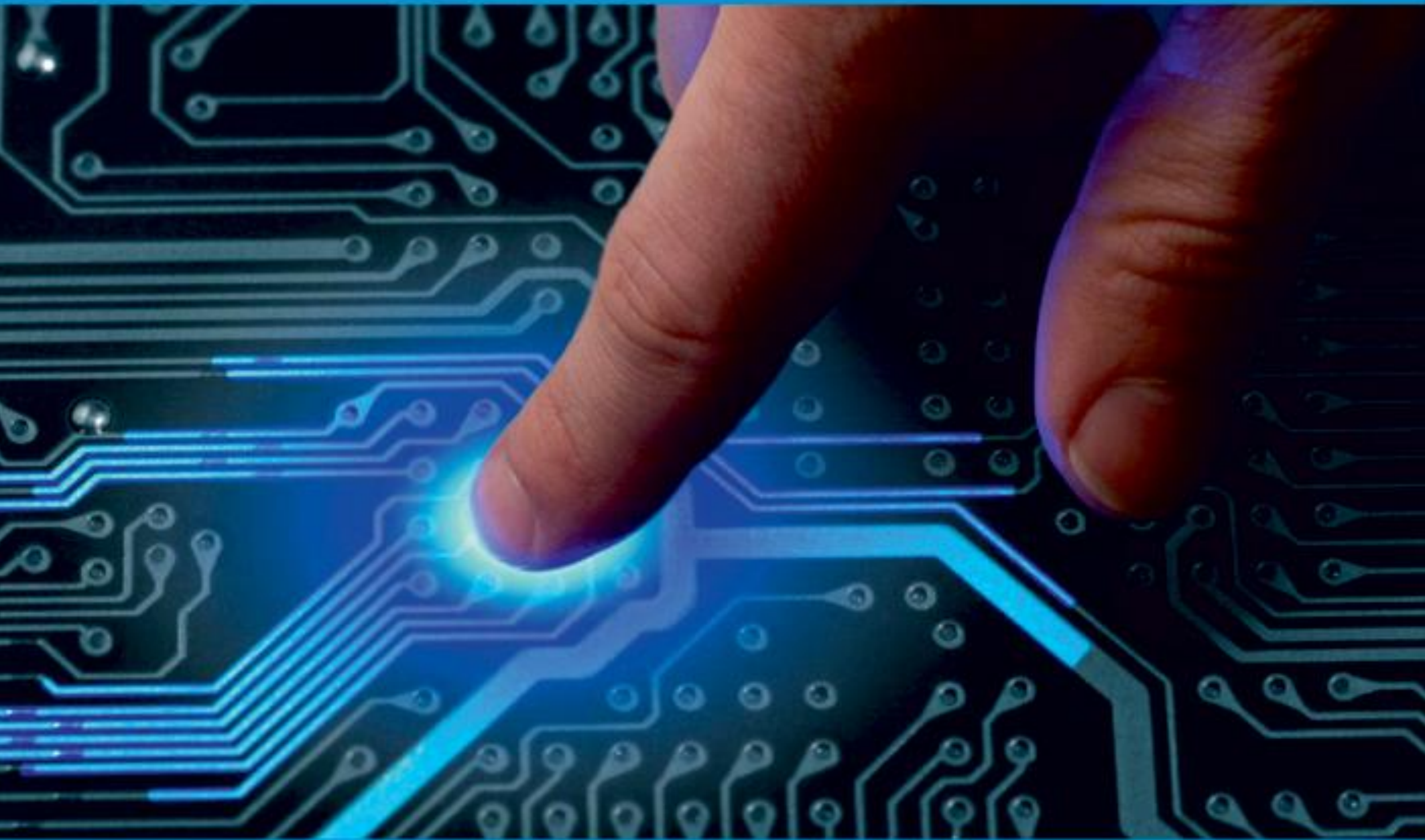
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 5, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105162 |

Deep Learning Techniques to Recognize Stress using Facial Expression

S.N.Swetha[1], P.Padma Prapoorna[2], and G. Nagalakshmi[3]

[1,2] PG student, Msc Computer and Sanskrit Language Technology, National Sanskrit University ,Tirupati, Andhra Pradesh, India.

[3] Assistant Professor, Department of computer science, National Sanskrit University, Tirupati, Andhra Pradesh, India.

ABSTRACT : Stress is how human will react when they feel under pressure or threatened. It usually occurs when unexpected things that are happen in a situation that we don't feel how to manage or control. We all feel stressed at times, but what one person finds stressful may be very different from what another finds stressful. The stress can be easily identified by recognize through human face by conveying emotion or reaction. The human face has a great store and variety of expressions. By using image processing techniques facial expressions can be captured which is used for creating data set. A data set is a collection of related, discrete items of related data that may be accessed individually or in combination or managed as a whole entity. Deep learning plays a very vital role in resolving this issue of learning and analyzing big data. It also supports the nature of deep learning algorithms, which requires large amount of training data. A data classification is a supervised machine learning process it also involves in predicting the class of given data. It also can be targets , labels or categories. For example the image of a person will be given. It should able to recognise the person whether he/she is in "Stress" or "Not in Stress". There are some common Machine Learning Algorithms includes Decision Trees, Linear Regression, KNN, Naive Bayes.

KEYWORDS: Stress, Facial Expressions, Parameters, Image processing, Deep Learning, Classification, Machine Learning, Testing, Training, Big data.

I.INTRODUCTION

Stress is a regular reaction to the body when changes occurs in physical, emotional and intellectual response. For stress management yoga , meditation and few exercises can help you to get relief and keeps healthy.

Stress is a common act that occurs to everyone during hectic situations. Every human body is prepared and designed experience stress and to react to it. Due to stress our body will automatically get tuned to mental [1] and physical response. It also help us to face any problems at any situations. It will be positive and alert us to avoid danger. Example: i) For students the common reason for stress is Exam. ii) For men stress [2] will be mostly occurred due to Work in Office. iii) For women family problem will be the main stress. Physical Symptoms due to Stress are:

i) Aches and Pains.

ii) Chest pain, BP(Blood Pressure)

iii) Dizziness , Insomnia.

iv) Stomach Inflammation.

v) Lack of Immunity.

The above said are the major problems frequently facing by human beings in day today life. These emotions can be easily recognised through Facial Expressions. [3] Facial Expressions & action plays a major role for communicating with people. Through Facial Expressions some emotions are expressed universally. They are also called Universal Facial Expressions for emotions. The major five emotions are Anger, Disgust, Fear, Happiness, Sadness.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105162 |

NEUTRAL: ANGER: DISGUST:

FEAR: HAPPY: SAD:

The relationship between emotions and moods are agreed by some Scientists. Stress is not only identified by the Facial Expressions it also identified by the voice. There are no universal facial expression of emotion for stress and there are studies (Dinges et al.,2005),(Lerner et al.,2007) examined that feeling of stress is increase in cortisol levels and cardiac activity and it was confirmed that there is a relationship between facial expressions and stress. From these studies some negative emotions are related anger, disgust, fear, happiness and sadness.

With the collection of images, blood tests, saliva Automatic stress detection has been studied for many years. Gao et al., 2014 is one of the study of those works where a camera is fixed inside a car and collected some images of a driver face for detecting [4] stress.(SVMs) Support Vector Machines were taught in public facial expression datasets and collected images and then categorized into one of six facial expressions. An Algorithm is introduced to find whether the person is in stress or not. An algorithm can able to count the classification within a given time window. The percentage of facial expressions is outstripped to a specific threshold, found that the driver was under stress. Let us see the parameters of the image and how it processed. Let us take a image of a girl and see how she feels i.e., whether she is in anger or happy or sad or in some other way. A value is assigned to five facial expressions for example , A=Anger, D=Disgust, H=Happiness, S=Sad, F=fear. We are taking these five major facial expressions for parameters. A image is taken normally if we see that we can assume that she may be happy but there are some calculations for those expressions. In that picture all values will be counted and will checked the face count. When we are in anger our forehead will shrink and there is a count for it. Likewise for each parameters there will be a count for all expressions. If A has more value than D, F, H, S and the result will be concluded as :The image of a girl is in Angry mood. Remaining parameters will be calculated as the above given way. We can solve these parameters by using Deep Learning techniques. The Deep Learning has a good learning ability. It also extract the features and classification process for the image classification test.

II.METHODOLOGY

The user shows signs of stress using Facial Expressions detected by video images. The use of web cams when working with computers is advantageous and accesses for to detect personal stress. When the user is showing signs of stress the program will run in the background monitors the and notify him/her. The framework's most memorable module will be answerable for catching pictures continuously, from the PC's webcam, and sending these pictures to the subsequent module. In the subsequent module will be resolved the client's face's area, utilizing [12] a Haar -like element determination method and edited. The face will then be resized to 299x299 pixels and standardized, by separating the worth of all pixels by 255 so that all have values somewhere in the range of 0 and 1. The Feeling Order module, comprising of a prepared grouping model, will characterize the face and return a rundown of seven likelihood scores, one for every look. The look with the most noteworthy likelihood will then, at that point, be taken care of to the fourth module (stress evaluation). The different orders made over the long haul will be recorded and in view of the boundaries gave to the program, it will then decide if the client is under pressure.

The program just requires three boundaries:

1) the "recurrence" of the program, which decides the time stretch, in milliseconds, between each picture extraction from video's webcam, and subsequently, the leftover modules' execution;

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

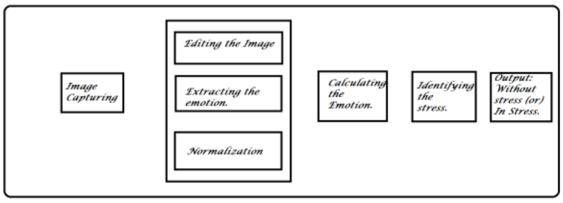
|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105162 |

2) the "time window" like a flash, which shows the time of past characterizations that will be viewed as in the pressure evaluation process;

3) a "edge" that demonstrates the level of pessimistic feelings expected to decide, inside the time window, assuming that the client is under the pressure.

For instance, assuming the time window is 900s (or 15 min.) and the limit is 75%, that's what it intends assuming as of now, 75% or a greater amount of the groupings are for unpleasant feelings, then the framework establishes that the client gives indications of stress and it will be shown a notice cautioning for that reality. Assuming the client wishes, he has the likelihood to affirm this notice. Likewise, in the event that the client allows, the pictures gathered by the webcam will be saved money on plate and those emphatically grouped by the model will be named with pressure and its timestamp. The objective to gather these pictures is twofold, first to make a genuine dataset of pictures marked with pressure/non stress, supposedly, from the exploration completed, there are n´t any open dataset with such qualities available; and further-more make these pictures accessible to be investigated by specialists.



III.FACIAL EXPRESSION RECOGNITION MODELS

The algorithms for recognising facial expressions come from both machine learning and deep learning, and a few of them are listed here.

- Closest Neighbours in K Each of us uses a certain set of facial expressions to convey our feelings. For instance, when we smile, the zygomaticus major muscle raises our cheeks, causing our lips to u-shape. Hence, by utilising KNN to identify the grin pattern and contrasting it with a test image, we can identify the happy feeling. The same is true for other expressions.
- Support Vector Machines (SVM): SVM is another widely used technique for FER. In order to find the plane with the greatest margin separation, it seeks to find a hyperplane in an N-dimensional space that divides the different types of data points. The method of feature extraction determines how effectively recognition works.
- Deep Belief Network: This technique uses the greedy method in many layers to address the issue of recognising facial expressions, however it does require faces to be properly aligned in an image.
- Multi-Layer Perceptron (MLP): MLP uses the pixel values of images to determine the pattern concealed in them for each emotion and depends on a backpropagation technique to assess various hyperparameters.
- Convolutional Neural Network: The convolution operation enables us to work with images without sacrificing the precision of the system while using less processing. Similar to MLP, it develops a facial expression recognition system using pixel values.

BUILDING THE MODEL:

It is now time to move on to the convolutional [10] neural network (CNN) model, which is the most intriguing stage of this Face Emotion Detection project's pipeline.

Generally, a CNN architecture consists of three different kinds of layers:

1)CONVOLUTION LAYER:

Two inputs are needed for the mathematical action of convolution. Assume that x and f are the two inputs. Convolution uses the values of the filter input, f, to turn a specific portion of the input into Y.

Y = x f

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105162 |

In our example, f is a 2D matrix and x is a 2D pixel array of an image. Each of these filter matrices' contents will be used to assess the output image Y's content. How are these numbers assessed? We don't choose that, though. We left that up to our model. We simply list the quantity and sizes of these filters. These sizes are symmetrical and always odd values for mathematical computation.

Convolutional layer illustration:

Conv2D (32, (3, 3), activation='relu', input shape (48, 48, 1)) model.add

The model employs 32 filters that are placed one on top of the other and are each 3x3 in size. Hence, a total of 288 parameters, or 3x3x32, will be learned by the model.

2)POOLING LAYER:

There will be no model parameter changes made by this layer. Simply put, it will scale down the prior image by the specified amount. There are other pooling techniques, but for our deep learning project, we'll choose the most popular "max-pooling method," which keeps just the largest number and discards the rest. This layer aids in cutting back on the overall training parameters. We can also take a look at this approach to lessen input variance.

as in, a model.

maxPooling2D((2, 2)) adds

Here, the layer cuts the input image's width and height in half.

3)FULLY CONNECTED LAYER:

Similar to other neural networks, this one is. It has a layer of neurons, all of which are trained throughout the training process. A fully connected layer with dimensions corresponding to the categories in the classification issue will make up the model's top layer.

Model.add(Dense(7, activation='softmax'), for instance)

The aforementioned layer, representing seven categories, will be the final layer of our model.

IV.EXPERIMENTATION

To handle the pressure recognition issue there were two prospects. Address the issue with information named pressure or non-stress (ideal methodology), or with information characterized with the seven looks. Sadly, datasets grouped straightforwardly as pressure or non-stress are not many and not openly accessible. Hence, attaching in thought the immediate relationship between's particular looks and upsetting circumstances, two datasets were mentioned.

The Karolinska Coordinated Close to home Faces (KDEF) (Lundqvist et al., 1998) is a dataset made by the Karolinska Establishment, uniting 4900 pictures of seven human looks, 700 for every one. The looks gathered correspond with the five looks of feeling acknowledged by mainstream researchers, in addition to unbiased and shock. Notwithstanding we will consider these seven feelings, just displeasure, revulsion and dread will be related with pressure, since just these are acknowledged as connected with pressure. Albeit complete and well formally dressed, KDEF is an exceptionally homogeneous dataset, where the subjects are all of a similar age bunch, same race, with practically no facial change like glasses or facial hair, and the pictures were totally caught in a controlled climate

Thusly, to check this homogeneity, it was chosen to get the CK+ dataset and furthermore to make a new dataset with pictures gathered by us from the web (here called Net Pictures), to prepare the models with additional heterogeneous information. With the utilization of these three datasets, getting a more practical preparation and evaluation of the models and ideally nearer to this present reality is normal.

The CK+ dataset is the consequence of an augmentation of the CK dataset (Kanade et al., 2000) that expected to advance examination into naturally distinguishing individual looks (Lucey et al., 2010). The CK+ dataset has a sum of 327 groupings of pictures and the quantity of every look in the dataset changes a little introducing a circulation of 45 successions of outrage, 18 of hatred, 59 of revulsion, 25 of dread, 69 of bliss, 28 of misery and 83 of shock.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105162 |

The dataset made by us, Net Pictures, comprises of twenty pictures for every inclination (same feelings as KDEF) in a sum of 140 pictures. These pictures were gathered from look through on two web search tools (Google and DuckDuckGo) and free stock pictures locales (unsplash.com; pexels.com; shutterstock.com; freepik.com). Then, at that point, were chosen pictures where the face was apparent, and the inclination was obviously present. We attempted to get exceptionally heterogeneous pictures, from individuals of various ages, races, and with and without a facial hair growth and the equivalent for glasses. Pictures with watermarks, with noticeable picture release and in which the look could be deciphered as a blend of feelings, as referenced in (Ekman and Friesen, 2003) were kept away from.

DATA PREPARATION:

After collecting the datasets, the models were given the images for training purpose. The first change is to adapt KDEF to the problem. The user uses the computer webcam to capture users Facial Expression in this project , the half left profile, the half right profile, straight images to train the neural network. The full profile images that contains less information about Facial Expressions , which can obstruct the learning for the neural network is the first reason. The web cams will be pointing at the users face from capturing mainly the straight , half profiles and the front by default. The data was reduced to () images , () each emotion of this adaptation. The sequences images they recognized by two reasons in the CK+ dataset. The contempt Facial Expression, that is not presented in other datasets which does not have the separation of neutral Facial Expression images. Recognition consist of

- To form a new neutral class remove the first image of every sequence (always a neutral facial expression).

- As the intermediate images which do not have any pronounced facial expressions elimination of all images except the last 6 images of each sequence is denoted.

The CK+ which includes 270 images of anger, 414 of happiness, 354 of disgust, 168 of sadness, 150 of fear and 309 of neutral in total 1665 of Facial Expressions.

After the adaptation of dataset all the images were cropped around the scale and face to 299/299 pixels. The distribution classes in CK+ is very quiet different with a certain imbalance between class which is contrary to the KDEF dataset. Finally the following hold-out method the training, the validation and the test sets were created.80% training, 10% validation, 10% test of the datasets were merged and divided. It is considered that the works (Gao et al.,., 2014; Viegas et al., 2018) where the models showed an ability to adapt the way they express their emotions or to adapt 2 peoples faces for the partitioning of the dataset. The divisions were made in such a way that images of specific person only existed in one of the dataset for the separation of the datasets in training, validation and test data. The models will be tested and trained not only with different images but also with different persons , has images of same person only exist in one of the datasets as a results.

MODELLING :

Convolution neural networks (CNN) are explored to the classification of facial expressions. Because many image are not present to train CNNs, it is already decided to use only pre-trained neural networks and also we can apply transfer learning.

In the given all, the three selected neural networks were created it also based on the imageNet Large Scale Visual Recognition Challenge (ILSVRC) (RUSSAKOVSKY et al ., 2015). Neural networks were selected the VGG16 and VGG19 proposed in the (Simonyan & Zisserman. 2015) and also the InceptionResNet V2 (Szegedy et al., 2016). VGG16 and VGG19 are selected because these two are so simple and easy and it is also very straightforward architecturally and it is also referred to main network to used for transfer learning . InceptionResNetV2 it is also selected because it is mentioned in (Hung et al., 2019)it give us good result in presenting the facial expression classification tasks.

Pre-trained two network VGG16 and VGG19 they have only the differences is the no. of layers of the convolutional base. Convolution base is prepared by convolutional and max-pooling layers. Convolutional layers is used the3*3 pixels for filter, and for padding and stride 1 pixel is used. Max-pooling layers is used 2*2 pixels windows is used and for stride is 2 is used. In very layers ReLU activation function is used, only for Dense layer output to the classifier SoftMax is used.

In an architectural level, Inception-ResNet V2 it is very complex network than VGG, it is wider and the feltier with of different sizes, and is also organised in different blocks. The Inception-ResNet modules were prepared by convolution layers ,the reduction modules were prepared by the both convolution layers and max-pooling layers because it reduce the picture or image size around the network . activation function were used in net work is ReLU, so the exception of output layer is used SoftMax and remaining some layer were used in the Inception-ResNet modules it does not use the activation function.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105162 |

Because of the size of a train dataset , it is very necessary while the training process applied the data augmentation for the training the image. This data augmentation are :

- Rotation of 20degrees.
- Width and Height has 10% and 15% translations receptively.
- Brightness will change between the 0.2 and 1.
- 10%of zoom out and 20%of zoom in .
- It has horizontal flips.

Data augmentation is no need to image in test set.

TRANSFER LEARNING :

Transfer learning makes use of skills discovered while addressing a general issue as a springboard for addressing further issues. Using this method, it is possible to use learning to answer a problem faster than it would take without the benefit of prior knowledge. The usage of pretrained networks—networks that have been trained on a sizable benchmark dataset—is the most frequent scenario in computer vision.

There are two components to these pre-trained networks: the convolutional base and the classifier. The convolutional base seeks to extract characteristics from the image by typically consisting of stacks of convolutional and pooling layers. Feature extraction is the method in question. The classifier, which is frequently made up of fully connected layers, categorises the image using the attributes that the convolutional base has extracted.

In computer vision [9], the typical transfer learning workflow is as follows: I Choose a pre-trained network that addresses a problem related to the one that is intended to be solved (ii) Replace the classifier with a new one that will be trained on the new dataset (iii) Freeze the convolutional base and train the neural network [5] on the new dataset.

The architecture of the classifiers used in the transfer learning process was defined as the initial phase in the modelling process. Two distinct strategies, one a classifier based on a global average pooling (GAP) layer and the other using a convolution layer, were tested.

Only a GAP layer and a fully linked layer with seven neurons were employed for the Global Average Pooling technique, and Softmax was used as the activation function. There are no hyperparameters to define with this design.

The second method consists of two fully linked layers, two convolution layers, a 50% dropout to prevent overfitting, and a flatten layer that reshapes all the filters into a single array of one dimension. The final layer, which acts as the output layer and uses a SoftMax function with seven neurons, is made up of the activation function ReLU. There are a number of hyperparameters that must be chosen for this method.

In order to choose the best values, a variety of classifiers were trained.

In essence, we aimed to find the optimal number of neurons for the penultimate densely linked layer and the number of filters for the convolution layer. presenting four different setups:

64 filters with 256 neurons, 64 filters with 512 neurons, 128 filters with 256 neurons, and 128 filters with 512 neurons are the possible combinations.

Four configurations were trained and tested with each of the pre-trained networks in order to determine which one worked best. An optimizer had to be chosen in order to train these classifiers. The optimizer Adam was initially attempted, but intermittently he lost the knowledge he had gained up to that point and sank to significant loss values.

FINE TUNING:

It is first required to choose how much of the convolutional base to train before beginning the fine-tuning phase. The target and source domains' similarity as well as the quantity of the classification dataset are critical considerations (Marcelino, 2018).

Despite the fact that the ILSVRC, the source domain, offers photos for over 1000 different classes, none of them relate to individuals or their facial expressions, which is our target domain. It will be required to disregard any similarities between the two domains, which will support the usage of fine-tuning. Given that the train dataset is rather tiny compared to the target

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105162 |

domain's dataset, most of the layers of the pre-trained network must be adjusted. The segmentation of the layers to train followed this structure because both networks are block-structured.

The final three convolution blocks, conv5, conv4, and conv3, for the VGG, as well as Inception-ResNet [11] V2's Inception-resnet-C, Reduction-B, and Inceptionresnet-B, were fine-tuned.

EVALUATION:

Confusion matrices were employed to assess how well the models performed. With this set of data is at hand, it is then possible to extract a number of indicators to assess the models. To forecast the seven facial expressions, a multi-class classification will be used, and a binary classification will be used to categorise stress versus non-stress. As there aren't many differences in the distribution of the seven facial expressions across the three datasets, the models will be assessed using the accuracy and F1-score (harmonic mean of precision and recall) metrics.

MULTI CLASS EVALUATION:

The six models were tested using confusion matrices displaying each of the seven facial emotions after being trained and adjusted with the validation dataset for the best outcome. Inception-Resnet V2 is outperformed by the GAP network, and for both of these techniques, the convolution layer works better than the global average pooling layer. VGG16 will be utilised since it is less complex than VGG19 despite the fact that there is not much of a performance difference between the two best models, VGG16 and VGG19.

The total accuracy of the multi-class VGG16 model is 89.6%.

BINARY EVALUATION:

Stress detection is a binary problem, despite the fact that the models have been trained for multiclass classification, hence it is important to assess how well they perform when separating stress from nonstress. Anger, contempt, and terror were associated with stress in this binary evaluation, while the remaining test images were classified as non-stress.

When the classifier flags stress, precision reflects how often this classification is accurate. Hence, 91.8% accuracy indicates that 8.2% of people who were marked as stressed out were not. Recall is a gauge of the model's usefulness in identifying stressors. The majority of stressed-out people are actually tagged (our recall is high at 88.4%), and precision is prioritised over recall in the stress application.

V.CONCLUSIONS

We created a system that can take real-time pictures of the user's face and, using a classifier for facial expressions, determine whether the user is showing indications of stress and alert them in that event.

Transfer learning and fine-tuning were utilised to create the classification model. With respect to two alternative classifier designs, the pre-trained networks VGG16, VGG19, and Inception-ResNet V2 were taken into consideration. The best model, which had an accuracy of 92.1%, relied on the correlation between stress and facial expressions to make up for the lack of a dataset that was specifically categorised as stress- or nonstress-related. Further evidence and case studies are required to support this association. So, as part of our ongoing work, we will gather user input regarding the alerts that are sent to them by our system. This will make it possible to confirm if the system categorises stressful circumstances appropriately, boosting confidence in the link between stressful situations and unpleasant feelings. The enhancement of the categorization models [8] by training them with more information gathered from our users is another suggestion for future work. It will also be thought about moving the classification module to a server so that it may benefit from centralised processing with graphics cards and lessen the load on users' devices.

VI.ACKNOWLEDGEMENTS

The authors are grateful to the designers of KDEF [6] and CK+ for allowing us to use their datasets.

REFERENCES

1. Dewa, C. (2015, July). Barriers to the Utilization of Mental Health Services Among Depressed Employees and Work Productivity. The Journal of Occupational and Environmental Medicine is referred to as J Occup Environ Med. https://doi.org/10.1097/JOM.0000000000000472

2. Among the authors are Dinges (D. F.), Rider (R. L.), Dorrian (J.), McGlinchey (E. L.), Rogers (N. L.), Cizman (Z.), Goldenstein (S. K.), Vogler (C.), Venkataraman (S.), and Metaxas (D. N.) (2005). ICEIS 2021 - 23rd International Conference

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105162 |

on Enterprise Information Systems 262 related stress brought on the performance demands by optical computer recognition of facial expressions. 76(6 Suppl), B172–182 in Aviation, Space, and Environmental Medicine.
3. P. Ekman, W. V. Friesen, & Co. (2003). Unmasking the Face: A Manual for Interpreting Facial Indicators of Emotion.
4. Giannakakis, G., Marias, K., Simos, P. G., Tsiknakis, M., Pediaditis, M., Manousos, D., Kazantzaki, E., Chiarugi, F. (2016). Detecting stress and anxiety using video-based facial clues. 31, 89–101, Biomedical Signal Processing and Control. the following URL: 10.1016/j.bspc.2016.06.020.
5. Hung, J. C., Lai, N. X., & Lin, K.-C. (2019). utilising transfer learning and convolutional neural networks to recognise emotion while learning. 84, 105724, Applied Soft Computing. J. Asoc. 2019:105724 (doi:10.1016/j.asoc.2019.105724)
6. D. Lundqvist, A. Flykt, & A. Hman (1998). The Karolinska Directed Emotional Faces (KDEF), CDROM from the Psychology division of the Department of Clinical Neuroscience at the Karolinska Institute, ISBN 91-630-7164-9.
7. Pruski, A., Bousefsaf, & Maaoui, C. (2015). Using photoplethysmographic data from webcams, automatically identify human stress. Mechanics in Medicine and Biology Journal, 16. The following URL: 10.1142/S0219519416500391
8. Marcello, P. (2018). Learning transfer from previously taught models. https://towardsdatascience.com/transfer learning-from-pre-trained-models-f2393f124751
9. The following individuals are involved: Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L. (2015). Large-Scale Visual Recognition Task for ImageNet. 115(3), 211-252, International Journal of Computer Vision.
https://doi.org/10.1007/s11263-015-0816-y
10. K. Simonyan and A. Zisserman (2015). Large-Scale Picture Recognition Using Extremely Deep Convolutional Networks. "Cs" in ArXiv: 1409.1556.http://arxiv.org/abs/1409.1556
11. Alemi, A., Vanhoucke, V., Ioffe, S., & Szegedy, C. (2016)., The Effect of Residual Connections on Learning with Inception-v4, Inception-ResNet, and Inception-v4. "Cs" in ArXiv:1602.07261. http://arxiv.org/abs/1602.07261
12. Maxion, R., Hauptmann, A., Lau, S.-H., and Viegas, C. (2018). A Dependent Model Utilizing Facial Action Units for Independent Stress Detection. https://doi.org/ 10.1109/CBMI.2018.8516497. 2018 International Conference on Content-Based Multimedia Indexing (CBMI), 1-6