

How to Cite:

Adithya, A., Anisa, H., Monika, J. S., Thomas, T. S., & Kumar, A. (2022). Facial expression detection and classification using SVM, CNN and decision tree algorithm. *International Journal of Health Sciences*, 6(S4), 3954–3962. <https://doi.org/10.53730/ijhs.v6nS4.9770>

Facial expression detection and classification using SVM, CNN and decision tree algorithm

Adithya A

Department of Information Technology, KCG College of Technology, Chennai, India
Corresponding author email: 18it01@kcgcollege.com

Anisa H

Department of Information Technology, KCG College of Technology, Chennai, India
Email: 18it03@kcgcollege.com

Monika J S

Department of Information Technology, KCG College of Technology, Chennai, India
Email: 18it26@kcgcollege.com

Ms. Tina Susan Thomas

Department of Information Technology, KCG College of Technology, Chennai, India
Email: tinaparayil@gmail.com

Mr. Ashwin Kumar

Senior Security, Consultant, Genpact, India
Email: aswinvhss@gmail.com

Abstract---The human face is often used as a visual representation of information, which is why facial expression recognition is very important in terms of human-machine interaction. It can be used for various applications such as detecting mental disorders and understanding human behavior. Despite the advantages of facial expression recognition technology, the high recognition rate to be achieved by a computer is still challenging. Two commonly used methods are geometry and appearance. Machine learning methods like CNN, Decision tree and SVM were applied to identify the human emotions like happiness, fear, disgust, anger, surprise, sadness and neutrality.

Keywords---facial expression, machine learning, human emotions, convolutional neural networks, support vector machine, decision tree.

Introduction

Our desires have risen higher along with the growth of modern technology and we are not constrained by limits. There is a great deal of research being done at present in the field of recognising digital emotion and digital image. The process of moving forward has become exponent and is constantly developing. Image Processing being one of the largest research facilities across the world, its usage is spreading widely. Image processing comes under signal processing field where the signals are input and images are output. The major concerned area in image processing is facial expression recognition. Our face depicts the emotions, Hence facial expression is crucial in human communication. Facial expressions are considered as non-verbal act of science. In AI and Robotics, facial emotion plays a vital role. So it is a key factor for upcoming generation. Some areas of FER includes Videophone, Computer and people interaction, Artificial Monitoring, Cosmotology, Teleconferencing and applications like forensic and personal identification.

The purpose of this work is to build an Facial Emotion Recognition System that can captures frame from video of human's face that takes input as specific expressions and detect and separate seven different emotion categories such as: I. Happy II. Sadness III. Anger IV. Surprise V. Neutral VI. Disgusting VII. Fear The goal is not only to get the system of Automatic Facial Emotion Recognition better but also to intensify the precision of the system when assessed with other existing systems.

Related Work

According to Adrian Rosebrock et.al (2017), various image descriptors have been used by facial emotion recognition which is relied on the characteristics known as Local Binary Patterns (LBPs). In many disparities of LBP, none of those variations can be automatically familiarise to the training data model till now. A new simplification of LBP that can learn the most discriminative features similar to LBP for each region in face in a overseen manner. Trials on general datasets that are used for face recognition shows the dominance of DT-LBP where several state characteristics in face are used extensively in face recognition applications. According to Afshan Jamil et.al (2019), an approach to hybrid RNN-CNN is applied for modelling the spatiotemporal information of human facial emotion. Different modals are performed and combined at feature level and decision level to achieve better accuracies. According to Alex Christian Llano Casa, Nora La Serna Palomino et.al (2020), a wide-ranging analysis of the work on facial recognition techniques permits to the reduction of the picture dimensionality, and Support Vector Machines, as classification of pattern. The class is calculated where the element belongs from the values of the data. The databases that have been used are reviewed and the results obtained are also reviewed and the important research movements are identified. A median filtering has been implemented during optimization of intermediate flow fields which improves the robustness of classical models. (Deqing Sun et.al, 2016), it actually precedes to improved higher results. Using filtering the median to remove the noise at every other step is the key to accuracy improvement, and this increases the vitality of the final result. The connections in between L1 based denoising and

median filtering, the algorithms based on the median filtering are overall enhancing a changed principle that generalize the flow field over a huge spatial neighborhood. According to Feryal Cubukcu et.al (2016), a method that involves two versions for attaining high prediction accuracy with restricted samples is taken forward. Local and global features of images which consists of facial emotion were obtained using Haar Cascade and Gabor wavelets. Using Nonlinear principal component analysis (NLPCA), reduction is done for dimensionalities of extracted features. For combining the local and global features and biased fusion methods have been approached.

The SVM system with differentiators are implemented to identify the 6 general facial emotions or FAUs – (Facial Action Units) sets are used. (Irene Kotsia et.al, 2016), the information which are statistical has been integrated by a unique class for SVMs of the classes is also proposed. The Candide grid nodes are initialized by the user on the facial image which will be displayed at the first frame of the image sequence. On the basis of deformable models, the tracing system has been used that tracks the facial expression as differences will be found over time, by changing the Candide grid, on the other hand producing the grid. According to Li Kuan et.al (2017), the use of histogram pyramid of oriented gradients and appearance features (LPQ) is done for detection of expressions in human face. The negligence of using edges with related feature points and appearance are first traced by a constrained tracker based on local model. The result of face parts are gathered and implemented for selection of key frames where expression feature is being taken. The final achievement of facial recognition is done using a face detector rather than using the facial pixel points by tracing. At last, expressions are identified using Support Vector Machines and highest margin neighbor classifiers which is nearest. According to Mohammed Saaiddia, Narima Naili Zermi et.al (2019), the Neural network method is used in to perform facial emotion recognition. The faces which are detected were treated to execute the characterization part that will be processed through vectors of Zernike moments. At final stage, a back propagation neural network will be processed to find difference between the states of emotion.

A thorough examination of the information transfer of knowledge among the student scholars has been taken (Isaac A.C. et.al, 2020), this study by the student researchers, evaluates different boundaries such as motivation for research, knowledge transfer in various fields and research department. The work has mainly attempted to go through the various aspects of the information transferring model among the students with the aid of software called Pajek and programming language R. The key interest is to isolate and group competencies at a separate step and to make changing synergetic effects. The formation of information sharing system for the researchers have been done and the network features that are centralised among the departments are analysed which focuses on whom to appease and whom to circumvent. According to Baral R et.al (2021), the study focuses on three differing variants of the process in Knowledge hiding. First, the external and internal elements from which the hidden process is initiated which is relied. Then, the key implementors of knowledge hiding is brought out. Finally, the conceptual framework is brought out that underscores the vital importance of neural science. And at last, this study determines the distinctive

need for memory and frontal lobe reliability in moderating the events as knowledge hiding.

Algorithm

We will be using the 2013 dataset fer publicly available on Kaggle. contains 48 * 48 pixels of grayscale face images and their emotional labels. Here this dataset contains 7 feelings and labels are given for it.

1 (anger), 2 (panic), 3 (Joyful), 4 (unhappy), 5 (astonished), 6 (Neutral), 7 (Disgust).

Start by importing pandas and other important libraries and then uploading the database.

This database carries 3 columns one is pixels other is emotions, and usage. In emotion column has the full number of coded emotions, the pixel column contains pixels which is in the format of a series separated by spaces, and the usage states that the data is used whether for training or testing. Image enhancements used for performance and capability of model. So using a data development is a good practice before transferring it to the model, which can be done using ImageData Generator provided by the cameras.

Rescale: Makes the pixel value normal when it divides by 255.

Horizontal flip: This horizontally flips the image.

Fill mode: Fills an image when the image is not available after a certain crop.

Rotation range: It rotates the image by 0 to 90 degrees.

Here image is first captured from video then this image is sent for segmentation. it includes the division of the video into segments to analyse the image in a simple way. A segment represents an object or part of an object and it contains a set of pixels. segmentation is performed with the help of top-down approach which divides the system into subsystems. Then we extract the feature of the segmented image. to convert the gradient value to gray scale an averaging method is used. The formula for the averaging method is,

$R+G+B$ divided by 3

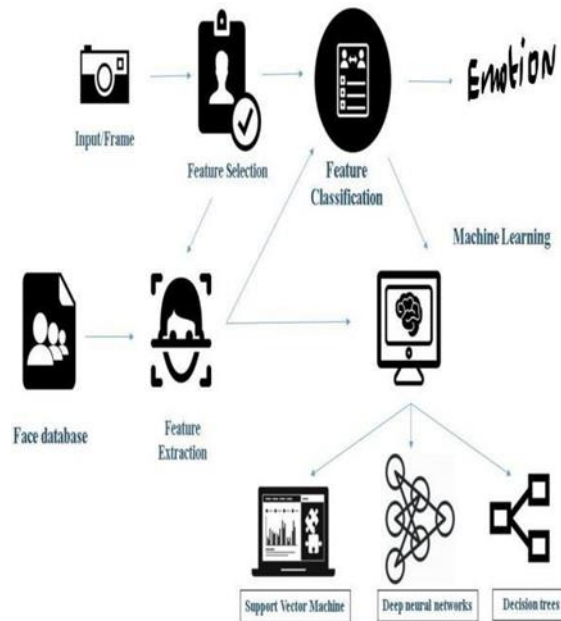
R for red, G for green, and B for blue components of pixel.

Now the extracted image passed to svm classifier. The image is trained many times, once this is done compare the trained image with the image you want to test. The output is displayed in the screen as a label and voice.

Architecture Diagram

SVM has become popular when we use pixel mapping as it provides accuracy in comparison to other sophisticated used algorithms. In SVM a kernel function is applied to every instance of data to map the original non linear data point to the higher dimensional space where it can be linearly separated. By using a kernel

function , A dot product is obtained between the two products such that every points fits into a high dimensional space. There are mainly four types of kernel function:



1. linear
2. polynomial
3. radial basic function
4. sigmoid.

We are going to go with radial basic kernel function. The expression for radial basic function is.

$$e^{-\gamma \|u-v\|^2}$$

where y is the kernel argument of the radial basic function and u is the test vector and v is the vector support.

Convolutional Neural Network

Architecture Diagram Support Vector Machine

Support vector machine is a prediction tool for regression and classification which maximizes the predictive accuracy by automatically avoiding overfit of the data. SVM can also be used as a system which defines the premise space into a linear function in a high-dimensional characteristic space. In the initial stage, this algorithm is applied to the neural information processing systems and at present it is a crucial part of machine learning research used by many developers around the world. When it comes to neural networks, a convolutional neural network is the most important class used in snapshot identification and in a snapshot training and object identification. An object detector recognizes objects and faces. CNN was inspired because of the connectivity pattern in biological process among the other neural resembles.

This system depicts an input image as pixel point array and this will rely upon photo decision, the resolution of picture will identify height*width*size. Convolutional Neural Network is made up of three layers, Convolutional, pooling, and the last one is completely linked layer. The CNN consists of three layers: Convolution layer, Pooling Layer, and completely linked layer.

Convolutional layer

It is a first layer in CNN which extracts the feature from the given input image. It conserves the relationship between the pixels by comprehending the feature of an image by accessing small-scale input squares. This is a mathematical equation where 2 inputs are passed, first is image matrix and then filter or a kernel. Dimensions are followed as, for image matrix, height*width*size and for filter,

Filter of (h)*filter of (w) *filter of (d).

The outcome of this operation would be $(hfh+1)*(wfw+1)*(dfd+1)$.

The above output is called object mapping. Output $\Rightarrow$ filter of (x) = max (0, x)

Pooling Layer

The implementation of pooling layer is needed where the frames are in big size and difficult to compress the parameters. Pooling is also defined as down-sampling or sub-sampling that helps to reduce the dimension every map however it facilitates significant information. Different types of Spatial pooling,

- Average pooling
- Max pooling
- Sum pooling

It considers the huge element from the corrected feature map. By choosing the huge element, it can provide average pooling. The sum pooling is defined as the summation of elements in the feature map.

Fully Connected Layer

The compression of matrix into vector is done and this will be fed into a connected layer with a neural network. This will act as a human neuron, which will interconnect layers with one another for transmitting the information. Every feature map from the pooling section is interconnected for providing the trained output. Then this trained image is used to compare with the new image. After the comparison, the matching expression is recognized.

In this way, you can get the expression you want and you can find 7 universal expressions. 1 (anger), 2 (panic), 3 (Joyful), 4 (unhappy), 5 (astonished), 6 (Neutral), 7 (Disgust). It is used to detect multiple faces simultaneously on one screen. This applies to both linear and non-linear separable data. For the high dimensional data spaces and the area where computation becomes complex, SVM and CNN can be applied.

Conclusion

If the predicted emotion, by the model is incorrect, the appropriate label will be the second option. The physical characteristics in a human face such as fear, happiness, anger, sadness, disgust, neutral and surprise are linked with geometrical structures which is being put back as bottom matching for the emotion recognition system. A robust face recognition study based on the behavioural characteristics which are mapped with physiological biometric characteristics has been contributed in this facial emotion recognition system is projected in this work. The property base to the system is related to the characteristic backing different expressions as the property base. These are differentiated as hidden and exposed category in genetic algorithm. The uniqueness of emotions in individual faces is evaluated by the training model and provide a related expression recognition model.

Future work

Due to the lack of predictability in highly composed system, we were not able to proceed to the dense neural network which is deeper as the routine lags and we will work on these areas in future. We like to process larger datasets in the system to build the accuracy of model but the sources is being a barrier within the way and we must develop in various sectors for errors to be resolved and get improved accuracy. The studied methods are dealt with emotion variation; in upcoming period, it is going to be examined with seriousness with additional focus related to the problem with face classification and ideal merge of depth with colour information. Additional study will be carried by genetic factor matching to the symmetrical features of the facial emotions. The evolution framework of facial expressional system which comprises of genetic properties will also be studied to enhance the requirement of different models of security like security breaches, government confidential, criminal detection, etc.

References

1. D. Sun, S. Roth, and M. J. Black, 2014, "A quantitative analysis of current practices in optical flow estimation and the principles behind them," in Springer International Journal of Computer Vision, vol. 106, no. 2, pp. 115–137.
2. Gory Seth, Al-Khassaweneh Mahmood, Szezurek Piotr, 2020, "Machine Learning Approach for Facial Expression Recognition", in IEEE International Conference on Electro Information Technology (EIT), pp 032-039.
3. Hamprecht Fred A Schnorr Christoph, Jahne Bernd, 2007, "Pattern Recognition", Springer, Chapter 22, pp 214-223.
4. Hu Min, Yang Chunjian, Zheng Yaqin, Wang Xiaohua, He Lei, Ren Fuji, 2019, "Facial Expression Recognition Based on Fusion Features of Center Symmetric Local Signal Magnitude Pattern", IEEE, pp 118435-118445.
5. Issac A.C., Thomas, T.S., 2020, "Whom to appease and whom to circumvent: analyzing Knowledge sharing with social networks", in Scopus, Vol 69, pp 75-93.
6. Issac A.C., Issac T.G., Baral R, Bednall T.C., Thomas T.S., 2021,

7. "Why you hide what you know: Neuroscience behind Knowledge hiding" in Scopus, Vol 28, pp 266-276.
8. Kim Ji-Hae, Kima Byung-Gyu, Roy Partha Pratim, Jeong Da-Mi, 2019, "Efficient Facial Expression Recognition Algorithm Based on Hierarchical Deep Neural Network Structure", in IEEE, vol 4 pp 2169-3536
9. Kotsia Irene, Pitas Ioannis, 2007, "Facial Expression Recognition in Image Sequences using Geometric Deformation Features and Support Vector Machines", IEEE Transactions on Image Processing, vol 16, pp-172-187
10. Li Kuan, Jin Yi, Akram Muhammad Waqar
11. Han Ruize, Chen Jiongwei, 2019, "Facial expression recognition with convolutional neural networks via a new face cropping and rotation strategy", in Springer.
12. Liu Yong-Jin, Zhang Jin-Kai, Yan WenJing, Wang Su-Jing, Zhao Guoying Fu Xialon, 2015, "A Main Directional Mean Optical Flow Feature for Spontaneous Micro-Expression Recognition", IEEE Transactions on Affective Computing, pp 1949-3045.
13. Liu Yunfan, Hou Xueshi, Chen Jiansheng, Yang Chang, Su Guangda, Dou Weibei
14. , 2014, "Facial expression recognition and generation using sparse autoencoder", in IEEE International Conference on Smart Computing, pp 125-130.
15. Liu Yanpeng, Cao Yuwen, Li Yibin, Liu Ming, Song Rui, Wang Yafang, Xu Zhigang, Ma Xin, 2016, "Facial expression recognition with PCA and LBP features extracting from active facial patches" IEEE International Conference on Real-time Computing and Robotics, pp 368-373.
16. Saaidia Mohammed, Zermi Narima, Ramdani Messaoud, 2014, "Facial Expression Recognition Using Neural Network Trained with Zernike Moments", in IEEE International Conference on Artificial Intelligence with Applications in Engineering and Technology, pp 187-192.
17. Siddiqi Muhammad Hameed, Ali Rahman, Khan Adil Mehmood, Park Young-Tack, Lee Sungyoung, 2015, "Human Facial Expression Recognition Using Stepwise Linear Discriminant analysis and Hidden Conditional Random Fields", IEEE Transactions on Image Processing, Vol 24(4), pp 1386-1398.
18. Valstar M F, Mehu M, Bihan Jiang, Pantic M, Scherer K, 2012, "MetaAnalysis of the First Facial Expression Recognition Challenge", in IEEE, vol 42(4).
19. Zhang Z, Li M, 2020, "Research on Facial Expression Recognition Based on Neural Network", in IEEE International Conference on Computer Network, Electronic and Automation (ICCNEA)
20. Zhang T, 2017, "Facial Expression Recognition Based on Deep Learning: A Survey", In Springer advances in Intelligent Systems and Interactive Applications, vol 686, pp 345-352.
21. Suryasa, I. W., Rodríguez-Gámez, M., & Koldoris, T. (2021). The COVID-19 pandemic. *International Journal of Health Sciences*, 5(2), vi-ix. <https://doi.org/10.53730/ijhs.v5n2.2937>
22. Suryasa, I. W., Rodríguez-Gámez, M., & Koldoris, T. (2022). Post-pandemic health and its sustainability: Educational situation. *International Journal of Health Sciences*, 6(1), i-v. <https://doi.org/10.53730/ijhs.v6n1.5949>
23. Kusumawati, A. H., Wulan, I. R., & Ridwanuloh, D. (2020). Formulation and physical evaluation sheet mask from red rice (*Oryza Nivara*) and virgin

coconut oil (Cocos Nucifera L). *International Journal of Health & Medical Sciences*, 3(1), 60-64. <https://doi.org/10.31295/ijhms.v3n1.148>