WWW.IJRPR.COM

International Journal of Research Publication and Reviews

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

Face Identification and Expression Recognition using Deep Learning Techniques

Parimi Leela Sri Srujan

Student, Rajam, Vizianagaram, 532127, India.

ABSTRACT

In recent years, there has been a lot of research on facial detection and recognition. Identity verification, biometric applications, and numerous security cameras all rely heavily on facial recognition technology. A deep learning application is facial recognition (DL). Convolution neural networks (CNN) have developed into a reliable facial recognition technique. Due to the great degree of structural similarity between features, it is challenging to recognize comparable faces in the field of face recognition. To advance in this area, the next two elements are required: large scale identical face training datasets are readily available, and a fine-grained face recognition technique is also available. To recognize and analyse various images, a variety of algorithmic procedures like image eigenvalue extraction, recognition, and convolution are performed. Face identification in CNN uses a traditional feature-based classifier that combines features from the Haar cascade and Viola jones algorithms with the open CV library.A face expression identification approach based on a convolutional neural network (CNN) and an image edge detection method is proposed in order to eliminate the labor-intensive procedure of explicit feature extraction in traditional facial expression recognition. Finally, picture edge detection is used to recognize facial expressions. The image is normalized, and during the convolution process, the edge of each layer of the image is retrieved. To avoid the edge structure information of the texture image, the extracted edge information is placed on each feature image. The maximum pooling method is then used to process the dimensionally reduced retrieved implicit features.

Keywords:Analysis, Crowd management, Urban data, Algorithms, Validations

1. Introduction

A recurrent tendency and intrinsic ability to discriminate between various faces is human computer interaction. Computer vision difficulties were very difficult until recently, but the development of modern technology has significantly reduced the difficulty of issues with changing light, altered by age, hair, and other accessories. The aforementioned approach still has some issues with recognising identical faces, though. Applications for face recognition are used to make it easier to recognise and authenticate individuals based on their facial traits. Therefore, it is important to interpret facial features and their activities. As a result of these characteristics and expressions, human facial emotions can be categorised. Numerous techniques are needed to identify and categorise human faces, however deep learning technology excels them due to its massive dataset capabilities and quick computation speed. Typically, the procedures involved in face recognition and classification include preprocessing, detection, orientation, feature extraction, and emotion categorization. Although the CNN algorithm has improved in the area of recognising facial expressions, it still has certain drawbacks, such as a lengthy training period and a poor identification rate in complex backgrounds. A facial expression identification approach based on CNN and image edge detection is proposed in order to circumvent the challenging procedure of explicit feature extraction in conventional facial expression recognition. Deep learning can readily complete these jobs.

2. Literature Survey

In paper [1] The proposed method reduces the incompleteness brought on by artificial design characteristics and can automatically learn pattern features in comparison to conventional methods. The suggested method uses a training example image to directly input the image's pixel value. In this research, they suggest a method for CNN data-based face expression recognition. Autonomous learning can subtly pick up on the image's abstract feature expression. In complicated backdrop contexts, the suggested model's convergence speed is significantly faster. Additionally, the suggested technique increases recognition rates. The retrieved implicit features' dimension is decreased using the maximum pooling method,which can lower the convolutional neural network model's training time. The suggested algorithm's rate of recognition is 88.56, compared to 79.34 and 70.63 for R-CNN and FRR-CNN, respectively.

In paper [2]This Paper offers two new insights. The first one suggests a method for producing a sizable dataset of similar faces. We make use of the LFW and CASIA-Web Face data sets. The second step offers a model (IE-CNN) that improves both the internal and external aspects of the face, significantly enhancing the accuracy of face matching. The true positive rate for the recognition task of comparable face images can be effectively increased using the IE-CNN. Two branches make up the entire algorithm: the trunk branch and the IE-CNN branch. IE-CNN distinguishes the face features by separating the implicit and explicit features on the trunk branch. Finally, it compares the suggested model's accuracy to that of the v2 madel on the LFA and CISIA datasets.

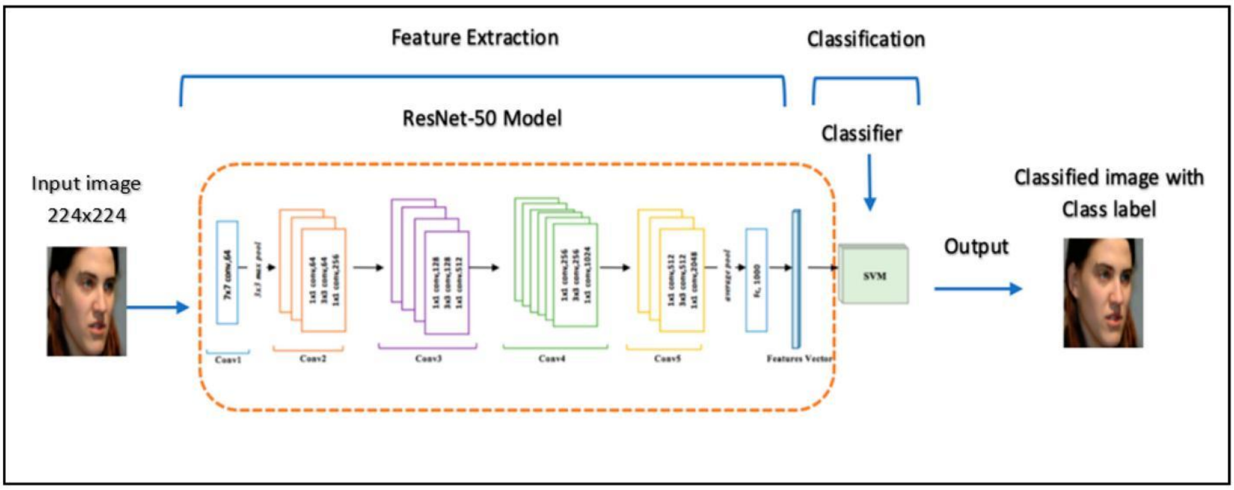
In paper [3] Two branches make up the entire algorithm: the trunk branch and the IE-CNN branch. IE-CNN distinguishes the face features by separating the implicit and explicit features on the trunk branch. Finally, it compares the suggested model's accuracy to that of the v2 madel on the LFA and CISIA datasets. A CNN model is created in this study to increase the precision of facial picture classification. The model's structure is comparable to that of the traditional LeNet-5 model, but several of its parameters—such as input data, network width, and entire connection layer—are different. Finally, they evaluated the model's accuracy using various numbers of training and test samples.

In paper [4] The weaknesses of conventional data augmentation techniques are examined and compiled in this work. The paper enhances Cycle GAN, suggests a technique of facial expression identification based on constraint cycle consistent generation to resist network, andintroduces class constraint condition and gradient penalty rule to address the issue of class imbalance in the existing face expression database. To begin with, this paper's expression identification and data augmentation are based on the static image can only capture a person's expression state at that certain moment in time, whereas emotional shifts in real life happen at specific times. The next project will concentrate on improving video sequence data. Second, although the expression state of a human in a real situation can be changed at will, neutral expression photographs are utilised as the source domain and other expression images as the target domain in the process of data enhancement.

In paper [5] This study suggests a face identification and recognition algorithm based on convolution neural networks (CNN) that performs better than the established methods. To verify the effectiveness of the suggested algorithm, an automatic attendance system has been expected in order to reduce human error that occurs in the traditional attendance taking system. To produce the cutting-edge findings, faster region convolution neural networks and edge computing techniques are used. Out of 35 detected faces, the system was able to identify 30 faces. By capturing clearer pictures of the kids, the accuracy that has been attained can be increased. Although the system is getting more accurate, its biggest drawback is distance because, by nature, as the distance grows, the image becomes blurry, leading in certain instances to erroneous findings for faces that are indistinct. Edge computing approaches have been used to increase data latency and reaction time between devices. The suggested approach is safe, dependable, and simple to apply. The proposed system can be used without any additional hardware or software.

3.Methodology

For human detection, a Haar classifier is used in this study. The Haar classifier is trained using the AdaBoost algorithm in conjunction with tiny characteristics that are similar to Haar and an integral graph technique. A popular texture descriptor is "Haar-like," the linear, edge, centre, and diagonal are its primary characteristics. The main goal of AdaBoost, a Boosting algorithm upgrade, is to create a strong classifier by repeatedly iterating weak classifiers. In the development of face detection, the Viola-Jones detector represents a turning point. Due to its rapid detection and excellent efficiency, it has been frequently used. This technique employs an integral graph to quickly calculate Haar-like features while also filtering out key features from a huge set of Haar-like features. It uses the Haar-like to extract facial features. The weak classifier is then trained and integrated into a strong classifier using the AdaBoost algorithm. Finally, to increase the precision of face detection, many powerful classifiers are cascaded in sequence. It is simple for illumination, shadows, and other elements to alter the actual picture collecting process, resulting in a state of uneven light and shade distribution in the gathered image, which makes feature extraction more challenging. In order to improve the contrast of the image, it is important to average the grey level. The Histogram Equalization (HE) approach is utilised to process photos in this paper.The place where the gradient information of an image varies dramatically is frequently where the edge information of the image is reflected. People have a greater visual perception at the image's edge. Therefore, throughout the texture generation process, it is impossible to ignore the image's edge information. A portion of the image's edge information is lost, which causes blurred edge information in the final synthesis result and has an impact on the table. This study preserves the edge structure information of the texture image by first extracting the edge of each layer of the image during the convolution process, and then superimposing the retrieved edge information on each feature map



CNN-BASED MODEL

The core of the deep learning approach is to build a deep neural network with characteristics of the human brain that can learn more complex feature expressions of data layer by layer via multiple-hidden non-linear structures. The collected features are given a more thorough characterization of the data thanks to this approach for automatically learning the internal laws of massive datasets, considerably improving the classification outcomes. The neural network model can classify a two-dimensional picture input directly within the model to produce recognition results after layer-by-layer interpreting it from the pixels initially recognised by the computer to edges, portions, outlines of things, and objects known by the human brain.

Three crucial properties of CNN include local perception, weight sharing, and down sampling. Due to these traits, the typical feedforward neural network's issue with too many parameters and complex calculation in high-dimensional input is resolved, and the model is improved.attain specific distortion, rotation, and translation invariance.Common sense dictates that people typically perceive the outer world on a local to global scale. There is a certain spatial link in the image. While the distant pixels exhibit little connectivity, the nearby pixels are highly connected. In order to obtain the global information at the top level, neurons just need to receive the local pixels and then integrate the local information at the bottom. This idea of local perception significantly lowers the number of parameters required for learning.

The weight sharing principle states that if the statistical characteristics of one aspect of an image are similar toTherefore, features may be extracted from the image at all points using the same convolution kernel. However,To learn the features, using a single convolution kernel is insufficient. In order to increase the variety of feature mapping, several convolution kernels are utilised during the actual training of convolution neural networks. Every type of convolution may obtain the mapping plane for many picture characteristics. Weight sharing allows for the acquisition of large amounts of visual data while also drastically reducing the number of parameters required for network training. As long as the network is reasonably controlled Convolutional neural networks' capacity for generalisation can be improved by adding structure. Convolution-based feature extraction can be used directly to train classifiers, although it still presents significant computing problems. After the convolution procedure, a down sampling step is suggested in order to further minimise the parameters.

POOLING LAYER

The dimension reduction is the primary goal of the pooling procedure. The dimension of the subsequent feature map can be cut in half by using a pooling window with a 2 2 step size. Although the number of training parameters is not directly decreased, half the dimension of the feature graph will result in a significant decrease in the computational complexity of the convolution operation, which will significantly speed up training.

CONNECTION LAYER, FULL

While the output of the prior pooling layer S2 is a two-dimensional array, the input of the full connection layer must be a one-dimensional array. Prior to connecting 128 one-dimensional arrays in succession to a feature vector with 51200 dimensions (20 20 128 = 51200), the two-dimensional arrays corresponding to each feature graph are first converted into one-dimensional arrays.

Experimental comparisons between the suggested approach and the R-CNN model are made to confirm the performance of the proposed algorithm. The identical experimental setting and experimental data were employed in the experiment. The association between iterations and Accuracy of training sets of the two models is discovered through simulated studies In complicated backdrop situations, the model runs much more quickly. The suggested approach also achieves a greater recognition rate.

7. Results and Discussion

METHOD	TRAINING TIME(S)	TEST TIME(S)	RECOGNITION RATE(%)
THE PROPOSED ALGORITHM	178	24.89	88.56
R-CNN ALGORITHM	256	33.97	79.34
FRR-CNN ALGORITHM	148	17.92	70.63

According to Table the suggested algorithm takes less time to train and test than the R-CNN algorithm and the FRR-CNN method. Conclusion: The collected implicit features' dimension is decreased using the maximum pooling method, which can lower the convolutional neural network model's training time. Additionally, the suggested method gets an average recognition rate of 88.56%, which is the highest.

8.CONCLUSION

We suggest a technique for recognising facial expressions that makes use of a CNN model that successfully extracts face features. The proposed method reduces the incompleteness brought on by artificial design characteristics and can automatically learn pattern features in comparison to conventional methods. The suggested method uses training sample picture data to directly input the image pixel value. Autonomous learning can subtly pick up on the image's abstract feature expression. The initialization of weights is done correctly during the training phase of the suggested method, which has a significant impact on the updating of weights. Our thorough experimental analysis demonstrates that, as compared to the prior literature, the suggested method can somewhat increase the identification rate of facial expressions in complicated backgrounds. The suggested model's rate of convergence is faster than that of the FRR-CNN and R-CNN models.

REFERENCES

1. 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.

2. A. Chen, H. Xing and F. Wang, "A Facial Expression Recognition Method Using Deep Convolutional Neural Networks Based on Edge Computing," in IEEE Access, vol. 8, pp. 49741- 49751, 2020.

3. A. Song, Q. Hu, X. Ding, X. Di and Z. Song, "Similar Face Recognition Using the IE-CNN Model," in IEEE Access, vol. 8, pp. 45244-45253, 2020.

4. Peng Lu, Baoye Song & Lin Xu Human face recognition based on convolutional neural network and augmented dataset, Systems Science & Control Engineering, 9:sup2, 29-37,2021.

5. A. Mallol-Ragolta, S. Liu and B. W. Schuller, "The Filtering Effect of Face Masks in their Detection from Speech," 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, pp. 2079-2082, doi: 10.1109/EMBC46164.2021.9630634.

6. A. -P. Song, Q. Hu, X. -H. Ding, X. -Y. Di and Z. -H. Song, "Similar Face Recognition Using the IE-CNN Model," in IEEE Access, vol. 8, pp. 45244-45253, 2020, doi: 10.1109/ACCESS.2020.2978938.

7. H. Chen, Y. Chen, X. Tian and R. Jiang, "A Cascade Face Spoofing Detector Based on Face Anti-Spoofing R-CNN and Improved Retinex LBP," in IEEE Access, vol. 7, pp. 170116-170133, 2019, doi: 10.1109/ACCESS.2019.2955383.

8. B. Jin, L. Cruz and N. Gonçalves, "Deep Facial Diagnosis: Deep Transfer Learning From Face Recognition to Facial Diagnosis," in IEEE Access, vol. 8, pp. 123649-123661, 2020, doi: 10.1109/ACCESS.2020.3005687.

9. M. Z. Khan, S. Harous, S. U. Hassan, M. U. Ghani Khan, R. Iqbal and S. Mumtaz, "Deep Unified Model For Face Recognition Based on Convolution Neural Network and Edge Computing," in IEEE Access, vol. 7, pp. 72622-72633, 2019, doi: 10.1109/ACCESS.2019.2918275.

10. B. Kocacinar, B. Tas, F. P. Akbulut, C. Catal and D. Mishra, "A Real-Time CNN-Based Lightweight Mobile Masked Face Recognition System," in IEEE Access, vol. 10, pp. 63496-63507, 2022, doi: 10.1109/ACCESS.2022.3182055.