SRCNN: Stress Recognition Using Face Images Using Convolutional Neural Network

E.Jayaramnaveen[1], M.Aswanth Kumar[2], R.Gowtham[3]
[1, 2, 3]Dept of ECE
[1, 2, 3]Sri Krishna College of Engineering and Technology, Coimbatore

Abstract- *Human psychological tension and human emotion are inextricably linked. The connection between stress and emotions is the key to understanding human behavior in computational psychology. Human-computer contact will be more natural if computers can understand and recognize nonverbal human communication such as emotions. Although many methods to recognizing human emotions based on facial expressions, voice, or textual data have been proposed, it is essential to improve the accuracy and robustness of the emotion identification system. Based on the findings, the method presented in this article will identify the emotion and offer appropriate suggestions. Several emotions will be identified using real-time pictures of human faces, including sorrow, happiness, rage, fear, surprise, and neutral mood. Deep learning techniques such as CNN will be used to record detailed facial movements and identify suitable emotion. Single and multi-layered networks will be evaluated to find the best classifiers for detecting certain emotions. Pictures depicting faces in various resolutions, as well as images including areas of the lips and eyes, will be included. We proposed the SRCNN Method in this article. A neural network cascade will be suggested based on the test findings. The cascade will identify six fundamental emotions as well as neutral expression. Depending on the resulting feeling, the module will recommend songs or activities to cheer up a person and improve his or her mood. Sentiment analysis using CNN will assist in recommending certain activities linked with a specific mood. As a result, it will assist a person in being positive and cheerful. For the emotions studied, the findings show that the system based on facial expression outperforms the systems based on textual data or auditory information. For 5 layer CNN and 65 iterations, our technique obtained 87.76 % accuracy.*

Keywords- Deep Learning, CNN, SRCNN, Stress Recognition

I. INTRODUCTION

Stress Management has become a major issue in the twenty-first century. It is a set of techniques aimed at controlling a person's condition of stress, particularly chronic stress, with the goal of improving day-to-day living. Stress causes a variety of physical and mental illnesses, which vary depending on the individual's environmental circumstances [1]. These may include both health problems and depression. One of the most significant ways for people to communicate their feelings is via their facial expressions. Though no words are said, there is much to learn about the signals and indications we give and receive via nonverbal communication, such as facial expressions [3]. The practice of stress management is regarded as one of the cornerstones to happiness, satisfaction, and success in life. Although life may provide a variety of difficulties that might be tough to confront, stress management offers a variety of methods for managing anxiety and maintaining general well-being. Although humans can identify the facial expressions of people they know without trouble or delay, computer expression identification remains a problem [4]. Several studies have been conducted in recent years on feature extraction, face recognition, and techniques for expression categorization, but developing an automated system that performs this job is challenging [9]. To investigate emotions, this research employed a method that used Convolutional Neural Networks (CNN) over facial pictures. Our system receives a webcam picture as input, and then we use CNN to predict the emotion label, which should be one of the following: anger, fear, happiness, or sorrow. In this paper, we continue the line of research that uses convolutional neural networks (CNNs) and image recognition. Our goal is to analyze emotions using pictures of the individual. An image dataset is fed into the convolution neural network model. Hundreds of training datasets are used to train it. The pictures used for training are essentially Google photographs that have been labeled according to their kind. We get four distinct kinds of emotions after passing through various layers of CNN for image processing: angry, fearful, joyful, and sad. This approach offers to assist stressed persons by allowing others near to them to detect when they are in the grip of depression, assist them in breaking free from it, and generate natural good emotions to decrease their stress. Our project may contain numerous modules, but for this article, we have completed three. These are the creation of an image data set, the creation of a model trained on this data set, and the use of a camera-captured picture as the testing image for the output. Essentially, we are hoping to develop a platform that can operate on a device and allow it to function based on a person's mood by processing its Macro and Micro expressions, as stated in [8].

Figure 1 Sample Image dataset

As a result, we can identify a person's feelings via their facial expressions and provide suitable recommendations to improve their mood. The purpose of this article is to provide an AI/ML module that takes a picture of a human face as input and returns the emotion expressed by the face. Single and multi-layered networks will be evaluated to find the best classifiers for detecting certain emotions. Pictures with varying resolutions of faces, as well as images with areas of the lips and eyes, will be provided. If computers can detect these emotional inputs, they will be able to offer users with clear and relevant help in ways that best fit the user's requirements and preferences, keeping them happy and satisfied [10].

The remainder of the paper as follows, Section II represents existing research methods, Section III proposed system model, section IV presents the results and discussion and finally concludes represents the section V.

II. BACKGROUND STUDY

Boussaad, L., & Boucetta, A. **[2]** Nowadays, “ageing” is a difficult issue for face recognition algorithms. The difficulty is exacerbated by age-related biological changes, which may result in substantial differences in face features between two pictures of the same individual taken at various ages. Because the face is the most impacted by ageing, there is an increasing need to extract strong face characteristics for age-invariant face identification, especially in the presence of significant age variations between the same person's face pictures. The purpose of this article is to investigate the efficacy of deep-learning-based techniques as a feature extraction tool for age-invariant face recognition. The authors utilize K-NN, discriminant analysis, and support vector machines (SVM) classifiers to assess five famous pre-trained deep-convolutional neural network (CNN) models, namely AlexNet, GoogleNet, Inception V3, ResNet50, and SqueezeNet, using a widely used face-aging database, FG-NET. A statistical analysis test is then run to validate the statistical significance of the findings. Experiment findings on this database indicate that utilizing Convolutional Neural Networks (CNN) for face identification throughout age progression has promise.

Hills, P. J. et al. [4] it was discovered that being watched had a detrimental impact on facial recognition. This is particularly noticeable while learning new upright faces (or inverted familiar faces). The authors hypothesise that this is due to social observation making participants nervous and the Yerkes-Dodson rule leading individuals to perform badly in a moderately tough activity. The scientists, however, found no evidence that physiological stress influenced the link between observation and facial recognition. As a result, the authors believe that observation has a direct impact on how individuals encode faces. This has a negative impact on identifying moderately challenging face recognition tasks.

Jeon, T. et al. [5] the authors present a stress recognition system that makes use of facial pictures and landmarks. As a consequence of the trial, the scientists verified that utilizing facial landmarks enhanced stress recognition ability even more. Because they enable you to better comprehend eye, mouth, and head movements, facial landmarks are better at detecting stress. The scientists also discovered that utilizing a grey facial picture of the right size enhanced performance by better detecting stress-related information.

Jahanbin, S. et al. [6] the authors offer a novel set of 3D face recognition characteristics that combine multiresolution surface representation with PCA. Range data from a hexagonal region of interest is dissected using Barycentric wavelet kernels in this method. At each resolution level, principal component analysis is performed to reduce the dimensionality of the obtained coefficients (PCA). On a set of 206 range images, recognition rates as high as 94.17 percent were achieved using just 35 PCA coefficients in a simple nearest neighbor classifier.

Lu, H. et al. [7] the authors proposed a multilinear method for face identification based on picture texture. The LBP algorithm first extracts texture characteristics from face pictures. The derived texture characteristics may effectively represent face pictures and decrease the size of the original data. The HOOI algorithm is then used to do face analysis and recognition. Through the trials, the authors may also get the optimal parameters for feature selection as well as the optimum and reduced factor-specific mode matrix.

III. STRESS RECOGNITION USING IMAGES USING CNN

Our algorithm has two major steps: facial expression identification and stress detection based on the decoded emotions. We utilized Convolutional Neural Network (CNN) to determine the likelihood that a facial expression represents a certain emotion for all emotions. (As shown in figure).

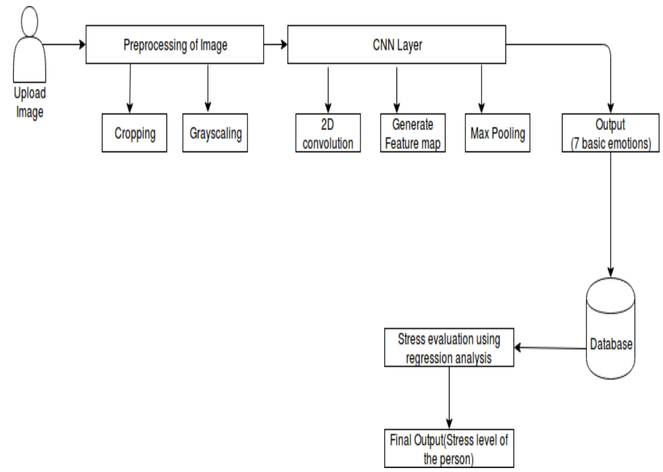

Figure: 2. Architectural diagram of our model

3.1 DATASET

We have used FER2013 dataset from the Kaggle Facial Expression Recognition Challenge to train our model. The dataset consists of 35,887 grayscale images, each image of 48-by-48 pixel labeled with one of the 7 emotion categories: Angry, Disgust, Fear, Happy, Sad, Surprise and Neutral. We used 80% of the dataset for training and the remaining 20% for testing.

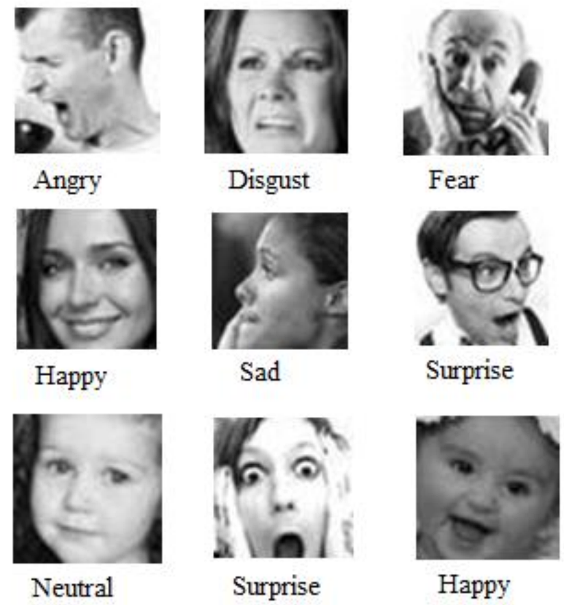

Figure: 3 sample of images from FER2013 dataset with their corresponding emotions

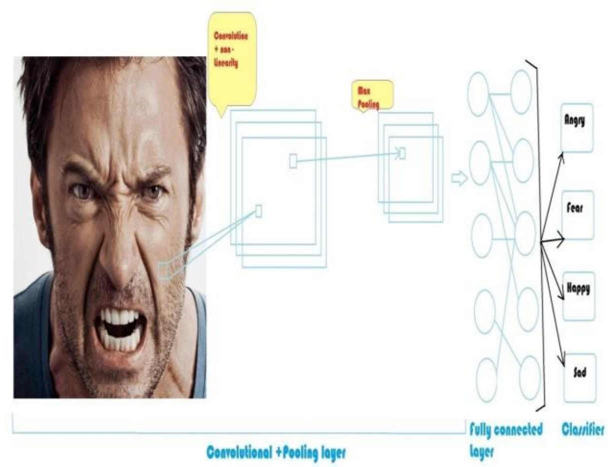

Figure 4 Working of CNN (Convolutional neural network)

3.2 STEP 1

We preprocessed the obtained picture in the input layer as part of emotion recognition. Cropping and grey scaling of pictures were part of the preprocessing procedure in order to produce images with a normalized size or intensity. The preprocessed picture was then put into a NumPy array.

The NumPy array was then sent through convolutional layers, which included three convolution layers and a set of filters, to create the feature maps. In lieu of completely linked layers, we used a mix of depth-wise separable convolutions and residual modules. The ADAM optimizer was used to train the model. We utilized Global

Average Pooling to entirely eliminate any fully linked layers. This was accomplished by using the same number of feature maps in the final convolutional layer as the number of classes and applying a softmax activation function to each reduced feature map.

Finally, in the output layer, the softmax function displayed the result as a probability for each emotion class.

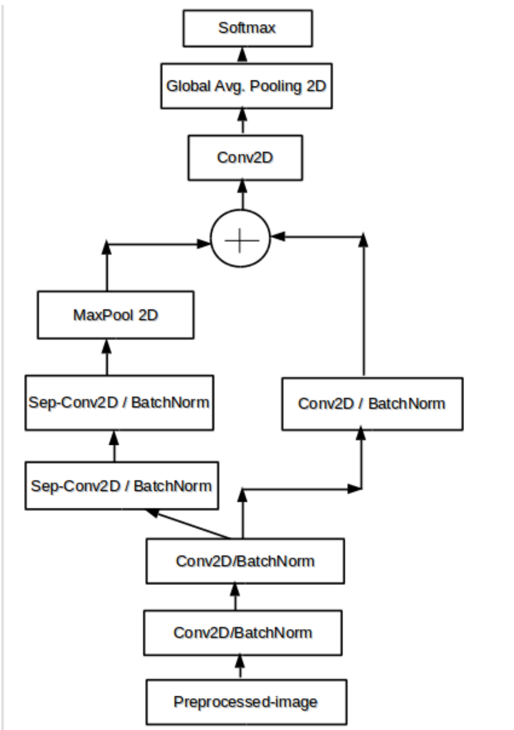

Figure: 5. CNN Implementation

3.3 STEP 2

We assessed stress levels after deciphering emotion information from facial expressions and performed a poll among individuals that connects emotion degrees to stress levels. The survey was carried out without the use of any pictures. Respondents were given percentages of emotion intensity ranging from 0 to 100 in increments of 20, and they were asked to map each level of the fundamental emotion with a stress level ranging from 0 to 9. We got 108 replies in all.
Following the completion of this survey, we utilized the survey findings as well as the decoded emotion information data to forecast a regression model that best fits the stress and emotion data. For forecasting stress levels using emotion degrees or percentages, we presented five alternative linear and non-linear regression models. The best model was chosen based on criteria such as quality of fit and root mean square errors. We discovered that the logarithmic fit was the best model among all of them, and it also followed the Weber-Fechner rule. The logarithmic model equations were then given by finding the coefficients of unknown variables using regression analysis.

Finally, the above-mentioned method may be used to assess visually detectable stress levels (in terms of likelihood) as observable from a subject's face in terms of seven fundamental emotions, and the emotions can be decoded using information from facial muscle activity.

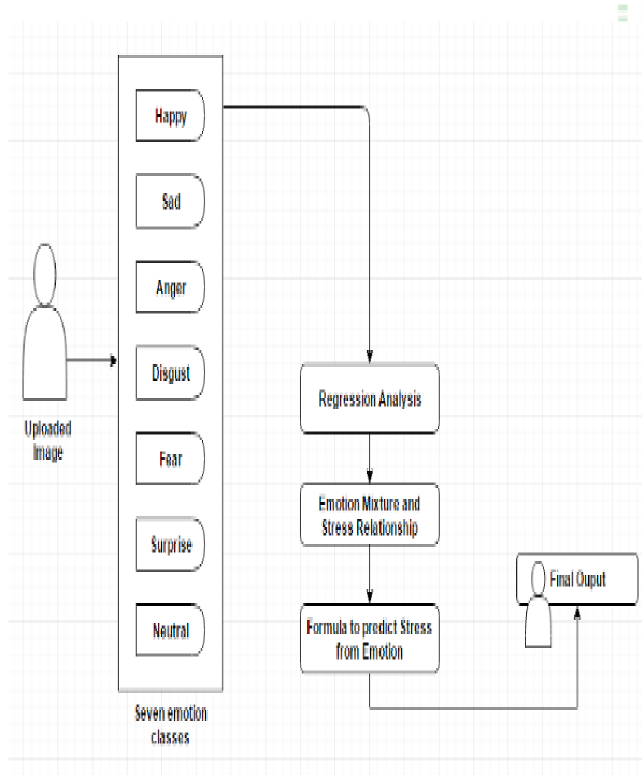

Figure: 6. Stress Evaluation Model

3.4 VERY LOW-STRESS LEVEL

Very low stress indicates that a person is healthy. It indicates that the person is happy. Hence in the output screen, a happy emoji is also displayed.

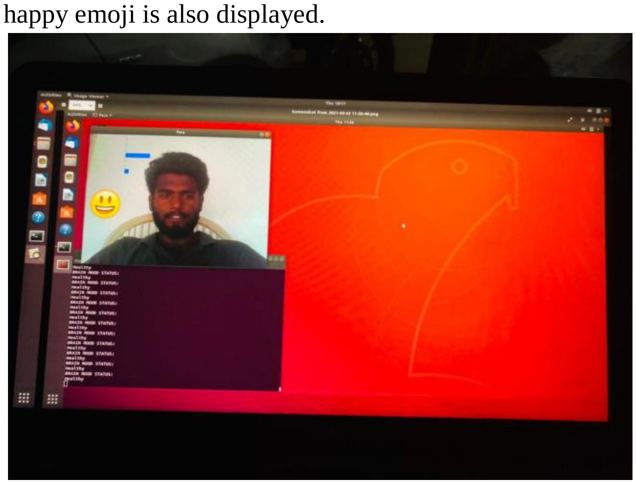

Figure 7: Output Screen when the person is happy and healthy

3.5 LOW STRESS LEVEL OR NORMAL

Low stress level indicates a normal mood. Every person has a normal stress level which is an indication the state of a normal person. This low stress level indicates a normal level.

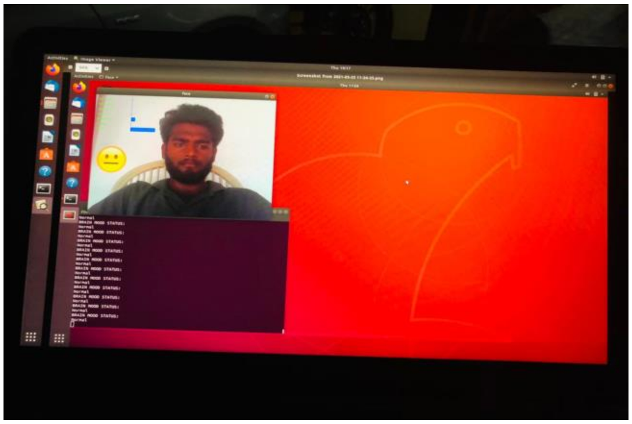

Figure 8: Output screen of normal stress level

3.6 Medium stress Level

Medium stress level indicates some kinds of mood swings in a person. These changes can be easily identified and then also can be changed easily. This mood swing indicates mild tension.

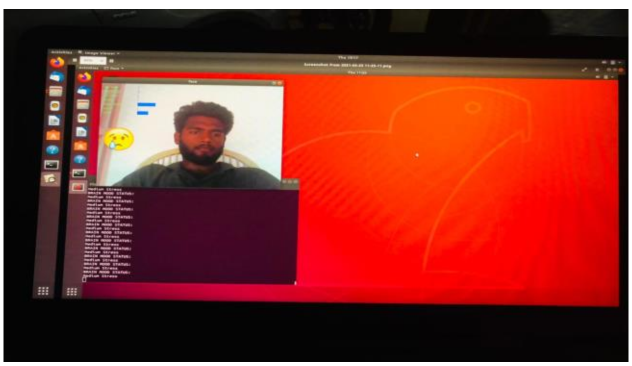

Figure 9: Output screen of medium stress

3.7 HIGH STRESS LEVEL

High Stress level is shown in this figure. The high stress level indicates that a person is in severe trouble or in need of help. This person must free his/her mind by doing activities that make them happy.

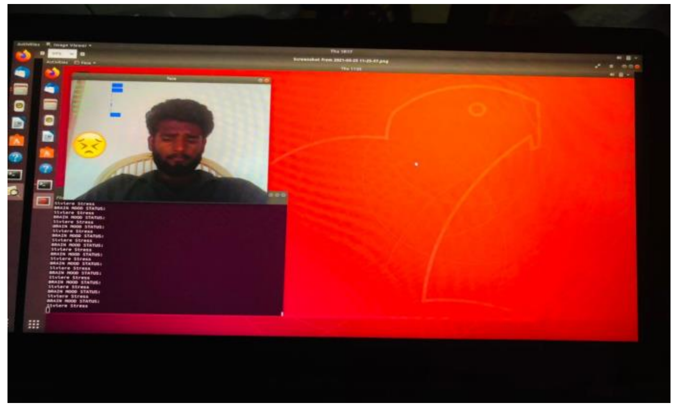

Figure 10 Output screen of high-stress level

IV. RESULTS AND DISCUSSION

We are aware that the following limitations may exist in our present implementation of the stress detection model. If someone attempts to create false emotions, or if a beard or spectacles are utilized in the picture, our model may produce incorrect results. Also, since the training on the pictures is limited, emotions such as fear, sadness, or contempt may not be properly recognized. The stress findings are also based only on a single poll, which may occasionally provide inaccurate stress likelihood.

Figures 11 and 12 depict the emotion categorization task and the detection of stress levels. These figures show the matching picture, as well as its identified emotion label, likelihood, and stress probability.

The final curve fitting and parameter estimates of the logarithmic model for all emotions were performed in Python using all 108 data per emotion.

Figure 13 depicts the logarithmic curve for Stress vs. Anger, which indicates that the rise in stress level is sharp from 0 to 20%, but from 20% onwards, the slope is extremely mild and positive. The initial steepness may be represented as a person's angry outburst, which will grow with the continuous existence of the cause of Anger, but not at the same pace as it was originally. In real-world situations, this eruption is occasionally apparent but is often repressed at will owing to environmental constraints or personal tendencies.

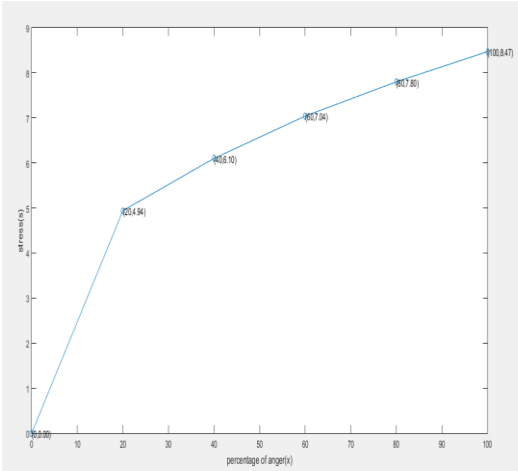

Figure: 11. Stress vs. Anger logarithmic curve

Figure 12 shows the logarithmic curve for Stress vs. Disgust which is close to linear without and sharp uprisings or spikes. The stress levels rises for Disgust smoothly but with continued presence of the causal factors it can reach a moderately high peak value of around 6.

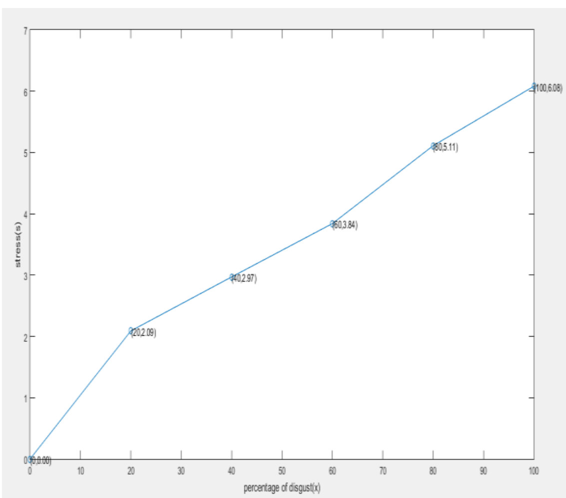

Figure: 12. Stress vs. Disgust logarithmic curve

Figure 13 depicts the logarithmic curve for Stress vs. Fear, which displays similar characteristics to anger but is greater in magnitude. The chart shows that the stress level rises sharply from 0 to 6 in the first 20% period. This may be described as the unexpected emergence of a causal factor in one's surroundings or in one's thinking that causes dread. After 20%, with the persistence of the causative factor, the stress level for dread gradually increases, reaching a maximal stress level of approximately 8.5.

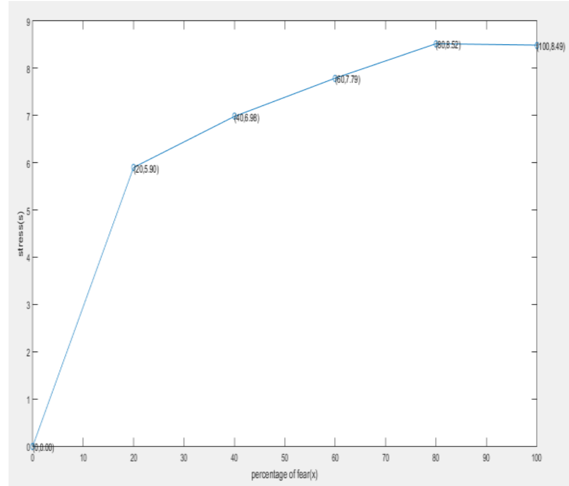

Figure: 13. Stress vs. Fear logarithmic curve

Figure 14 shows the logarithmic curve for Stress vs. Happy is the most unique among all the stress curves relating individual basic emotions to stress. In the interval of 0 to 40 percent it is at a stress level of 0 and in beyond 40 percent it almost resembles a straight line with a very moderate positive slope. For Happiness, the peak value reached is just over 2.5, much lower than any other basic emotion.

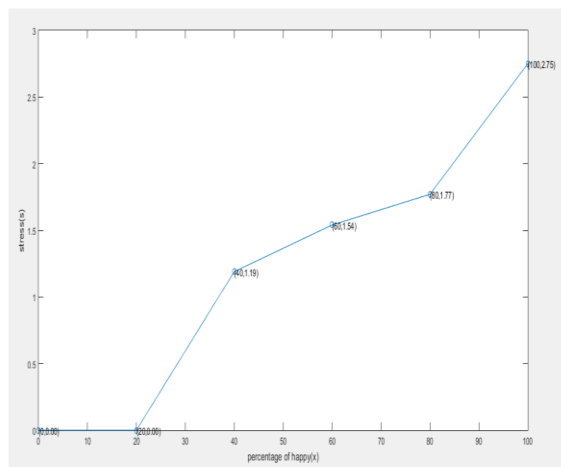

Figure: 14. Stress vs. Happy logarithmic curve

Figure 15 shows the logarithmic curve for Stress vs. Sad and it is easily observable that the peak stress level for sadness reaches near 8. In the initial 20 percent interval it steeply raises 4 units and climbs steadily to the peak. An example situation for the initial steep climb can be given. Suppose a person suddenly comes to know that a close relative died, he experiences the initial shock of sadness, hence the steep climb of stress level.

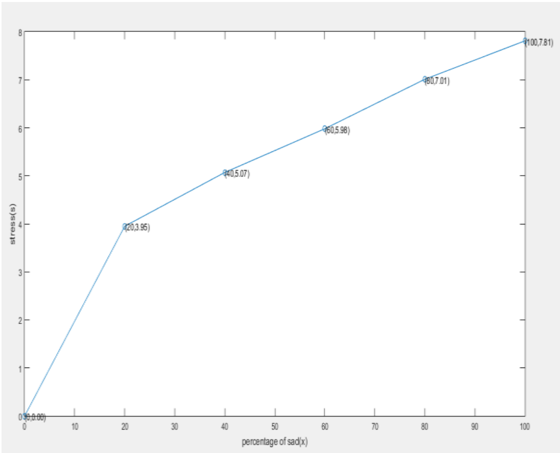

Figure: 15. Stress vs. Sad logarithmic curve

Figure 16 shows the logarithmic model curve for Stress vs. Surprise appears similar to that of Anger but for Surprise the peak stress is marginally lower. Surprise is the emotion that ranks third in severity of stress response following fear and Anger.

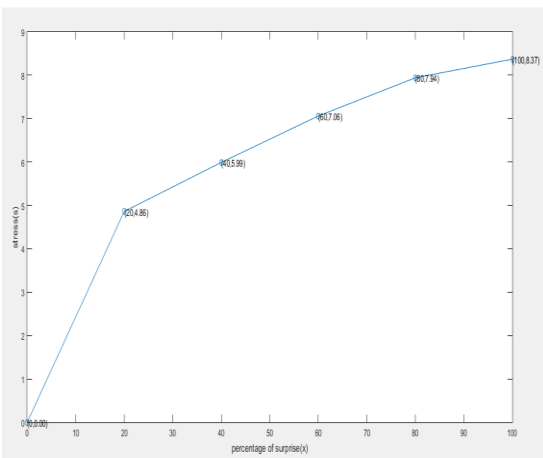

Figure: 16. Stress vs. Surprise logarithmic curve

Figure 17 shows the logarithmic model curve for Stress vs. Neutral appears to be steady and constantly increasing, indicating that the stress levels of a neutral person are in the range of 4-5.

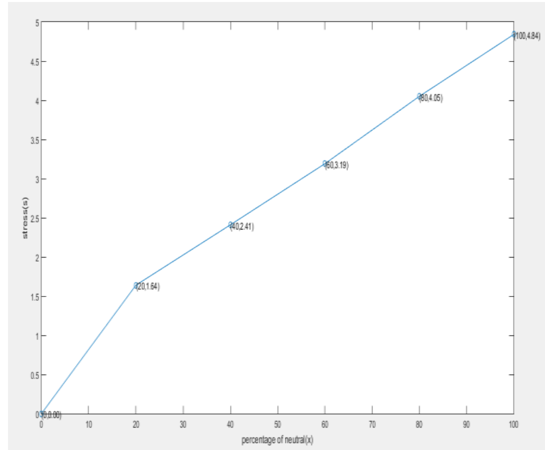

Figure: 17. Stress vs. Neutral logarithmic curve

V. CONCLUSION

The SRCNN technique for stress detection from facial expression recognition using CNN was described in this article. As shown by the findings, when compared to the other techniques that utilize the same facial expression database, our method employs CNN, which works better for pictures and provides a simpler solution. The suggested method has an accuracy of 87.76 percent.

We want to try this technique in additional databases and conduct cross-database validation in the future. In addition, by training our model on a GPU or in the cloud, we may improve its accuracy. Furthermore, we want to broaden our study to include real-time stress monitoring of employees in companies.

REFERENCES

[1] Andre Teixeira Lopes, Edilson de Aguiar, Thiago Oliveira-Santos, A Facial Expression Recognition System Using Convolutional Networks.

[2] Boussaad, L., & Boucetta, A. (2020). Deep-learning based descriptors in application to aging problem in face recognition. Journal of King Saud University - Computer and Information Sciences. doi:10.1016/j.jksuci.2020.10.002

[3] D. C. Ali Mollahosseini and M. H. Mahoor. Going deeper in facial expression recognition using deep neural networks. IEEE Winter Conference on Applications of Computer Vision, 2016.

[4] Hills, P. J., Dickinson, D., Daniels, L. M., Boobyer, C. A., & Burton, R. (2019). Being observed caused physiological stress leading to poorer face recognition. Acta Psychologica, 196, 118–128. doi:10.1016/j.actpsy.2019.04.012

[5] Jeon, T., Bae, H., Lee, Y., Jang, S., & Lee, S. (2020). Stress Recognition using Face Images and Facial Landmarks. 2020 International Conference on Electronics, Information, and Communication (ICEIC). doi:10.1109/iceic49074.2020.9051145

[6] Jahanbin, S., Choi, H., Bovik, A. C., & Castleman, K. R. (2007). Three Dimensional Face Recognition using Wavelet Decomposition of Range Images. 2007 IEEE International Conference on Image Processing. doi:10.1109/icip.2007.4378912

[7] Lu, H., Chen, H., & Chen, Y. (2008). Multilinear analysis based on image texture for face recognition. 2008 19th International Conference on Pattern Recognition. doi:10.1109/icpr.2008.4761491

[8] F. Dornaika, B. Raducanu, “Efficient facial expression recognition for human robot interaction”, In Proceedings of the 9th International Work-Conference on Artificial Neural Networks on Computational and Ambient Intelligence, San Sebastián, Spain, 20–22 June 2007; pp. 700–708.

[9] S. Hickson, N. Dufour, A. Sud, V. Kwatra, I. A. Essa, “Eyemotion: Classifying facial expressions in VR using eye-tracking cameras”, arXiv 2017, arxiv:1707.07204

[10] M. A. Assari, M. Rahmati, “Driver drowsiness detection using face expression recognition”, In Proceedings of the IEEE International Conference on Signal and Image Processing Applications, Kuala Lumpur, Malaysia, 16–18 November 2011; pp. 337–341.