DeepQ Residue Analysis of Computer Vision Dataset using Support Vector Machine

Rahama Salman[1*], Arshia Arjumand Banu[2]

[1*]Lecturer, Department of Information Technology and Security, College of Computer Science and Information Technology Jazan University, Jazan, Kingdom of Saudi Arabia.
rabdol@jazanu.edu.sa, ORCID: https://orcid.org/0000-0003-4105-5019

[2]Lecturer, Department of Computer Science, College of Computer Science and Information Technology, Jazan University, Jazan, Kingdom of Saudi Arabia. abanu@jazanu.edu.sa, ORCID: https://orcid.org/0009-0009-7000-3809

Received: October 30, 2022; Accepted: December 22, 2022; Published: February 28, 2023

Abstract

A vision-based human computer interface is used to automatically recognize human mood. Image processing techniques used include a web camera for eye detection. Appearance tracking method (ABT) is measured face identification and K means Nearest Neighbor (K-NN) is used for eye detection. DWT - Discrete Wavelet Transform and DCT - Discrete Cosine Transform are suitable to extract features of eye and SVM is used to classify eye expressions. Classification of eye expressions includes anger, fear, happiness, disgust, neutral and sad. Experimental results confirm that the proposed method recognized facial expressions with higher accuracy.

Keywords: Face Detection; KNN Classifier, Computer Vision Dataset, Support Vector Machine, Deep Learning, Residue Analysis.

1 Introduction

In recent years, the development of the human-computer interface has attracted the attention of researchers worldwide (S. Manikandan, 2022). Realization of human emotions is one of the important parts of communication between the user and the system. Human interacted facial movements play an important role in trying to understand human emotion values and behaviors for human computer interaction (HCI) (S. Manikandan, 2022) (Jharnama jumdar, 2018). This paper proposes a system for automatic human mood recognition. When users use the system for any purpose, their facial expressions are recognized and identified by this proposed system. Here, the webcam acts as an interface between the user and the system. Section II discuss about related works. Section III discusses the paper's proposed methodology. Section IV describes the experimental results of the work. Section V presents conclusions and discusses improvements that can be made in future work.

Psychological evidence has shown the eyelids and eyebrows values are important for recognizing suitable facial expressions and human emotions. It was achieved an average recognition of 97% ± 1%, confidence of 96% ± 2% for clarity between 20% and 100% (Javier Orozco, 2019). Facial expressions

Journal of Internet Services and Information Security (JISIS), volume: 13, number: 1 (February), pp. 78-84.
DOI: *10.58346/JISIS.2023.I1.008*

*Corresponding author: Lecturer, Department of Information Technology and Security, College of Computer Science and Information Technology Jazan University, Jazan, Kingdom of Saudi Arabia.

is used to selected non-verbal communication channels for interacting with other features. Computer assisted recognition of human emotions using face expression is an interesting feature. Cubic Bezier curves were used to define emotions.

There are seven different emotions there and the recognition success rates range between 95% to 65% (Rembiyekandemir, 2017). Simina Emmerich describes a multilingual emotion recognition system using face expressions and input signals using a support vector and k means nearest neighbor (Simina Emerich, 2019). A system of medical diagnosis was implemented to identify specific mental processes by facial expression (Ryan, A., 2019) (P. Ekman, 2015). Michael describes a robust real-time face detecting application framework that is capable of handling fast image processing with high detection rates (Michael J. jones, 2014). A simultaneous eye tracking and blink detection system uses two illumination-invariant motion detection algorithms to detect moving objects (Angelinjayanthi A, 2012).

2 KNN Classifier – Proposed Approach

In this proposed work, the face region is identified using a tracking method. Appear Based Tracker (ABT) is a reliable and accurate method to select and extract the facial movements from image sequences. The human face is detected using a well-trained ABT (Appearance Based Tracker) classifier (I.K. Zafeiriou, 2018) (S. Zhou, 2014). This classifier uses features similar to ABT, which are defined as the ratio of intensities taken from a neighboring rectangle. The face detector is trained on several highlights. Thus, the selected shape and texture form an active appearance model that can detect the faces of people standing in front. The K-nearest-neighbor (KNN) algorithm (Yang Song, 2007) (M.I. Ahamed, 2021) (Irshad Ahamed, 2019) (Irshad Ahamed, 2022) (Ahamed, M.I., 2019) (P.J.S. Babu, 2021) is used to detect eyes and eyebrows in the face region in a video sequence. It selects the distance between a query region and a set of regions in the dataset. The distance between two regions is calculated using the Euclidean distance function as shown below:

$$d(x, y) = \sum_{i=1}^{N} \sqrt{{x_i}^2 - {y_i}^2} \qquad (1)$$

The KNN algorithm used in the proposed work is given as follows:

Output values stored in nearest neighbors' function from optimized query region

r={r^1,…,r^m} repeat loop M times,

Go:s^i the data set,

From i iteration index from domain {1,..,P}

If

q <d(q, s^i) : q ← d(q, s^i), t → o^i

Loop i = P when stop i=0

q to c and t r assign the values

Arithmetic mean of r as follows:

$$\bar{r} = \frac{1}{M} \sum_{i=1}^{M} r_i \qquad (2)$$

$\bar{r}$output value from q.

DWT is used to select the signal in different frequency bands with different resolution by detailed information scanner (M.I. Ahamed, 2021) (Irshad Ahamed, 2019). Each eye image is passed through a filter containing four parts such as CA: coefficient of approximate values, CH: horizontal values, CV:

vertical values and CD: diagonal values. The value approximation (CA) coefficients of the eye images at each of the decomposed levels are given better values compared to other coefficients.

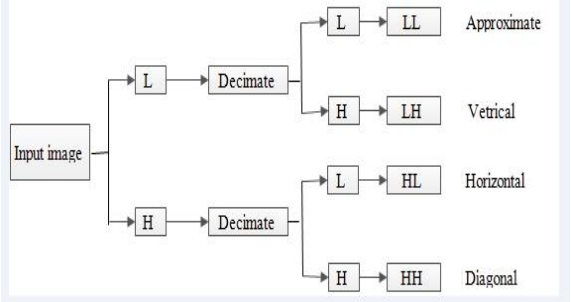

Figure 1: DeepQ Residue Analysis Selection of Dataset

The input of eye image in the proposed work size is 45 X 110. Eye image after filtering is reduced to 26 X 58 at the first level, 16 X 32 at the second level, 11 X 19 at the third level. , 9 X 13 on the fourth level, 8 X 10 on the fifth level and 7 X 8 on the sixth level. A set of 56 (7 X 8) coefficients of DWT is extracted from each eye image for classification training.

Discrete Cosine Transform Feature Extraction

DCT is suitable for feature extraction because it achieves greater energy compaction. After applying DCT to the eye image, some of the coefficients with the highest energy are selected in a zigzag fashion to construct a feature vector. Mathematically, 1D and 2D-DCT images of size N x M are defined by Equations (3) and (4), respectively. The equation for one dimensional DCT is

$$F(u) = \sqrt{\frac{2}{N}} \sum_{i=0}^{N-1} A(i) * \cos\left(\frac{u(2i+1)\pi}{2N}\right) * f(i) \quad (3)$$

Where,

$$A(i) = \begin{cases} \frac{1}{\sqrt{2}} & for\ u = 0 \\ 1 & otherwise \end{cases}$$

f (i) - input sequence. The equation for two dimensional DCT is

$$F(u,v) = \sqrt{\frac{2}{N}}\sqrt{\frac{2}{M}} \sum_{i=0}^{N-1} A(i) * cos\left(\frac{u(2i+1)\pi}{2N}\right) * \sum_{j=0}^{M-1} A(j)\, cos\left(\frac{v(2j+1)\pi}{2M}\right) * f(i,j) \quad (4)$$

So

$$A(i) = \begin{cases} \frac{1}{\sqrt{2}} & for\ u = 0 \\ 1 & otherwise \end{cases}$$

$$A(j) = \begin{cases} \frac{1}{\sqrt{2}} & for\ v = 0 \\ 1 & otherwise \end{cases}$$

if (i, j) is the 2D input sequence. The highest energy 30 DCT- co-efficient are extracted in the zigzag manner from the top left corner of the 2D image.

Support Vector Machine

SVM model is used to estimate the choice perform exploitation nonlinear category boundaries supported support vector values. If the info is partitioned off linearly, the linear SVM machine finds the best hyperplane partitions with less error rate and at the utmost distance between the hyperplane and therefore

the nearest points. For linearly non-separable data, it maps the input pattern house X to a four-dimensional feature space Z employing a non-linear function □(x). The SVM then finds the optimal hyperplane as a choice surface to separate the samples of the 2 categories within the feature space. associate degree example of mapping two-dimensional data into three-dimensional house exploitation the perform

F(x) = is shown in Figure a pair of input values and estimated output

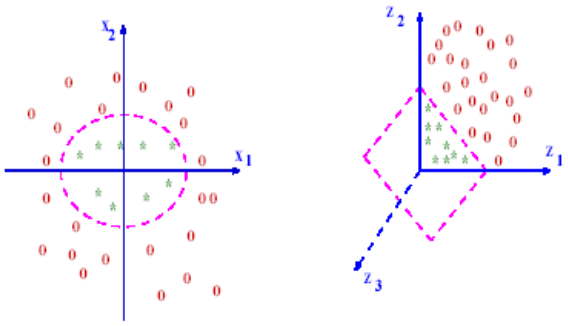

Figure 2: Mapping of Two-Dimensional Data into Three-Dimensional Data

SVM is trained with 2 sorts of modes similar to binary mode and multimode. Binary mode is employed for two-class categorification to differentiate objects with class labels "0" and "1" within the coaching set. Multi model function is used for selecting multiclass classification to distinguish objects with labels as {0, 1, 2} and so forth this text acknowledges 6 types of expressions such as anger, disgust, fear, happiness, neutrality and sadness. A multi-class SVM is shown below.

0 if anger
1 if disgust
1 = 2 if fear
3 if happiness
4 if neutral
5 if sadness

For testing, the DWT and DCT feature are extracted from regions and gives as input to the SVM model and The distance is calculated as every of the feature vectors. SVM hyperplane is obtained from expression index. The facial expressions are recognized as anger, disgust, fear, happiness, neutral and unhappiness {if the scores are 0, 1, 2, 3, 4 and 5 respectively}

3 Experimental Setup

As will be seen in fig. three the camera is put in as a computer program between the system and also the user. this text focuses on pc vision pursuit and eye expression recognition. Recognizing a person' mood is one among the most tasks of e-commerce. client sentiment is extremely necessary for business those that run e-business online. Entrepreneurs or suppliers should effectively attract customers by marketing their product supported the wants and satisfaction of users. whereas the purchasers are browsing the products being sold, the marketing mode ought to be modified in keeping with the mood of the customer. If customers are unhappy or sick of or Associate in Nursinggry, the system ought to place alternative promoting methods.

The performance of recognition system is obtained for ten samples. For the training, an AVI file of three hundred frames as 50 for every of the six cuts is recorded at one hundred sixty × a hundred and twenty employing a Logitech Quick cam Pro5000. Counter fact is recognized for all frames within the check video. The SVM is trained with 56-dimensional DWT feature vectors and 30-dimensional DCT feature vectors, and its performance is analyzed and compared. A coaching file with 3000 feature vectors (300 per subject) is given to coach the SVM. AVI files containing the subject' countenance are used for testing. The ensuing screenshots are shown in fig. 4. Average performance is measured victimization SVM with polynomial, mathematician basis perform, and colon kernel functions. the tactic provides optimum performance with a Gaussian basis function kernel as shown in Tables one and 2.

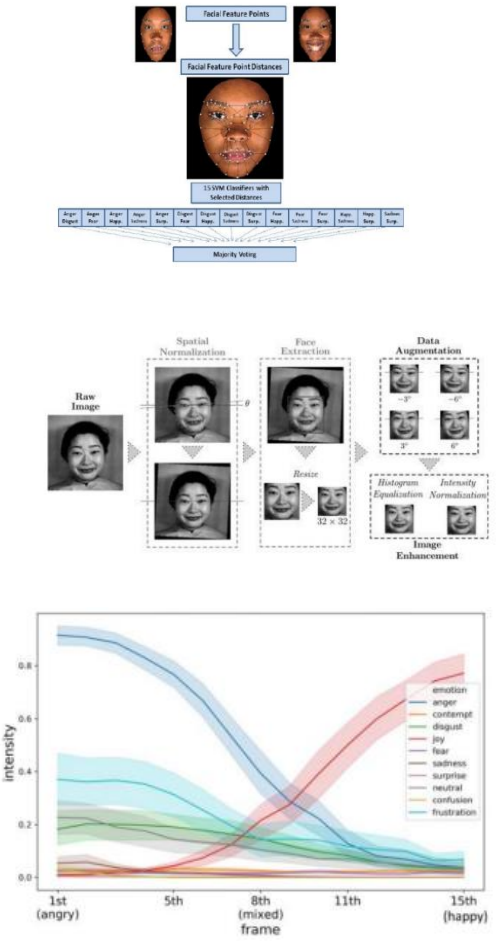

Figure 3: Computer Vision Dataset Recognition System –Matlab

Table 1 and 2: Comparison Features

Eye Expressions	Anger	Disgust	Fear	Happy	Neutral	Sad
Anger	93%	4%	1%	1%	1%	0%
Disgust	7%	82%	7%	0%	2%	2%
Fear	8%	7%	83%	1%	1%	0%
Happy	1%	2%	1%	92%	2%	2%
Neutral	1%	0%	1%	5%	90%	3%
Sad	4%	1%	5%	0%	7%	83%

Eye Expressions	Anger	Disgust	Fear	Happy	Neutral	Sad
Anger	95%	2%	1%	1%	1%	0%
Disgust	5%	85%	6%	2%	1%	1%
Fear	7%	2%	83%	5%	2%	1%
Happy	4%	1%	0%	92%	2%	1%
Neutral	1%	1%	1%	5%	91%	1%
Sad	5%	1%	5%	2%	3%	84%

4 Conclusion

This paper used to find a real-time system for human emotion recognition. The proposed work was developed using MATLAB's Image Processing Toolbox. We used our own database as well as YouTube databases for testing. Two different extraction methods are DWT and DCT. The performance of each feature extraction method is evaluated and compared based on the SVM classifier. The result of the proposed work shown in Table 1 shows that the anger expression was correctly recognized with a success rate of 93.0% using DWT. Table 2 proves that the anger expression was correctly recognized with a success rate of 95.0% by DCT. In the future, other features such as Haar can be used for feature extraction. Also, other classifiers like AANN, RBFNN, GMM, HMM, fuzzy logic, etc. can be used for feature modeling. The proposed system is sensitive to pose variations and needs improvement to handle them.

References

[1] Manikandan, S., Radhika, K.S.R., Thiruvenkatasuresh, M.P., & Sivakumar, G. (2022). Deepq: Residue analysis of localization images in large scale solid state physical environments. *In AIP Conference Proceedings*, *2393*(1), 1-6.

[2] Manikandan, S., Dhanalakshmi, P., Rajeswari, K.C., & Rani, A. (2022). Deep Sentiment Learning for Measuring Similarity Recommendations in Twitter Data. *Intelligent Automation & Soft Computing*, *34*(1), 183–192.

[3] Jharnama jumdar, & Ramya Avarbhrith, (2018). Human Face Expression Recognition. *International Journal of Emerging Technology and Advanced Engineering*, *4*(7).

[4] Orozco, J., Rudovic, O., Roca, F.X., & Gonzalez, J. (2008). Confidence assessment on eyelid and eyebrow expression recognition. *In 8th IEEE International Conference on Automatic Face & Gesture Recognition*, 1-8.

[5] Kandemir, R., & Özmen, G. (2013). Facial expression classification with haar features, geometric features and cubic bã‰ zier curves. *IU-Journal of Electrical & Electronics Engineering*, *13*(2), 1667-1673.

[6] Emerich, S., Lupu, E., & Apatean, A. (2009). Emotions recognition by speechand facial expressions analysis. *In IEEE 17th European signal processing conference*, 1617-1621.

[7] Ryan, A., Cohn, J.F., Lucey, S., Saragih, J., Lucey, P., De la Torre, F., & Rossi, A. (2009). Automated facial expression recognition system. *In IEEE 43rd annual 2009 international Carnahan conference on security technology*, 172-177.

[8] Ekman, P., Matsumoto, D., & Friesen, W.V. (1997). Facial expression in affective disorders. *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*, *2*, 331-342.

[9] Viola, P., & Jones, M.J. (2004). Robust real-time face detection. *International journal of computer vision*, *57*, 137-154.

[10] Angelinjayanthi A, (2012). Eye Faction Along with Glisten Psychiatry using Human Computer Interface. *International Journal of Computing Algorithm, 1*(2).

[11] Kotsia, I., Zafeiriou, S., & Pitas, I. (2008). Texture and shape information fusion for facial expression and facial action unit recognition. *Pattern Recognition*, *41*(3), 833-851.
[12] Zhou, S.K., Chellappa, R., & Moghaddam, B. (2004). Visual tracking and recognition using appearance-adaptive models in particle filters. *IEEE Transactions on Image Processing*, *13*(11), 1491-1506.
[13] Song, Y., Huang, J., Zhou, D., Zha, H., & Giles, C.L. (2007). Iknn: Informative k-nearest neighbor pattern classification. In *Knowledge Discovery in Databases: PKDD 2007: 11th European Conference on Principles and Practice of Knowledge Discovery in Databases, Warsaw, Poland. Proceedings 11*, 248-264. Springer Berlin Heidelberg.
[14] Ahamed, M.I., Ahamed, M., Sivaranjani, A., & Chockalingam, S. (2021). Energy bandgap studies on copper chalcogenide semiconductor nanostructures using cohesive energy. *Chalcogenide letters*, *18*(5), 245-253.
[15] Ahamed, M.I., & Kumar, K.S. (2019). Studies on CuSnS quantum dots for O-band wavelength detection. *Materials Science-Poland*, *37*(2), 225-229.
[16] Ahamed, I., Ahamed, M., Kumar, K.S., & Sivaranjani, A. (2022). Comparative energy bandgap analysis of zinc and tin based chalcogenide quantum dots. *Revista Mexicana de Física*, *68*(4), 1-8.
[17] Ferreira, A.P., Gupta, C., Inácio, P.R., & Freire, M.M. (2021). Behaviour-based Malware Detection in Mobile Android Platforms Using Machine Learning Algorithms. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)*, *12*(4), 62-88.
[18] Ahamed, M.I., & Kumar, K.S. (2019). Modelling of electronic and optical properties of CuSnS quantum dots for optoelectronics applications. *Materials Science-Poland*, *37*(1), 108-115.
[19] Babu, P. J. S., Padmanabhan, T. S., Ahamed, M. I., & Sivaranjani, A. (2021). Studies on copper indium selenide/Zinc sulphide semiconductor quantum dots for solar cell applications. *Chalcogenide Letters*, *18*(11), 701-715.

Authors Biography

Rahama Salman is currently a **Research Scholar** in the Department of Computer Science at SAM GLOBAL UNIVERSITY(SGU), BHOPAL, MP, INDIA. She is currently working on her PH. D Thesis work in the area of Data Mining and Machine Learning as major specialization and with AI and Networking as specific area of Interest. She received Masters degrees in Computer science Applications and Bachelor's degree in Mathematics, Statistics and Computers from Osmania University, Hyderabad, India. She published a number of papers in preferred Journals and chapters in books, and participated in International Conferences on a range of subject areas like Deep learning, IOT and Data Mining. She has 14 years of teaching experience in academics. Her research activities are based on: To predict chronic diseases with improved accuracy and precision using efficient Hybrid Voting Deep learning and ML classifiers.

Arshia Arjumand Banu received the Bachelor of Science (B. Sc) degree in Maths, Statistics & Computer Science from R.B.V.R.R Women's College, Osmania University, Telangana, India, in 1998. Persuade Master of Computer in Computer Science from Osmania University, Telangana, India in 2001. Currently working as lecturer in department of Computer Science, College of Computer Science & Information Technology, and Jazan University. Her research interests are Artificial Intelligence, Machine Learning, Image Processing and IoTs.