International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 3 Mar. 2021, pp: 605-611 www.ijaem.net ISSN: 2395-5252

Facial Expression Recognition from Facial images using RNN Algorithm

Devendra Kumar R.N.[#1], Bhupesh A.R.[#2], Akash S[#3], Hari Prakash M[#4]

#1 Assistant Professor, Department of CSE, Sri Ramakrishna Institute of Technology
#2, #3, #4 Students, Department of CSE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, India

---

Submitted: 01-03-2021 Revised: 15-03-2021 Accepted: 18-03-2021

ABSTRACT - The human face has peculiar and specific characteristics, therefore it becomes difficult in understanding and identifying the facial expressions. It is easy to identify the facial expression of particular person in any image sequence. If we look to automated recognition system, however, the systems available are quite inadequate and incapable of accurately identify emotions. The area of facial expression identification has many important applications. It is an interactive tool between humans and computers. The user, without using the hand can go-ahead with the facial expressions. Presently, the research on facial expression are on the factors i.e. sad, happy, disgust, surprise, fear and angry. This paper aims to detect faces from any given image, extract facial features (eyes and lips) and classify them into 6 emotions (happy, fear, anger, disgust, neutral, sadness). The training data is passed through a series of filters and processes and is eventually characterized through a Support Vector Machine (SVM), refined using Grid Search. The testing data then tests the data and their labels and gives the accuracy of classification of the testing data in a classification report. Various approaches, including passing the training images through Gabor filter, or transforming images using Histogram of Oriented Gradients (HOG) and Discrete Wavelet Transform (DWT) for better classification of data are implemented. The best result achieved so far is by passing the training images through Histogram of Oriented Gradients (HOG), followed by characterization by SVM, which gives an average precision of 85%.

Keywords–Facial, emotion, expression, detection, facial feature extraction, facial movement coding machine, recurrent neural network, rnn architecture

I. INTRODUCTION

Facial expression is one of the most powerful, herbal and generic signals for people to bring their emotional states and intentions. Several studies had been conducted on automatic facial expression analysis today's its practical significance in sociable robotics, medical treatment, driver Fatigue surveillance, and many different human-pc interplay systems. Inside the area modern computer vision and device studying, various facial features recognition (FER) structures had been Explored to encode expression data from facial representations. As early as the 20 th century, Ekman and Friesen defined six fundamental feelings based totally on cross-lifestyle take a look at, which indicated that humans perceive certain simple emotions inside the same way irrespective of lifestyle. these prototypical facial expressions are anger, disgust, worry, happiness, disappointment, and surprise. Contempt was ultimately delivered as one of the fundamental emotions. Recently, superior studies on neuroscience and psychology argued that the version modern day six simple feelings are subculture-specific and not generic. although the affect model based on basic feelings is restricted inside the ability to represent the complexity and subtlety contemporary our each day affective shows and different emotion description fashions, inclusive of the Facial movement Coding machine (FACS) and the continuous version the usage of affect dimensions, are considered to symbolize a wider variety state-of-the-art feelings, the specific model that describes feelings in phrases modern discrete fundamental emotions is nevertheless the most popular angle for FER, contemporary its pioneering investigations along with the direct and intuitive definition trendy facial expressions. And in this survey, we can limit our discussion on FER primarily based on the explicit model. FER structures can be divided into two primary categories in keeping with the function representations: static image FER and dynamic collection FER. In static-based totally strategies the feature representation is encoded with simplest spatial statistics from the modern single picture, while dynamic-based totally techniques do not forget the temporal relation amongst contiguous frames in the enter facial expression series. Based on those vision based strategies, other modalities, including audio and physiological channels, have also been utilized in multimodal structures to help the recognition cutting-edge expression. Trendy cutting-edge the conventional methods have used handcrafted features or shallow trendy (e.g., neighborhood binary patterns LBP on three orthogonal planes (LBP-pinnacle) non-poor matrix factorization (NMF) and sparse studying) for FER. But, on account that 2013, emotion recognition competitions including have amassed quite sufficient training facts from difficult real-world eventualities, which implicitly sell the transition today's FER from lab-managed to in-the-wild settings. In the in the meantime, modern-day the dramatically extended chip processing skills (e.g., GPU gadgets) and nicely-designed community structure, research in various fields have all started to transfer to deep contemporary methods, which have achieved the 49a2d564f1275e1c4e633abc331547db recognition accuracy and surpassed previous effects by using a massive margin. Likewise, given with greater powerful schooling records modern day facial features, deep trendy techniques have latest been carried out to deal with the hard factors for emotion recognition inside the wild. Figure 1 illustrates this evolution on FER within the component trendy algorithms and datasets. Exhaustive surveys on computerized expression evaluation have been posted in latest years. These surveys have mounted a hard and fast modern-day general algorithmic pipelines for FER. But, they cognizance on conventional methods, and deep trendy has rarely been reviewed. Very currently, FER based totally on deep gaining knowledge state modern has been surveyed in, that is a short review without introductions on FER datasets and technical information on deep FER. Consequently, on this paper, we make a systematic research on deep present day for FER tasks based on both static images and videos (photo sequences). FER. Despite the powerful function learning potential of deep learning, issues remain when implemented to FER. First, deep neural networks require a large amount of education facts to keep away from overfitting. But, the present facial expression databases aren't enough to train the 9aaf3f374c58e8c9dcdd1ebf10256fa5 neural network with deep structure that carried out the maximum promising effects in object reputation duties. Moreover, excessive inter-concern variations exist because of special personal attributes, together with age, gender, ethnic backgrounds and stage of expressiveness. Further to concern identification bias, variations in pose, illumination and occlusions are not unusual in unconstrained facial features situations. Those factors are nonlinearly coupled with facial expressions and therefore give a boost to the requirement of deep networks to cope with the massive intra-class variability and to learn effective expression-precise representations. On this paper, we introduce current advances in studies on fixing the above issues for deep FER. We have a look at the state of-the-artwork effects that have no longer been reviewed in preceding survey papers. The rest of this paper is prepared as follows. Often used expression databases are delivered in phase 2. Segment three identifies three main steps required in a deep FER machine and describes the associated heritage. phase four presents an in depth overview of novel neural community architectures and unique community schooling hints designed for FER based totally on static pix and dynamic photograph sequences. We then cover extra related problems and other realistic scenarios in phase 5. Segment 6 discusses some of the challenges and possibilities on this discipline and identifies capability future guidelines.

II. LITERATURE SURVEY

"Neural Network Rough Contour Estimation Routine (RCER) (Own Database) 92.1% recognition Rate". In this paper, they describe radial basis function network (RBFN) and a multilayer perception (MLP) network. "Principal Component Analysis (FACE94) 35% less computation time and 100% recognition". Useful where larger database and less computational time they want to repeat their experiment on larger and different databases. " PCA and Eigen faces (CK, JAFFE) 83% Surprise in CK, 83% Happiness in JAFFE " , Fear was the most confused expression Compared with the facial expression recognition method based on the video sequence, the one based on the static image is more difficult due to the lack of temporal information. Future work is to develop a facial expression recognition system, which combines body gestures of the user with user facial expressions. "2D Gabor filter (Random Images) 12 Gabor Filter bank used to locate edge" .Multichannel Gabor filtration scheme used for the detection of salient points and the extraction of texture features for image retrieval applications. They work on adding global and local colour histograms and parameters connected with the shapes of objects within images. "Local Gabor Filter + PCA + LDA (JAFFE)". Obtained 97.33% recognition rate with the help of PCA+LDA

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 3 Mar. 2021, pp: 605-611 www.ijaem.net ISSN: 2395-5252

features. They conclude that PCA+LDA features partially eliminate sensitivity of illumination. "PCA + AAM (Image sequences from FG-NET consortium)". The performance ratios are 100 % for expression recognition from extracted faces, the computational time and complexity was also very small. Improve the efficiency Extend the work to identify the face and it's expressions from 3D images.

III. EXISTING SYSTEM

This paper proposes another provincial energy mindful bunching technique utilizing secluded hubs for WSNs, called Regional Energy Aware Clustering with Isolated Nodes (REAC-IN). In REAC-IN, CHs are chosen dependent on weight. Weight is resolved by the remaining energy of sensor and the local normal energy of all sensors in each group. Inappropriately planned disseminated bunching calculations can make hubs become secluded from CHs. Such detached hubs speak with the sink by burning-through abundance measure of energy. To drag out organization lifetime, the provincial normal energy and the distance among sensors and the sink are utilized to decide if the separated hub sends its information to a CH hub in the past round or to the sink. The reproduction aftereffects of the current investigation uncovered that REAC-IN outflanks other bunching calculations.

IV. PROPOSED SYSTEM

This paper offers an entire system for facial expression recognition. The face model is used alongside a discovered goal function for face version fitting. The ensuing sequence of model parameters is then supplied to a recurrent neural community for class. The gain of the use of a recurrent community is that the temporal dependencies present inside the photo sequences may be taken under consideration during the category. Since the complete system is automated, and the recurrent networks can be used to make on line predictions, the device could be ideal for actual-time reputation. This would make it appropriate for the conference situation, wherein visitors have to be diagnosed and served by using robotic waiters. With the assist of this venture, a person who is meant to screen the humans may be seated in a far off vicinity and still can monitor efficaciously and supply instructions for this reason.

I

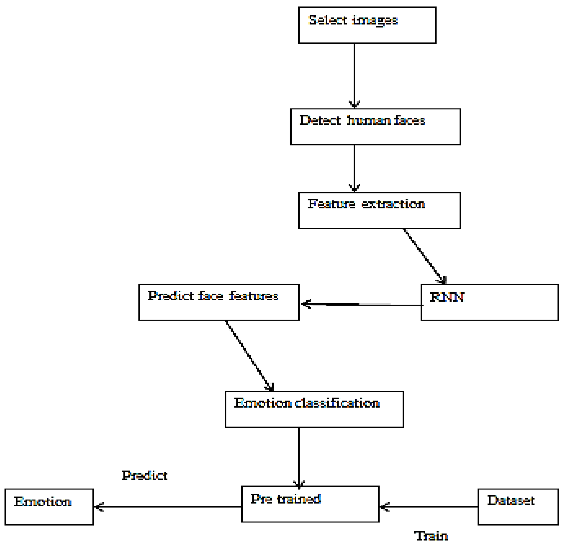

Fig 1 block diagram of proposed system

In this paper, a multichip bunching calculation (MHC) is proposed for energy saving in remote sensor organizations. In MHC, the sensor is chosen as bunch head as per the two boundaries remaining energy and hub degree. Likewise bunch heads select their individuals as indicated by the two boundaries of sensor the leftover energy and the distance to its group head. MHC is done in three stages, introductory, various levelled, and last stages. This calculation plays out the underlying stage just in the start of organization bunching and the last stage subsequent to completing organization grouping. Notwithstanding, the calculation rehashes the progressive stage from the primary level until

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 3 Mar. 2021, pp: 605-611 www.ijaem.net ISSN: 2395-5252

last level progressively finishes the bunching of the whole organization. In information assortment stage, sensors contrast assembled natural information and its adjoining information. On the off chance that information was comparative, the sensor stores ID of message sender in the rundown of its neighbours and checks the quantity of adjoining and set Ni variable. In introductory stage at start of grouping, BS that as a bunch head of first level, send a "Start" message experiencing significant change scope of sensors, and advise beginning of grouping to all. Just sensors that are near BS, get this message. Various levelled stage is done in four stages progressively that entire sensors of organization can bunched.

Stage 1.

On this phase, we describe our proposed gadget to investigate students' facial expressions the use of a recurrent neural network (RNN) architecture. First, the gadget detects the face from input image and these detected faces are cropped and normalized to a length of forty eight forty eight. Then, these face photos are used as enter to RNN ultimately, the output is the facial expression popularity consequences (anger, happiness, unhappiness, disgust, surprise or impartial). Parent 1 affords the shape of our proposed approach. In this phase, we describe the three principal steps that are common in automated deep FER, i.e., pre-processing, deep characteristic gaining knowledge of and deep function classification. We in brief summarize the extensively used algorithms for every step and recommend the existing kingdom-of the-art great practice implementations according to the referenced papers.

Pre-processing

Variations which are beside the point to facial expressions, inclusive of distinct backgrounds, illuminations and head poses, are fairly common in unconstrained situations. Consequently, earlier than education the deep neural network to examine significant functions, pre-processing is required to align and normalize the visible semantic data conveyed through the face.

Face alignment

Face alignment is a conventional pre-processing step in lots of face related recognition tasks. Given a chain of training information, the first step is to come across the face and then to eliminate background and non-face areas. The face detector is a classic and extensively employed implementation for face detection, which is strong and computationally simple for detecting close to-frontal faces. Although face detection is the simplest fundamental system to allow feature getting to know, in addition face alignment the usage of the coordinates of localized landmarks can significantly decorate the FER performance.

This step is important because it is able to lessen the variation in face scale and in-aircraft rotation.

RECUURENT NEURAL NETWORK (RNN) is the early paintings which predicts landmarks in a cascaded manner based on this, responsibilities-constrained Deep Recurrent community and Multi-mission RECUURENT NEURAL NETWORK (RNN) further leverage multi-venture studying to enhance the performance. In trendy, cascaded regression has end up the maximum famous and state-of the-art techniques for face alignment as its excessive velocity and accuracy.

Neural Network

We started with a Recurrent neural network has multiple Recurrent layers, activation units and Max Pooling Layer.

The primary purpose of a recurrent layer is to extract features from the input image. Convolution preserves the spatial relationship between pixels by learning image features using small squares of input data. The more number of filters, the more image features get extracted and the better the network becomes at recognizing patterns in unseen images.

Hyper Parameters:

Hyper parameters for recurrent layers

If the stride value is more, then there would be a chance of missing important parts of an image. If it is too small the system will take more memory and time.

Filter size and number of filters at each conelayer: Filters are usually chosen to be of the size odd number squares i.e. 3x3, 5x5, etc. . . . Number of filters would be in the powers of 2 i.e. 16, 32, 64 as it will be easier for the computation. Number of filters in the layer gives the number of feature maps obtained from that layer.

Padding: The size of the padding depends on filter size that we are applying. It's a good practice not to converge the image quickly. So we use padding so that input size remains same even after applying filters. Input size is only reduced in max pool layer.
Dropout probability: It is chosen so that the neurons are dropped at each layer during training.

Max pooling: Frame size is 2x2. Each max pooling

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 3 Mar. 2021, pp: 605-611 www.ijaem.net ISSN: 2395-5252

layer reduces the size to half if we use 2x2 frame.

Application Module:
An images feed into the neural network. The network subsequently classifies the emotion showed by the subject based on the facial expressions. An output corresponding to the emotion is displayed on the screen.

Data Augmentation:
As mentioned above, Deep learning requires large amounts of data for training. So for training, we used dataset which has images in range of tens of thousands.

The dataset has images in 48x48 pixels which are already cantered around the face, so pre-processing for fern data is minimal. We used FERC as training data, the reason is that, once trained on the pixelated images of the FERC dataset, then emotions from 'clean' images can be easily recognised, but not vice versa. We have tested the final model on 'clean' images of dataset. In, we chose only pictures with frontal faces as these are highly represented in the FERC training dataset and chose to discard others. Then the faces from these images are extracted and resized to 48x48 and then fed to the network.

V. WORKING PRINCIPLE

The given block diagram shows the pictorial representation of the working principle of this tool. The system comprises of four different modules namely acquisition module, segmentation module, feature extraction module and classification module.

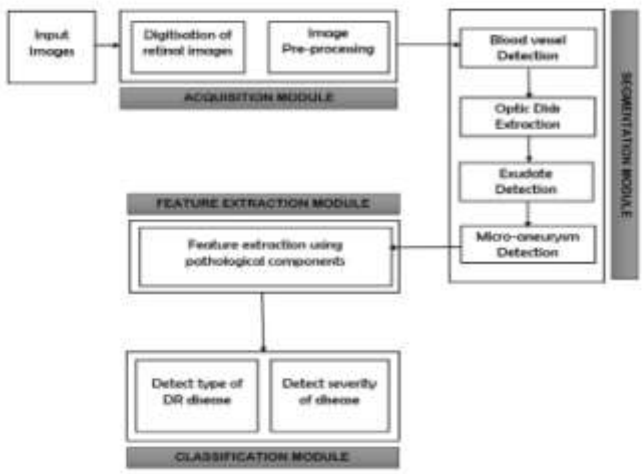

Fig 5: Block Diagram

At the initial module, that is the Acquisition module, the input images are digitized and pre-processed to attain perfect resolution. Then these images are stored for undergoing the further processes.

In the next module, that is the Segmentation module, segmentation process takes place in order to detect and differentiate based on the important classifying characteristics such as micro-aneurysm, exudates, hemorrhages and optic disks. These morphological characteristics are helpful in segmenting the images using an unsupervised blood vessel segmentation methodology. Also in some exceptional cases, some morphological operations such as Dilation (process of adding pixels to the image; makes the object to be more visible) and Erosion (remove pixels from the object boundaries) are done to get much better accurate results.

In the Feature Extraction module, the important features are extracted using the pathological components and then it is classified into the types and severity of the defect whether it's in Proliferative or Non-Proliferative Diabetic Retinopathy stage. After all these stages, the end result is shown to the user.

VI. EXPERIMENTAL SETUP

In this experiment, Python IDE version 3.9 or higher should be installed and executable. High package libraries such as Koras, Tensor flow, Matplotlib and Sickie should be installed in a runnable interface. Also the CNN network

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 3 Mar. 2021, pp: 605-611 www.ijaem.net ISSN: 2395-5252

VGG16 model should be installed which is used for feature extraction. The input images are stored in a working directory file and set ready for training procedure.

VII. RESULTS

As a result, our system can able to fetch the images from the database and predict the result, about the emotion of the person.
Each of the emotion is categorized and the live time detection is being found for each of the person.

The above is a standard face expression for which the emotion is determined as Neutral. Each of the emotion of the person on the video streaming is lively monitored and expressed in each of the way.

VIII. CONCLUSION

In this project, we built a recurrent neural network to recognize emotion from grayscale pictures of faces. We experimented with different models, achieving highest test accuracy of on a RNN trained from scratch.

IX. FUTURE WORKS

In future we would like to improve the accuracy of the system in detecting the Diabetic Retinopathy defect so that the results are more accurate and clear.Also,we would like to make the whole toolaccessible as a web or mobile application so that it provides a much more user friendly and integrated platform.

ACKNOWLEDGEMENT

The authors are deeply grateful to The Honourable Principal and Faculties of Sri Ramakrishna Institute of Technology, Coimbatore for providing the necessary support, guidance and facilities for the preparation of this paper.

REFERENCES

[1] A. Kołakowska, A. Landowska, M. Szwoch, W. Szwoch, and M. R. Wrobel, ´ Human-Computer Systems Interaction: Backgrounds and Applications 3, ch. Emotion Recognition and Its Applications, pp. 51–62. Cham: Springer International Publishing, 2014.
[2] P. Ekman, Universals and cultural differences in facial expressions of emotion. Nebraska, USA: Lincoln University of Nebraska Press, 1971.
[3] Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pages 1097-1105, 2012.

[4] T. Ahsan, T. Jabid, and U.-P. Chong. Facial expression recognition using local transitional pattern on gabor filtered facial images. IETE Technical Review, 30(1):47-52, 2013.
[5] Gudi. Recognizing semantic features in faces using deep learning. arXiv preprint arXiv:1512.00743, 2015.
[6] "Dataset: Facial Emotion Recognition (FER2013)" ICML 2013 Workshop in Challenges in Representation Learning, June 21 in Atlanta, GA.
[7] Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D.H.J., Hawk, S.T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. Cognition & Emotion, 24(8), 1377—1388. DOI: 10.1080/02699930903485076.
[8] Mart´ın Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mane, Rajat Monga, Sherry Moore, Derek Murray, ´ Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viegas, Oriol Vinyals, ´ Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng Google Research∗-TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems (Preliminary White Paper, November 9, 2015).
[9] TFlearn. Tearn: Deep learning library featuring a higher-level api for tensorow. URL :http://tflearn.org/.
[10] OpenSourceComputerVision. Face detection using haar cascades. URL: http://docs.opencv.org/master/d7/d8b/tutorial_py_face_detection.html.
[11] P. E. Trahanias and A. N. Venetsanopoulos, "Color image enhancement through 3-D histogram equalization," in Proc. 15th IAPR Int. Conf. Pattern Recognition, vol. 1, pp. 545–548, Aug.-Sep. 1992.
[12] https://cs231n.github.io/convolutional-networks/
[13] Yann LeCun, Koray Kavukcuoglu, and Cl´ement Farabet. Convolutional networks and applications in vision. In Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on, pages 253–256. IEEE, 2010. 2,4
[14] Dropout : Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. J. Machine Learning Res. 15, 1929–1958 (2014)
[15] J. D. Boucher and P. Ekman. Facial areas and emotional information. 1975.
[16] Zeiler, M. D. and Fergus, R. Visualizing and understanding convolutional networks. CoRR, abs/1311.2901, 2013. Published in Proc. ECCV, 2014.
[17] Yichuan Tang. Deep learning using support vector machines. CoRR, abs/1306.0239, 2013. 2

International Journal of Advances in Engineering and Management

ISSN: 2395-5252

IJAEM

Volume: 03 Issue: 03 DOI: 10.35629/5252

www.ijaem.net

Email id: ijaem.paper@gmail.com