Literature Review on Emotion Recognition System

E. Kodhai
Department of Computer Science & Engineering
Sri Manakula Vinayagar Engineering College
Puducherry, India
kodhaiej@gmail.com

A. Pooveswari
Department of Computer Science & Engineering
Sri Manakula Vinayagar Engineering College
Puducherry, India
pooveswariarul@gmail.com

P. Sharmila
Department of Computer Science & Engineering
Sri Manakula Vinayagar Engineering College
Puducherry, India
sharmi2898@gmail.com

N. Ramiya
Department of Computer Science & Engineering
Sri Manakula Vinayagar Engineering College
Puducherry, India
ramiya1998@gmail.com

Abstract—**Emotion plays a significant role in human beings daily lives. Humans can easily sense a person's emotions. But in some cases devices need to sense people's emotions. Machine learning is a sub-part of artificial intelligence that produces robots handling tasks like us. Emotion recognition is a small module that can be easily achieved by machines using machine learning algorithms. This paper describes the various algorithms used to recognize the facial expressions of a person such as happy, angry, sad, disgust, neutral, fear. Gabor filters and Local Binary Pattern Operators (LBP) are discussed for the process of feature extraction. Different types of classification algorithms such as Support Vector Machines, K-Nearest Neighbors are discussed. The training of the image data is carried by comparing various neural networks including Attentional Neural Network, Convolutional neural network, shallow neural network etc.**

Keywords— emotion recognition, gabor filter, support vector machines, convolutional neural network, LBP

I. INTRODUCTION

Artificial Intelligence, otherwise called machine intelligence is one of the top current technologies. The machines or intelligent agents are trained to perform a task or make decisions replicating a natural human being. Machine learning[26] is a part of Artificial Intelligence which makes a computer trying to learn from its knowledge. Machine learning focuses on different branches of technologies like computer vision, robotics and motion, natural language processing, etc. Data such as Texts, Images, audio & video is the key tool for any AI model based on its purpose.

Most of the Artificial Intelligence projects incorporate neural networks for learning, analyzing, processing and training data. [15] A neural network helps in identifying the relationship between the data by processing like a normal human brain. It is made up of interconnected nodes. [24] Input layer, hidden layer and output layer are the three main layers of neural network. When there are a large number of hidden layers a deep neural network can be created.

A part of machine learning is deep learning where artificial neural networks similar to the human brain, learn from large amounts of data. [11] Like how we learn from experience, the deep learning algorithm would perform a task repeatedly. Each time it will improve a little to reach the outcome. Computer image data is trained to model for identification, verification or controlling purposes. To assess the patterns and features embedded in the images, it uses deep neural learning networks.[25] These are often referred to as computer vision. Applications such as Face recognition systems, Object recognition system, emotion recognition systems, video surveillance system, anomaly detection systems, and driver drowsiness recognition are categorized under computer vision projects.

The face recognition systems are mostly used for attendance purposes in schools, universities and government or private sectors etc. [27] These systems will recognize the person by extracting and analyzing the inherent patterns in the human face. Similar to face recognition emotion recognition is also one among the computer vision projects.

The various facial expressions [11] of the people can be recognized. Many models have been developed for the recognition including CNN, Hidden Markov model. The emotions of the people can be detected both in real time and also using sentiment analysis. In sentimental analysis the emotions can be detected by analyzing the text [28] like comments that are posted by people in social media. Keyword spotting is one of the methods for detecting the presence of matching words in a text. However later, learning models have been implemented for automatically recognizing the emotion directly from the text. Emotions can be recognized also from speech using the [29] Natural Language Processing (NLP) technique.

Driver drowsiness systems are also a kind of emotion recognition system that is a very useful application which will detect whether an automobile driver feels sleepy or not. It is recognized by analyzing the eye movements and positions.

Similar object detection, anomaly detection are also great examples of computer vision.

II. Literature Survey

The facial expressions of people are often recognized by systems by teaching it through machine learning techniques. The familiar models that are being used are the Support Vector Machines (SVM) and Convolutional Neural Networks for classification purposes.

A. Feature Extraction

The feature extraction is the initial process that is carried in any emotion recognition model. This process focuses on extracting the important parts and information from the images. These parts include the eyes, mouth, positions and some facial features. [21] The feature extraction in emotion recognition is performed mostly through Gabor and log-Gabor filters and Local Binary Pattern Operator (LBP).

1) Gabor Filter

The Gabor filter is considered as one of the superior feature extractor used in data such as images. The computer vision technology uses the Gabor filters in a large proportion. The Gabor filter will extract the features and patterns present in the data and pass it to the training model. Just the local and small pieces of data are generally extracted from the images. All these pieces are then combined to make a neighborhood which will be recognized. Ketki R. Kulkarni et.al [5] proposed Facial Expression Recognition. The facial features are obtained by the Gabor filters by the system. Hypotheses show that the accuracy of recognition increases with the use of Log Gabor filter, while the time needed for processing is greater when using Log Gabor filter than Gabor filter. [23] Sometimes the whole frontal face image can be processed in order to end up with the classifications of facial expression. Also instead of using the face images as a whole, it is divided into subsections and processed further. Comparison of results for a local segmentation of the face around the mouth, eyes, and eyebrows with global segments of the whole face indicated that classification of the whole face yields greater accuracy

2) Local Binary Pattern Operator

The Local Binary Pattern Operator [22] is also used for feature extraction and classification of information from an image. The LBP is known to be a texture descriptor or texture classifier. Initially it will label the pixels of the images locally. And, after a whole array, it combines all the neighboring pixels to create a global pixel value. The LBP transformation to binary values from pixel values is shown in Figure 1. These values are in binary numbers. Zahir M. Hussain [16] suggested an emotion detection program using the LBP and Gabor filters. The facial regions are made from 1000 small parts. Then, each of these portions are merged into a single histogram. The important edges, corners, spots information are carried by these histograms.
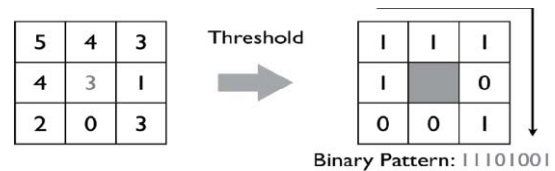

Fig.1 LBP Conversion

B. Classifiers

Classifiers are the main part of a supervised learning model. It is used for categorizing the objects according to the labels by analyzing the features of the images. Some of the classifiers used in emotion recognition are support vector machines, KNN etc.

3) Support Vector Machines

The support vector machine is a classifier that classifies the emotions based on the weights. The SVM is one of the powerful classifiers. Tuhin Kundu et.al [1] suggested familiar techniques in emotion recognition. He measures the five primary emotions or moods typically recorded on human face-containing photographs: anger, joy, normality, somnolence, automatic machinery surprise. [12] The feed forward neural network is used to supply data to the SVM. He also suggested a further method using the Oriented Quick and Rotated (ORB) to extract texture information on a single frame of the image, and SVM completes the classification. Likewise, Ma Xiaoxi [9] also suggested multiple algorithms like support vector machines. Dumas, Melanie [17] proves that classifying the emotion using SVM gives accuracy of about 85% .

4) K-Nearest Neighbors

Another kind of basic classifier used in machine learning models is the KNN. The KNN can categorize unlabeled data with the labeled data getting the nearest majority vote. Dr Poonam Tanwar[7] used models of KNN and Hidden Markov to identify emotions. The KNN is used as the initial classifier where the obtained value is passed to the Hidden Markov Model for further deep classification. M. Murugappan also used KNN [18] as the classifier for facial expression identification. The overall accuracy was about 82.87 %.
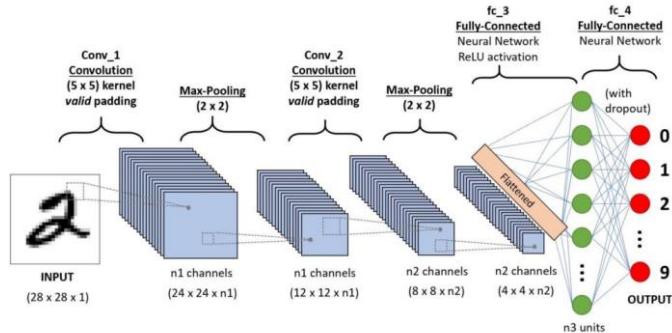
C. Neural Networks

Neural Networks are the set of neurons interconnected for identifying certain patterns in the data. The neural networks can also be used in recognizing the facial expressions of people. In most researches the Convolutional neural network is used.

5) Convolutional Neural Network

D Y Liliana [2] used CNN for emotion recognition. She suggests that the detection can be processed by the occurrences of Facial Action Units (FAU). [24] The FAU is a one of the sub-classifications of Facial Action Coding System (FACS).The FACS method can identify the emotion based on motions of facial expression. The usage of CNN proved to

reduce the over fitting and obtained result accuracy rate of 92.81%. Figure 2 shows the overview of CNN.

Rohit Pathar et al compared two neural networks for emotion recognition. A shallow CNN and a Deep CNN is compared. [19] The shallow CNN only has one convolutional layer and three fully connected layers. Whereas the deep CNN has eight convolutional layers. On experimental results, the Deep CNN has a higher accuracy of 89 %. The shallow CNN had only 45 % accuracy. The more the layers of convolutional layers and fully connected layers, the more the accuracy rate of the model.

Akash Saravanan et.al [8] proposed facial emotion recognition using algorithms from the Convolutional Neural Networks. He grouped facial expressions into each of seven emotions, using different models on the FER-2013 dataset. Models that were experimented with include feed neural networks, decision trees and smaller networks of convolution before reaching the proposed model, and also explored the effects of various hyper-parameters. Using the Adam optimizer with adjusted hyper parameters the final accuracy was 0.60.

Fig 2. Convolutional Neural Network

6) Attentional Neural Network

A face expression recognition system using an attentional convolutional network is proposed by Shervin Minaee, Amirali Abdolrashidi [20]. Traditional approaches to this problem rely on hand-made features such as SIFT, HOG, and LBP, followed by an image or video database-trained classifier. Several studies in recent years have suggested an end-to-end system for recognition of facial expression, using deep learning models. He used an attentional convolutional network that focused on important parts of the face, and achieved significant improvement over previous models on multiple datasets, including FER-2013, CK+, FERG, and JAFFE.

CONCLUSION

Facial expression recognition is a complex task in machine learning. Many methods and techniques have been developed to achieve this recognition task. Attentional CNN, Deep CNN, are discussed and provide varied output accuracy. From all the above discussions, the convolutional neural network is used by most of them since it gives high accuracy. Similarly, the classification process is the final stage of any supervised learning model. While several classification algorithms have arrived, such as SVM, KNN, the Support Vector Machine has a higher precision rate than the K-Nearest Neighbor.

REFERENCES

[1] Tuhin Kundu, Chandran Saravanan,"Advancements and recent trends in Emotion Recognition using facial image analysis and machine learning models", 2017 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT)

[2] D Y Liliana, "Emotion recognition from facial expression using deep convolutional neural network" Journal of Physics: Conference Series, Volume 1193, 2018 International Conference of Computer and Informatics Engineering 12–13 September 2018, Bogor, CC

[3] Prof. V.D. Bharate, Suhas Panchbhai, Vishal Pawar, Shubham Sunil Phadatare, "Emotion Detection using Raspberry Pi", International Research Journal of Engineering and Technology (IRJET)

[4] P.C., Vasanth & K.R., Nataraj. (2015), "Facial Expression Recognition Using SVM Classifier", Indonesian Journal of Electrical Engineering and Informatics (IJEEI). 3. 10.11591/ijeei.v3i1.126.

[5] Ketki R. Kulkarni,"Facial Expression Recognition", 2015 International Conference on Information Processing (ICIP)

[6] Kerkeni, Leila & Serretou, Youssef & Raoof, Kosai & Cléder, Catherine & Mahjoub, Mohamed & Mbarki, Mohamed. (2019). Automatic Speech Emotion Recognition Using Machine Learning. 10.5772/intechopen.84856.

[7] Dr.Poonam Tanwar, "Facial Expression Detection using Hidden Markov model".

[8] Akash Saravanan et.al, "Facial emotion recognition Using Convolutional Neural Networks"

[9] Ma Xiaoxi, Lin Weisi,"Facial Emotion Recognition", 2017 IEEE 2nd International Conference on Signal and Image Processing.

[10] Omkar Pawar , Prathamesh Lonkar , Randhir Singh , Vivek Salunke , Prof. D.M. Ujlambkar, "Door Lock System using Facial Recognition ", International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: Volume 7 Issue III, Mar 2019.

[11] https://medium.com/themlblog/how-to-do-facial-emotion-recognition-using-a-cnn-b7bbae79cd8f

[12] https://medium.com/machine-learning-101/chapter-2-svm-support-vector-machine-theory-f0812effc72

[13] https://arxiv.org/abs/1902.01019v1

[14] https://machinelearningmastery.com/introduction-to-deep-learning-for-face-recognition/

[15] https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

[16] Lajevardi, Seyed & Hussain, Zahir, "Facial expression recognition using log-Gabor filters and local binary pattern operators", International Conference on Communication, Computer and Power (ICCCP'09) MUSCAT, February 15-18, 2009

[17] Dumas, Melanie. (2001). "Emotional Expression Recognition using Support Vector Machines".

[18] M. Murugappan, "Human emotion classification using wavelet transform and KNN," 2011 International Conference on Pattern Analysis and Intelligence Robotics, Putrajaya, 2011, pp. 148-153

[19] R. Pathar, A. Adivarekar, A. Mishra and A. Deshmukh, "Human Emotion Recognition using Convolutional Neural Network in Real Time," 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), CHENNAI, India, 2019, pp. 1-7.

[20] Minaee, Shervin & Abdolrashidi, Amirali. (2019). Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network.

[21] https://medium.com/@anuj_shah/through-the-eyes-of-gabor-filter-17d1fdb3ac97

[22] https://www.pyimagesearch.com/2015/12/07/local-binary-patterns-with-python-opencv/M. M. Taghi Zadeh, M. Imani and B. Majidi, "Fast

Facial emotion recognition Using Convolutional Neural Networks and Gabor Filters," 2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI), Tehran, Iran, 2019, pp. 577-581.

[23] https://imotions.com/blog/facial-action-coding-system/

[24] https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148

[25] https://www.learnopencv.com/computer-vision-machine-learning-artificial-intelligence-consulting/

[26] https://www.sas.com/en_in/insights/analytics/machine-learning.html

[27] https://machinelearningmastery.com/introduction-to-deep-learning-for-face-recognition/

[28] Shivhare, Shiv Naresh & Khethawat, Saritha. (2012), "Emotion Detection from Text", Computer Science & Information Technology. 2. 10.5121/csit.2012.2237.

[29] https://towardsdatascience.com/speech-emotion-recognition-with-convolution-neural-network-1e6bb7130ce3