International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 3 Issue 5, August 2019 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Human Emotion Recognition using Machine Learning

Prof. Mrs. Dhanamma Jagli[1], Ms. Pooja Shetty[2]

[1]Professor, [2]Student

[1,2]MCA, Vivekanand Education Society's Institute of Technology, University of Mumbai, Maharashtra, India

How to cite this paper: Prof. Mrs. Dhanamma Jagli | Ms. Pooja Shetty "Human Emotion Recognition using Machine Learning" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-3 | Issue-5, August 2019, pp.101-103, **https://doi.org/10.31142/ijtsrd25217**

IJTSRD25217

ABSTRACT

It is quite interesting to recognize the human emotions in the field of machine learning. Using a person's facial expression one can know his emotions or what the person wants to express. But at the same time it's not easy to recognize one's emotion easily its quite challenging at times. Facial expression consist of various human emotions such as sad, happy , excited, angry, frustrated and surprise. Few years back Natural language processing was used to detect the sentiment from the text and then it took a step forward towards emotion detection. Sentiments can be positive, negative or neutral where as emotions are more refined categories. There are many techniques used to recognize emotions. This paper provides a review of research work carried out and published in the field of human emotion recognition and various techniques used for human emotions recognition.

KEYWORDS: *deep-learning, deep-neural-networks, Tensor Flow, Inception, transfer learning, convolutional neural networks.*

I. INTRODUCTION

The human face conveys an intricate blend of information including age, gender, ethnicity, identity, personality, intentions, and emotions. In addition, speech articulation greatly affects the facial appearance Facial expressions are a form of nonverbal communication. Any human gestures can be identified by observing the different movements of mouth, nose, eyes and hands.

In this proposed system it is focusing on the human face for recognizing expression using machine learning. There are most of the datasets which are labelled as Valence – Arousal scores to capture emotion. Five years back training classifiers would have been used to make emotion word list, deciding what features to use to classify and then train SVM. However these features are becoming past due to Deep Learning, which can do feature extractions automatically, this is how we can built our Emotion Classifier at Parallel Dots. Deep Learning makes it easier by converting the problem into classification problem by identifying what exactly you want to predict. This vision of the future motivates the research for automatic recognizing of nonverbal actions and expression. Human emotion recognition has increased the attention in computer vision, pattern recognition, and human-computer interaction research communities. While having face-to-face conversion it is easy to identify the facial expression of a human being like blink rate can reveal how nervous or at ease a person may be. Raised eyebrows combined with a slightly forward head tilt indicate what is being expressed is a yes or no question. Lowered eyebrows are used for what was the questions . People use the muscles around the mouth area for talking and eating, and especially speech articulation. But using machine learning Emotion we have to create a dataset of emotions which is then fed to the neural network and trained accordingly. Reorganization is considered to be a key requirement in many applications such as affective computing technologies, intelligent learning systems, Biometrics, Facial recognition systems, video surveillance, Human computer interface , patient wellness monitoring systems, etc. Human emotion varies from person to person. Therefore human emotion detection is more challenging task in computer vision. Therefore reliable human emotion detection is required for the success of these applications.

II. CHOICE OF NEURAL NETWORK

There are multiple options for implementing the algorithm. Convolutional Neural Networks(CNN) and Recurrent Neural Network(RNN) are the two options any Data Scientist will have while solving the text classification problem. RNN is used for longer context and Convents is used for feature detection task.

Neural Network is trained until we reach a creditable accuracy.

III. IMPLEMENTATION

III.I. Dataset

Given an image/picture, detecting the human face is a complex task due to the possible variations of the face. The various shapes, angles and different poses that human face might have within the image cause such variation. The dataset contains a picture of human facial expressions of emotion. This material was developed in 1998 by Daniel Lindquist, Anders Flykt and Professor Arne Ohman at Karolinska Institute, Department of Clinical Neuroscience, Section of Psychology, Stockholm, Sweden.

III.II. Tensor Flow

Tensor Flow is an open source library for machine learning which is written in Python and C++. Tensor Flow is developed by Google Brain Team. Google is already using Tensor Flow to improve the task on several products. These task include speech recognition, search features in Google Photos. Some design decision in TF have led to this framework to be early adopted by a big community. It is easy to move from prototype to production. There is no need to compile or to modify the code to use it on a product. A key component of the library is the data flow graph. The sense of expressing mathematical computations with nodes and edges is a TF trademark. Nodes are usually the mathematical operations, while edges define the input / output association between nodes. The information travels around the graph as a tensor, a multidimensional array. Finally, the nodes are allocated on devices where they are executed asynchronously or in parallel when all the resources are ready.

III.III. Inception

Inception is a pre-trained deep neural network, for identifying patterns in images, It was designed by Yann LeCun and his colleagues. Inception takes images as its input. It can process only JPEG format images. The recommended resolution is 299*299. If the image is of higher resolution, it will be compressed automatically. It produces a private class of array of the image as its output. It was developed as a part of ImageNet Large Sale Visual Recognition challenge 2014. It can classify almost every day-to-day objects. Inception consists of 22 layers. The penultimate layer is called as "Bottlenecks". The final layer is called as softmax layer. This is the layer that can be retrained to classify the required image group.

Over the past few decades various approaches have been introduced for classification of emotions. Six universal emotions are classified using these approaches. Any good classifier should be able to recognize emotions independently.

III.IV. PUTTING THE PIECES TOGETHER

Once we are sure that inception is correctly installed and is working correctly we retrain the model for dataset. Modern object recognition models take weeks to get fully trained. Transfer learning takes a fully-trained model to shortcut a lot of work for a set of categories like ImageNet, and retrains from the existing weights for new classes. How it is done?

Features that are extracted from the activation of a deep convolutional network is evaluated to check whether it is trained in a fully supervised fashion on a large, fixed set of object recognition tasks that can be repurposed to novel generic tasks. Originally trained tasks may be different from the generic task and there may be insufficient labelled or unlabeled data to conventionally train or adapt a deep architecture to the new tasks. A set of images is required to teach the network about the new classes you want to recognize before any training is started. A dataset of emotions is gathered from a variety of sources, which we use. Once you have the images, from the root of your Tensor Flow source directory you can begin the training process. The pre-trained Inception v3 model is loaded, the old top layer is removed, and train a new one on the emotion photos that is downloaded. Transfer learning is useful because the lower layers have been trained to distinguish between objects that can be reused for many recognition tasks without any alteration.

In the first phase all the images on disk is analyzed and the bottleneck value for each of them is calculated. Penultimate layer is trained to output a set of values that is used by the classifier to distinguish between all the classes it's been asked to recognize. All the images have been reused multiple times while training and calculating each bottleneck takes a significant amount of time, it speeds things up to cache these bottleneck values on disk so they don't have to be repeatedly recalculated and if you rerun the script they'll be reused so you don't have to wait for this part again. After the completion of bottleneck, the training of the top layer of network begins. By default, it will run 4,000 training steps. At each step ten images are chosen randomly from the training sets and their bottlenecks are found from the cache and then they are fed into the final layer to get predictions. Then these predictions are compared with the actual labels to update the final layer weights through back-propagation process. As the process continues the reported accuracy improves, and after the completion of all the test, a final test is done on a set of images that is kept separately from the training and validation pictures.

IV. PROBLEM

We are able to recognize human emotions using facial expressions but reliable facial expression recognition by computer interface is still a challenge. An ideal emotion detection system should recognize expressions regardless of his/her gender, what age he or she is . Such a system should also be invariant to various distractions like glasses, their hair styles, facial hairs, their complexion etc. It should have the ability to reconstruct a whole face if there are some missing parts of the face due to various distractions. It should also perform good facial expression analysis regardless of changes in viewing condition and rigid. For more better recognition rates most current facial expressions recognition methods require some work to control image processing conditions like position and orientation of the face with respect to the quality of camera as it can result in wide variability of image views.

V. CONCLUSION

In this paper, a novel way of classifying human emotions from facial expressions is explored. After the training process, we provide the retrained model with the image we wish to classify. The system can identify only the images it is trained for just like humans, seeing something we have never seen before we shall not be able to identify it.

VI. FUTURE ENHANCEMENTS

The future enhancement can be an action that is done when an emotion is recognized. A system should play a sad song when we get a sad emotion. The next step of AI can be a system which can understand, comprehend the user's feeling, emotions and react accordingly. This bridges the gap between machines and humans. We can also have an interactive keyboard where the users can just use the app and the app will then identify the emotion and convert that emotion to the emoticon of choice.

ACKNOWLEDGMENT

The authors wish to thank Mrs. Dhanamma Jagli for her constant support in preparing the journal paper for publication.

REFERENCES

[1] IJCSNS International Journal of Computer Science and Network Security, VOL.14 No.7, July 2014.

[2] Irene Kotsia Ioannis Pitas, " Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines" in IEEE Transactions on Image Processing 16(1): pp.172-187, 2007

[3] International Journal of Computer Applications (0975 – 8887) Volume 139 – No.11, April 2016

[4] P. Saudagare, D. Chaudhari, "Facial Expression Recognition using Neural Network – An Overview", International Journal of Soft Computing and Engineering (IJSCE), ISSN: 2231-2307, Volue-2, Issue-1, March 2012.

[5] Brunelli, R., and Poggio., T., Face recognition: features versus templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(10):1042–1052, 1993.

[6] Dumasm, M., Emotional expression recognition using support vector machines. Technical report, Machine Perception Lab, Univeristy of California, 2001.

[7] Nilsson, M., Nordberg, J., and Claesson, I., Face detection using local smqt features and split up snow classifier. In IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2007.

[8] Pantic, M., Rothkrantz, L. J. M., and Koppelaar, H., Automation of nonverbal communication of facial expressions. In: EUROMEDIA 98, SCS International, pages 86–93, 1998. 14) Pham, T., and Worring, M., Face detection methods: A critical evaluation. ISIS Technical Report Series, University of Amsterdam, 11, 2000.

[9] Pham, T. V., Worring, M., and Smeulders, A. W. M., Face detection by aggregated Bayesian network classifiers. Pattern Recogn. Lett., 23(4):451–461, 2002.

[10] Gu, H., Su, G., and Du, C., Feature points extraction from faces. 2003. 11) Lisetti, C. L., and Rumelhart, D. E., Facial expression recognition using a neural network. In Proceedings of the Eleveth International FLAIRS Conference. Menlo Park, pages 328–332. AAAI Press, 1998.

[11] Yang, M. H., D. Kriegman, J., Member, S., and Ahuja, N., Detecting faces in images: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24:34–58, 2002.