



Machine Learning Analysis of Anxiety and Depression in children

**Gudepu Sandeep Reddy, Mohammed Saquib², Nalabolu Nagaraju³, Mohammed Shahid⁴,
Mrs. Farahana Kousar⁵**

Atria Institute of Technology, Visvesvaray Technological University, India

Abstract: Psychological issues among students, including as depression and anxiety, are mostly caused by a failure to check students' psychological well-being on a regular basis. If depression is discovered early enough, it can be treated successfully.

Image processing advancements have resulted in the development of successful algorithms that can identify emotions from facial pictures in a much more basic manner. As a result, for successful diagnosis of depression, we require an automated system that records and analyses student facial photos. In the proposed system, image processing techniques are used to analyse student facial features and depression can be predicted.

To predict depression, a video of the student is gathered, and the student's face is identified by use of haarcascade classifier. The Mini Xception classifier is used to classify these facial traits. After quantifying the system, we will be able to implant it in mobile devices in the future.

1. INTRODUCTION

Depression among college students is a consequence of cultural change brought on by the online, cell phones, and numerous social networking sites. Because of the social stigmas associated with depression, as well as peer pressure, the majority of students conceal their psychiatric difficulties. Some students are completely unaware of their psychological issues, and as a result, they are unable to receive any assistance that would be beneficial to their mental health. The counsellor faces a challenging task in keeping a record of the significant events that happens in children as a outcome of depression in a large number of students.

As a result, we require an automated system that records and analyses photos of kids in order to detect depression. The most essential nonverbal communication tool for expressing a person's emotional or mental condition is facial expressions. A large number of research are currently being undertaken on 'Facial feature analysis' for emotion detection from images, that could aid in the prediction of human mental wellbeing. By analysing frontal facial photos of college students, this study suggests an automated technique for detecting depression levels in students.

(Pampouchidou, A., O. Simantiraki, C-M. Vazakopoulou) proposes that facial landmark points can be used to identify depression. The distances between them were calculated using the geometrical and city blocks distance methods. This stage retrieves, fuses, and categorises both visual / voice features.

(Wang, Yutai, Xinghai Yang, and Jing Zou) suggest using voice signal processing and facial training recognition to identify emotion. Phonation characteristics from audio signal and facial appearance from video signals were extracted and classified in parallel and the findings of both classifiers are combined utilizing 'Bimodal' integrations for the concluding expression recognition result.

2. DATA ACQUISITION

Every machine learning-based problem-solving strategy demands a substantial amount of data collection, which is a crucial decision factor. In this work, in addition to detecting emotion, I also detected a subject's gender, for which I used two datasets from Kaggle:

- FER2013 ((Facial Expression Recognition 2013) emotion classification datasets:
Fer2013 comprises almost 30,000 face RGB photos with various emotions with a size limit of 4848 pixels. It consists of the classes that follow:
- Surprise
- Fear
- Disgust
- Happy



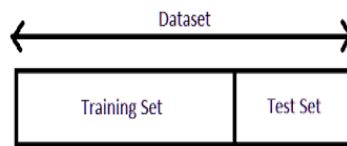
- Sad
- neutral
- Angry
- IMDB categorization based on gender datasets

It consists of following classes : Man and Woman

The datasets collected from the aforementioned sources are in raw format, and must be pre-processed before being fed into the opencv model.

Data pre-processing:

1. Images of both the dataset were normalized in the range [0..1] and resized.
2. Converted class labels to a single hot encoder, which is needed for inference.
3. Split the dataset into a train-test set in the ratio 80:20.



3. METHODOLOGY

To accomplish this task I used 3 computer vision models they are:

- Haarcascade Classifier (for Face extraction)
- Mini Xception (for emotion classification)
- Simple CNN (for gender classification)

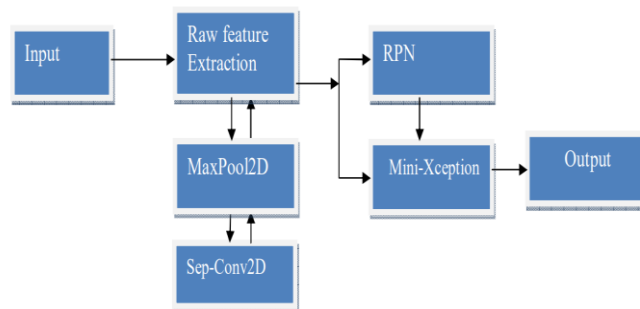
Haarcascade Classifier :

The HaarCascade classifier will be used to implement our use case. In their 2001 publication "Rapid Object Detection using a Boosted Cascade of Simple Features," Paul Viola and Michael Jones together created the HaarCascade classification like an effective object identification solution.

A haarcascade feature is developed out of a large number of sample images like (positive and negative) in this Machine learning technique. Based upon this training, algorithm will be brought into use to recognise objects in other photos. It features an extensive number of individual xml format, each of which corresponds to a highly specific application. Using the haarcascade frontal face default.xml, we will attempt to detect the faces of individuals in this use-case.

The facial predictor, that is an element preloaded with haarcascade frontal face default.xml, we're using an inbuilt function named detect MultiScale. The feature can aid us all in identifying the characteristics and location of the different picture. It will be using all of the characteristics out from face classifier objects to identify the characteristics of the proposed image.

Mini Xception :



We had implemented the recommended model for emotion recognition, which was based on improved deep learning algorithms. The Xception paradigm was used to create a mini - Xception for sentimental analysis. The Xception algorithm



influenced our proposed model. This structure employs residual modules and context separable convolutions. Residual modules alter the expected projection across two succeeding layers, resulting in training data that deviate from the initial and expected feature maps. As an outcome, the desirable features $H(x)$ are tweaked in able to fix a simplified training problems $F(X)$ as it follows:

$$H(x) = F(x) + x$$

Using depth-wise separable convolutions, the features of the suggested model are minimized from the Convolution layer. Depth-wise separable convolution are divided into 2 layers: depth-wise convolution and point-wise convolutions. In the presented work, which contains four residual depth-wise separated convolution layers, every convolution is being preceded by a batch normalization procedure and a ReLUs activation function. The last layer employs average global pool as well as a soft-max activation function to produce a forecast. In this structure, there are around 60 thousand parameters.

Simple CNN :

It's a straightforward convolutional neural network that determines if a person is a male or a female. A pre-trained mini-Xception model was employed in this case as well, but the weights were fine-tuned for the downstream task (gender classification). Except for the last few layers, this model's layers are all frozen.

Steps to take while making an inference:

1. Load pre-trained weights of haarcascade classifier, mini-xception and simple CNN.
2. Define bounding box hyper-parameters (emotion and gender offsets).
3. Get the input (image) shape of these computer vision models.
4. Initialize a counter variable to count angry, sad, depression and fear face counts.
5. Open the webcam and use opencv to capture a video in order to recognise emotions and gender, then iteratively repeat the following steps:
 - A. Convert the captured video frame (image) from BGR to Gray and RGB scale.
 - B. Detect face and extract its contours (face coordinates) from grayscale image using haarcascade classifier.
 - C. Extract faces from an image and feed it to the model to get the emotion and gender of a subject.
 - D. Draw a bounding box around the person's face and display the emotion and gender on the top of the video frame along with its count.
6. When a user taps the 'Q' key on the keyboard, the camera is released.

4. RESULTS

Performance of Mini-Xception model

Expression	Anger	Disgust	Fear	Happy	Sad	Surprise	Neutral
Anger	60.0	0.1	0.0	0.0	25.0	0.0	10
Disgust	0.0	55.0	0.5	0.7	11.9	25.0	0.0
Fear	12.9	0.01	65.78	0.5	0.12	17.8	0.0
Happy	0.02	0.5	0.02	95.55	0.08	14.5	0.78
Sad	25.0	11.9	0.12	0.08	85.0	0.05	7.50
Surprise	0.0	25.0	17.8	14.5	0.05	73.68	11.9
Neutral	10.0	0.0	0.0	0.78	7.50	11.9	89.66
Average Accuracy =95.60							

To calculate accuracy, the confusion matrix is employed. For quantifiable comparison with experimental results, the essential performance metrics, including precision, are investigated (AC). True Negative, True Positive, False Negative and False Positive are the variables we use to do this false negative.

On the testing dataset, it had a 95.6 percent predictive performance.

Performance of Haarcascade classifier :

Mostly on testing dataset, it had a precision of 96.24 percent. The aim of this exercise was seeing how much depression was present in various videos. The ability to recognise depression in videos is just one part of the total approach of diagnosing depression. Students who may not be depressed at the moment may experience depression in the coming years. As a result, all their extracurricular activity must be closely monitored at all times. As a result, more footage of the same



student captured at different times can be considered in future studies. This may aid in the analysis and comparison of the student's previous and present mental states, as well as providing more information to the depression level detection procedure.

5. CONCLUSION

Anger, disgust, anxiety, joy, sad, surprised, and neutral were among the seven emotion expressions we attempted to recognise. Many recent efforts have used deep learning models to provide an end-to-end framework for emotion expression identification. Even if recognising emotions is a difficult endeavour, This appears to be an area where there is a lot of potential for development. The presence of 'Joy,' 'Neutral,' (positive emotion) and 'disgust,' and 'anxiety,' (negative emotion) facial expressions were discovered and analysed in depression films. The training and testing datasets were recorded separately, and facial features were identified using Cascade classifier and mini xception model. The positive and negative emotions in each video were analysed, as well as videos was assigned to one of three categories: 'Extreme Depression,' 'Moderate Depression,' or 'Low Depression.' The classifier predicted the outcomes with the highest accuracy.

The classifier predictions will be even more precise if there are enough training samples. The training videos were recorded in over a thousand frames, but only the first few were used in the assessment. By identifying the video's important frames and extracting them only with the use of a key frame extraction method for following work, this procedure can be performed for the full video. The current research is limited to the student's most recent video . However, for a more accurate detection of depression, the student's past should be taken into account.

6. REFERENCES

- [1]. Cohn, Jeffrey F., Tomas Simon Kruez, Iain Matthews, Ying Yang, Minh Hoai Nguyen, MargaraTejera Padilla, Feng Zhou, and Fernando De la Torre. "Detecting depression from facial actions and vocal prosody." In *Affective Computing and Intelligent Interaction and Workshops*. third International Conference on ACII, pp. 1-7. 2009, IEEE.
- [2] Najmeh Samadiani , Guangyan Huang , Borui Cai , Wei Luo , Chi-Hung Chi , Yong Xiang and Jing He "A Review on Automatic Facial Expression Recognition Systems Assisted by Multimodal Sensor Data". *International Congress of Information and Communication Technology (ICICT 2017)*.
- [3] Owayjan, Michel, Roger Achkar, and Moussa Iskandar. "Face Detection with Expression Recognition using Artificial Neural Networks." In *Biomedical Engineering (MECBME)*, 3rd Middle East Conference on, pp. 115-119. IEEE, 2016
- [4] Venkataraman, D., Parameswaran, N. S. "Extraction of Facial Features for Depression Detection among Students." *International Journal of Pure and Applied Mathematics, International Conference on Advances in Computer Science, Engineering and Technology*, pp. 455-462, 2018.
- [5] Andrej Karpathy, George Toderici, Sanketh Shetty, Thomas Leung, Rahul Sukthankar and Li FeiFei. Large-scale video classification with Convolutional neural networks. In *Proceedings of International Computer Vision and Pattern Recognition (CVPR 2014)*, 2014
- [6] Meng, Hongying, Di Huang, Heng Wang, Hongyu Yang, Mohammed AI-Shuraifi, and Wang Yunhong". "Partial least square regression is used to recognise depression based on dynamic facial and voice expression variables." In *Proceedings of the 3rd ACM international workshop on Audio/visual emotion challenge*, pp. 21-30. ACM. 2013.
- [7] Asim Munira, Ayyaz Hussaina,d, Sajid Ali Khan, Muhammad Nadeema, Sadia Arshid "Illumination invariant facial expression recognition for real-world images using selected merged binary patterns". <http://dx.doi.org/10.1145/2663204.2666274>
- [8] Junkai Chen, Zenghai Chen, Zheru Chi, and Hong Fu "Facial Expression Recognition in Video with Multiple Feature Fusion", *IEEE Transactions on Affective Computing • Optik 158 (2018) 1016–1025*.