International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Edge Computing-Enabled Stress Detection through Emotion-Classified CNN

K.N Apinaya Prethi*[1], S.M Nithya[2], T. Jayanthi[3], S. Hariharan[4]

Submitted: 07/02/2024 **Revised**: 15/03/2024 **Accepted**: 21/03/2024

Abstract: Modern society is extremely stressful. As well, the environment in which we live does nothing to aid individuals; rather, it pushes us over the brink and adds to our stress levels. Increased stress can lead to mortality in certain extreme cases; therefore an image-based stress detection system was created for Cloud-Edge computing that is non-invasive. A person's stress is expressed through facial expressions. Hence, in this paper deep learning algorithm is employed on facial photos to classify emotions for stress detection. The proposed neural network for emotion classification achieved an accuracy of 88%. The classified emotions were then fed into a stress detection module which detects the subjected individual as stressed if more than 75% of the classified emotions fall under the stressful emotions such as anger, sadness, disgust, and fear. These emotions are identified as high priority tasks which will help to provide personalized treatment by using edge devices.

Keywords: *Stress detection, Cloud-Edge computing, Facial Expressions, Deep learning, Emotion classification, Personalized treatment*

1 Introduction

Make sure the headings are correctly formatted throughout the article. Stress is a natural reaction in individuals when they are confronted with a threat. When the human are under stress, our bodies release hormones. It usually affects our body reactions like breathing, cardiac activity, and our blood level tends to slow down or high up. Anxiety and depression are two mental health conditions that can result from excessive stress. It also interferes with our ability to work, which could lead to a decrease in results. According to much research, students are more stressed than others. People with demanding occupations are more prevalent in the modern world, and as a result, stress levels are increasing. Stress can cause difficulties if it is not noticed early, thus early identification is crucial. Along with heart rate and other elements, facial expressions are one of the defining criteria in diagnosing stress. Stress can be identified through one's face as it leaves a mark on our face. Some of them include dry skin, wrinkles, and acne. The hormones which is released in our body when the human are under stress results in physiological changes, which in turn gives back a negative impact on our face and also it may lead to bad habits like grinding our teeth and biting our lips. Thus selected facial expressions to detect stress in people as a non-invasive method at edge layer is used.

[1]*Kumaraguru College of Technology, Coimbatore, Tamil Nadu, E-mail:abipreethu@gmail.com*
[2]*Kumaraguru College of Technology, Coimbatore, Tamil Nadu, E-mail:nithitechnologist@gmail.com*
[3]*Dr.NGP Institute of Technology, Coimbatore, Tamil Nadu, E-mail:jayanthiparu@gmail.com*
[4]*Government Polytechnic College, Coimbatore, Tamil Nadu, E-mail: hariharanaccet@gmail.com*

Personalized medicine is crucial in addressing biological factors and effectively treating individuals in a cost-efficient manner. Its significance lies in accommodating the diverse health needs of individuals, underscoring the essential role of personalization in healthcare[2].

The existence of universal facial expressions was proven in [3]. More than 30 studies working on the judgments of facial expressions have proven the universal recognition of emotion in the face [5]. Taking the results of these studies into account, considered face expressions of seven basic emotions namely anger, fear, disgust, happiness, sadness, surprise, and neutral to detect stress. However, other characteristics such as attitude, expression, location and orientation, the color of skin, and pixel values, as well as the existence of spectacles, might make detecting faces in an image increasingly difficult. As a result, attempted to use deep learning at the edge of the cloud to effectively predict expressions and detect the presence of stress.

Many prevalence studies have been conducted in the midst of the COVID-19 epidemic to assess any potential psychological effects on medical students. However, it is crucial to combine accurate prevalence statistics in order to have a more comprehensive knowledge of the true burden. In order to provide a thorough picture of the prevalence of stress, anxiety, and depression among medical students during the pandemic, the systematic review and meta-analysis were proposed [4].

Pushing the resources to the edge known as Edge computing. Cloud computing can be supplemented with the edge computing paradigm to handle this massive amount of data[1]. This greatly reduces on bandwidth consumption and latency. IoT devices come in a variety of hardware, software, and functions, and they operate in a variety of network architectures. It's also known as near-user processing because it works in close proximity to the user [6], [7].

Face detection algorithms typically start with the human eye since it is one of the quickest elements to detect. Subsequently, the algorithm will identify the iris, mouth, nose, nostrils, and eyebrows. The system runs additional tests to confirm that it has truly detected a face after determining that a facial region has been found. Local binary patterns were chosen as our face detection technique because they work consistently and reliably over low-resolution videos captured in real-time [8].

Because emotions are directly linked to what happens inside a person, it is necessary to distinguish emotions from the face. Facial expressions reveal a lot about a person's emotions. People's emotions are divided into seven categories, and then they are classified as stressed or unstressed based on their percentage of negative (anger, disgust, fear, sad) and positive (happy, surprise, neutral) emotion.

The data is sent to the edge layer, which categorises stress emotions as a high-priority task and unstressed emotions as a low-priority task. High-priority tasks must be handled quickly at the edge layer, whereas low-priority tasks may be routed to the central cloud for processing. The edge devices execute high-priority tasks and can recommend stress-relieving choices to a stressed user. Listening to music, which is freely available and accessible to everyone, is another stress-relieving choice. Furthermore, music aids in the diversification or forwarding of one's mind to other activities rather than one's own ideas. As a result, we are optimistic that research like these can help in the timely identification of high priority tasks like stress and reduce the cost of intervention in cloud-edge systems, resulting in a more productive working environment.

The other sections of this document are constructed in the following manner: Section II presents the most recent research related to task prioritization and the requirement for stress detection. Section III outlines the proposed technique. While Section IV discusses the outcomes,the final section provides a summary of this research.

2 Related Work

Automatic stress detection has been under study for several years. In [9] users' action cues were taken along with their facial expressions for stress detection. For which a two-levelled stress detection network (TSDNet) was proposed. The face level and action level representations were learned separately and then their results were fed into the stream weighted integrator which used local and global attention for stress identification. The stressed and unstressed video clips were collected by making the participants watch the relaxed and neutral video clips. The proposed solution achieved a detection accuracy of 85.42% and an F1 score of 85.28%.

A cascade classifier with haar-like features was used for face detection in the input image and a convolutional neural network was proposed for emotion classification in [10]. In [11], the Haar Cascade algorithm was used for face detection. A convolutional neural network architecture consisting of 2 convoluted layers with ReLU function, 3 fully connected layers, and a softmax activation was proposed for emotion analysis.If the emotion is determined to be sadness, a chatbot pop-up appears on the screen, allowing the user to communicate his or her feelings with the Tkinter library-based chatbot. "Stress detection in IT professional by image processing and machine learning" [12], KNN was used with haar-like features for face detection. A convolutional neural network was proposed for emotion recognition.

In [13] images were captured in daily life and the location of the face was determined using haar like feature selection technique. Transfer learning on trained neural networks like VGG16, VGG19, and Inception-ResNet V2 was used for facial expression classification. The classifications were recorded and then sent to the stress detection module. Net Images, CK+, and The Karolinska Directed Emotional Faces (KDEF) datasets were used for training various models. The convolutional bases of each pre-trained network were frozen but the classifier parts were replaced with two different configurations, one with a global average pooling layer and the other with a convolutional layer. The pre-trained networks were then trained for each of the configurations and their accuracies were recorded. The Inception-ResNet V2 yielded a max accuracy of 82.49% thus giving the best performance. The multi-class (7 classes of emotions) fine-tuned VGG16 model gave an accuracy of 89.6%. In binary evaluation (2 classes namely 'stress' and 'not stress'), the VGG16 model gave an accuracy 92.1%.

A Multi-task Cascaded Convolutional Neural Network (MTCNN) was used with one input layer, four convolutional layers, two fully-connected layers, and an output layer for emotion detection were used in [14]. If more than two-thirds of consecutive images were classified as stress-related facial expressions, then the person is considered to be under stress and a warning is made to take a break.

In [**15**], A convolutional neural network architecture with three convolution layers with a set of filters, depth-wise separable convolutions, and residual modules, ADAM optimizer, was proposed for emotion classification in "A deep learning approach for human stress detection." Regression analysis was used to find the coefficients of unknown variables, which was used to present the logarithmic model equations for evaluating the stress levels.

The group **[16**], proposed a Cascade Multi-view Hourglass Model. 3D Morphable Models were used for appearance features and ResNet50 network architecture with ADAM optimizer was another method proposed in this paper. A multi-channel binary classification with a binary label indicating the existence or non-existence of each AU is obtained from an Action unit classification employing fully connected layers. Pairwise transformation and normalization, feature selection and relevance with Radial basis function (RBF) kernel, and the SVM hyper parameters optimization were some of the other methods proposed in this paper.

3 Semantic based Retrieval using Meta data

3.1 Dataset

FER2013, CK+ & Ryerson emotion datasets from Kaggle are used. In edge, for training the boosted cascade of weak classifiers to detect the face from an image samples used with human frontal faces obtained from the above-mentioned datasets as positive images and the Multi-Salient-Object dataset from Kaggle as negative images. The FER2013, CK+ dataset has human faces with various emotions like anger, sad, disgust, fear, happy, surprise, and neutral. Ryerson emotion dataset is a video dataset from which frames were extracted, cropped the faces and resized them into dimensions 48 x 48, and converted them to grey scale. Tab. 1 shows the samples of input data.

Priority	Facial Expressions	Training	Validation
High	Anger	4172	1231
	Disgust	1000	162
	Sad	4511	1787
	Fear	4185	1259
Low	Neutral	4504	1747
	Happy	7285	2532
	Surprise	3454	196

Table 1:Number of Images used for Training and Validation

3.2 Face Detection

Focusing on automatic stress detection using facial expressions at edge layer, started with detecting emotion using face. The boosted cascade of weak classifiers with LBP features to detect the face. LBP features give integer precision. Hence both training and detection with LBP features will be faster than that with HAAR features (which gives floating point precision). LBP operates in a 3X3 block, with the center pixel serving as a threshold for the surrounding pixels. Cascade classifiers area concatenation of many classifiers.Each classifier is a stage in the cascading pipeline that acts as a filter to eliminate false positives. To ensure accuracy, this classifier was trained on large data sets with hundreds of thousands of positive and negative samples. The algorithms' ability to identify faces in an image is improved by the training.

This classifier was trained on massive data sets containing hundreds of thousands of positive and negative samples to ensure accuracy. The training enhances the algorithms' capacity to determine whether and where there are faces in an image. Images with frontal human faces obtained from the FER2013 and CK+ were labelled as 'positive'samples while thosetaken from the Multi-Salient-Object dataset, not containing the target object(face), were labelled as 'negative' samples. Then a text file was created for both negative and positive images followed by the generation of a vector file for positive images. Finally, the trained the cascade classifier. Fig.1 describes the flow of proposed architecture. Fig.2 shows the local binary pattern calculation. The LBP features are calculated with equations (1) and (2).

$$\text{LBP} = \sum_{i=0}^{P-1} s(n_i - G_c)2^i \quad (1)$$

$$\text{S(x)} = \{1,\ if\ x > 0 \quad 0,\ otherwise \quad (2)$$

where, P is the total number of pixels which are nearer (neighbor), n_i denotes the neighbour pixel at ith place,and c denotes the pixel at the center.

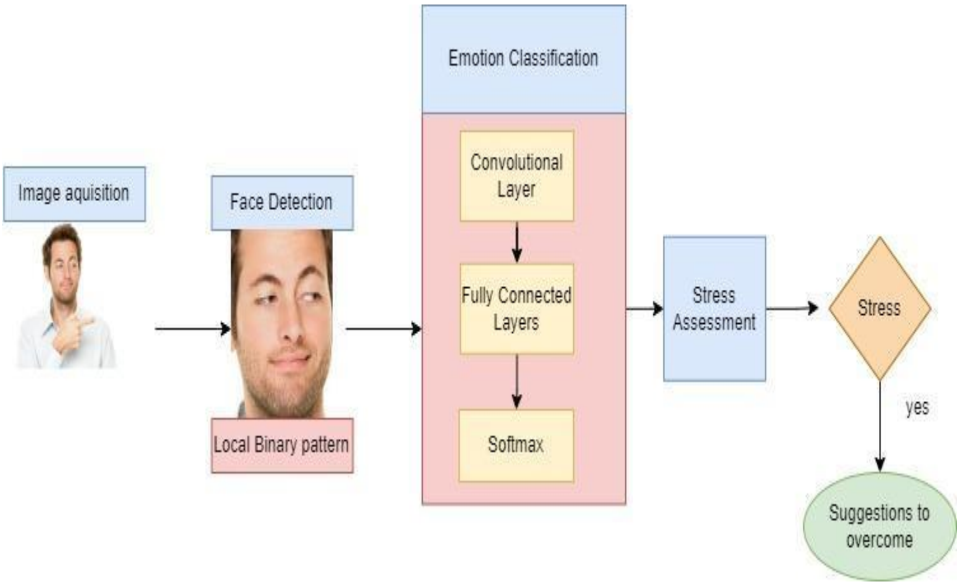

Figure 1: Proposed Architecture

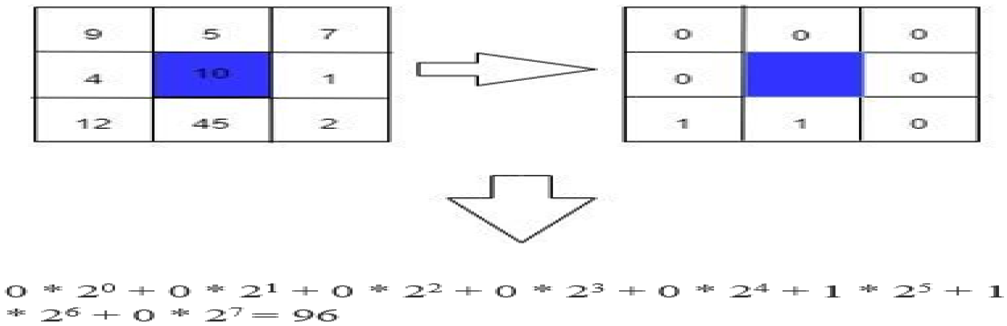

Figure 2: Local Binary Pattern Calculation

3.3 motion Classification

FER2013, CK+, and Ryerson emotion datasets were utilized for this module. All of the photographs were cropped and made to size 48 X 48 pixels. A CSV file was generated for all the images present in seven different folders. As our dataset was highly skewed in certain categories as represented in Table 1, used random oversampling, which duplicated the photos at random. Convolutional neural networks were employed in this research's edge layer. Convolution neural networks are deep learning techniques that take an image as input and use weights to identify distinct parts of the image. As shown in Fig. 3, our model consists of fifteen layers for feature extraction, excluding an input layer, a layer of convolution with batch normalization, a Rectified Linear Unit activation layer, a dropout layer, and a layer of max-pooling.

The convolutional layer, also known as the kernel or filter, moves with a certain stride value until it has crossed the entire width of the image. The batch normalization layer was then employed to allow each layer to learn more independently, making learning easier and hence speeding up the training. The result was then fed to the activation function. By converting its inputs to outputs with a specific range, the activation function aids in the learning of complicated patterns in data. The Rectified Linear Unit is selected, which is noted for its ease of use and sensitivity. The pooling layer was then utilized to minimize the size of features, which helps to reduce computing resources by extracting the most important information. The output of this network was then flattened and fed into the dense neural network having 1 fully connected layer followed by a dropout of 0.6 and an output layer with a softmax activation function which gives the probability scores for each of the seven different classes of emotions like angry, anxiety, disgust, happy, sorrow, unexpectedness and neutral.The neural network was trained, which resulted in a validation accuracy of 88% and accuracy of 96%.

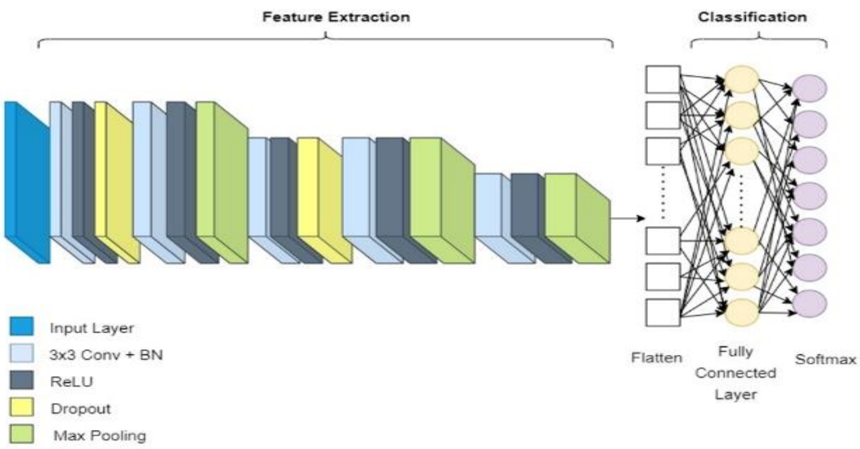

Figure 3 : Proposed neural network for emotion classification

Stress assessment was carried out by considering two parameters, namely:

(i) ‘Window size’ which shows the number of past classifications taken into consideration for our stress assessment.

(ii) ‘Threshold’ to determine the percentage of negative emotions needed for a person to be considered as stressed.

When the threshold value exceeds 75%, the task is categorised as negative emotion (stress) and is regarded high-priority. Negative emotions are processed instantly by edge devices, and the remaining may be sent to the server cloud. Stress must be treated immediately because it can lead to a variety of physical and mental difficulties. The subjected individual's stress state was updated on a regular basis in a CSV file with a timestamp in edge devices. If an individual is determined to be under persistent stress, included an extra feature that provides some brief stress-relieving remedies such as watching a funny video, listening to calming music, and so on.Individuals with a medical history of blood pressure issues may benefit from treatment that takes into account their current emotional state in order to provide more effective care.
Personalized treatment is essential for patients, particularly those with a history of blood pressure issues, as it enables tailored care to address their specific needs and current emotional state, ensuring effective treatment.

4 Result and Discussion

The proposed neural network yielded a maximum validation accuracy of 88%. The results of training our model for emotion recognition are shown in Table 2 and Table3, in which the seven different emotions namely anger, disgust, fear, happy, neutral, sad, and surprise are labelled from 0 to 6 respectively. Figure 4 and 5 shows the precision, recall and f1-score and support of training set whereas Figure 6 and 7 shows for testing set.The performance metrics represented in Table 2 and Table 3 such as precision, recall, f1-score, and accuracy can be calculated using the following formulae.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FN+FP} \quad (3)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (5)$$

$$F1 = 2 \times \frac{Precision * Recall}{Precision + Recall} \quad (6)$$

where,

True Positive, or TP, is the output that the model recognized as belongs to the positive class. False Positive, or FP, is the output that the model wrongly identified as belongs to the positive class. True Negative, or TN, is the output that the model correctly identified as corresponding to the negative class. False Negative, or FN, is the output that the model misclassified as corresponding to the negative class.

Table 2:Precision, Recall, F1-Score for Training Set

Emotion Recognition	Precision	Recall	F1-Score	Support
Anger	0.98	0.95	0.97	8657
Disgust	1	0.99	1	8600
Fear	0.96	0.958	0.96	8678
Happy	0.97	0.99	0.98	8607

Neutral	0.95	0.96	0.96	8656
Sad	0.94	0.96	0.95	8657
Surprise	0.99	0.98	0.99	8688

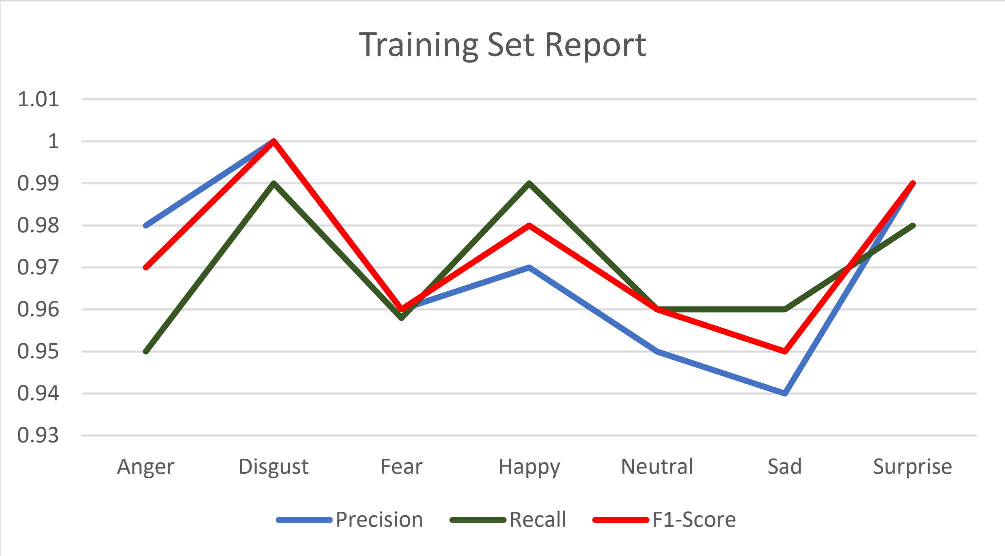

Figure 4:Precision, Recall, F1-Score for Training Set

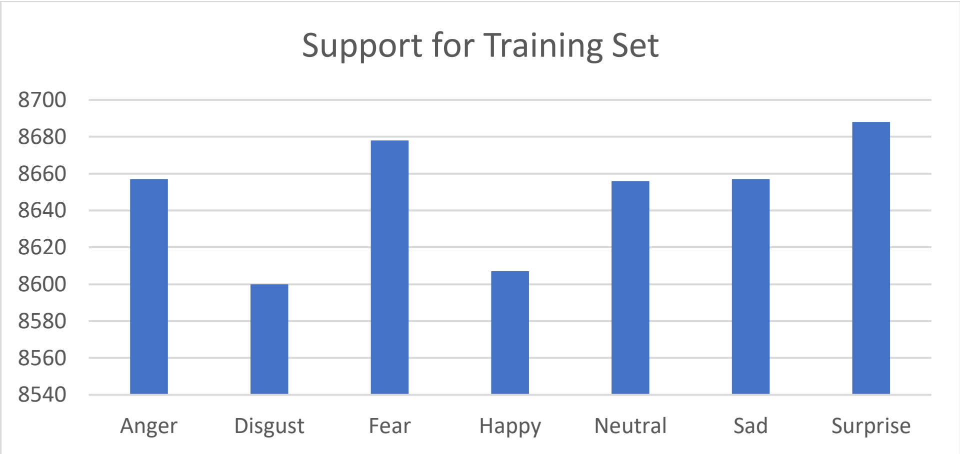

Figure 5:Support for Training Set

Table 3 :Precision, Recall, F1-Score for Testing Set

No.	Precision	Recall	F1-Score	Support
0	0.89	0.84	0.86	953
1	0.99	1	0.99	1010
2	0.87	0.86	0.86	932
3	0.91	0.83	0.87	1003
4	0.8	0.84	0.82	954
5	0.77	0.83	0.8	953
6	0.95	0.97	0.96	922

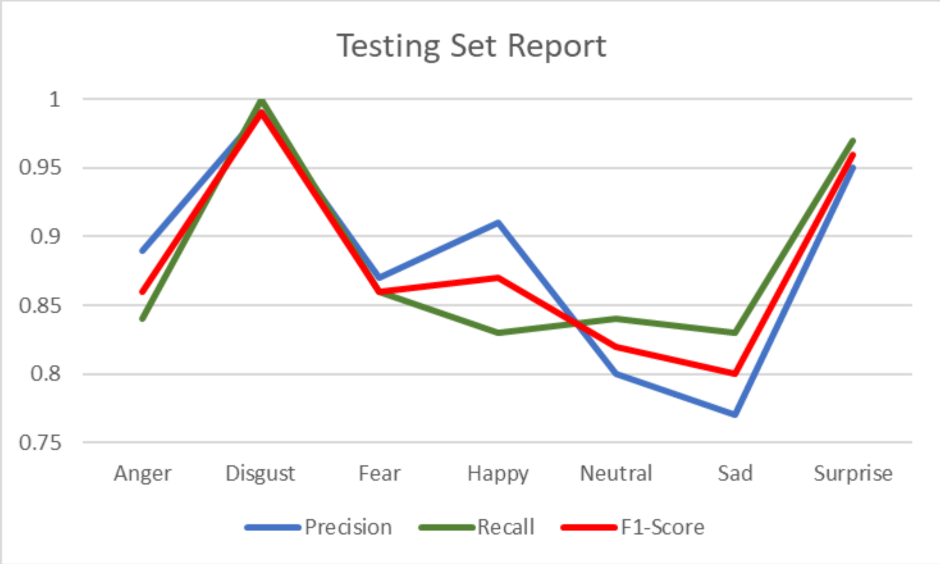

Figure 6: Precision, Recall, F1-Score for Testing Set

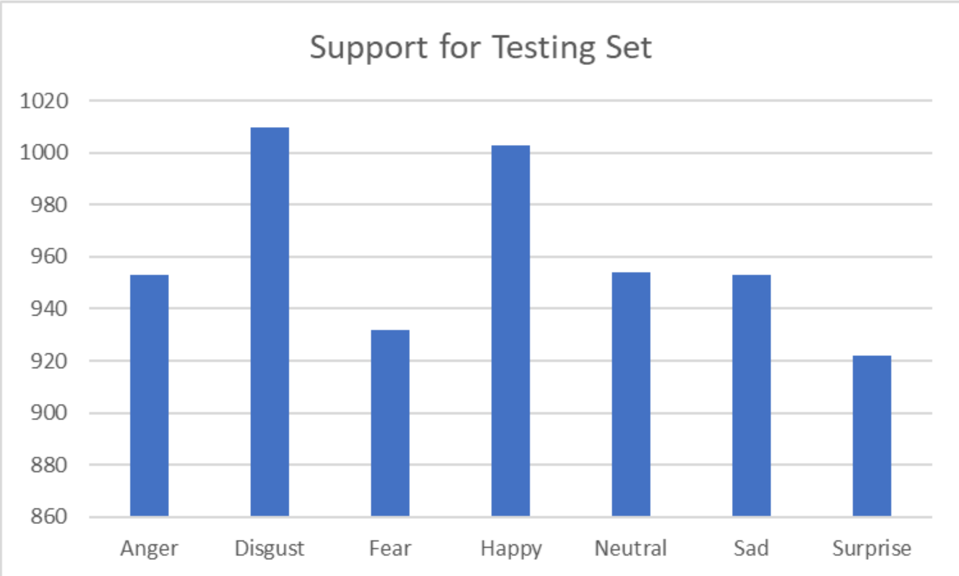

Figure 7: Support for Testing Set

5 Conclusions

Even if a server cloud fails or is overloaded, the Cloud-Edge system continues to function. Edge devices at the cloud's edge are given limited resources. As a result, in the proposed system, high-priority activities are categorised and executed using deep learning techniques at the edge layer. The identification of high-priority tasks is crucial to the convolutional neural network employed for stress detection in cloud-edge computing (stress) for personalized care. Consequently, the accuracy of the emotion identification module directly influences the efficacy of our stress detection system, enhancing our ability to provide optimal treatment.

6 Future Work

In next few years, pre-trained networks can help to improve the accuracy of the emotion recognition module. When processing high-priority jobs, the load on the edge devices can also be taken into account. In summary, our solution significantly enhances the detection of human stress.

References

[1] S. Nithya, M. Sangeetha, K. N. Apinaya Prethi, Kshira Sagar Sahoo , Sanjaya Kumar Panda, and Amir H. Gandomi.:SDCF: A Software-Defined Cyber Foraging Framework for Cloudlet Environment ,vol.17, no.4, December 2020.

[2] Goetz, Laura H., and Nicholas J. Schork.:Personalized medicine: motivation, challenges, and progress.,Fertility and sterility 109, no. 6, 952-963 2018.

[3] D. A. Sauter., F. Eisner, P. Ekman, S. K. Scott.:Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations" Proc. Natl. Acad. Sci. U.S.A. 107, 2408–2412 2010.

[4] Ayubi, E., Bashirian, S., Jenabi, E., Barati, M., & Khazaei, S. Stress, anxiety and depression among medical students during COVID-19 pandemic: A systematic review and meta-analysis. Personalized Medicine in Psychiatry, 41, 100108 2023.

[5] Matsumoto, D.: Culture and Emotion. In D. Matsumoto (Ed.), The handbook of culture and psychology. Oxford University Press. 171–1942001.

[6] Weisong Shi, Jie Cao, Quan Zhang, Youhuizi Li, Lanyu Xu .:Edge Computing: Vision and Challenges. IEEE Internet Of Things Journal 3, 5 2016.

[7] Wei Yu, Fan Liang, Xiaofei He, William Grant Hatcher, Chao Lu, Jie Lin, And Xinyu Yang. : A Survey on the Edge Computing for the Internet of Things. IEEE Access, Special Section On Mobile Edge Computing, March 2018.

[8] Di Huang, Caifeng Shan, Mohsen Ardabilian, Yunhong Wang, and Liming Chen. 2011. Local Binary Patterns and Its Application to Facial Image Analysis: A Survey. IEEE transactions on systems, man, and cybernetics—Part C: Applications and Reviews. 41. 6 2011, doi:10.1109/TSMCC.2011.2118750

[9] Huijun Zhang , Ling Feng, Ningyun Li, Zhanyu Jin and Lei Cao. :Video-Based Stress Detection through Deep Learning. Sensors. 20, 5552 2020; doi:10.3390/s20195552

[10] Mohammad Failzal, Nikhil V, Nikhil V, Prajwal C, Aruna Rao B P. : Recognition of stress using face image and facial landmarks. International Advanced Research Journal in Science, Engineering and Technology. 8, 5 2021.

[11] Neha S, Nivya, Pooja HC Shekar, Keerthana S Kumar, Asha VG. 2: Emotion Recognition and Depression Detection using Deep Learning. International Research Journal of Engineering and Technology (IRJET) 07, 08 2020.

[12] Prof. Vishal R. Shinde, Mr.Shubham Memane, Miss.Priya Thakare, Miss.Vibha Vishe. 2021. Stress detection in IT professional by image processing and machine learning. International Journal for Research in Engineering Application & Management .7 2021.

[13] José Almeida and Fátima Rodrigues.: Facial Expression Recognition System for Stress Detection with Deep Learning. In Proceedings of the 23rd International Conference on Enterprise Information Systems. 1(256-263) 2021.

[14] Xue Mei, Huan Liu, Shenqiang Yuan, Tiancheng Qian, Jin Zhang.: Detecting Negative Emotional Stress Based on Facial Expression in Real Time. IEEE 4th International Conference on Signal and Image Processing, 2019.

[15] Akanksha Nogaja.:A deep learning approach for human stress detection. International Journal of Computer Engineering and Applications,Volume XIII, Issue VIII 2019.

[16] Giorgos Giannakakis, Mohammad Rami Koujan, Anastasios Roussos, Kostas Marias.: Automatic stress analysis from facial videos based on deep facial action units recognition, 2019. DOI:10.1007/s10044-021-01012-9