Stress detection using adaptive threshold methodology

Srivasthasva Srinivas Machiraju
Department of Computer Science and Engineering
Amrita Vishwa Vidyapeetham
Amritapuri, India
machirajusrinivas1999@gmail.com

Nikhil Konijeti
Department of Computer Science and Engineering
Amrita Vishwa Vidyapeetham
Amritapuri, India
konijetinikhil@gmail.com

Akhil Batchu
Department of Computer Science and Engineering
Amrita Vishwa Vidyapeetham
Amritapuri, India
akhilbatchu123@gmail.com

Naveen Tata
Department of Computer Science and Engineering
Amrita Vishwa Vidyapeetham
Amritapuri, India
tatanaveen18@gmail.com

***Abstract*—A stress detection system that can identify if a person is suffering from stress or strain over a long period and advises him for a breather to reduce stress. Earlier stress detection was carried out using a defined threshold but using a fixed threshold value is not appropriate as the threshold varies from person to person. In this manuscript, we are using an adaptive threshold methodology to overcome this issue. The proposed method uses a derived set of equations which accepts facial emotions as parameters and generate the threshold value for stress detection. The methodology followed in this manuscript is compared with static threshold methodology to find the superior process.**

***Index Terms*—Stress Detection System, Sentiment Analysis, Machine learning, Convolutional Neural Network**

I. Introduction

According to [1], Stress is an inclination of mental weight and tension. Low degrees of stress may be wanted, helpful, and even suitable for well-being of the individual but elevated levels of stress could result in biological, mental, and social issues and even causes serious harm to individual. Stress may be caused mainly in two ways i.e. either through extrinsic environment or through the intrinsic perceptions of individuals, which can cause negative emotions that results in anxiety and pressure that deteriorates the health condition.
Stress can be broadly classified into 2 categories:

1. Short-term or Acute stress
2. Long-term or Chronic stress

The effect on human well-being due to short-term stress is less contrasted to long-term stress, and it sometimes helps in generating positive energy which allows us to complete the challenge. Short-term stress is caused by the emission of hormones that alert the user. However, long term stress is developed over a constant period and can negatively affect your body. A person who is under stress for an extended period goes through the following phases [2]-

1. Alarm Reaction
2. Stage of Resistance
3. Exhaustion

These stages are considered as various set of reactions an individual experience and this is considered as General Adaptation Syndrome (GAS) [2]. On the off chance that this cycle proceeds, the exorbitant degree of hormones during the resistance phase may upset homeostasis and damage the internal organs leaving the organism liable to diseases. During these stages, the body's energy resources are completely depleted, and breakdown happens.

Chronic stressors have a more drawn out span and are not promptly distinguished as stressors since they are regularly vague and intangible. They possess a serious health risk if they are not recognized and properly managed. According to [8], it is evident that facial expressions can be used to identify long term stress. Our work proposes a derived approach which includes set of equations which accepts one among the set of expressions; anger, fear, sad, disgust, happy, surprise along with neutral expression as proposed by Ekman [3] as input and monitor stress.

This whole manuscript has a total of 5 sections, which includes Related Work, Proposed Work, Result Analysis, Conclusion and References. Related work summarizes the existing work done on stress detection system and how our methodology varies from other, Proposed work is subdivide into multiple sections which includes System overview and Approach and Methodology. System overview details the working architecture of our model. Approach and Methodology section explains the complete working process of adaptive threshold methodology. Result analysis section mainly describes the comparison between static threshold methodology and dynamic threshold methodology. Finally, the conclusion section details our research and the future work.

II. Related work

The ideology of monitoring human stress is used in various applications. In [4] the idea is to mainly focus on road safety and according to [5], [6] emotional status (e.g. stress, impatience) of the driver may as well endanger the safety. As described in [7] if proper stress detection methodologies are not followed it might lead to problems in work life and personal life of the individual.

Lot of research is going on to apply the concepts of Machine learning, Image processing in the field of medical science to make the life easy for common man. In [11] random forest model was used to predict the possibility of coronary heart disease considering multiple risk factors. In [12] Alzheimer's disease detection was done using Gaussian models and support vector machine whereas in [13] convolutional neural networks were applied to automate the process of detecting the malaria infected red blood cells. Significant amount of work was done in the areas of monitoring and managing stress. In [10] stress detection was mainly carried out using physiological signals monitored by sensors. The system followed in [10] has a total of three stages which involves setting up the sensors, using the data pre-processing techniques on the extracted data and passing the processed data to a learning system. As per the work done in [14] stress monitoring is done using the data generated from ECG, galvanic skin response and accelerometer. But using all these physical devices is difficult in a real work scenario. So, we propose a system that can use the facial emotions of the user as its input to monitor his stress. The idea of considering facial expressions for stress detection makes the model applicable to the real world.

III. Proposed Work

A. System Overview

Fig. 1 details the complete overview of our working module. The total system is divided into two individual modules. They are 1. Facial emotion detection module and 2. Stress detection module.

1) Facial Emotion Detection: Classifying facial expressions is very crucial because all the further calculations are based on results produced from those facial emotions. The model architecture shown in Fig. 2 is implemented, trained and evaluated on the Cohn-Kanade-images dataset [18], [19] with a different accuracy calculation. The model architecture implemented was adopted from [15]. Instead of using seven class classification, three class classification is applied. Positive emotions which include happiness and surprise are considered as one class, and negative emotions which include fear, sadness, disgust, angry is considered as one class, and the neutral expression is considered as one class. The model has a total of four convolutional layers, six dropout layers, four max pooling layers, one flatten layer and three dense layers. This model has a total of 4,470,000 trainable parameters.

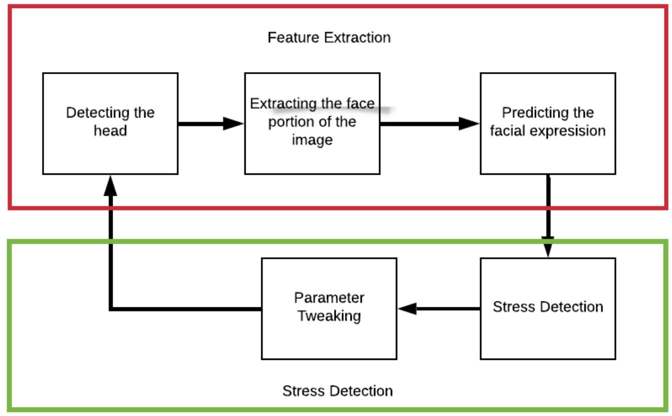

Fig. 1. Software modules of the Adaptive stress detection system. Facial emotion detection module and Stress detection module

2) Stress Detection: Once the facial emotions are predicted the output is used for monitoring the stress. If the system detects that the stress level of the person under surveillance exceeds a threshold value, it alerts the user and the parameters are reinitialized. The way in which parameters are reinitialized is described in the further sections. This process of analyzing the facial emotions, monitoring the stress and tweaking the parameters follows an iterative process. If the system is in running state it keeps on updating itself. The longer it runs the better it will be able to analyze the user behavior and the better it tunes the parameters.

B. Approach and Methodology

1) Initial Setup: The input is divided into individual frames and using Open CV (Head Cascade) the specific portion of the image is extracted, and the corresponding facial emotions are predicted. Since neural networks have outperformed many of the conventional methods, we used the convolutional neural network, and one of the seven Emotions is being predicted and stored for further usage [17]. This process continues until the number of predicted emotions exceeds the window size, the state of the user is predicted in this timestamp using the following calculations.

$$State = data.count('Emotion') * Emotionweight$$

$$TotalStress = TotalStress + State$$

; where data is a list that stores the captured expression within a window, emotion weights are the weights allotted for each of the predicted emotion, and total stress is a variable that's being updated after each window. Weights to the emotions are allotted in such a way that the positive emotions which includes happy and surprise updates the total stress in the negative direction and negative emotions which includes anger, fear, sad, disgust updates the total stress in a positive direction. Neutral expression has no impact on the stress level.

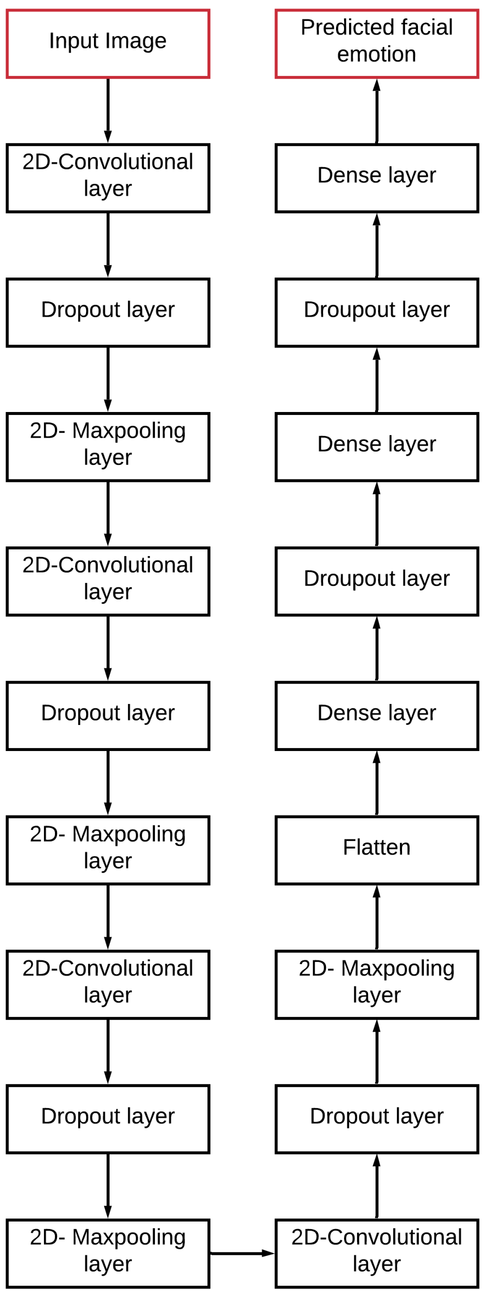

Fig. 2. Convolutional neural network architecture of the implemented model from [15]

The equations are formed in such a way that the total stress will never fall below zero because positive emotions have no impact on total stress when it is 0. The reason for selecting a window size instead of processing every input sample is to avoid outliers since the model used for predicting the facial expressions are not 100% accurate. So, processing a bunch of input data at a time helps the system to overcome the flaws and optimize the processor usage of the host machine.

2) *Using Adaptive Threshold:* Earlier works relied heavily on using static threshold for stress detection. But according to [9] stress level varies from person to person. As per the work done in [9], people are divided into two different categories as Type A and Type B. Various stressors were imposed on them, and their responses were recorded. The results show that the way Type A people respond under various stressors differs when compared to Type B. From this result we were able to understand that the threshold value cannot remain constant for every user.

As described in [16], we chose to implement the adaptive threshold methodology in a similar manner. In this method the initial step followed is the same as that of the static threshold methodology but if the total stress crosses the threshold for the first time the parameters are updated based on the records from the previous iteration. This process continues for every window of the further iterations. Initially total stress is initialized to 0 and a timer is started when the total stress crosses threshold, the timer is turned off and the run time for this iteration and the no. of windows passed in this iteration are recorded. Now the new threshold is implemented using the below equation.

$$T_{dyn} = T_{low} + (T_{high} - T_{low}) * f(t) \tag{1}$$

; where T_{dyn} is the new threshold calculated, T_{low} is the threshold of the previous iteration, T_{high} is the total time taken for an iteration to complete and f(t) is the function multiplier.

$$f(t) = \alpha^{\frac{-t}{t_{max}}} \tag{2}$$

; where alpha is greater than 1, t_{max} is the average of all the windows sizes in every iteration and t is the current window position.

$$t^n_{max} = \frac{\sum_{k=0}^{n-1} t^n_{max}}{n-1} \tag{3}$$

Additional modifications are made in order to apply these equations to our work which includes re-initialization of all the parameters such that the system does not suffer from long run problem. The long run problem is a scenario where the system is more reliable on the previous iterations than the existing one. Since the parameters are updated for every window this system removes the chances of suffering from long run and short run problems. A short run problem is a scenario where the system is completely dependent on the current iteration and neglect all the previous iterations. As shown in Fig. 6 even though the threshold dropped to a very low value it will keep on updating itself based on the input data. This suggests that even though the system had fallen for some sudden changes it would be able to come to normal states if the further iterations are normal.

Table I briefs the reason behind the fact that how this methodology will survive from short run problems. Assume initially T_{low} is 1000, T_{high} is 800, α is 5 and number of windows in previous iteration are 10. t_{max} is the average of all the previous window sizes so t_{max} gets a value of 10 then Table I briefs the results in the second iteration if the total stress is greater than 0. After window 5, total stress falls back to 0 and the parameters are reinitialized to their recent standard values and start timer is also reinitialized. When the total Stress crosses threshold new T_{dyn} is calculated same as that of the previous one but the change is that t_{low} will be the average values of all the T_{dyn} values in the previous iteration. As the system proceeds further it calculates the parameters more precisely and produces better results. Fig. 3 describes the change in threshold value based on the same scenario explained above. As you can see even though the threshold value reduced to a great extent, the system is capable to get back to a normal value. Multiple α values are considered and the trend generated by various alpha values is same even though there is a change in magnitude value.

TABLE I
TABLE GENERATED FROM THE PARAMETERS IN SCENARIO 1

Total Stress	Windows	f(t)	T_{dyn}
0	0	1	800
50	1	0.85	830
80	2	0.72	856
154	3	0.61	878
78	4	0.52	896
35	5	0.44	912
0	0	1	800

Similarly assume that T_{low} is 1000, T_{high} is 1200, α is 5 and number of windows in previous iteration are 8. Table II briefs the results in the second iteration when total stress is greater than 0. As shown in Fig. 4, even though the threshold reached to a high value it keeps on descending slowly. This prevents the system from running over a long time to reach the threshold. Multiple α values are considered and the trend generated by various alpha values is same even though there is a change in magnitude value.

IV. RESULT ANALYSIS

A. Facial Emotion Detection

The reason behind changing the accuracy calculation is to reduce the mismatch between emotions that comes under same cluster. As per the system point of view it doesn't matter if there is a mismatch between the classification of emotions that are under same class. The model described above got an accuracy of 86% when evaluated with three class classification. Fig. 5 details the predictions made by the model on some of the images from Cohn-Kanade-images [18], [19] dataset. First row briefs the prediction on positive emotions, second row briefs the prediction on neutral emotions and the third row briefs the prediction on negative emotions.

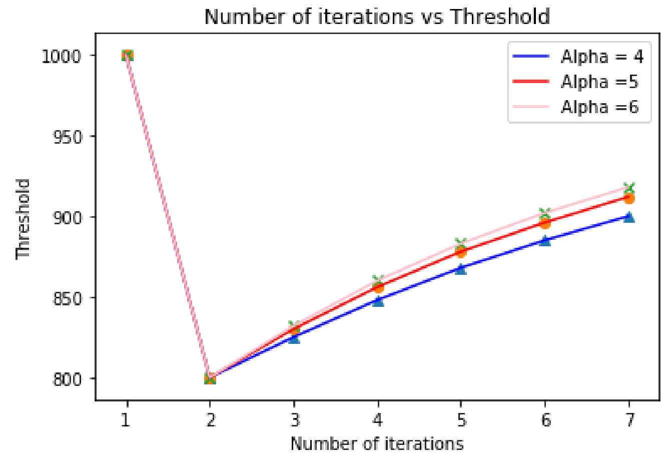

Fig. 3. Number of iterations vs Threshold using the parameters considered in scenario 1

TABLE II
TABLE GENERATED FROM THE PARAMETERS IN SCENARIO 2

Total Stress	Windows	f(t)	T_{dyn}
0	0	1	1200
60	1	0.81	1162
140	2	0.66	1132
124	3	0.54	1108
157	4	0.44	1088
.	.	.	.

B. Comparison Between Static and Adaptive Methodology

6 videos of duration 40 minutes, 1 hour 15 minutes, 1 hour 50 minutes, 2 hour 30 minutes, 3 hours 05 minutes

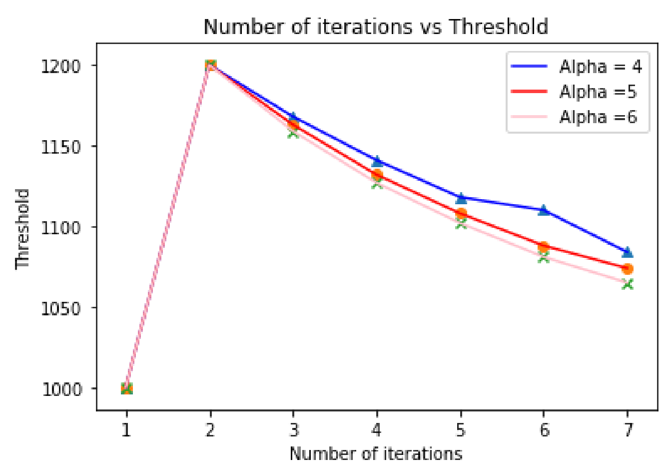

Fig. 4. Number of iterations vs Threshold using the parameters considered in scenario 2

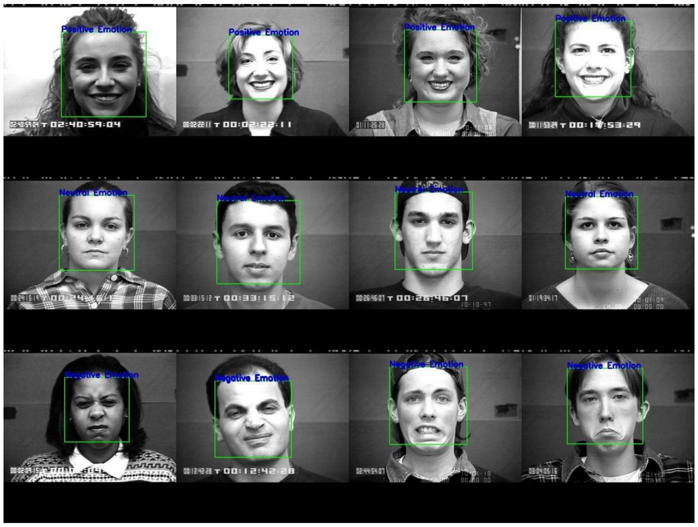

Fig. 5. Predictions made by our model on images from Cohn-Kanade-images dataset [18], [19]

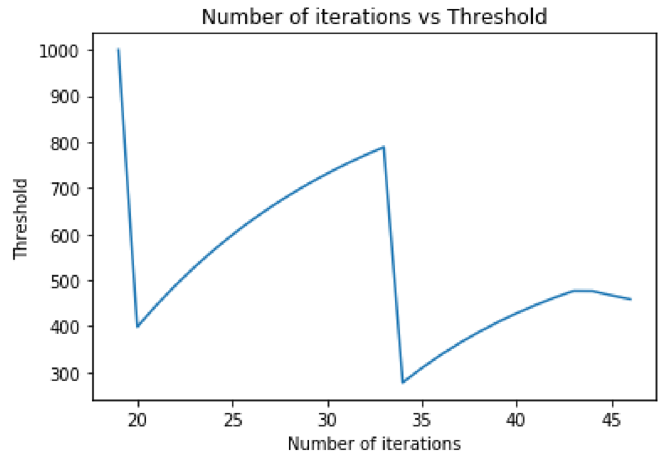

Fig. 6. Number of iterations vs Threshold plot that shows the wide variations on the threshold

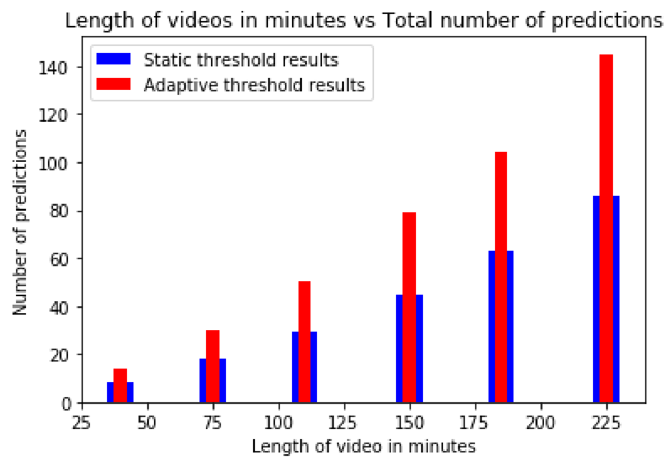

Fig. 7. Length of videos in minutes vs Results generated by the models

and 3 hour 45 minutes are generated using the Cohn-Kanade-images dataset [18], [19]. Majority of the content portrayed in these videos are negative and neutral emotions. The reason for choosing such videos is because both static and adaptive threshold methodology work in a similar manner with positive emotions. The difference between the working mechanism of static and adaptive threshold methodology can be traced out when they are dealing with negative emotions. The results produced by both the methodologies are shown in Table III. As you can see there's a huge difference in static and adaptive threshold methodologies while alerting the user. Since only negative emotions are portrayed in the videos, we expect the system to produce more alert points whereas in real world the system is provided with mixed emotions. A graph is plotted between the threshold value and number of iterations from the data generated in video 6. As you can see from Fig. 6, in adaptive threshold methodology the threshold value keeps on changing unlike in static threshold methodology. In 6 even though there is a sudden drop in threshold it didn't remain there for a long time. The change in the values of the threshold varies from person to person, and the parameters are tuned accordingly. Fig.7 shows the results generated by static and adaptive threshold methodologies on various videos.

TABLE III
PREDICTION ACCURACY'S OF STATIC AND DYNAMIC METHODOLOGIES

Video length	**Results obtained from static methodology**	**Result obtained from adaptive methodology**
40 minutes	8	14
1 hour 15 minutes	18	30
1 hour 50 minutes	29	50
2 hour 30 minutes	45	79
3 hour 05 minutes	63	104
3 hour 45 minutes	86	145

V. Conclusion

In this paper, the drawbacks of using static threshold over the adaptive threshold are explained. In static threshold methodology threshold remains constant, but adaptive threshold methodology threshold value changes from person to person and time to time accordingly. In our methodology we initialized the parameters for the first iteration and parameters are reinitialized using the specified calculations from the further iterations. Updating weights for emotions are not considered in any of the related work, which can be taken into consideration for future works.

ACKNOWLEDGEMENT

Authors would like to thank Dr. Vivek Menon from the department of computer science of Amrita Vishwa Vidyapeetham, Amritapuri and the Computer Science department of Amrita Vishwa Vidyapeetham, Amritapuri for providing their valuable guidance and resources for our work.

REFERENCES

[1] Amir Mohammad Shahsavarai, Esfandiar Azad Marz Abadi, Maryam Hakimi Kalkhoran, “Stress: Facts and Theories through Literature Review”, International Journal of Medical Reviews, Spring 2015, 230-241

[2] George Fink, Stress: Concepts Definition, History, Elsevier Inc, 2017,1–9
,

[3] Paul Ekman and Wallace V. Friesen, Maryam Hakimi Kalkhoran,Maureen O’Sullivan et al., Detecting emotional stress from facial expressions for driving safety, Journal of Personality and Social Psychology, November 1987, Vol. 5 3, No. 4,712-717

[4] H. Gao, A. Yüce and J. Thiran, ”Detecting emotional stress from facial expressions for driving safety,” 2014 IEEE International Conference on Image Processing (ICIP), Paris, 2014, pp. 5961-5965.

[5] J. A. Healey and R. W. Picard, ”Detecting stress during real-world driving tasks using physiological sensors,” in IEEE Transactions on Intelligent Transportation Systems, vol. 6, no. 2, pp. 156-166, June 2005.

[6] H. Kataoka, H. Kano, H. Yoshida, A. Saijo, M. Yasuda and M. Osumi, ”Development of a skin temperature measuring system for non-contact stress evaluation,” Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol.20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No.98CH36286), Hong Kong, China, 1998, pp. 940-943 vol.2.

[7] Sriramprakash.S , Prasanna Vadana. D, O. V. Ramana Murthy,Stress Detection in Working People, 7th International Conference On Advances In Computing Communications, ICACC 2017, 359–366

[8] Liang Zhang,Shaozheng Qin, Zhuxi Yao et all, ”International Journal of Psychophysiology” Volume 109, November 2016, Pages 138-146

[9] Janjhua Y, Chandrakanta. Behavior of personality type toward stress and job performance: a study of healthcare professionals. J Family Med Prim Care. 2012;1(2):109–113.

[10] J.Zhai and A. Barreto, ”Stress Detection in Computer Users Based on Digital Signal Processing of Noninvasive Physiological Variables,” 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, 2006, pp. 1355-1358.

[11] Ani R., Augustine, A., Akhil, N. C., and Dr. Deepa Gopakumar O. S., “Random Forest Ensemble Classifier to Predict the Coronary Heart Disease Using Risk Factors”, in Proceedings of the International Conference on Soft Computing Systems, 2016.

[12] Jyothisha J. Nair and Mohan, N., “Alzheimer’s disease diagnosis in MR images using statistical methods”, in Proceedings of the 2017 IEEE International Conference on Communication and Signal Processing, ICCSP 2017, 2017.

[13] Gopakumar G, Swetha, M., Siva, G. Sai, and Subrahmanyam, G. R. K. S., “Automatic Detection of Malaria Infected RBCs from a Focus Stack of Bright Field Microscope Slide Images”, Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing. ACM, New York, NY, USA, 2016

[14] Sun FT., Kuo C., Cheng HT., Buthpitiya S., Collins P., Griss M. (2012) Activity-Aware Mental Stress Detection Using Physiological Sensors. In: Gris M., Yang G. (eds) Mobile Computing, Applications, and Services. MobiCASE 2010. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 76. Springer, Berlin, Heidelberg

[15] Mariana-Iuliana Georgescu Radu, Tudor Ionescul, Marius Popescu1,”Local Learning with Deep and Handcrafted Features for Facial Expression Recognition”, arXiv:1804.10892v6 [cs.CV] 13 May 2019

[16] Sungju-Youm and Woosaeng-Kim, ”Dynamic threshold method for scene change detection,” 2003 International Conference on Multimedia and Expo. ICME ’03. Proceedings (Cat. No.03TH8698), Baltimore, MD, USA, 2003, pp. II-337.

[17] A. S. Kumar and Jyothisha J. Nair, “Pair wise training for stacked convolutional autoencoders using small scale images”, Journal of Intelligent and Fuzzy Systems, vol. 36, pp. 1987-1995, 2019

[18] Kanade, T., Cohn, J. F., Tian, Y. (2000). Comprehensive database for facial expression analysis. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG’00), Grenoble, France, 46-53.

[19] Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression. Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010), San Francisco, USA, 94-101.