

Survey of Emotion Detection Based on Text and Facial Modalities

Archi Agarwal¹, Harshita Dadhich² and Anushree Shrivastava^{3*}

¹Department of Computer Science & Engineering, SET, Mody University of Science and Technology, Laxmangarh, Sikar, Rajasthan, India. Email: archiagrawal362@gmail.com

²Department of Computer Science & Engineering, SET, Mody University of Science and Technology, Laxmangarh, Sikar, Rajasthan, India. Email: hrs.dadhich@gmail.com

³Department of Computer Science & Engineering, SET, Mody University of Science and Technology, Laxmangarh, Sikar, Rajasthan, India. Email: shrivastava.anushree@gmail.com

*Corresponding Author

Abstract: Humans express their feelings directly or indirectly through their facial expression, speech, writing or gestures. In the current era, people express their feelings through social media, news, articles and micro blogs. It means of emoting varies from location, culture, gender etc. Due to advancement computer robotics, new angles in human-computer interactions have been brought about which are affecting our day to day life. Real time face to face communication requires quick and accurate assessments to make computers comprehend human needs and improve on its ability to communicate. An important basis for these assessments is human emotion. To tackle emotion detection problem more precisely, researchers all around the world are finding effective ways through text, speech, psychology and facial modality. This Survey covers few of the already existing emotion recognition model datasets, their techniques and features. The paper focuses on reviewing emotions based on text and facial modalities. It recaps the achievements in the field of emotion detection and has highlighted extensions for better outcome.

Keywords: Deep learning, Emotion detection, Facial recognition, Machine learning.

I. INTRODUCTION

Emotion is a strong feeling that a person derives from their circumstances, mood, or relationships with another person. The expression of emotions varies in humans according to their personality, gender, tradition, circumstances, along with many emotional, social and personal parameters [1]. A person feelings and emotions drive their decision making and moral judgement of their life [2]. Furthermore, emotional states can implicitly

affect human communication, interest, and the personalised capacity to memorize information [3]. While the understanding and explanation of emotional states comes naturally to humans, the tasks of emotion detection is challenging to computational routines [4].

Emotion detection is a process of identifying human emotions with one's personal skills or speculation. It can be detected through various inputs: facial, psychological, text and speech. Conventional ways of recognizing emotion and their classification relate certain values to basic emotions such as happiness, joy, neutrality, wonder, sadness, disgust, anger, or worry.

Detection of emotion for a person is difficult from a piece of text, speech and facial expressions. With the development in the field of technology, there is an increasing human-machine interaction. Therefore, it is essential to build a software that can understand human emotions in the same way humans do [5].

Traditional knowledge says that 70–90% of conveyance among humans is nonverbal [6]. The research carried out by Albert Mehrabian in 1967 set up the 7%–38%–55% rule, also referred to as the “3V rule”: 7% of the conversation is verbal, 38% exchange is vocal and 55% exchange is visual [6]. Consider a scenario where a driver is drowsy and he needs a break but as he has to travel a lot of distance in a short period of time, he is forced to risk his and his co-passenger's life. In a survey it was said that 20% of the car accidents take place due to fatigue. Therefore, there is a need of a software which can detect emotional state of the driver and recommend him to take a break or play music to reduce his fatigue. This will in turn make car driving safer [7]. The four methods through which emotions can be detected are shown in Fig. 1.

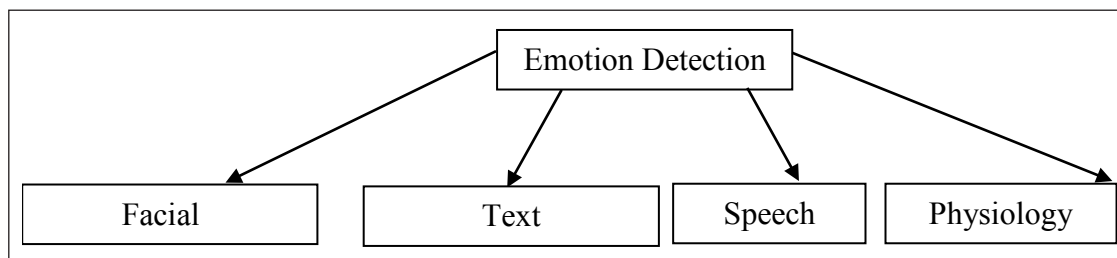


Fig. 1: Classification of Emotion Detection

This paper presents a short survey of emotion detection through facial and text. Using text-based approach, posts on social media, blogs, articles, or any other way of text can be used as an input. The rest of the paper is structured as follows - Section II work done on text based emotion detection by various researchers, Section III involves work based on emotion detection through facial recognition and Section IV is conclusion.

II. FACIAL

Facial expression is among the most important means of communication, to show emotions and intentions. It is very useful for human-machine interaction to recognize the emotional state like fury, contempt, disgust, fear, happiness, sadness and wonder. It is very difficult to state the emotion through facial expression as different humans express the same emotions in different ways. Hence, many approaches are proposed to automatically recognize facial expression. It can be classified into two approaches: feature based approach and template-based approach [8-9].

Allen Joseph *et al.* gave a method in which the authors used Discrete Wavelet Transform along with Fuzzy Combination to amplify an image. The author used an Eye-Map, Mouth-Map and Voila Jones algorithms to identify face geometry [10]. Happy *et al.* used Japanese Female Facial Expression [JAFPE] along with Cohn-Kanade [CK] [11-12] to recognize expression with the help of features extracted from facial patches. Malakar *et al.* proposed facial emotions based on Histogram along with Cohn-Kanade [CK] database. They used support vector machines (SVM) to classify their features [13-14]. Osamah *et al.* gave the way to predict emotions through Cascade Classifier in Open CV library which contain large number of algorithms used in Computer Vision [15-16]. This method was given by Voila and Jones [17]. They used the following data sets: the CK [13], JAFPE [18] and the Karolinska Directed Emotional Faces Database (KDEF) [19]. The authors experimented on these data sets by dividing them into training and testing data sets and using different combinations of the database. The two datasets used by the authors were of the same individual but had distinct images taken at distinct times to make sure that there is no identical data. They used three different methods to evaluate the three different combinations of dataset i.e. SVM, K-Nearest Neighbour (KNN) and Stochastic Gradient Descent (SGD) algorithms. They found out that the accuracy rate of KNN was

85%, SVM was 26% and SGD was 70%. Out of all the three algorithms KNN had the highest accuracy rate [15]. Nianyin *et al.* also proposed for facial learning via learning deep sparse autoencoders, the authors used feature-based approach, which means information is extracted out from geometrical features or appearances [20-21]. Deep Sparse Auto-encoder based Deep Learning Framework was used to recognize facial expression with greater accuracy by using features from extended CK+ database. The features that were used by them were happy, angry, contempt, disgust, fear, sadness, surprise and neutral. The result of this evaluation depended upon the emotion and had different accuracy for each emotion. The calculated accuracy rates were found to be: Anger 86.12%, Contempt 75%, Disgust 92.38%, Fear 78%, Happy 97.83%, Sadness 76.79 %, Surprise 96.99% and Neutral 91.44%. The highest accuracy was of happy and surprise followed by anger, disgust and neutral having satisfactory results. Whereas contempt, fear and sadness have lowest result. Facial emotion can also be detected through some popular commercial packages that offer image analysis, recognition, attribute analysis and face tracking. Some examples are:

- The Emotion API by Microsoft Azure
- Micro Expression Training Tool
- IntraFace (IF) [22]

However, emoting through facial expressions is not the only way of interpreting emotions. In the current generation, texting is the most common form of communication and text can often be misinterpreted due to lack of clarity and context. This leads to text-based emotion detection being another challenge in this field.

III. TEXT

There is a significant share of text-based communication in daily life of people. Predicting such emotions through text is more complex than other means of emotion detection like facial, psychology and sound because of the outlook of the data. Detecting emotions is easier if the emotions like happy, angry, sad and surprise are explicitly mentioned in the text but most of the time this is not the case. Some words can be ambiguous, synonyms and antonyms of each other. The text can have spelling mistakes, can be written in multiple languages or have grammatical mistakes, making the task of automatic emotion detection all the more difficult [1].

Hasan *et al.* collected data from tweets by stating a twitter key-word query. They categorized the given tweets into three categories:

- General Tweets: It included tweets which do not have any specific key-word or hashtags.
- Sad Tweets: It contained tweets having sad hashtags.
- Happy Tweets: In contained tweets having happy hashtags.

Emotex model was used to classify the tweets into emotion classes: happy-active, happy-inactive, unhappy-active and unhappy-inactive. Thus, analysing the happy, sad and anger emotions [23].

The model was able to provide the highest accuracy when the event was happy and people were happy too. Gao *et al.* studied both sentiment and emotion using data sets which were a combination of both emotion and sentiments. They trained two separate classifiers for both sentiments and emotions. During the testing section sample is segregated by using both classifiers and the probability of belonging to any one of the sections i.e. emotion or sentiment is labelled [24]. Wang *et al.* proposed a method in which he used a combination of sentiments along with emotion and used Integer Linear Programming [25]. Sen *et al.* proposed a representation of jointly learnt emotion-sentiment specific word embedding and ensured the improvement in performance of jointly learnt embeddings in the emotion detection against the discretely learnt emotions or sentiments with specific word embeddings [26]. Jainet *et al.* gave a model in which they used Naïve Bayes algorithm and SVM for emotion classification. They used data sets of political elections, health care and sports for analysis. For the purpose of defining feature sets to categorize emotions in multilingual environment three lexical resources can be used: Word-Net-Affect (WNA), Hindi Word-Net-Affect (HWNA) and Senti-Word-Net. The author used four feature groups:

- Group 1: Contains the features of WNA
- Group 2: Contains the features of WNA and Senti-world Effect
- Group 3: Contains the features of WNA with Hindi WNA
- Group 4: Contains the combination of WNA + HWNA + Senti-word Net

For the election dataset, Group 1 has accuracy for Naive Bayes Classifier has 71.65% while through SVM it has 71.33%. Group 2 has 72.16% for Navies Bayes Classifier while for SVM 70.58%. Group 3 has 72.70% for Navies Bayes while 73.89 % for SVM and Group 4 have 74.80 % for Navies Bayes while 75.61% for SVM. The experiment showed that SVM classifier gave better performance than Navies Bayes Classifier in maximum cases in the given dataset [27].

IV. CONCLUSION

By conferring the model approaches that are already in existence, we saw how detection as well as analysis of emotions

are done in its two forms: text and facial expressions. The study found that it was possible to analyse emotion through computerized processes and can be utilized to reach to the exact result of computer/machine reaction. From our study we found that many emotions could be analysed: happy, angry, sad, disgust, surprise, fear, neutral and contempt. With the help of such an approach there should be a quicker development of user-friendly systems for everyday life and work. Due to lack of knowledge and experience we don't rely on Artificial Intelligence emotion measurement approaches. Our further research will constitute the next two methods of emotion detection i.e. Speech and Psychology.

REFERENCES

- [1] K. Sailunaz, M. Dhaliwal, J. Rokne, and R. Alhaji, "Emotion detection from text and speech: A survey," *Social Network Analysis and Mining*, pp. 1-8, 2018.
- [2] J. Greene, and J. Haidt, "How (and where) does moral judgment work?" *Trends in Cognitive Sciences*, vol. 6, no. 12, pp. 517-523, 2002.
- [3] N. Derakshan, and M. W. Eysenck, "Introduction to the special issue: Emotional states, attention, and working memory," *Cognition and Emotion*, vol. 24, no. 2, pp. 189-199, 2010.
- [4] S. Poria, E. Cambria, R. Bajpai, and A. Hussain, "A review of affective computing: From unimodal analysis to multimodal fusion," *Information Fusion*, vol. 37, pp. 98-125, 2017.
- [5] Y. Ma, Y. Hao, M. Chen, J. Chen, P. Lu, and A. Košir, "Audiovisual emotion fusion (AVEF): A deep efficient weighted approach," *Information Fusion*, vol. 46, pp. 184-192, 2019.
- [6] A. Mehrabian, S. R. Ferris, "Inference of attitudes from nonverbal communication in two channels," *Journal of Consultancy Psychology*, vol. 31, no. 3, pp. 248-252, 1967.
- [7] J. Hernandez, D. McDuff, X. Benavides, J. Amores, P. Maes, and R. W. Picard, "AutoEmotive: Bringing empathy to the driving experience to manage stress," *In: Designing Interactive Systems (DIS)*, 2014.
- [8] B. Fasel, and J. Luetttin, "Automatic facial expression analysis: A survey," *Pattern Recognition*, vol. 36, no. 1, pp. 259-275, 2003.
- [9] M. Pantic, and L. J. M. Rothkrantz, "Automatic analysis of facial expressions: The state of the art," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1424-1445, 2000.
- [10] A. Joseph, and P. Geetha, "Facial emotion detection using modified eyemap-mouthmap algorithm on an enhanced image and classification with tensorflow," *The Visual Computer*, pp. 1-11, 2019.
- [11] S. L. Happy, and A. Routray, "Automatic facial expression recognition using features of salient facial

- patches,” *IEEE Transactions on Affective Computing*, vol. 6, no. 1, pp. 1-12, 2015.
- [12] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, “Coding facial expressions with Gabor wavelets,” In: *Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 200-205, 1998.
- [13] T. Kanade, J. F. Cohn, and Y. Tian, “Comprehensive database for facial expression analysis,” *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 46-53, 2000.
- [14] U. Mlakar, and B. Potocnik, “Automated facial expression recognition based on histograms of oriented gradient feature vector differences,” *Signal, Image and Video Processing* vol. 9, no. 1, pp. 245-253, 2015.
- [15] O. M. Al-Omair, and S. Huang, “A comparative study of algorithms and methods for facial expression recognition,” *2019 IEEE International Systems Conference (SysCon)*, pp. 1-6, 2019.
- [16] “Introduction — OpenCV 3.0.0-dev documentation,” Docs.opencv.org, 2018. [Online]. Available: <https://docs.opencv.org/3.0beta/modules/core/doc/intro.html>
- [17] P. Viola, and M. Jones, “Rapid object detection using a boosted cascade of simple features,” *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1-9, 2001.
- [18] M. J. Lyons, S. Akamatsu, M. G. Kamachi, and J. Gyoba, “Coding facial expressions with Gabor wavelets,” *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 200-205, 1998.
- [19] D. Lundqvist, A. Flykt, and A. Öhman, “The karolinska directed emotional faces - KDEF, CD ROM,” Department of Clinical Neuroscience, Psychology section, Karolinska Institute, 1998.
- [20] M. Pantic, and I. Patras, “Dynamics of facial expression: Recognition of facial actions and their temporal segments from face profile image sequences,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 36, no. 2, pp. 433-449, 2006.
- [21] Y. L. Tian, T. Kanade, and J. F. Cohn, “Facial expression analysis,” *Handbook of Face Recognition*, Springer, ch.11, pp. 247-275, 2005.
- [22] C. Marechal, D. Mikołajewski, K. Tyburek, P. Prokopowicz, L. Bougueroua, C. Ancourt, and K. Węgrzyn-Wolska, “Survey on AI-based multimodal methods for emotion detection,” *High-Performance Modelling and Simulation for Big Data Applications*, Springer, pp. 307-324, 2019.
- [23] M. Hasan, E. Rundensteiner, X. Kong, and E. Agu, “Using social sensing to discover trends in public emotion,” *2017 IEEE 11th International Conference on Semantic Computing (ICSC)*, pp. 172-179, 2017.
- [24] W. Gao, S. Li, S.Y. M. Lee, G. Zhou, C. R. Huang, “Joint learning on sentiment and emotion classification,” *Proceedings of the 22nd ACM International Conference on Conference on Information & Knowledge Management*, pp. 1505-1508, 2013.
- [25] S. Feng, D. Wang, G. Yu, W. Gao, and K. F. Wong, “Extracting common emotions from blogs based on fine-grained sentiment clustering,” *Knowledge and Information Systems*, vol. 27, no. 2, pp. 281-302, 2011.
- [26] A. Sen, M. Sinha, S. Mannarswamy, and S. Roy, “Multi-task representation learning for enhanced emotion categorization in short text,” *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Springer, pp. 324-336, 2017.
- [27] V. K. Jain, S. Kumar, and S. L. Fernandes, “Extraction of emotions from multilingual text using intelligent text processing and computational linguistics,” *Journal of Computational Science*, vol. 21, pp. 316-326, 2017.