Recognition of Emotional State Based On Handwriting Analysis and Psychological Assessment

S. V. Kedar, Shilpa Rokade

Abstract. Emotions describe the physiological states of an individual and are generated subconsciously. They motivate, organize, and guide perception, thought, and action. Emotions can be positive or negative. Negative emotions manifest in the form of depression, anxiety and stress. It is necessary to identify negative emotions of an individual who might be in the need for counseling or psychological treatment. Body signal analysis, handwriting analysis, and psychological assessment are some mechanisms to measure them. In this paper, emotional state is being measured through the person's handwriting sample analysis and psychological assessment. Psychological assessment is done by using the results of DASS questionnaire attempted by the individual. Convolutional Neural Network (CNN) algorithm is used to find the emotional state of an individual from his/her handwriting sample. Comparative analysis is performed to suggest counseling/medication if required. The final CNN model is formed by using the ensemble method over cross-validation models. The accuracy achieved by the CNN model over the test dataset is 91.25%.

Keywords : Emotion Recognition, Depression, Anxiety, Stress, Handwriting, Convolutional Neural Network

I. INTRODUCTION

For a human being, health is a very important aspect. Among the various reasons for causing illness, some of them can be lot of tension, stress and depression. It is important to know if a person is suffering from depression or stress, which are negative emotions. Emotions are generated subconsciously and they describe physiological states. The positive emotions are joy, happiness, excitement etc. while negative emotions are sadness, anger, disgust, fear, depression, anxiety and stress. The different methods to recognize emotions are, by using handwriting analysis, psychological assessment, analysis of speech and video [1-4]. There are number of applications of emotion recognition like education, gaming, colleges, offices etc. [5]. Like in the field of education, to understand the level of concentration of the learner, it is helpful to detect the movements of eyes and head. In the field of gaming, capturing the emotions of

players could be used in interactive games for various purposes. While emotion recognition in colleges and offices can help to understand the person's negative emotions who might need psychological treatment.Machine learning approaches such as random forests, neural networks, Bayesian networks, k-NN and SVM are implemented for recognition of emotional state [1, 6, 7] and human behavior prediction [8]. For the behavioral signals like speech or body gestures and facial expressions, the features are extracted to be given as input to a classifier. For the task of handwritten text recognition, deep learning approach has been used [9]. The authors in [2-4] have done research to recognize emotional state using audio, video and body gestures.

During direct interaction with a person, audio is the primary means of commsunication used by humans. The mood, emotion or mental state of the person is expressed by audio. Tarunika et. al [10] have implemented Deep Neural Network (DNN) and knearest neighbor (k-NN) for emotion recognition using audio, particularly when the person's mind is in scary state. For detection of emotion of an individual from the video, the characteristics analyzed are facial expressions, movements and activities, using machine learning. Samira Ebrahimi Kahou et. al. [11] have implemented the deep learning algorithm, convolutional neural network for facial expressions recognition. Facial expressions are captured from individual frames in a video. The validation accuracy obtained by the network is 43.7%.

Emotions can also be predicted through psychological analysis and graphology. Psychology is the study of human mind and its functions, feelings and thoughts. Psychological assessment is widely used for emotion recognition of an individual. It is concerned with assessing depression, anxiety and stress using different approaches. Depression is a feeling of sadness, anger, loss of interest, and feeling low from time to time, which affects day to day activities like behavior, sleep, appetite and social life [1]. Anxiety is a feeling of nervousness and worry about something with an uncertain outcome. Stress is a state which can have positive impact or negative impact on the mind and body. While in stress people feel tensed, angry, nervous because of demanding situation.

Graphology is the analysis of the physical characteristics and patterns of handwriting. Graphologist uses number of handwriting analysis techniques to know personality, emotions and other aspects. Graphology can be used to determine a complete personality and character profile of any person [12].

The graphologists study handwriting, doodles, drawings, sculptures, and paintings which give details of the physical, mental and emotional state of the person using handwriting traits. Various handwriting traits are baseline, slant, size, margin, pressure, spacing, zones.

These traits are the features of handwriting which makes one handwritten sample different from another [13].

In this paper, emotional state is recognized from handwritten text samples as well as psychological assessment. DASS scale is used for psychological assessment, it results into various scores and categories for the emotions like anxiety, stress and depression. CNN is used to extract the features from handwritten text sample. Convolutional Neural Network (CNN) is a popular technique used for image recognition [14-15] document analysis [16], handwritten numeral recognition [4]. CNN is one of the various algorithms in deep learning technology, a part of machine learning which is used successfully in many real-world applications [17-18]. For facial analysis, text analysis and video analysis, deep learning-based algorithms have performed very well to identify the face, emotion or various other activities [11].

The remainder of this paper is organized as follows. In Section II, related work is described, followed by Section III, which provides details on the methodology. Section IV presents the results and discussions. Finally, in Section V some conclusions are drawn.

II. RELATED WORK

Literature has shown various techniques for recognition of emotions using different modalities. The combination of various modalities like facial expressions and audio signals, gestures and audio, handwriting analysis and psychological analysis, gives better results [2-3]. The authors in the paper [19] has surveyed on gesture recognition especially facial expressions and hand gestures, while Sebe et. al [20] present research for recognition of emotion from facial expressions, audio signals, and physiological signals. They have treated these modalities independently. The combination of psychological assessment and handwriting analysis is presented in the paper [1]. Likforman-Sulem et. al. [1] have created a dataset in which handwriting and drawing are related to emotional state. The handwriting and drawing samples are taken from 129 users, who attempted the DASS questionnaire. The emotional states, like anxiety, depression, and stress, are evaluated by the DASS test. Random forest classifier is used for feature extraction. They have used cross-validation experiments to obtain the different evaluation measures including accuracy. Anxiety and Stress are recognized more accurately than depression. The study of handwriting feature for emotion recognition is presented by authors in the paper [2]. It is stated that through handwriting analysis various characteristics in personality can be detected. Controlling the emotions is one of them. It is helpful for the counselors to identify the emotions of their counselee. For feature extraction from handwriting, they have used the fuzzy technique. The level of emotion is identified using baseline or slope of the handwriting with the help of Mamdani inference.

The study of psychological assessment using DASS scale is presented by John and Julie [21], they have proposed that the concealed composition of the DASS was obtained from conceptual and experimental studies. Evaluation of DASS is done using confirmatory factor analysis. It was concluded that the DASS scale proves reliability and validity for intended purpose.

Apart from these, there are number of other research work on handwriting [14], relating handwriting to depression and anxiety [22], handwriting to personality [8, 23], combination of audio-video-body gestures [2, 10, 11]. To the best of our knowledge this is the first approach for recognition of emotion using the combination of psychological assessment with handwriting analysis using deep learning.

III. METHODOLOY

In this system, the emotional state of an individual is recognized from his/her handwritten text sample and psychological assessment. There are two modules in this system:

Module 1: Emotion recognition using handwritten text sample.

Module 2: Emotion recognition using psychological assessment which gives anxiety, stress and depression values

The first module predicts emotional state using handwriting sample and CNN. For this, the database of handwritten text samples is created by collecting the handwriting samples from 1600 individuals. The individuals were asked to write a paragraph on blank A4 size paper using a blue cello gripper 0.7mm ball pen. The samples are separated and named according to the results of module 2. They are scanned and stored as a dataset for training and testing.

The features in handwriting are baseline, slant, size, margin, pressure, spacing between characters, words and lines, zone. Images in different classes have different combinations of features. The features observed in the class depression are wide spacing between words/lines, ascending baseline, uneven left and right margin, less top margin. For class moderate anxiety, features are more pressure, left and top wide margin, mixed slant, wide spacing between lines, more curves of letters like y, g, f. For class severe anxiety, the features are right slant, wide spacing between lines/characters, mixed baseline. For class depression anxiety, the features are uneven baseline, uneven margin, long curves for letters like y, g, f etc., large size of letters, wide left margin, wide spacing between words. For class anxiety stress, the features are wide top/left margin, mixed/left slant, descending baseline, uneven spacing between letters. For class depression anxiety stress, the features are more pressure, right slant, mixed baseline, small size, ascending baseline, uneven margin, large curves for letters like s, g, y f etc. for class normal, the features are even margin, straight baseline, straight/right/left slant, small-medium size. The second module identifies emotional state using psychological assessment based on DASS test.

Retrieval Number F8960088619/2019©BEIESP
DOI: 10.35940/ijeat.F8960.088619
Journal Website: www.ijeat.org

4396

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Depression Anxiety Stress Scale (DASS) is a rating scale with 42 items in the questionnaire, developed by Lovibond and Lovibond in 1995 [24]. Emotions like anxiety, stress and depression are identified by this scale. The official website for DASS [25] gives all the details. There is a score for each item and there are set of items each for anxiety, stress and depression. The participant answers the questions depending on his/her emotional state
.

Table- I: DASS categories and score

Interpretation	Depression (D)	Anxiety (A)	Stress (S)
Normal	0-9	0-7	0-14
Mild	10-13	8-9	15-18
Moderate	14-20	10-14	19-25
Severe	21-27	15-19	26-33
Extremely Severe	>28	>20	>34

Accordingly, the severity levels such as normal, mild, moderate, severe and extremely severe are calculated and displayed for scores of emotions anxiety, stress and depression which are shown in in Table I. The psychometric properties (technique of psychological measurement) of the DASS are useful in variety of fields like corporate offices and companies where employee's health is taken care [12] [26]. The participants attempt the DASS scale, according to the results obtained, the handwriting samples collected in module 1 are separated. Calculations for the score are done depending on the rating for each item.

A. System Description

The block diagram of the proposed system is shown in Figure 1. The writer gives the handwriting sample as well as attempts the DASS test. According to the results obtained from the DASS test, the emotions database is created. This involves various classes in the database. Then the scanned image is preprocessed by using normalization and image resizing. The convolutional neural network algorithm is trained on the training dataset and validated on the validation dataset. Then it is tested on the testing dataset. Once this is done a new input image is given to CNN model, the model predicts the emotional state of the image based on features. The predicted emotion is compared with DASS analysis to find out the accuracy. Accordingly, psychological counselling/medication is suggested if required.

B. Algorithm

In this system, CNN is used for recognition of emotional state from handwritten text sample. CNN architecture in Figure 2 shows the three layers in CNN, convolutional layer (CL), pooling layer (PL) and fully connected layer (FCL). The number of convolution and pooling layers varies with the application.

Convolution Layer: Given an input image I of dimensions I1 x I2 and a filter (kernel) K of dimensions k1 x k2, a bias b, the output is a feature map 'F' of dimensions i x j is produced, the convolution operation is given by:
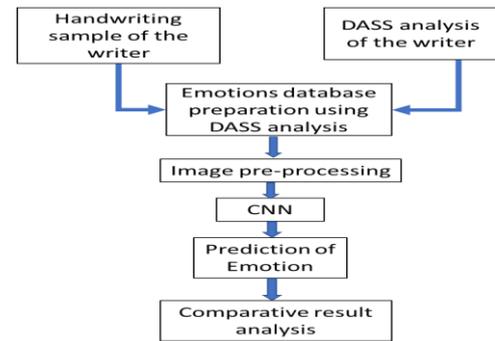

Fig 1. Block Diagram of Proposed System

$$F(I * K)_{ij} = \sum_{p=0}^{k1-1} \sum_{q=0}^{k2-1} I_{i+p,j+q} K_{p,q} + b \qquad (1)$$

Size of feature map F = [(n-w+2p)/s] + 1 where,

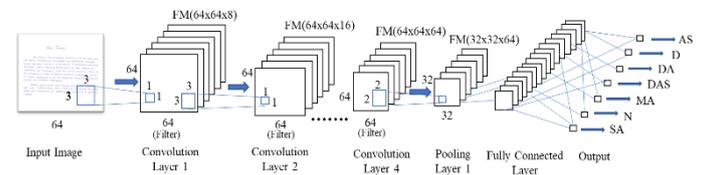n is image size; w is filter size; p is padding; s is stride;

Fig 2. Convolutional Neural Network Architecture

After every Convolution operation nonlinearity needs to be introduced, it is done by Rectified Linear Unit (ReLU) which changes the values of the negative pixel to zero in the output of the convolution layer that is feature [27]. It is given by the equation: $f(x) = \max(x,0)$.

Pooling Layer: It reduces the dimensionality of the convolution layer output that is each feature map. Among various pooling types, here max pooling is implemented which takes the largest element from the matrix of the rectified feature map depending on the size specified.

Fully Connected Layer: After the combination of CL and PL, there is FCL, which classifies the input image to correct class using the high-level features from CL/PL.

Softmax Activation Function: It is the activation function in the last layer, which flattens the outputs from each unit to the values from 0 to 1. The softmax function is given by the equation 2 below, where exponential value of each element of z is taken, j is the index for output unit.

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \quad \text{for j=1…K} \qquad (2)$$

Categorical Cross Entropy Function: It measures the performance of a classification model by calculating the loss at the output layer. There are two distributions m(x) and n(x) which are the true distribution and the estimated distribution respectively. m(x) and n(x) are defined over x which is the discrete variable. It is given by:

$$H(m,n) = -\sum_{\forall x} m(x)\log(n(x)) \qquad (3)$$

Backpropagation: The error which is calculated at the output layer, needs to be reduced. The Backpropagation algorithm tries to reduce the error by going backward through the layers by changing the values of weights and biases. This is done by gradient descent technique.
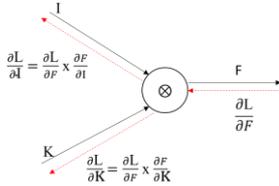
Fig 3. Convolution operation of input I and filter K to get feature map F as output

As shown in figure 3, input is I, filter is K and after applying convolution of K on I, output F is obtained.

L is the loss obtained from the next layer which needs to be back propagated to update the parameters. Partial derivative of L is calculated with respect to F i. e. $\frac{\partial L}{\partial F}$. This is back propagated towards the earlier layer through I as $\frac{\partial L}{\partial I}$ and through F as $\frac{\partial L}{\partial F}$. For every element in the filter K,

$$\frac{\partial L}{\partial K_i} = \sum_{l=1}^{M} \frac{\partial L}{\partial F_l} \ X \ \frac{\partial F_l}{\partial K_i} \qquad (4)$$

This is the convolution between the input and the loss gradient. $\frac{\partial L}{\partial F}$
This is used to change the filter using the learning rate α, given by,

$$K_{updated} = K - \propto \ \partial L / \partial K_i \qquad (5)$$

For every element of I_i,

$$\frac{\partial L}{\partial I_i} = \sum_{l=1}^{M} \frac{\partial L}{\partial F_l} \ X \ \frac{\partial F_l}{\partial I_i} \qquad (6)$$

This is the convolution between filter K and the loss gradient $\frac{\partial L}{\partial F}$. This becomes the gradient of the previous layer.

Convolution Neural Network Algorithm
Step 1: for each fold repeat steps 1 to 11
Step 2: Divide the dataset into training and validation datasets.
Step 3: for the training dataset follow the step 4 to step 10.
Step 4: Initialize parameter like weights and biases and all filters.
Step 5: The input image is preprocessed to get the required size.
Step 6: In forward propagation convolution operation as given in Equation (1) and Pooling operation are executed.
Step 7: For each class, the output probabilities are calculated using softmax activation function given by Equation (2).
Step 8: The total error at the output layer is calculated using Equation (3).
Step 9: In backpropagation step:
The gradients of the error are calculated by taking the partial derivative with respect to all weights using Equation (4).
The parameter values are updated for minimizing the output error using Equations (5) and (6).
Step 10: Repeat step 5 to step 9 for all training set images.

Step 11: Find out the loss and accuracy of the model on the validation data.
Step 12: Save the best weights of the model.
Step 13: Create an ensemble model from the cross-validation models.
Step 14: Find out the loss and accuracy of the ensemble model on the testing dataset.

Data Analysis
For evaluating performance of the CNN model, the following metrics are used.
1) *Cross Validation Accuracy*, the result of five folds CV experiments
2) *Test Accuracy*, the performance of the model on unseen data
3) *Recall, Precision,* and *F1-score*, the classification metrices on the class predictions
4) *Confusion Matrix*, the table which describes the classification performance

C. Cross Validation and Ensemble Method
The Cross Validation is a technique for evaluating machine learning models by training on subsets of dataset and evaluating on complementary subsets of dataset. Literature has shown various models which use cross validation [27, 28, 29]. Kuang Liu et. al. [30], have proposed facial expression recognition with CNN ensemble. An early work in literature by Lars Kai Hansen and Peter Salmon, Neural Network Ensembles [31] uses cross-validation as a tool for optimizing network parameters and architecture. In this CNN model, k-fold cross-validation divides the dataset into k parts of equal size, one part is kept for validation and remaining parts are used for training. Here 5-fold cross validation is used, the following steps are followed:
1) Divide the entire dataset into similar size k folds.
2) Build the model on k-1 folds.
3) Test the model on the kth fold.
4) The error/loss, accuracy is saved for each fold.
5) Repeat until all the folds are used as testing dataset.
6) The average accuracy is calculated using accuracies saved in k folds.

Ensemble learning improves the performance of various models in machine learning by combining several models [32]. In ensemble learning algorithms, a "base learning algorithm" is executed several times to form a vote from the hypotheses which are derived [33]. Anders and Jesper [34] have shown how to estimate the optimal weights of the ensemble members using unlabeled data. In different research work on the neural networks, "ensembles" of neural networks has been studied by several authors. Cross validation was used, by training the individual members of the ensemble on different training sets by holding out some examples for each individual during training. The authors in the paper [35] states that Ensembles are widely used to obtain highly accurate classifiers from the combination of classifiers which gives less accurate results.

Elad and Liar [35] employ CNNs using gradient boosting and selective sampling techniques of ensemble method. David et. al. [36] study the ensemble strategies: boosting, cross-validation and bagging. The multilayer perceptron is used as a base classifier. The CV ensemble and other strategies achieved statistically significant reduction in error.

Retrieval Number F8960088619/2019©BEIESP
DOI: 10.35940/ijeat.F8960.088619
Journal Website: www.ijeat.org

4398

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Li and John [37] states how can ensemble learning be applied to deep learning systems to achieve greater recognition accuracy. In this paper, to increase the accuracy after cross-validation, ensemble method is used where the best weights of the 5 CNN models are used while creating an ensemble model.

IV RESULTS AND DISCUSSIONS

All the experiments in this study were conducted using a laptop with Intel® Core™ i5-7200U CPU @ 2.50 GHz x 4, x64-based processor, 8GB RAM and NVIDIA GetForce GTX 940MX graphics card with 2 GB of dedicated memory and 8GB of DDR4 RAM, GPU.

The dataset consists of 1600 images of .jpg format. Seven classes such as anxiety-stress, depression, depression-anxiety, depression-anxiety-stress, moderate anxiety, normal and severe anxiety are formed using these samples. The training dataset contains 1296 images, testing dataset contains 160 images and validation dataset contains 144 images. The Python libraries Keras and Tensorflow are used to construct the CNN model. Cross validation is used on the dataset to divide into training and validation dataset. CNN requires equal number of samples in the classes, so the results of DASS test were used to obtain various classes, out of which the classes used are shown in the Table II.

Table- II: CLASSES AND NUMBER OF SAMPLES

Class	Samples #
Anxiety Stress (AS)	228
Depression (D)	228
Depression Anxiety (DA)	228
Depression Anxiety Stress (DAS)	229
Moderate Anxiety (MA)	229
Normal (N)	229
Severe Anxiety (SA)	229

The CNN model consists of five CLs where each CL is followed by an activation layer then batch normalization layer. The type of layer, the output shape and the number of parameters is shown in Table III. This model is created for each fold, but the training data is different. ReLU activation function is used in this model. It is used to obtain non linearity after convolutional layer. The batch normalization normalizes the activations of the previous layer for each batch. It provides the mean value to zero and the standard deviation value to 1 for the activations. Max pooling is used to reduce the dimensionality. The second column is output shape, which shows the dimension of the input image 64 x 64 pixels. 'None' is the batch size which is decided during training. The fourth value is the number of filters used.

After cross validation, the five models are combined using CV ensemble method. The summary of the ensemble model is presented in the Table IV. The first column depicts the type of layer, second column shows the output shape, third column shows the number of parameters. The first layer is the input layer with output shape of size 64 x 64 pixels. Other layers are the cross-validation models which are of sequential type. Last layer is the average of these cross-validation models. In second column, seven is the number of classes. Last column shows the connection of the layers.

Table- III: Model Summary

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 64, 64, 8)	224
activation_1 (Activation)	(None, 64, 64, 8)	0
batch_normalization_1 (Batch N)	(None, 64, 64, 8)	32
conv2d_2 (Conv2D)	(None, 64, 64, 16)	1168
activation_2 (Activation)	(None, 64, 64, 16)	0
batch_normalization_2 (Batch N)	(None, 64, 64, 16)	64
conv2d_3 (Conv2D)	(None, 64, 64, 32)	12832
activation_3 (Activation)	(None, 64, 64, 32)	0
batch_normalization_3 (Batch N)	(None, 64, 64, 32)	128
conv2d_4 (Conv2D)	(None, 64, 64, 64)	100416
activation_4 (Activation)	(None, 64, 64, 64)	0
batch_normalization_4 (Batch N)	(None, 64, 64, 64)	256
max_pooling2d_1 (MaxPooling2)	(None, 32, 32, 64)	0

Table- IV: Model Summary of Ensemble Model

Layer (type)	Output Shape	Param #	Connected to
input_1 (Input Layer)	(None, 64, 64, 3)	0	
model_1 (Sequential)	(None, 7)	383455	input_1[0][0]
model_2 (Sequential)	(None, 7)	383455	input_1[0][0]
model_3 (Sequential)	(None, 7)	383455	input_1[0][0]
model_4 (Sequential)	(None, 7)	383455	input_1[0][0]
model_5 (Sequential)	(None, 7)	383455	input_1[0][0]
average_1 (Average)	(None, 7)	0	model_1[1][0]
			model_2[1][0]
			model_3[1][0]
			model_4[1][0]
			model_5[1][0]

The ensemble model is evaluated on the testing dataset, which results in the accuracy of 91.25%. Confusion matrix for the testing dataset is shown in Figure 4. The diagonal values in the matrix shows the true positive values. The accuracy obtained for class anxiety-stress is 90.47%, depression is 82.6%, depression-anxiety is 92%, depression-anxiety-stress is 100%, moderate-anxiety is 84.6%, normal is 100%, severe-anxiety is 90.9%.

Loss function is an important part in any neural network model. It shows how predicted value differs from actual value. The loss and validation accuracy of the cross-validation models on the validation dataset for each fold is given in Table V. The mean accuracy obtained is 81.29%. The ensemble method improves the accuracy as compared to cross validation result.
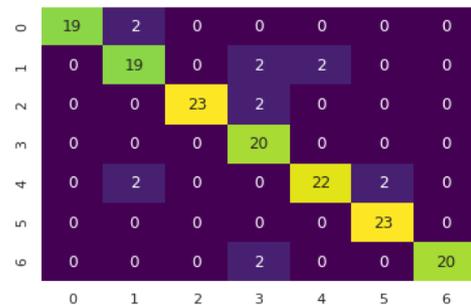
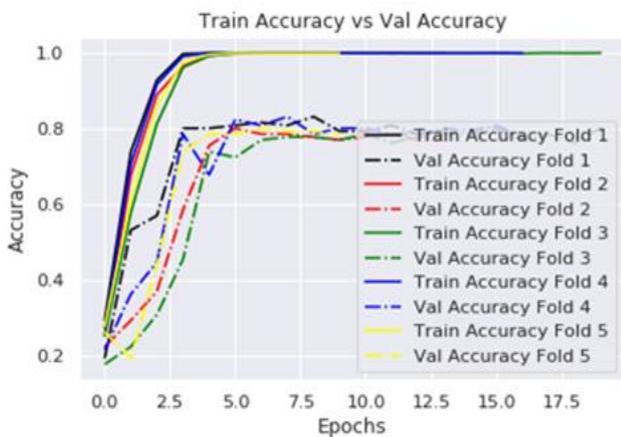**Fig. 4. Confusion Matrix on test data**

Table V: Loss and Accuracy for the CV folds

Fold #	Loss	Validation Accuracy (x100%)
1	0.56	0.8214
2	0.96	0.7571
3	0.51	0.8428
4	0.68	0.8285
5	0.68	0.8142

Figure 5 shows the accuracy for training dataset and validation dataset obtained by using cross validation method. X axis and Y axis show number of epochs and accuracy respectively. The dashed lines, validation accuracies, increases with epoch values.

Table VII: The comparison of CNN based CNN based emotion recognition model with the contemporary methods

Author, reference and year	Modality	Technique/algorithms	Test set accuracy
Hatice Gunes, Massimo Piccardi [11] (2006)	Face and upper body gestures	BayesNet	Face-76.40,Body-89.90
George Caridakis et. al. [12] (2007)	Facial expressions, speech, gestures	Bayesian classifier	Facial-59.6,Speech-70.8, Gestures-83.2
Michael Fairhurst et. al. [16] (2014)	Handwriting samples	kNN,SVM	Happy-80,Stressed-70
kuang Liu et. al. [30] (2016)	Facial expressions	CNN	Surprise-81%
Laurence Likforman-Sulem et. al. [1] 2017	Handwriting, drawing	Random forest approach, DASS scale	D-72.8,A-59.7,S-55
Marco Giannini et. al. [6] 2018	Handwriting samples	BDI-II and analysis by graphologists	80%
Proposed Method	**Handwriting samples**	**CNN, DASS scale**	**91%**

Fig. 5. Accuracy for cross-validation folds

Along with accuracy, other important metric for evaluation are precision-recall. Table 6 shows the values of precision, recall and F1 score for seven classes in our system. The value of precision is one for classes anxiety-stress, depression-anxiety and severe-anxiety, which indicates what proportion of the classes predicted are actually correct. While the value of recall is one for the class normal, which indicates what proportion of handwriting samples indicated normal class actually was the normal class. Precision and recall are represented by F1 score. As can be seen from the figures the convolution layer extracts feature from the input image stage by stage. In the Figure 6(b) the convolutional layer 2 is shown, there are 8 filters of size 3 x 3 applied on the input image of size 64 x 64. The features like edges and curves are extracted initially, while layer 2 extracts the features at the higher-level.

The last convolutional layer extracts the highest-level features. As can be seen from Figure 6(a) to Figure 6(d) more details are learnt, as convolutional layer increase. Table VII shows the comparison of CNN based emotion recognition model with the contemporary methods. As can be seen from the table, for various modalities for emotion recognition the percentage of accuracy is better in the proposed system.

To understand the visualizations of the convolution layers and activation layers, an example is considered. The size of input image is 64 x 64 x 3 (3 refers to the RGB values).

Table VI Precision, Recall and F1-score for all the classes on the test data

Class Number	Class Name	Precision	Recall	F1-score
0	AS	1.00	0.90	0.95
1	D	0.83	0.83	0.83
2	DA	1.00	0.92	0.96
3	DAS	0.77	1.00	0.87
4	MA	0.92	0.85	0.88
5	N	0.92	1.00	0.96
6	SA	1.00	0.91	0.95

The image is scaled to values between 0 to 1. This input image is passed to the CNN model. There are five convolution layers, each followed by activation layers and batch normalization.

Figure 6(a) to Figure 6(d) shows the output of convolution layer1, the output of convolution layer 2, the output of convolution layer 3, the output of convolution layer 4 respectively.
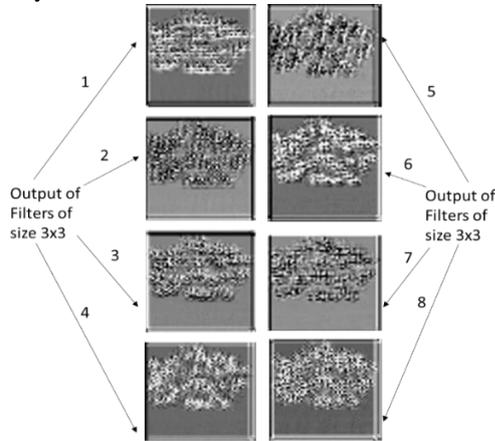

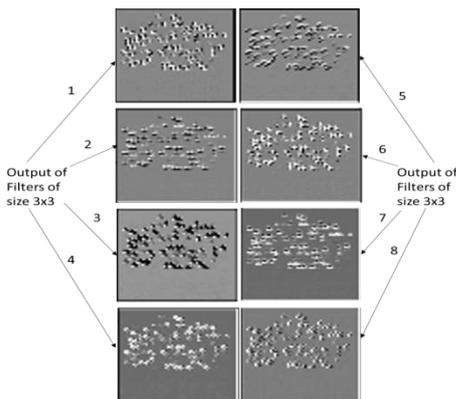**Figure 6b Visualization of Convolution layer 2**

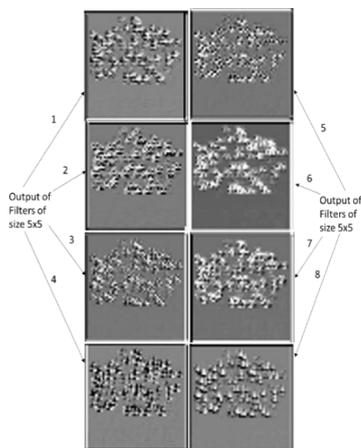**Figure 6a Visualization of Convolution layer 1**

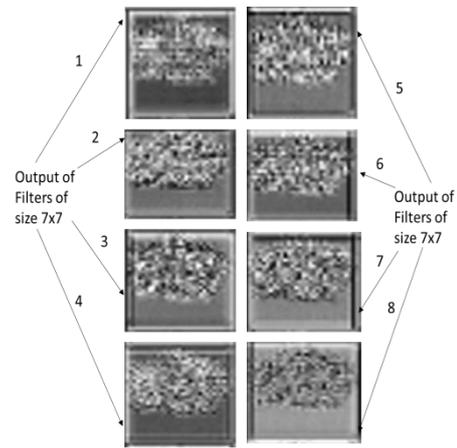**Figure 6c Visualization of Convolution layer 3**

Figure 6d Visualization of Convolution layer 4

V CONCLUSION

This paper has focused on convolution neural network for recognizing emotional state using handwritten text analysis and psychological assessment. For handwriting analysis, database of handwriting samples is created. Convolutional Neural Network algorithm is used for extracting features from the handwriting samples and classification to appropriate class. Psychological assessment is done using the 42 items DASS scale. The training time required by the CNN model is reduced due to the use of GPU. The function train_test_split splits the dataset into two parts to train and test the model, which gives less accuracy. When cross-validation is used, the number of training samples increases which gives a better accuracy. The best weights of the cross-validation models are saved during training. Ensemble method is used to combine the cross-validation models with best weights. The model gives the accuracy of 91.25% on the test data after ensemble method is used. While the accuracy obtained after only cross validation is 81.29%. This shows that the classifier accuracy improves after doing ensemble of the best weights of the cross-validation models. Finally, we conclude that, based on handwritten text, we can predict negative emotions such as depression, anxiety and stress.

REFERENCES

1. Likforman-Sulem, Laurence, Anna Esposito, Marcos Faundez-Zanuy, St´ephan Cl´emenc¸on and Gennaro Cordasco. . "EMOTHAW: A novel database for emotional state recognition from handwriting and drawing." IEEE Transactions on Human-Machine Systems 47, no. 2 (2017): 273-284
2. Gunes, Hatice, and Massimo Piccardi. "Bi-modal emotion recognition from expressive face and body gestures." Journal of Network and Computer Applications 30, no. 4 (2007): 1334-1345
3. Caridakis, George, Ginevra Castellano, Loic Kessous, Amaryllis Raouzaiou, Lori Malatesta, Stelios Asteriadis, and Kostas Karpouzis. "Multimodal emotion recognition from expressive faces, body gestures and speech." In *IFIP International Conference on Artificial Intelligence Applications and Innovations*, pp. 375-388. Springer, Boston, MA, 2007.

Retrieval Number F8960088619/2019©BEIESP
DOI: 10.35940/ijeat.F8960.088619
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

4401

4. Zhao, Huijuan, Ning Ye, and Ruchuan Wang. "A Survey on Automatic Emotion Recognition Using Audio Big Data and Deep Learning Architectures." In 2018 IEEE 4th International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing,(HPSC) and IEEE International Conference on Intelligent Data and Security (IDS), pp. 139-142. IEEE, 2018

5. Kołakowska, Agata, Agnieszka Landowska, Mariusz Szwoch, Wioleta Szwoch, and Michal R. Wrobel. "Emotion recognition and its applications." In Human-Computer Systems Interaction: Backgrounds and Applications 3, pp. 51-62. Springer, Cham, 2014

6. Mutalib, Sofianita, Roslina Ramli, Shuzlina Abdul Rahman, Marina Yusoff, and Azlinah Mohamed. "Towards emotional control recognition through handwriting using fuzzy inference." In Information Technology, 2008.

7. Fairhurst, Michael, Meryem Erbilek, and Cheng Li. "Study of automatic prediction of emotion from handwriting samples." IET Biometrics 4, no. 2 (2015): 90-97.

8. Champa, H. N., and K. R. AnandaKumar. "Artificial neural network for human behavior prediction through handwriting analysis." International Journal of Computer Applications (0975–8887) Volume (2010).

9. Mane, D. T., and U. V. Kulkarni. "Visualizing and Understanding Customized Convolutional Neural Network for Recognition of Handwritten Marathi Numerals." Procedia Computer Science 132 (2018): 1123-1137.

10. Tarunika, K., R. B. Pradeeba, and P. Aruna. "Applying Machine Learning Techniques for Speech Emotion Recognition." In 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1-5. IEEE, 2018.

11. Ebrahimi Kahou, Samira, Vincent Michalski, Kishore Konda, Roland Memisevic, and Christopher Pal. "Recurrent neural networks for emotion recognition in video." In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, pp. 467-474. ACM, 2015.

12. Nieuwenhuijsen, K., A. G. E. M. De Boer, J. H. A. M. Verbeek, R. W. B. Blonk, and F. J. H. Van Dijk. "The Depression Anxiety Stress Scales (DASS): detecting anxiety disorder and depression in employees absent from work because of mental health problems." Occupational and Environmental Medicine 60, no. suppl 1 (2003): i77-i82.

13. Antony, D. John, and O. F. M. Cap. "Personality Profile Through Handwriting Analysis." (2008).

14. Balci, Batuhan, Dan Saadati, and Dan Shiferaw. "HandwrittenText Recognition Using Deep Learning." CS231n: Convolutional Neural Networks for Visual Recognition, Stanford University, Course Project Report, Spring (2017).

15. Zeiler, Matthew D., and Rob Fergus. "Visualizing and understanding convolutional networks." In European conference on computer vision, pp. 818-833. Springer, Cham, 2014

16. Simard, Patrice Y., David Steinkraus, and John C. Platt. "Best practices for convolutional neural networks applied to visual document analysis." In Icdar, vol. 3, no. 2003. 2003.

17. Arel, Itamar, Derek C. Rose, and Thomas P. Karnowski. "Deep machine learning-a new frontier in artificial intelligence research." IEEE computational intelligence magazine 5, no. 4 (2010): 13-18.

18. LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521, no. 7553 (2015): 436.

19. Mitra, Sushmita, and Tinku Acharya. "Gesture recognition: A survey." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 37, no. 3 (2007): 311-324.

20. Sebe, Nicu, Ira Cohen, Theo Gevers, and Thomas S. Huang. "Multimodal approaches for emotion recognition: a survey." In Internet Imaging VI, vol. 5670, pp. 56-68. International Society for Optics and Photonics, 2005.

21. Crawford, John R., and Julie D. Henry. "The Depression Anxiety Stress Scales(DASS): Normative data and latent structure in a large non-clinical sample." British journal of clinical psychology 42, no. 2 (2003): 111-131

22. Giannini, Marco, Pietro Pellegrini, Alessio Gori, and Yura Loscalzo. "Is Graphology Useful in Assessing Major Depression?." Psychological reports (2018): 0033294118759667.

23. Dazzi, Carla, and Luigi Pedrabissi. "Graphology and personality: an empirical study on validity of handwriting analysis." Psychological reports 105, no. 3 suppl (2009): 1255-1268

24. Lovibond, Peter F., and Sydney H. Lovibond. "The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories." Behaviour research and therapy 33, no. 3 (1995): 335-343.

25. Lovibond S. H. Depression Anxiety Stress Scales, www2.psy.unsw.edu.au/Groups/Dass/.

26. Brown, Timothy A., Bruce F. Chorpita, William Korotitsch, and David H. Barlow. "Psychometric properties of the Depression Anxiety Stress Scales (DASS) in clinical samples." Behaviour research and therapy 35, no. 1 (1997): 79-89.

27. Agarap, Abien Fred. "Deep learning using rectified linear units (relu)." arXiv preprint arXiv:1803.08375 (2018).

28. Kagaya, Hokuto, Kiyoharu Aizawa, and Makoto Ogawa. "Food detection and recognition using convolutional neural network." In Proceedings of the 22nd ACM international conference on Multimedia, pp. 1085-1088. ACM, 2014.

29. Palaz, Dimitri, and Ronan Collobert. Analysis of cnn-based speech recognition system using raw speech as input. No. REP_WORK. Idiap, 2015.

30. Liu, Kuang, Mingmin Zhang, and Zhigeng Pan. "Facial expression recognition with CNN ensemble." In 2016 international conference on cyberworlds (CW), pp. 163-166. IEEE, 2016.

31. Hansen, Lars Kai, and Peter Salamon. "Neural network ensembles." IEEE Transactions on Pattern Analysis & Machine Intelligence 10 (1990): 993-1001

32. Dietterich, Thomas G. "Ensemble methods in machine learning." In International workshop on multiple classifier systems, pp. 1-15. Springer, Berlin, Heidelberg, 2000.

33. Hansen, Lars Kai, and Peter Salamon. "Neural network ensembles." IEEE Transactions on Pattern Analysis & Machine Intelligence 10 (1990): 993-1001.

34. Dietterich, Thomas G. "Ensemble learning." The handbook of brain theory and neural networks 2 (2002): 110-125.

35. Walach, Elad, and Lior Wolf. "Learning to count with cnn boosting." In European conference on computer vision, pp. 660-676. Springer, Cham, 2016.

36. West, David, Scott Dellana, and Jingxia Qian. "Neural network ensemble strategies for financial decision applications." Computers & operations research 32, no. 10 (2005): 2543-2559.

37. Deng, Li, and John C. Platt. "Ensemble deep learning for speech recognition." In Fifteenth Annual Conference of the International Speech Communication Association. 2014.

AUTHORS PROFILE

Dr. Seema Kedar is working as a Professor and Head of Computer Engineering Department in Rajarshi Shahu College of Engineering, Pune. More than 35 papers published in various National and International Journals; published various books on Database and Object-Oriented Programming; Published 7 patents and 4 copyrights.

Shilpa Rokade is a second-year student of Master of Computer Engineering at JSPM's Rajarshi Shahu College of Engineering, Tathawade, Pune, affiliated to Savitribai Phule Pune University. She is graduated with Bachelor of Information Technology from G. H. Raisoni College, Nagpur. She has over 3 of industry experience and 4 years of teaching experience. She has published 2 research papers in reputed national and international conferences and filed 1 copyright. Her research interests are Machine Learning, Deep Learning and Artificial Intelligence.

Retrieval Number F8960088619/2019©BEIESP
DOI: 10.35940/ijeat.F8960.088619
Journal Website: www.ijeat.org

4402

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication