JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

EMPLOYEE EMOTION ANALYSIS USING COMPUTER VISION

Leema Roselin George
Assistant Profeesor
Department of computer Science Engineering
IFET College of Engineering
Villupuram, Tamilnadu, India.

A. Saravanakumar
Undergraduate Student
Department of computer Science Engineering
IFET College of Engineering
Villupuram, Tamilnadu, India.

Abstract — The employee's mental health has a direct impact on productivity, by factors including work-related challenges, lengthier job durations, toxic collaboration, and cold war amongst coworkers. Another important aspect for employee retention rates in any organization is the individual's mental wellness. In order to maintain their employees' mental health and level of engagement at work, every organization strives to provide a healthy atmosphere. However, the majority of the staff members used diplomacy to resolve their issues. A solution to aid the employees with their mental health is therefore desperately needed. A computer vision solution is proposed to detect the employees' mental health while they are at work and helping them become daily conscious of their mental health utilizing a data analytics dashboard. Additionally, our solution uses a machine learning recommendation system to help users determine when to work and when to rest for increased productivity.

I. INTRODUCTION (*HEADING 1*)

Emotions are well-known to have a significant impact on human life. Human faces can divulge how anyone is feeling or in what temper they are at a number of instances or moments. When communicating, human beings are efficient of performing hundreds of exceptional facial expressions that vary in complexity, intensity, and meaning. Subtle differences in one or greater awesome characteristics often sign emotion or intention. Its interpretation can also trade relying on whether or not one or extra facial movements are existing or absent. Additionally, no matter having a comparable ordinary morphology, a number of facial expressions can convey a range of meanings relying on their depth. I will employ an existing simulator that can read or compare facial expressions to be able to capture human emotions in order to represent the subtlety of face expression in non-verbal communication. This programme detects subtle differences in face expressions, extracts features and associated motion information, and estimates expression strength. A healthcare system is created that concentrates on the emotional components of the employee in order to cope with bad emotions in daily life and to improve mental states. In order to raise users' good feelings and decrease their negative emotions, such as sadness, anger, and fear, the system includes emotion detection to recognise users' present emotional states.

A better facial emotion recognition model to address the aforementioned issues. ResNet-50 serves as our network infrastructure. Convolutional neural networks, activation functions like Rectified Linear Unit, and batch normalization are employed to enhance the model's capacity for convergence. It can be confirmed by the simulation experiment below that the proposed face emotion recognition method performs recognition better than the most sophisticated method.

II. DATASET

The model has been trained using FER-2013. The training set of the FER-2013 dataset consists of 28,000 labelled images, whereas the validation set and test set each have 3,500 tagged images. The seven emotions that are included in the FER-2013 images—happy, sad, angry, terrified, surprise, disgust, and neutral—is assigned to a particular image, with joyful being the most common and providing a baseline of 24.4% for random guessing. Both posed and candid headshots are included in the FER-2013 pictures, which are grayscale and 48x48 pixels. By combining the results of Google image searches for each emotion and its synonyms, the FER-2013 dataset was created.

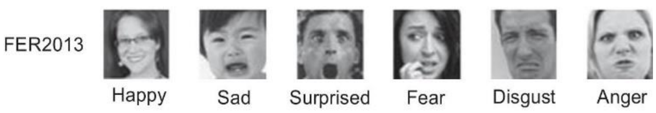

Figure 1. Examples of facial emotions from the FER-2013 dataset

III. METHODOLGY

Convolution

Convolution is an operation that creates a third function by combining the results of two functions (f and g). One of the most crucial processes in signal and image processing is convolution. It may work in 1D (for example, speech processing), 2D (for example, image processing), or 3D (for example, video processing).

A picture can be altered using the convolution approach, which applies a kernel to each pixel and the pixels around it over the entire image.. The size and values of the kernel, which is a matrix of values, control how the convolution process transforms data.

These steps are part of the convolution process.

(1) It lays the kernel matrix over each pixel of the image and multiplies each kernel value by the corresponding pixel it is over in order to make sure the entire kernel is contained within the image.

(2) After that, adds the values from the multiplication and returns the total as the updated value for the centre pixel.

(3) The entire image is subjected to this process repeatedly.

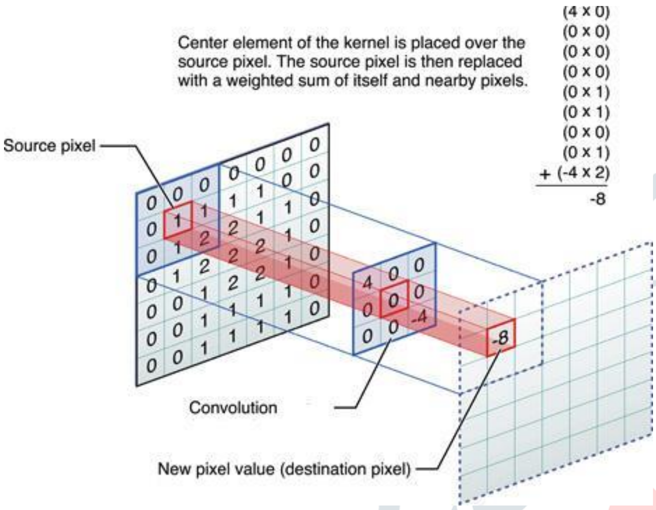

Figure 2. Example of Convolution Operation

Batch Normalization

The mean and variance of the inputs to each layer are fixed during a normalization stage in a neural network to achieve batch normalization. It is not practical to use the global information in this phase when stochastic optimisation techniques are being used. The normalization should ideally be carried out during the entire training set. As a result, normalisation is restricted to every mini-batch during the training phase.

Let B be used to represent a mini-batch of size m of the complete training set. The mean(μ) and variance(σ) of the mini-batch could be denoted as

$$\mu_B = \frac{1}{m}\sum_{\{i=1\}}^{m} x_i \quad (1)$$

$$\sigma_B^2 = \frac{1}{m}\sum_{\{i=1\}}^{m}(x_i - \mu_B)^2 \quad (2)$$

Use the mean and standard deviation discovered to normalize the hidden activations. To do this, we'll take the mean from each input and divide the total by the smoothing factor (ε) plus the sum of the standard deviation.

$$x = \frac{(x_i - \mu_B)}{\sigma_B + \varepsilon} \quad (3)$$

By preventing division by a zero value, the smoothing term (ε) ensures numerical stability inside the operation. Rescaling and offsetting of the input happen in the final operation. Here, γ(gamma) and β (beta), two elements of the Batch Normalisation algorithm, enter the scene. Re-scaling(γ) and shifting(β) the vector containing the results of the previous operations are accomplished using these parameters.

$$x_i = \gamma x_{i(norm)} + \beta \quad (4)$$

Rectified linear unit

A rectified linear unit (ReLU) is an activation function that addresses the problem of vanishing gradients and adds the property of non-linearity to a deep learning model. The most popular activation function used in deep learning models is the Rectified Linear Unit. If the function receives any value less than 0, the function returns 0, but if it receives any positive value x, it returns that value x. So it can be written as

$$f(x) = \max(0, x) \quad (5)$$

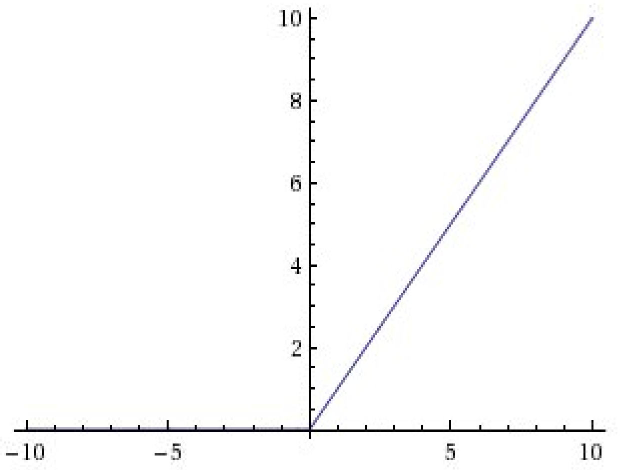

Figure 3. Rectified Linear Unit

This is comparable to having no hidden layers, which means that matrix multiplication applies to all layers that lack activation functions. Layers don't matter; all they do is multiply matrices. Therefore, the network's capacity to approximate is highly constrained. We chose to employ a nonlinear function as the activation function for the aforementioned reasons, making the deep neural network more powerful and able to mimic nearly any function rather than just a linear combination of inputs. Unilateral inhibition is the process by which ReLU, a piecewise function, converts all negative values to zero while leaving the positive values unchanged. To put it another way, if the input is negative, the function will produce zero, which will prevent the neuron from being engaged. Because just a small number of neurons are active at once, the network is sparse and hence very effective for computation. The advantages of the Rectified Linear Unit function are:

1) Both a saturation zone and a gradient disappearance don't exist.

2) The rectifier function's ability to provide a real zero value makes it useful.

3) The max() function is all that is necessary to implement the rectifier function.

ResNet50

ResNet-50 is a convolutional neural network that consists of 50 layers. Numerous computer vision applications use neural networks known as ResNet or Residual Networks, as their foundation. ResNet's major innovation allowed us to train extremely deep neural networks with layers upto 150.

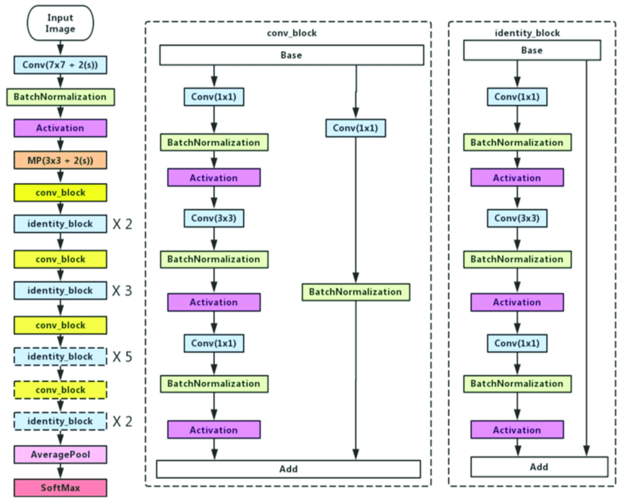

Figure 4. ResNet-50 Architecture

The 'Vanishing Gradient Problem' is a serious drawback that needs to be addressed in convolutional neural networks. Gradient value greatly reduces during back propagation therefore weights scarcely change at all. ResNet is employed to get around this. It uses the "SKIP CONNECTION" feature i.e. Adding the original input to the output of the convolutional block.

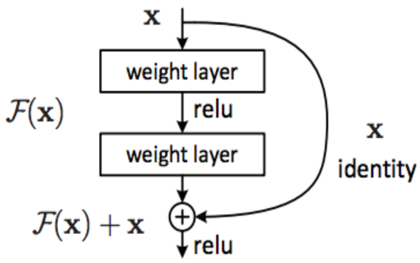

Figure 5. Skip Connection

A skip connection is a straight connection that skips some model layers.

IV. Results and Discussion

The performance of the model in the FER-2013 dataset was examined in terms of accuracy, loss graph in both training and testing data.

With the RESNET-50 model, accuracy of 69% is acquired which can be used in real time to obtain and record the emotions of the employee.

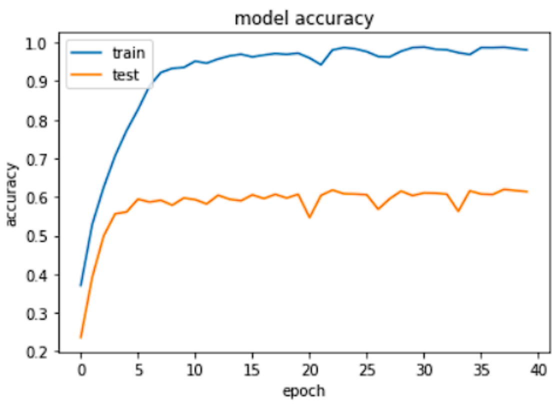

Figure 6. Model accuracy graph

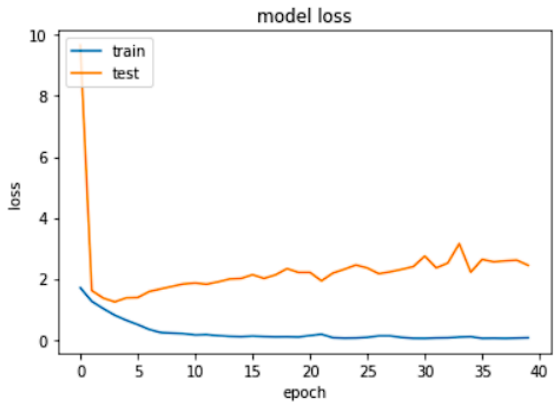

Figure 7. Model loss graph

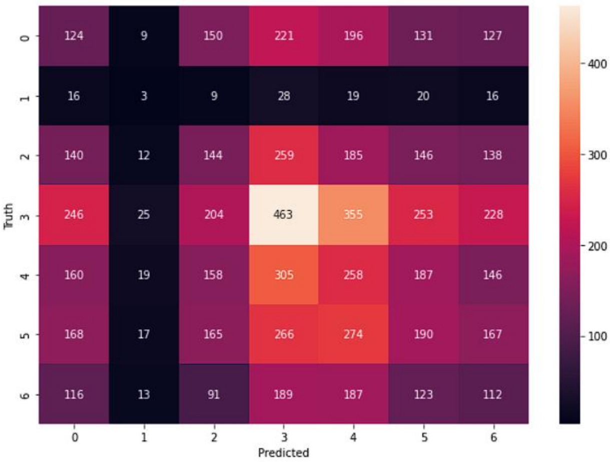

Figure 8. Heat Map

The system first loads the model then loads a pre-trained haar-cascade to detect faces in the image then we take a stream of images from the webcam then for each image we detect faces, extract them then feeds them to the model in order to predict the expression of the face then display a bounding box around the face with expression labeled on top of it. Then the results are displayed in the dashboard.

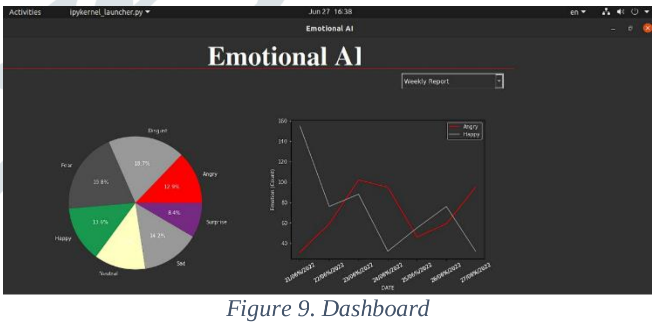

Figure 9. Dashboard

V. Conclusion

The idea of this project was to develop an emotional recognition system by applying computer vision techniques and utilizing the advanced feature extraction and face expression classification detection algorithm. This study examines a more effective face expression recognition system and proposes a deep residual network-based method for identifying facial expressions. ResNet-50 residual network, a commonly used convolutional neural network approach, is used since it did well in the multi-classification test. Through the testing of the data, the obtained results show that the method proposed forth in this research has good accuracy and a good emotion recognition that can be used for real time purposes. The data obtained from the model can be stored and used to present the data in form of visuals in a dashboard. Dashboard includes both weekly and monthly report of the employee emotions which can then be used to assess the mental health of the employee.

REFERENCES

[1] Ali, H. , Hariharan, M. , Yaacob, S. , & Adom, A. H. (2015, October). Facial emotion recogni- tion based on higher-order spectra using support vector machines. Journal Of Medical Imaging And Health Informatics, 5 , 1272–1277 .

[2] Ali, Y. , Janabi-Sharifi, F. , & Beheshti, S. (2021, February). Echocardiographic image seg- mentation using deep Res-U network. Biomedical Signal Processing and Control, 64 (Ar- ticle ID: 102248), 14 .

[3] Ashiquzzaman, A. , Lee, H. , Kim, K. , Kim, H. Y. , Park, J. , & Kim, J. (2020, November). Com- pact spatial pyramid pooling deep convolutional neural network based hand gestures decoder. Applied Sciences-Basel, 10 (Article ID: 7898), 22 .

[4] Belinschi, S. T. , Bercovici, H. , & Liu, W. H. (2021). The atoms of operator-valued free convolutions. Journal of Operator Theory, 85 , 303–320 Win .

[5] Benbahria, Z. , Sebari, I. , Hajji, H. , & Smiej, M. F. (2021, February). Intelligent mapping of irrigated areas from landsat 8 images using transfer learning. International Journal of Engineering and Geosciences, 6 , 41–51 .

[6] Bister, T. , Erdmann, M. , Glombitza, J. , Langner, N. , Schulte, J. , & Wirtz, M. (2021, March). Identification of patterns in cosmic-ray arrival directions using dynamic graph con- volutional neural networks. Astroparticle Physics, 126 (Article ID: 102527), 10 .

[7] Choi, S. H. , & Jung, S. H. (2020, March). Stable acquisition of fine-grained segments using batch normalization and focal loss with L1 regularization in U-Net structure. Interna- tional Journal of Fuzzy Logic and Intelligent Systems, 20 , 59–68 .

[8] Evans, F. (2017). Haar wavelet transform based facial emotion recognition. Advances in Computer Science Research, 61 , 342–346 2017/03 .

[9] Furlong, L. S. , Rossell, S. L. , Caruana, G. F. , Cropley, V. L. , Hughes, M. , & Van Rhee- nen, T. E. (2021, January). The activity and connectivity of the facial emotion pro- cessing neural circuitry in bipolar disorder: A systematic review. Journal of Affective Disorders, 279 , 518–548 .

[10] Ganguly, B. , Chaudhuri, S. , Biswas, S. , Dey, D. , Munshi, S. , Chatterjee, B. , et al. (2021, March). Wavelet kernel-based convolutional neural network for localization of partial discharge sources within a power apparatus. IEEE Transactions on Industrial Informat- ics, 17 , 1831–1841 .

[11] Garbin, C. , Zhu, X. Q. , & Marques, O. (2020, May). Dropout vs. batch normalization: An empirical study of their impact to deep learning. Multimedia Tools and Applications, 79 , 12777–12815 .

[12] Gonzalez-Yubero, S. , Lazaro-Visa, S. , & Palomera, R. (2021, January). How does emo- tional intelligence contribute to the study of personal protective factors for alcohol consumption in adolescence? Psicologia Educativa, 27 , 27–36 .

[13] Govindaraj, V. V. (2019). High performance multiple sclerosis classification by data aug- mentation and AlexNet transfer learning model. Journal of Medical Imaging and Health Informatics, 9 , 2012–2021 .

[14] Graumann, L. , Duesenberg, M. , Metz, S. , Schulze, L. , Wolf, O. T. , Roepke, S. , et al. (2021, January). Facial emotion recognition in borderline patients is unaffected by acute psychosocial stress. Journal of Psychiatric Research, 132 , 131–135

[15] Hajarolasvadi, N. , & Demirel, H. (2020, December). Deep facial emotion recognition in video using eigenframes. IET Image Processing, 14 , 3536–3546 .