Depression Level Calculation for Predicting Child Psychometric Retardation using DepressNet Approach through GPU Accelerated Google Cloud Platform

M.Sharmila[a], R.P.Dharshinie[b], A.Keerthana[c], K.Deepika[d] and T.Ananthi[e]

[a] Assistant Professor, Department of Information Technology, M.Kumarasamy College of Engineering, Karur, India.
[b]UG Scholar, Department of Information Technology, M.Kumarasamy College of Engineering, Karur, Tamil Nadu, India.
[c]UG Scholar, Department of Information Technology, M.Kumarasamy College of Engineering, Karur, Tamil Nadu, India.
[d]UG Scholar, Department of Information Technology, M.Kumarasamy College of Engineering, Karur, Tamil Nadu, India.
[e]UG Scholar, Department of Information Technology, M.Kumarasamy College of Engineering, Karur, Tamil Nadu, India.

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 20 April 2021

Abstract: From the past decades, depression has become a common threat for everyone. Doctors often ignore the intellectual disability of young children. This happens as a single discovery or as a part of the problem in 2% to 3% of the population. There are many reasons for the widespread syndrome or mental disability. Among the observed patient data 30% - 50% of cases, cannot be determined by the doctor about the cause despite through examination. The diagnosis mainly depends on a complete personal family medical history, a complete physical examination, and a careful evaluation of the child's development. Being able to deal with these intermediate stages of emotions will be an easier preventive measure. This measure will also help to build a healthy society. There must be a way that does not compromise people's privacy. Human-computer Interaction (HCI) along with Machine Learning algorithms paved the way for finding solution for certain children with depression. One can analyse these issues of depression prediction among children easily with the process of image detection. The goal at the beginning of the project was to create a reliable toolkit for children. Data sets from sources such as AVEC can be used to build real-time Depression prediction systems. There are many ancient systems for recognizing emotions, so this proposed ResNet Algorithm promotes good results when comparing to other detection approaches of depression. However, all this work aims to establish a system based on the analysis of depression. The researchers can rely on this system to make a good model of depression analysis in future and can solves the problem in hand.

Keywords: Machine Learning, Feature Detection, Depression Analysis, Pattern Identification, Retardation Rating Scale, Neural Network, Logistic Regression.

1. Introduction

Depression analysis with facial pattern focuses on the analysis of depression level of people. In recent times, there are a huge number of reports about the depression of child it is a recorded fact that one in five Chandigarh. Since the world is moving into a single room for every work and also due to the digitalization of the world there is an increase in depression among the people. There may be a various factor for the depression of people other than this. But most people move forward to work instead of looking upon for the cute of the depression. So we develop a model for the depression analysis of child. The main objective of our project is to create a model which identifier the depression of child and also the intensity of the depression by over viewing the existing the ideas. It can be extracting the features from the images and videos. The primary setup is done by creating deep regression network with the layers of convolutional neural network and there by an interaction within themselves. This model includes four basic steps, the primary step is feature extraction. Here we take the input dataset as images or frames of the video and pre-process them for the feature extraction. The second major task is to create a depress-net with layers of CNN. The third step is to create a depression activation map to identify the features in the different areas of our face. In second stage[1], the image can have top, middle, bottom and full face to increase the image localization capacity of our model. The final and the fourth part of our model are to analyze the results from the depression activation map. The final scores will give us the depression status of the input face model. This is the basic overview of our model. Various image recognition methods have been implemented to develop this model. We use keras in Tensor inflow to build the layers and networks of our model. It can be created with the help of simple software implementation. The purpose of this project is to analyze depression, its purpose is to identify the child's depression state, so as to improve their condition before the problem occurs, so that they understand their current situation. You can create interactive ways to improve image processing.

The complete organized system of our project is explained below. It [2] explains methods in the image localization related to our project which are the basic standards for us to do the image localization and also the depression analysis with the localized models. The following content explains the overall view of the methods used in the project including feature extraction, depress-net, score calculation, final analysis and prediction. And also the experimental results, evaluation metrics of the project was denoted. Finally, further development of the project and improvement in the future has also been elucidated.

DepressNet Approach

The Depression Analysis is done with the help of facial Recognition which has the combination of both regression and the classification. This Depression Analysis model[3] is used to identify the depression state of the children. This depression identification has several phases of face detection, depression score calculation, building of the Dam module and our final stage will be deployment. In order to find out the normal child, we should do the camera implementation section. So that we can perform the depression analysis without the prior data collected from the test subjects.

The workflow diagram (Fig 1) is the model for implementation the above concept.

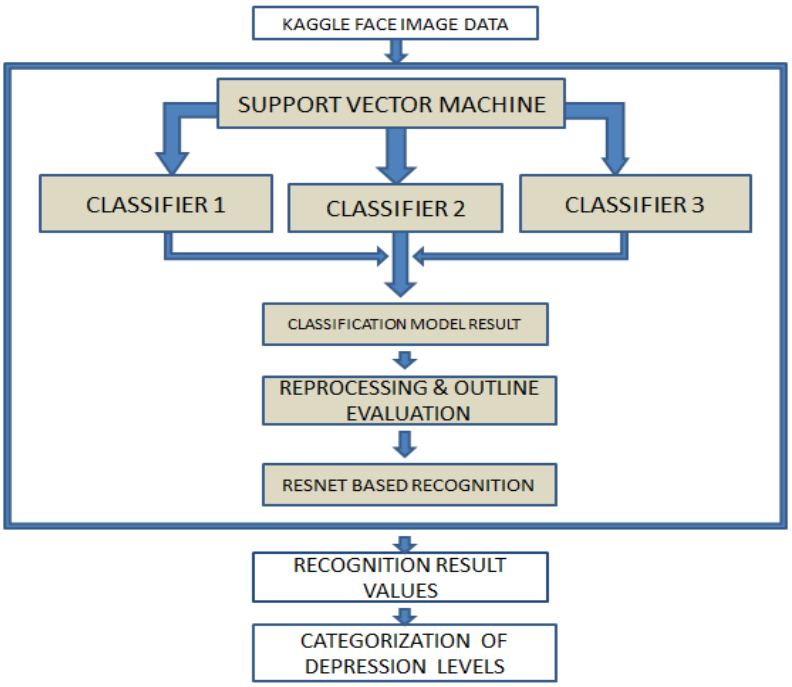

Figure 1 - Depression Analysis Process

Kaggle face image Data

Using Kaggle, customers can discover and post statistics sets, discover and construct fashions in a web-primarily based totally computing environment, collaborate with different statistics scientists and system gaining knowledge of engineers, and take part in statistics technology hassle-fixing competitions. The hassle description makes use of exceptional strategies to test and compete with every different to create the first-rate model. These paintings became posted on Kaggle Cores to higher take a look at and encourage new ideas. It may be dispatched through the Kaggle kernel, manually uploaded or thru the Kaggle API. In maximum competitions, the gadget at once ranks the application (primarily based totally at the accuracy of the prediction in comparison to the hidden answer file) and aggregates it within side the. At the end of this period, the competition organizer will acquire a coins prize in change for a global, permanent, irrevocable and loose license for the triumphing bid, that is, except in any other case stated, the algorithm, software program improvement and associated highbrow assets rights.

Face Detection

In the construction phase, our face detection will be the last module. But in the implementation section it will be exactly opposite[4], it will be first module. In order to take the input, our first goal is the face detection and follows the order of the workflow diagram.

Our model also has the option to prediction with single image and also with recorded video or live video capture method. Prediction with single image is an easy task where we can predict the depression levels of the test subject with a single image. Our motive is to build a tool which has the live video capture option. Below is the live video capture.

In this face detection module[5], we are taking inputs from the child through camera. Our first step is the code to turn on the camera and run it to infinite time. The user has the access to stop the camera recording. But there is a minimum timing of 10 seconds for which the video will be recorded, that they cannot stop the recording before 10 seconds. Since the minimum timing is required for the depression score calculation. So it is necessary to record the face for at least 10 seconds. If there is in any case, the video which was recorded below 10 seconds, the process will restart until the minimum timing criteria is achieved.

Our next phase is the picture calculation. Here we are taking the seconds and also the frames into account. Our motive is to have ten images. We will divide[6] our time or time or frames into ten sections. We will then pick randomly an image from one section for the selection unwanted noise. Finally, our data is ready to be given into the DepressNet module.

DepressNet and DepressScore

DepressNet is known to be as the significant module. It covers the major work in this depression analysis implementation process. It has important two sub-modules. These two sub-modules in the construction of building DepressNet were listed below. 'Identify Block' 'Convolution block' DepressNet Depress Score This module will cover calculation of the depression depth. Our motive is to reduce the loss as much as possible. In normal models[7], the output from one model (Y) is used for the error calculation and the same is added to the input of the next layer. Here the loss cannot be reduced in a major order. In our model, we pre-recorded the input to the layer. The output of the layer and the pre-recorded input to the layer are added to perform the F(X) function. The F(X) [8] function is an error calculated and loss minimized function which can be given as an input of the next of year. When we add both pre-recorded input and output values, the result will be more accurate. If the loss Y was low then it can be added with pre-recorded input, So that we can get an output F(X) with low variation. On the other hand, if the loss Y was high then it can be added with the pre-recorded input. So that the output variation will be high defined in the further progress.

Identity Block

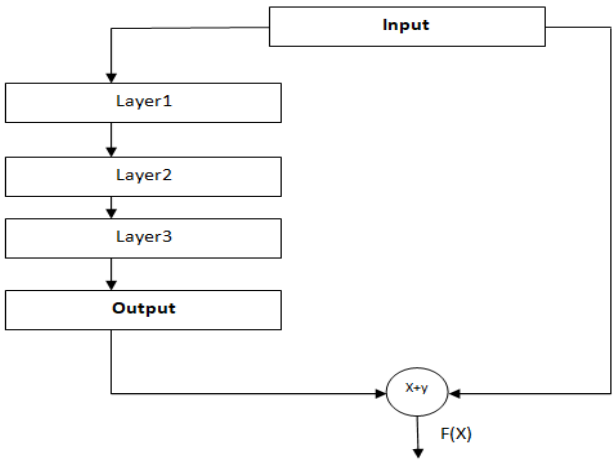

Figure 2 - Identity block diagram

In the Identity block, the input and the shape will remain same. So in identity block there is no complex construction. The flow of the model[9] will be sequential except the final output. Here we are alternating the flow, in order to achieve f(x)-x is nearly equal to zero. The input to the next model should have the same and the features which shall not be reduced while we are moving further. Identity block has three layers from which two layers have 1x1 filters and one layer has 3x3 filter. Each layer should be followed by the regularization and the activation except the final layer. In the final layer[10], after the regularization phase is completed we will add the output and Input of the identity block. This result will be activated with the help of final activation layer and returning the output.

Convolutional Block

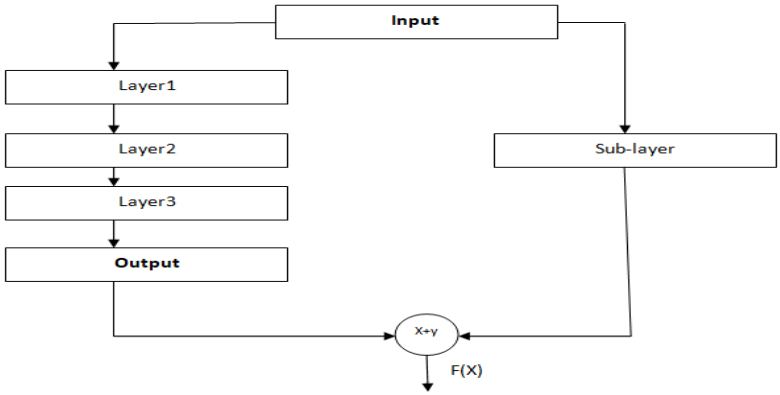

Figure 3 - Convolution block diagram

This is the block diagram of the convolution block. In the convolution block, the input and output shape of our will remain different, which requires some complex construction. This flow will be same as the identity model except for the final year. The filters can be given as an input of the convolution block. So that it can be used as dynamic filter in a specific area. Since the input of the convolution block and the output of the layer 3 remain different, we are performing this process to convolve the input of the convolution block. We will construct a layer in the path of input to output. Here[11] we will convolve the shape of the data from the input to the required shape for the output. Now the shapes will be equal so that we continue our regular process. These normal three layers

perform the operations in convolved result from the input and the normalized output of third layer will be added and combine. Finally they are activated before returning them to the DepressNet module.

DepressNet Prediction of Depression

It is the principle module so as to name each the identification block and the convolution block. DepressNet has absolutely 5 tiers. The first level is tough coding staging level and the opposite tiers are the mixture of the identification block and the convolution block. First is to feed the enter from to the X_Inp to create a dummy version. Our subsequent step is to create 0 padding. It is critical due to fact we're appearing such a lot of convolving, pooling and striding operations within side the future. In order to hold the threshold capabilities of our enter we're developing a 0 padding of 3x3 round our version. The output from those tactics may be fed as an enter to the X. After the ones pre-process[12], we can flow into the primary level. In this primary level, we can create our first convolution layer with sixty four neurons of 7x7 kernel sizes so as to be observed via way of means of the normalization layer, activation layer with the relu characteristic and sooner or later the pooling layer. It may have simplest one layer considering that it's miles a strategy planning stage and does a few pre-processing. Second level include one convolution block and identification blocks with the clear out of [64,64,256]. Each block has 3 layers. So 2nd level includes 9 layers. Then it's miles observed via way of means of the 1/3 level. Third level includes one convolution block and 3 identification blocks with the clear out of [128,128,512]. Since each block has 3 layers, this level incorporates absolutely twelve layers. Fourth level has the bulk of layers. Fourth level has one convolution and 5 identification layers. Since it has six blocks, it will likely be which include eighteen blocks with the clear out of [256,256,1024]. Final level[13] is the 5th level which has 3 blocks of 1 convolution block and identification blocks having the clear out of [512,512,2048] having nine layers. Finally we should take common pooling with the assist of common pooling layer with the stride of 2x2. Next step is to take international common pooling with the assist of world common pooling layer. This layer will lessen the peak and duration of the version to one. Then this end result is handed to a dense layer of seven neurons considering that we've got seven forms of classes. Then the ensuing version with the above structure is lower back to the user. This very last structure is the DepressNet.

Depress Score

The above DepressNet discusses about construction of the architecture. We have trained the model with the depressNet architecture of about 28,000 samples. Finally[14] we will predict with the above model, with single or several samples. We pass the prediction result to the depress score section. The depress score section will calculate the depression score of a single or several images and return the single output. This is a calculation phase. It has a simple plication and division with respect to the relations. It will calculate depress score for every image and finally it will return score of the image. This phase is to get the normalized result.

Deployment

The deployment part is the final module. Here the entire project is integrated and deployed. We finally get the depression state of the child and the high lightened areas of the face with help of heat map generation.

Classification of Depression Level

Based on output we can find the depression level. In addition to that it will return some statement which states the emotions of the person in the image. Some common statements are joy', 'sadness', 'shame', 'guilt', 'disgust'...etc. Figure x shows some common emotions.

Figure 4 - Common Emotions

Psychomotor retardation is a long-term melancholic lifestyle, and it can also additionally have vital scientific and therapeutic importance in treatment. Because it negatively influences the average traits of sufferers

with depression[15], we suppose it is inherently important. Successful interventions and associated techniques can enlarge consciousness of psychomotor problems and possible depression, and can additionally lead to greater research and greater diagnosis. This meditation involves the foremost traits of depression. Psychomotor retardation looks to be a key element in predominant depressive sickness (MDD). It is one of the 9 most important signs of MDD. Although the superiority of psychomotor retardation in MDD has lengthy been discovered, the traits and scientific importance of psychomotor retardation have no longer been. This examines natural correlations, dimensional results, and psychomotor retardation results. There is no consensus on whether or not the severity of melancholy is associated to the presence or diploma of psychomotor disorders. Various researches[16] have proven that the severity of despair is associated to delayed psychomotor development.

The Hamilton Depression Rating Scale (HDRS) has a high-quality correlation ($r=0.69$, $p=0.010$) with the Salpetriere Retardation Rating Scale (SRRS), a symptom severity scale special to psychomotor retardation. Other lookup states[17] that neuropsychological measures have a larger correlation with psychomotor retardation than despair severity. For example[18], the SRRS was once as soon as discovered in a analyze about to be notably correlated ($r=0.67$, $p<0.005$) with Posner's covert orientation of seen pastime test, motor responsibilities ($p<0.02$), and exceptional neuropsychological measures of psychomotor retardation. In one study[19], in distinction with healthful controls, MDD victims tested large bodily delays, even as victims with dyskinesias did not. However, effects indicated that subjects with melancholic selections had been more in all probability to private cognitive content cloth retardation than matters with excessive depressive symptoms. The discrepant findings[20] between lookup that show off a correlation with psychomotor retardation to despair severity or psychology measures have to relate to variants in methodology, sample sizes, and measure techniques of psychomotor retardation.

Empirical Measures

Depending on the dataset the machine learning accuracy can vary, the above table represents the highest accuracy that can be obtained with the classification models. In the above table DepressNet is the machine learning model that has been proposed in our system.

For Example:

Table 1- Machine Learning Model Percent of accuracy and precision

Machine Learning Model	Percent of accuracy	Precision
DepressNet	88.5	0.788
Support Vector Machine	87.0	0.65
Logistic Regression	78.0	0.56
SDA Neural Network	73.0	0.7
Random Forest	82.0	0.63
K Nearest Neighbors (KNN)	60.0	0.55
1d-CNN	69.0	0.64
DepAudionet	65.0	0.50

Support Vector laptop is used for the classification purpose, assist vector laptop produces outcomes in the binary structure whether or not the face in the photograph is depressed or not. Support vector computing device managed to get a end result of 87.0 percentage accuracy however SVM can't distinguish the quantity the melancholy that is measured when an picture is exceeded i.e, 87 percentage accuracy is bought solely if the classification is binary.

Logistic regression had the benefit to produce non-stop effects that is the quantity of despair in the picture however evaluating depressnet and assist vector machine, logistic regression algorithm supplied a decrease accuracy of about seventy eight percent, and additionally the accuracy of seventy eight will be got solely if the dataset includes appropriate examples of the photos that are used for training, or else the accuracy can limit to fifty eight percent.

SDA neural networks make use of the neurons and its activation features to produce a community which is known as neural community for the reason of classification. This is comparable to DepressNet which makes use of low layers when in contrast to DepressNet however as a change off the accuracy bought decreased to seventy three percent.

Random wooded area falls below the supervised getting to know strategies the place a tree is created, with guidelines as its branches in order to classify the data. The dataset used to create guidelines ought to be cleaned with greater pre-processing methods in order to keep away from overfitting the model. Even after that the random woodland managed to provide eighty two percentage and on the different hand the DepressNet used longer time to create the mannequin instead than pre-processing.

Pre-processing may range for exclusive sorts of snap shots consequently DepressNet have a correct chance to classify the facts in a accepted way. DepressNet managed to get a improve of 6 percentage in phrases of accuracy in assessment to random forest.

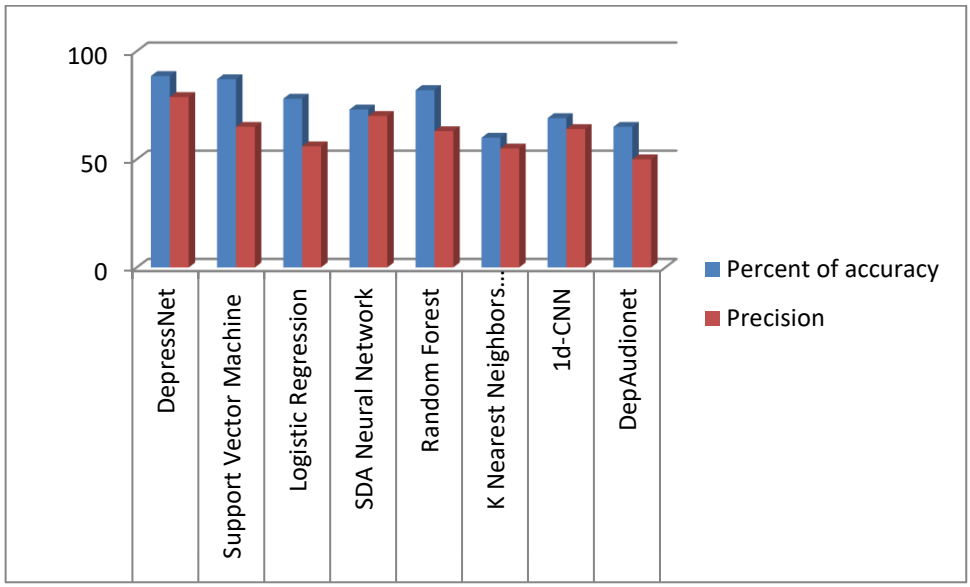

Figure 5 - Machine Learning Model Percent of accuracy and precision

The enter function set is given to a Support Vector Machine classifier for training. The Support Vector Machine is a mannequin that splits the two units in the nice feasible way. This is the excellent break up due to the fact it is the widest margin that separates the two groups. This line is known as the hyperplane. The nearest factors are known as the Support Vectors. Score will be calculated with the values of depress_data which is the predictions data in list by the listed formula.

score = (depress_data[0]*0.1) + (depress_data[1]*0.64) + sqrt(depress_data[2]) +
(depress_data[3]*0.5) + (depress_data[4]*0.3) + (depress_data[5]*0.73) +
(depress_data[6]*0.9) - 0.1 (1)

Variety of input information – We will give normal images to train a ResNet model for the face detection in the image of the given subject. We then will give the depressed videos or the images from the clinical diagnosis of the subject to train the depression analysis model.

Accuracy – We are calculating accuracy as the metrics for our problem. Our motive is to reduce the error as much as possible to achieve our required accuracy.

Experimental Results and Discussions

Different analysis has been identified and experimented in order to achieve efficient outcomes and results. Thousands of images are passed to find out the capacity of our model in order to increase the efficiency of our model.

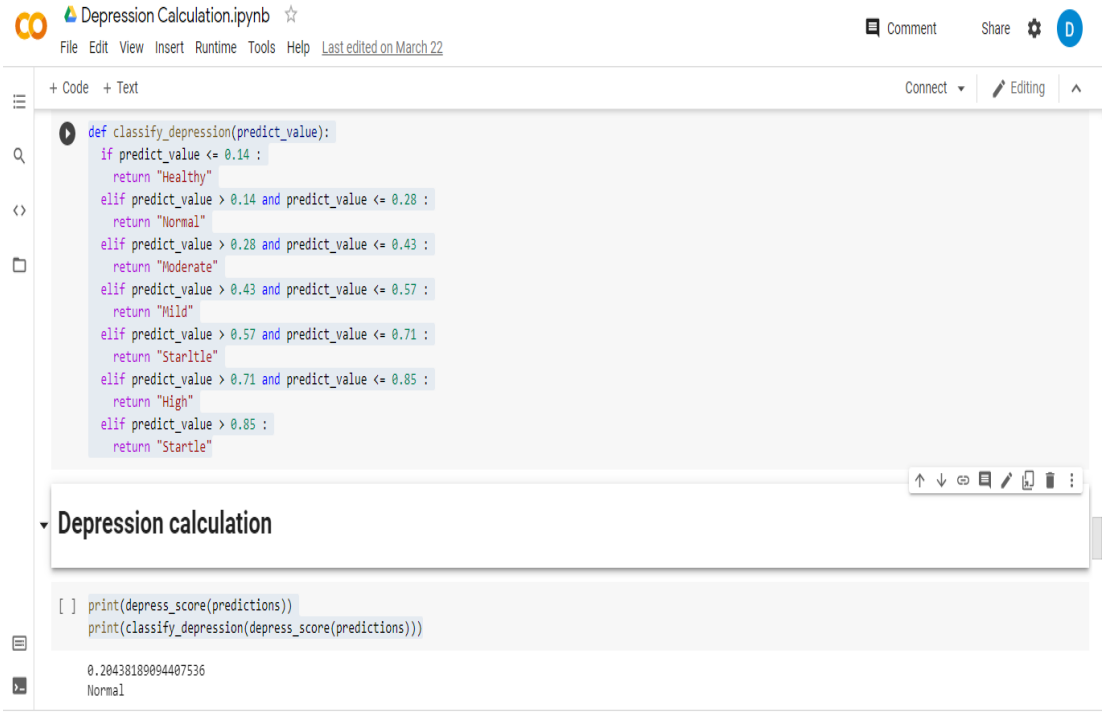

Figure 6 - Depression Calculation

Figure 7 - GPU Configuration

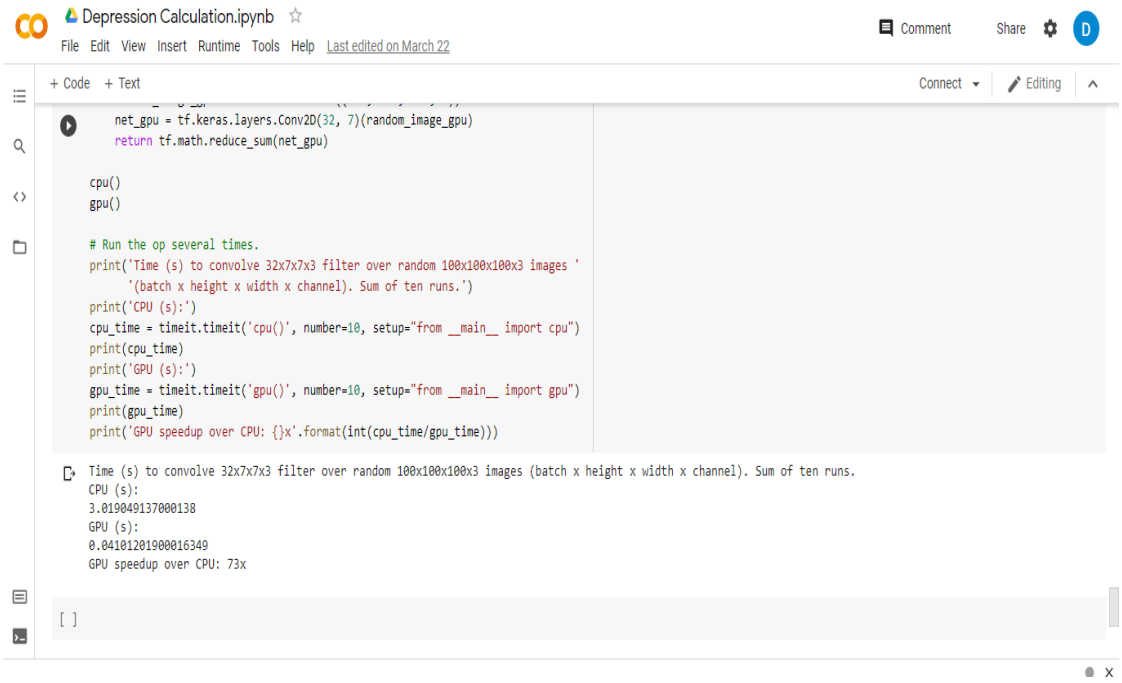

Figure 8 - Execution time of CPU and GPU

Thus the final heat map with image is generated to show the results. The depression score or the intensity of the given subject is identified with possible results.

Conclusion and Future works

Many general medical depression patients are eligible for treatment in several medical diagnostic institutions. Although the follow-up treatment is usually outside the scope of current practice guidelines, about half of primary care depression cases are recognized. To discover the state of depression, it is easy to negotiate with the counselor and bring a happy life to the client. In future, we plan to implement depression data from various models, including audio, video, and facial dynamics. Thus, the proposed DepressNet approach will also proves its efficiency in analyzing multiple forms of patient data for both youths and adults in upcoming days and improves its optimization.

References

1. Alex Krizhevsky, Ilya Sutskever and Geoffrey E.Hinton, ”ImageNet Classification with Deep Convolutional Neural Networks”, (2012) .
2. Sharifa Alghowinem, Roland Goecke, Michael Wagner, Julien Epps, Matthew Hyett, Gordon Parker and Michael Breakspear, ”Multimodal Depression Detection: Fusion Analysis of Paralinguistic, Head Pose and Eye Gaze Behaviors”, (2016) .
3. Asim Jan, Hongying Meng, Yona Falinie Abdul Gaus and Fan Zhang , ”Artificial Intelligent System for Automatic Depression Level Analysis Through Visual and Vocal Expressions”, (2017) .
4. Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva and Antonio Torralba, ”Learning Deep Features for Discriminative Localization”, (2015) .
5. Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed,Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke and Andrew Rabinovich, ”Going Deeper with Convolutions”, (2015) .
6. Zhongzheng Ren1, Zhiding Yu ,Xiaodong Yang, Ming-Yu Liu,Yong Jae Lee, Alexander G. Schwing and Jan Kautz, ”Instance-Aware, Context-Focused, and Memory-Efficient Weakly Supervised Object Detection”, (2020)
7. Naoto Inoue,Ryosuke Furuta,Toshihiko Yamasaki and Kiyoharu Aizawa, ”Cross-Domain Weakly-Supervised Object Detection through Progressive Domain Adaptation”, (2018) .
8. Spyros Gidaris and Nikos Komodakis, ”Object detection via a multi-region & semantic segmentation-aware CNN model”, (2015) .
9. Chan TH, Jia K, Gao S, Lu J, Zeng Z and Ma Y, ”PCANet: A Simple Deep Learning Baseline for Image Classification”, (2015) .

10. Jorge Alberto Marcial Basilio and Gualberto Aguilar Torres, ”Explicit image detection using YCbCr space color model as skin detection”, (2016).
11. Salvatore, Christian, Antonio Cerasa, Isabella Castiglioni, F Gallivanone, A Augimeri, M Lopez, G Arabia, M Morelli, M C Gilardi, and A Quattrone, “Machine Learningon Brain MRI Data for Differential Diagnosis of Parkinson’s Disease and Progressive Supranuclear Palsy”, (2014).
12. Sahli, “Multimodal Measurement of Depression Using Deep Learning Models”, (2017).
13. Cruz, Joseph A, and David S Wishart, “Applications of Machine Learning in Cancer Prediction and Prognosis.”, (2006).
14. Cohn, J F, T S Kruez, I Matthews, Y Yang, M H Nguyen, M T Padilla, F Zhou, and F De la Torre, “Detecting Depression from Facial Actions and Vocal Prosody.”, (2009).
15. Al-Mosaiwi, Mohammed, and Tom Johnstone, “In an Absolute State: Elevated Use ofAbsolutist Words Is a Marker Specific to Anxiety, Depression, and Suicidal Ideation”, (2018).
16. Bergstra, James, and Yoshua Bengio, “Random Search for Hyper-ParameterOptimization”, (2012).
17. Cohn, J F, T S Kruez, I Matthews, Y Yang, M H Nguyen, M T Padilla, F Zhou, and F De la Torre, “Detecting Depression from Facial Actions and Vocal Prosody”, (2009).
18. Wang, Peng, Frederick Barrett, Elizabeth Martin, Marina Milonova, Raquel E Gur, Ruben C Gur, Christian Kohler, and Ragini Verma, “Automated Video-Based Facial Expression Analysis of Neuropsychiatric Disorders”, (2008).
19. Marsland, Stephen, “Machine Learning: An Algorithmic Perspective. Chapman andHall/CRC”, (2011).
20. Kroenke, K, R L Spitzer, and J B Williams. 2001. “The PHQ-9: Validity of a Brief Depression Severity Measure”, (2001).