images showcasing these seven emotions.

Fig.2. Visual representation of sample images showcasing these seven emotions

Within the image pre-processing stage, the dataset comprises photos captured under various lighting conditions. To ensure consistent lighting conditions across all images, we have applied histogram equalization to all pictures. Furthermore, we utilized image normalization, a common image processing method that aims to adjust the intensity range of pixels. The main objective behind normalization is to transform an input image into a pixel value range that aligns better with human perception, thereby facilitating subsequent computations. This normalization process also standardizes the size and color of the photos.

The initial step in building an automated facial expression recognition system involves face detection, where the system determines the presence of a human face in an image. This face detection step is pivotal for subsequent processing. To detect faces within the images, we make use of OpenCV alongside its integrated Dlib [19] function, leveraging its capability to yield an object detector for identifying faces in images. Additionally, the Dlib library offers a face detection function called 'get_frontal_face_detector()', which produces an array of rectangle objects denoting detected faces within the image. Each rectangle object encapsulates the rectangular region containing the detected face, and it provides the coordinates of this region. Figure 3 offers displaying a visual representation of an image sourced from the database, showcasing the identified facial region.

Fig.3. Displaying a visual representation of an image sourced from the database, showcasing the identified facial region.

Following the successful face detection process, the subsequent step involves predicting facial landmarks. Facial landmarks refer to distinct key points on human facial images, characterized by their (x, y) coordinates within the image. These points serve as identifiable markers on the face. These markers play a pivotal role in identifying and depicting significant facial features, contributing to the classification of facial expressions.

To obtain the 68 (x, y) coordinates representing these facial landmarks, we utilize a pre-trained facial landmark detector available within the Dlib package [28]. These coordinates correspond to various facial structures on the detected face. Once the prediction of facial landmarks is complete, following the placement of these overlaid points on the face, the subsequent objective involves converting them into meaningful attributes suitable for input into the classifier [37]. Consequently, upon acquiring the coordinates of these facial points of interest, we compute a feature vector. This vector succinctly portrays an individual's emotional state by factoring in the spatial arrangements of these facial landmarks concerning each other.

Fig.4. Creating a visual representation of the facial landmarks overlaid onto an image from our dataset.

One method to achieve this is by computing the average of the two axes, providing a central reference point. We can then determine the positions of all the facial landmarks in relation to this central reference point. For a visual representation of these facial landmarks, please refer to Figure 4, creating a visual representation of the facial landmarks overlaid onto an image from our dataset.

3.2. Deep Learning Unravelled: Deciphering Human Emotions

Our CNN model [29] boasts a sophisticated design encompassing four distinct stages, each contributing to the progressive refinement of input data. As the model advances through these stages, the size of the input image undergoes a gradual reduction, ultimately leading to accurate emotion classification. Figure 5. CNN Model architecture is shown below diagram.

At the inception of the model, an input layer is primed to receive 48 x 48-pixel images. The journey of each layer in the first three phases commences with a convolutional operation and culminates with a dropout mechanism. In the initial phase, we employ 64 kernels to perform convolution on the input. Subsequently, batch normalization [30] is carried out to obtain inputs for the subsequent layer. This pattern of convolution followed by batch normalization is reiterated in the subsequent layers, ensuring a systematic extraction of features.

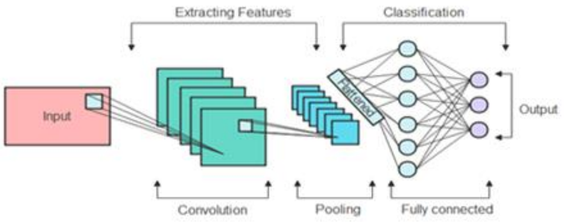

Fig.5. CNN Model Architecture

During the subsequent phase, a max-pooling operation is implemented using a pool size of 2 x 2, resulting in an output size of 24 x 24. Following this, a dropout is incorporated at a rate of 0.25%, aiding in the regularization of the model. This phase incorporates 128 kernels and maintains a dropout rate of 0.4. The resulting output size following max-pooling is reduced to 12 x 12. Transitioning into the third phase, the model employs 256 kernels with a dropout rate of 0.25 [31]. Another round of max pooling leads to a further reduction in output size, down to 6 x 6. The next layer introduces 512 kernels while maintaining a dropout rate of 0.25%, consequently reducing the output size to a more condensed 3 x 3.

In the ultimate phase, feature transformation commences with a flattening layer (Flatten) [32], followed by dense and output layers. This transformation ensures that the data is converted into a one-dimensional array, aligning with the model's specifications and enabling the classification of the seven distinct emotions. The flattened output is passed to the dense layer, which utilizes the SoftMax function [33] to calculate class probabilities. Subsequently, a final round of batch normalization is conducted, culminating in the output layer unveiling the predicted class probabilities, offering valuable insights into the emotional states captured within the input data.

3.3. VGG-16 model

The VGG-16 model is a renowned convolutional neural network architecture crafted by the Visual Geometry Group at the University of Oxford. Widely influential in computer vision, it's recognized for its simplicity and efficacy in image classification tasks. Here's a breakdown of its architecture: Input Layer: The VGG-16 model takes in images set at a fixed size of 224x224 pixels, accepting color images with three color channels (RGB).

Convolutional Layers (Block 1-5): VGG-16 comprises five blocks of convolutional layers. Each block contains several convolutional layers, succeeded by max-pooling layers. The count of convolutional layers escalates as one progress deeper into the network. Block 1: Two 3x3 convolutional layers with 64 filters, followed by a max-pooling layer (2x2). Block 2: Two 3x3 convolutional layers with 128 filters, followed by a max-pooling layer (2x2). Block 3: Three 3x3 convolutional layers with 256 filters, followed by a max-pooling layer (2x2). Block 4: Three 3x3 convolutional layers with 512 filters, followed by a max-pooling layer (2x2). Block 5: Three 3x3 convolutional layers with 512 filters, followed by a max-pooling layer (2x2). Fully Connected Layers: After the convolutional layers, Figure 6. VGG-16 Architecture structure diagram has three fully connected layers, often referred to as dense layers. In these layers are responsible for the final classification. Fully Connected Layer 1: 4096 neurons with a Rectified Linear Unit (ReLU) [34] activation function. Fully Connected Layer 2: 4096 neurons with ReLU activation. Fully Connected Layer 3 (Output Layer): The final layer consists of as many neurons as there are classes in the classification task, usually 1,000 for the original ImageNet version. It uses a softmax activation function to produce class probabilities. Flattening Layer: Before the fully connected layers, there is a flattening layer that converts the 3D feature maps from the last convolutional block into a 1D vector to be fed into the dense layers. The key characteristics of the VGG-16 model are its extensive use of small 3x3 convolutional filters and its depth, which allows it to learn intricate patterns and features from input images.

The VGG-16 model has been pre-trained on the ImageNet dataset [35], encompassing an extensive array of images across diverse categories. Its deep architecture and utilization of smaller filters empower VGG-16 to capture fine details in images, enabling strong performance across multiple computer vision tasks such as image classification, object detection, and feature extraction. Its impact extends to influencing the creation of subsequent CNN architectures [36].

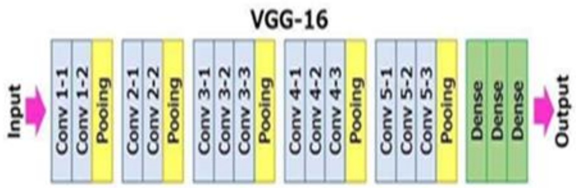

Fig.6. VGG-16 Architecture structure diagram

3.4. ResNet-50

Absolutely, ResNet-50 is indeed a robust architecture in the domain of computer vision, known for its depth and performance in image classification tasks. Here's a comprehensive breakdown of the ResNet-50 architecture: Input Layer: The typical input for ResNet-50 is a color image measuring 224x224 pixels, with three color channels (RGB). Convolutional and Pooling Layers (Initial Convolution): The network commences with a 7x7 convolutional layer equipped with 64 filters and a stride of 2.This is succeeded by a max-pooling layer featuring a 3x3 pool size and a stride of 2. Residual Blocks (Blocks 1-4): The hallmark of ResNet architecture lies in its residual blocks, integral for handling deeper networks.

ResNet-50 comprises four such blocks, each housing multiple convolutional layers and a skip connection (shortcut connection) that circumvents one or more convolutional layers. Block 1: Comprises three residual units, each containing two 3x3 convolutional layers with 64 filters. Block 2: Encompasses four residual units furnished with 128 filters. Block 3: Embraces six residual units equipped with 256 filters. Block 4: Consists of three residual units featuring 512 filters. Fully Connected Layer (Output Layer): Following the convolutional layers and residual blocks, ResNet-50 generally integrates a global average pooling layer followed by a fully connected layer designed for classification. The output layer's neuron count corresponds to the number of classes in the given classification task.

ResNet-50's architecture's prowess stems from its ability to handle deeper networks, introducing skip connections to mitigate the vanishing gradient problem, ultimately contributing to its outstanding performance in various image classification challenges.

The Global Average Pooling (GAP) Layer in ResNet-50 condenses spatial dimensions to 1x1, generating a feature vector for each channel in the final convolutional layer. One of ResNet-50's significant innovations lies in its use of residual connections, which facilitate the learning of the difference between the desired and actual outputs from the preceding layer. These skip connections aid in training very deep networks more effectively and address the vanishing gradient problem.

ResNet-50 is typically pertained on extensive image datasets like ImageNet and subsequently fine-tuned for specific tasks such as object detection, image segmentation, or facial recognition. This architecture has gained widespread popularity in deep learning due to its adeptness in capturing intricate image features and patterns, rendering it suitable for various computer vision applications. Figure 7 illustrates the ResNet-50 model's architecture.

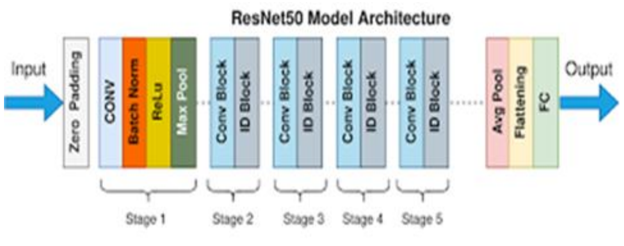

Fig.7. ResNet-50 model architecture.

4. Results and Analysis

Indeed, the training process for each model is carried out independently, and its duration hinges on multiple factors. Elements such as the dataset's size, the specific algorithm utilized, and the number of defined epochs for training significantly impact the duration required for effective model training. Larger datasets or more complex algorithms might necessitate longer training times to achieve desired performance levels. Similarly, adjusting the number of epochs, which denotes the number of times the entire dataset is passed forward and backward through the neural network, can also influence the duration of training.

In our specific case, we observe that the training times diverge significantly. In scenarios where the machine learning (ML) approach is applied, the training time is relatively shorter, typically taking a minimum of 45 minutes to complete. However, for the more complex deep learning (DL) approaches [16], especially when dealing with extensive datasets, the training duration can be considerably longer, extending up to approximately 3 hours. This time discrepancy is primarily due to the computational demands associated with larger datasets and the depth of the neural network architectures used.

It's essential to note that the number of epochs, which represent the number of times the entire dataset is processed during training, is a parameter that also influences the training time. Fine-tuning the number of epochs allows for a balance between model convergence and training time efficiency, ensuring that the model optimally captures the underlying patterns within the data.

Measuring the performance metrics of a machine learning or deep learning model is essential to evaluate its effectiveness and reliability. The choice of performance metrics depends on the specific task you are working on, such as classification, regression, or clustering. Here are some common performance metrics for various machine learning tasks: Classification Metrics (for tasks like image classification, sentiment analysis, etc.). Accuracy: Overall correctness of the model's predictions, calculated as the number of correct predictions divided by the total number of predictions [20].Precision: Quantifies the proportion of true positive predictions among all positive predictions. It helps assess the model's ability to avoid false positives. Recall (Sensitivity): Calculates the percentage of actual positives correctly predicted by the model. Crucial for identifying false negatives. F1 Score [21]: The harmonic mean of precision and recall. It balances precision and recall, offering a single metric to evaluate a model's performance. Area Under the ROC Curve (AUC-ROC): Primarily used for binary classification tasks, it measures the trade-off between true positive rate (TPR) and false positive rate (FPR).

Demonstrating the efficacy of the machine learning (ML) approach, our model proficiently classified test images into distinct emotions, namely surprise, happiness, disgust, sadness, fear, anger and contempt, achieving an impressive accuracy rate of 97.99%. Furthermore, the inference time for our model was notably efficient, with predictions generated in a mere 0.7 seconds. To provide a comprehensive perspective, we present a comparative Table 1 below, which elucidates the performance of our proposed method when juxtaposed with the method utilized on the CK+ Dataset [22].

Table.1. Contrasting the suggested approach with the GSDS method involves examining the differences and similarities between the two methodologies.

	GSDS	**Proposed**
Anger	0.781	0.851
Contempt	0.642	0.932
Disgust	0.934	0.893
Fear	0.81	0.85
Happiness	0.991	1
Sadness	0.65	0.92
Surprise	1	0.97
Total	0.888	0.984

The VGG16 [27] model underwent 50 iterations during training, utilizing a batch size of 128 and requiring around 2 hours and 34 minutes to finish the training process. The outcomes of this training process are visualized in Figure 8. In contrast, the transfer learning-based ResNet-50 model underwent training using a batch size of 128 and was trained over 50 epochs.

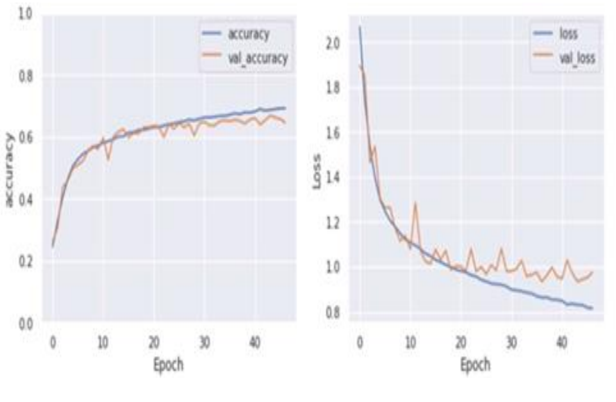

Fig.8. The outcomes concerning the training and validation accuracy as well as the loss for the VGG-16 model.

The training results for this ResNet-50 model are illustrated in Figure 9. The upcoming discussion will delve into an in-depth analysis of the detailed training results for each of these distinct architectural models. To facilitate a comprehensive understanding, the following Table 2 provides a comparative analysis of these models, considering a range of criteria.

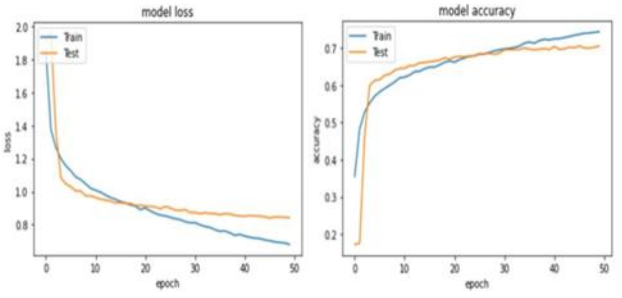

Fig.9. The training results for this ResNet-50 model.

An additional crucial criterion for this comparative analysis is real-time prediction performance, a pivotal factor given our system's intended use in VR systems [23]. We conducted a dedicated test utilizing an external camera to assess the real-time prediction capabilities of all the models. The test results unequivocally demonstrated that the ResNet-50 model outperforms the other models in the real-time prediction of all emotion classes.

Additionally, echoing previous literature sources [24–26], it's essential to highlight that the approach outlined in this paper allows for adaptability and implementation alongside image segmentation techniques utilizing wavelet transformations. Such integration has the potential to bolster the system's efficacy and precision, especially in real-time applications within immersive environments such as VR systems.

Table.2. Comparative analysis of model performances

Model	**Dataset**	**Accuracy**	**Model size**	**Compt. time**
VGG-16	FER2013	83%	15.3 Mo	0s 861ms
CNN	FER2013	75%	16.5 Mo	0s 413ms
ResNet-50 Transfer Learning	FER2013 JAFFE CK+	90.04%	186 Mo	1s 807ms

5. Conclusion

This research explores the feasibility of employing Machine Learning and Deep Learning techniques for the precise identification of facial expressions. In our investigation, we establish that leveraging a subset of 68 facial landmark points, as opposed to utilizing the entirety of the image's pixels, can effectively distinguish and forecast various facial expressions. A pivotal aspect of our experiment was feature extraction, wherein we found that the incorporation of additional distance and angle features yielded exceptional accuracy, reaching an impressive 98% on the CK+ database. Notably, the outcomes demonstrated a remarkable degree of similarity to those achieved using an alternative approach.

Moreover, our study emphasizes the notable performance advancements achievable by strategically employing transfer learning and data augmentation methods. Particularly, our experiments demonstrated the ResNet50 model's superior performance compared to other suggested architectures. This evaluation spanned various datasets: FER2013, JAFFE and CK+, emphasizing the resilience and adaptability of the ResNet-50 model in facial expression recognition.

Author contributions

A.Babisha: Conceptualization, Methodology, Software, **Dr.A.Swaminathan:** Field study, Data creation, Writing-Original draft preparation, **Dr. D Anuradha :** Software Validation., Field study **Dr.C.Gnanaprakasam:** Visualization, Investigation,**Dr**. **T.Kalaichelvi:** Writing-Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] L. Zahara, P. Musa, E. P. Wibowo, I. Karim, and S. B. Musa, "The facial emotion recognition (FER-2013) dataset for prediction system of micro-expressions face using the convolu-tional neural network (CNN) algorithm based Raspberry Pi," in 2020 Fifth international conference on informatics and computing (ICIC), pp. 1-9: IEEE, 2020. https://doi.org/10.1109/ICIC50835.2020.9288560

[2] P. S. Wijaya, A. A. Firdaus, A. P. J. Dwitama, and M. Mustiari, "Pengenalan Ekspresi Wajah Menggunakan DCT dan LDA untuk Aplikasi Pemutar Musik (MOODSIC)," Jurnal Teknologi Informasi dan Ilmu Komputer, vol. 5, no. 5, pp. 559-566, 2018. https://doi.org/10.25126/jtiik.201855935

[3] N. A. Chinchanikar, "Facial Expression Recognition Using Deep Learning: A Review," International Research Journal of Engineering and Technology (IRJET), vol. 6, pp. 3274-3281, 2019 [4].

[4] M. Wang, Z. Wang, S. Zhang, J. Luan, and Z. Jiao, "Face expression recognition based on deep convolution network," in 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), pp. 1-9: IEEE, 2018. https://doi.org/10.1109/CISP-BMEI.2018.8633014

[5] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," Computational intelligence and neuroscience, vol. 2018, 2018. https://doi.org/10.1155/2018/7068349

[6] P. R. Dachapally, "Facial emotion detection using convolutional neural networks and representational autoencoder units," arXiv preprint arXiv:1706.01509, 2017.

[7] D. Theckedath and R. Sedamkar, "Detecting affect states using VGG16, ResNet50 and SE-ResNet50 networks," SN Computer Science, vol. 1, no. 2, pp. 1-7, 2020. https://doi.org/10.1007/s42979-020-0114-9

[8] A. Qashlim, B. Basri, H. Haeruddin, A. Ardan, I. Nurtanio, and A. A. Ilham, “Smartphone Technology Applications for Milkfish Image Segmentation Using OpenCV Library”, Int. J. Interact. Mob. Technol., vol. 14, no. 08, pp. pp. 150–163, May 2020. https://doi.org/10.3991/ ijim.v14i08.12423.

[9] C. Gacav, B. Benligiray, and C. Topal, "Greedy search for descriptive spatial face features," in 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1497-1501: IEEE, 2017. https://doi.org/10.1109/ ICASSP.2017.7952406

[10] N. Özbey and M. B. Gülmezoğlu, "Facial Expression Recognition Using Proposed Geometric Features," EasyChair2516-2314, 2021.

[11] N. Gopalan, S. Bellamkonda, and V. S. Chaitanya, "Facial expression recognition using geometric landmark points and convolutional neural networks," in 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), pp. 1149-1153: IEEE, 2018. https://doi.org/10.1109/ICIRCA.2018.8597226

[12] L. Zhao, "A facial expression recognition method using two-stream convolutional networks in natural scenes," Journal of Information Processing Systems, vol. 17, no. 2, pp. 399-410, 2021.

[13] M. I. Rusydi, R. Hadelina, O. W. Samuel, A. W. Setiawan, and C. Machbub, "Facial Features Extraction Based on Distance and Area of Points for Expression Recognition," in 2019 4th Asia-Pacific Conference on Intelligent Robot Systems (ACIRS), pp. 211-215: IEEE, 2019. https://doi.org/10.1109/ACIRS.2019.8936005

[14] A. A. Pise, M. A. Alqahtani, P. Verma, D. A. Karras, and A. Halifa, "Methods for Facial Expression Recognition with Applications in Challenging Situations," Computational Intelligence and Neuroscience, vol. 2022. https://doi.org/10.1155/2022/9261438

[15] A. M. Ayyal Awwad, “Visual Emotion-Aware Cloud Localization User Experience Framework Based on Mobile Location Services”, Int. J. Interact. Mob. Technol., vol. 15, no. 14, pp. pp. 140–156, Jul. 2021 https://doi.org/10.3991/ijim.v15i14.20061

[16] Babu, G. H., Srinivas, M., Gnanaprakasam, C., Prabu, R. T., Devi, M. R., Ahammad, S. H., ... & Rashed, A. N. Z. (2023). Meander Line Base Asymmetric Co-planar Wave Guide (CPW) Feed Tri-Mode Antenna for Wi-MAX, North American Public Safety and Satellite Applications. Plasmonics, 18(3), 1007-1018

[17] I. J. Goodfellow et al., "Challenges in representation learning: A report on three machine learning contests," in International conference on neural information processing, 2013, pp. 117-124: Springer. https://doi.org/10.1007/978-3-642-42051-1_16

[18] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," in 2010 ieee computer society conference on computer vision and pattern recognition workshops, 2010, pp. 94-101: IEEE. https://doi.org/10.1109/CVPRW.2010.5543262

[19] Y. Celik, M. Talo, O. Yildirim, M. Karabatak, and U. R. Acharya, "Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images," Pattern Recognition Letters, vol. 133, pp. 232-239, 2020. https://doi.org/10.1016/j.patrec. 2020.03.011

[20] Krishnamoorthy, N. V., KH, S. M., Gnanaprakasam, C., Swarna, M., & Geetha, R. (2023, April). A Robust Blockchain Assisted Electronic Voting Mechanism with Enhanced Cyber Norms and Precautions. In 2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM) (pp. 1-8). IEEE.

[21] Swarna, M., Geetha, R., Saranya, G., KH, S. M., & Gnanaprakasam, C. (2023, April). An Empirical Design of IoT based Health Surveillance Scheme for Coronavirus Affected Patients. In 2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM) (pp. 1-8). IEEE.

[22] K. Li, Y. Jin, M. W. Akram, R. Han, and J. Chen, "Facial expression recognition with convolutional neural networks via a new face cropping and rotation strategy," The visual computer, vol. 36, no. 2, pp. 391-404, 2020. https://doi.org/10.1007/s00371-019-01627-4

[23] B. Houshmand and N. M. Khan, "Facial expression recognition under partial occlusion from virtual reality headsets based on transfer learning," in 2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM), 2020, pp. 70-75: IEEE. https://doi.org/10.1109/BigMM50055.2020.00020

[24] S, Ziani., El Hassouani, Y. A new approach for extracting and characterizing fetal electrocardiogram. Traitement du Signal, Vol. 37, No. 3, pp. 379-386, 2020. https://doi.org/10.18280/ts.370304

[25] S. Ziani and Y. El Hassouani, "Fetal-Maternal Electrocardiograms Mixtures Characterization Based on Time Analysis," 2019 5th International Conference on Optimization and Applications (ICOA), Kenitra, Morocco, pp. 1-5, 2019. https://doi.org/10.1109/ICOA.2019. 8727619

[26] S. Ziani, A. Jbari and L. Bellarbi, "QRS complex characterization based on non-negative matrix factorization NMF", 2018 4th International Conference on Optimization and Applications (ICOA), Mohammedia, Morocco, pp. 1-5, 2018. https://doi.org/10.1109/ICOA. 2018.8370548

[27] Geetha, R., Krishnamoorthy, N. V., Murugan, K. S., Gnanaprakasam, C., & Swarna, M. (2023, April). A Novel Deep Learning based Stress Analysis and Detection Scheme using Characteristic Data. In 2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM) (pp. 1-8). IEEE.

[28] Gnanaprakasam, C., Indumathy, M., Khilar, R., & Kumar, P. S. (2022). Artificial intelligence based optimization for mapping IP addresses to prevent cyber-based attacks. Measurement: Sensors, 24, 100508.

[29] Gnanaprakasam, C., Anand, S., Manoj Kumar, R., & Menaka, R. (2021). Facial Expression Image Analysis to Classify High and Low Level ASD Kids Using Attention Mechanism Embedded Deep Learning Technique. In Advances in Electrical and Computer Technologies: Select Proceedings of ICAECT 2020 (pp. 559-568). Springer Singapore.

[30] C. Gnanaprakasam, Manoj Kumar Rajagopal, Attention Residual Network for Micro-expression Recognition Using Image Analysis, in Journal of Advanced Research in Dynamical & Control Systems, 07-Special Issue,2020, Pages- 1261 – 1272.

[31] Gnanaprakasam, C., and Manoj Kumar Rajagopal. "Review on Facial Micro-Expression Detection." Int J Innov Technol Explor Eng 8 (2019): 1103-1115.

[32] [32] Sumathi, S., C. Gnanaprakasam, and R. RANIHEMA MALINI. "Face Recognition-Average-Half-Face Using Wavelets." IPCV 2010: proceedings of the 2010 international conference on image

processing, computer vision, & pattern recognition (Las Vegas NV, July 12-15, 2010). 2010.

[33] Gnanaprakasam, C., S. Sumathi, and R. RaniHema Malini. "Average-half-face in 2D and 3D using wavelets for face recognition." Proceedings of the 9th WSEAS international conference on Signal processing. 2010.

[34] Gnanaprakasam, C., and M. Rajagopal. "K.,(2023). Identification of Autism Spectrum Disorder using Residual Attention Network for Facial Image Analysis." J Current Trends Comp Sci Res 2.1: 31-39.

[35] Gnanaprakasam, C., et al. "Long-Range and Low-Power Automated Soil Irrigation System Using Internet of Things: An Experimental Study." Contemporary Developments in Agricultural Cyber-Physical Systems. IGI Global, 2023. 87-104.

[36] C, G. ., Rajagopal, M. K. ., Swaminathan, A. ., Livingston, L. M. M. ., & Ramkumar, G. . (2023). Cultivating Insight: Detecting Autism Spectrum Disorder through Residual Attention Network in Facial Image Analysis. International Journal on Recent and Innovation Trends in Computing and Communication, 11(11s), 333–340.

[37] Jayaram, Ramaprabha, S. Irin Sherly, and C. Gnanaprakasam. "Study of ECG Analysis based Cardiac Disease Prediction using Deep Learning Techniques." International Journal of Intelligent Systems and Applications in Engineering 11.4 (2023): 431-438.