IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Emotion Recognition Analysis On Real Time Video By Using The Concept Of Computer Vision And Video Processing

1Rewati Saha, 2Sachin Meshram

1Student, 2Professor

1Choksey engineering college bilaspur (CG),

2Choksey engineering college bilaspur (CG)

Abstract*— As we probably are aware we are residing in the period of advanced existence where everything depends on information investigation, presently a days after Coronavirus it's truly hard to do the examination on genuine world, so there is need of a calculation which can do the examination on virtual world, assume there is any application which can distinguish the client criticism in light of their feeling, so there is need of a clever calculation which is work on the idea of the feeling investigation, in this paper essentially we proposed a swift algorithm which is able to detection the person emotion in quick time, in this paper our main focus is to implement a algorithm which is able to give result in quick time, here we use the computer vision approach and also apply the logic of ML.

Key Words: Computer Vision, Machine Learning, Emotion Analysis, Deep Learning

I. INTRODUCTION

At times discourse or significant level scene setting can likewise be helpful to construe feeling. More often than not there is a significant cross-over between feeling classes making it a difficult classication task. In this paper we present a profound learning based way to deal with displaying different information modalities and to consolidating them to derive feeling marks from a given video succession. The Emotion acknowledgment in the wild (EmotiW 2015) challenge [9] is an expansion of a comparative test held in 2014 [8]. The undertaking is to anticipate one of seven inclination marks: furious, disdain, dread, blissful, miserable, shock and unbiased. The dataset utilized in the test is the Acted Facial Expressions in the Wild (AFEW) 5.0 dataset, which contains brief video cuts extricated from Hollywood films. The video cuts present feelings with a serious level of variety, for example entertainer personality, age, posture and lighting conditions. The dataset contains 723 recordings for preparing, 383 for approval and 539 test cuts. Conventional ways to deal with feeling acknowledgment depended close by designed highlights [17, 28]. With the accessibility of enormous datasets, profound learning has arisen as an overall way to deal with AI yielding cutting edge outcomes in numerous PC vision and regular language handling errands [22, 19]. The fundamental guideline of profound learning is to learn progressive portrayals of information to such an extent that the learned portrayals improve classi cation execution. The essential commitment of this work is to display the spatio-worldly development of looks of an individual in a video utilizing a Recurrent Neural Network (RNN) joined with a Convolutional Neural Network (CNN) in a hidden CNN-RNN engineering. Likewise, we additionally utilized an Autoencoder based action acknowledgment pipeline for displaying client movement and a basic Support Vector

Machine (SVM) based approach over energy and ghostly highlights for sound. We additionally present a brain network-based include level combination method to join di erent modalities for the nal feeling forecast for a brief video cut. Past work [18, 25] has accomplished cutting edge outcomes in the feeling acknowledgment challenge utilizing profound learning methods which incorporates our work that won the 2013 Emotion challenge. Vital writing overview related past research on leaf deficiency identification are given in II recognition based past work are given in segment ii though area III Research Gaps IV Proposed methadology. Trial results and its examination are given in area V. At long last, area VI closes the paper.

II. Litrecture Review

Encoding and understanding feelings is especially significant in instructive settings [3,31]. While eye to eye instruction with a fit, taught, and sympathetic instructor is ideal, it is additionally not generally imaginable. Individuals have been checking out instructing without educators since the time the innovation of books and with the new advances in innovation, for instance by utilizing recreations [43,66]. We have additionally seen huge advances in distance learning stages and frameworks [22,52]. In any case, while mechanization brings many benefits, like arriving at a wide populace of students or being accessible at areas where eye to eye training may not be imaginable, it additionally brings new difficulties [2,9,50,61]. One of them is the normalized look-and-feel of the course. One design doesn't fit all students, the speed of the conveyance ought to be dealt with, the assignments ought to shift contingent upon the level of the student, and the substance ought to be additionally aligned to the singular necessities of students. Full of feeling Agents: Some of these difficulties have been tended to by intuitive educational specialists that have been found powerful in improving distance learning [6,40,47,57]. Among them, vivified instructive specialists assume a significant part [12,39], in light of the fact that they can be effectively controlled and their conduct can be characterized by strategies regularly utilized in PC liveliness, for instance by giving satisfactory motions [25]. Educational specialists with passionate capacities can upgrade associations between the student and the PC and can further develop learning as shown by Kim et al. [30]. A few frameworks have been executed, for instance Lisetti and Nasoz [37] joined look and physiological signs to perceive a students feelings. DMello and Graesser [15] presented AutoTutor and they shown that students show an assortment of feelings during learning and they additionally shown that AutoTutor can be intended to recognize feelings and react to them. A virtual specialist SimSensei [42] takes part in meetings to evoke practices that can be consequently estimated and examined. It utilizes a multimodal detecting framework that catches an assortment of signs that survey the clients full of feeling state, just as to illuminate the specialist to give criticism. The control of the specialists emotional states essentially impacts learning [68] and affects student self-adequacy [30]. Be that as it may, a powerful academic specialist needs to react to students feelings that should be first identified. The correspondence ought to be founded on genuine contribution from the student, academic specialists ought to be sympathetic [11,30] and they ought to give passionate associations the student [29]. Different method for feeling discovery have been proposed, for example, utilizing eye-tracker [62], estimating internal heat level [4], utilizing visual setting [8], or skin conductivity [51] however a huge assortment of work has been zeroing in on distinguishing feelings in discourse [28,35,65]. Looks: While the previously mentioned past work gives very great outcomes, it may not be consistently appropriate in instructive setting. Discourse is frequently not needed while speaking with instructive specialists, and approaches that require appended sensors may not be great for the student. This leaves the discovery of looks and their examination as a decent choice. Different methodologies have been proposed to identify looks. Early works, for example, the FACS [16], center around facial definition, where the highlights are distinguished and encoded as an element vector that is utilized to track down a specific feeling. Late methodologies utilize dynamic forms [46] or other computerized strategies to identify the elements naturally. An enormous class of calculations endeavors to utilize math based methodologies, for example, facial remaking [59] and others distinguish notable facial elements [20,63]. Different feelings and their varieties have been considered [45] and ordered [24], and some attention on miniature articulations [17]. Novel approaches utilize robotized highlight recognition by utilizing AI techniques for example, support vector machine [5,58], however they share a similar reasonableness to the facial locator as the previously mentioned approaches (see likewise an audit [7]). One of the critical parts of these methodologies is a face global positioning framework [60] that ought to be fit for a vigorous identification of the face and its elements even in changing light conditions and for various students [56]. Nonetheless, existing strategies frequently require cautious adjustment, comparable lighting conditions, and the alignment may not move to different people. Such frameworks give great outcomes to head position or direction following, however they might neglect to distinguish unobtrusive changes in temperament that are significant for feeling location. Profound Learning: Recent advances in profound learning [34] brought profound neural organizations additionally to the field of feeling location. A few methodologies have been presented for powerful head pivot discovery [53], recognition of facial highlights [64], discourse [19], or even feelings [44]. Among them, EmoNets [26] recognizes acted feelings from films by all the while breaking down both video and sound transfers. This methodology expands on the past work for CNN facial identification [33]. Our work is roused by crafted by Burket et al. [10] who presented profound learning network called DeXpression for feeling location from recordings. Specifically, they utilize the Cohn-Kanade data set (CMU-Pittsburg AU coded information base) [27] furthermore, the MMI Facial Expression [45].Recurrent Neural Networks for Emotion Recognition in Video ,2015 In this work author present a complete system for the 2015 Emotion Recognition in the Wild (EmotiW) Challenge. We focus our presentation and experimental analysis on a hybrid CNN-RNN architecture for facial expression analysis that can outperform a previously applied CNN approach using temporal averaging for aggregation.

Deep Facial Expression Recognition: A Survey, 2018: In this paper, author provide a comprehensive survey on deep FER, including datasets and algorithms that provide insights into these intrinsic problems. First, we introduce the available datasets that are widely used in the literature and provide accepted data selection and evaluation principles for these datasets. We then describe the standard pipeline of a deep FER system with the related background knowledge and suggestions of applicable implementations for each stage. For the state of the art in deep FER, we review existing novel deep neural networks and related training strategies that are designed for FER based on both static images and dynamic image sequences, and discuss their advantages and limitations. Competitive performances on widely used benchmarks are also summarized in this section. We then extend our survey to additional related issues and application scenarios. Finally, we review the remaining challenges and corresponding opportunities in this field as well as future directions for the design of robust deep FER systems.

III. Research Issue & Future Scope

In this section basically we talk about research gap which need to be solved, as per the all-previous work there is no any researcher who solve the most important and critical factors and that are:

- Most of the time accuracy of emotion detection is very
- Quality of emotion analysis is low
- Time complexity is a main issue
- Lack in real time analysis

IV. Proposed Methadology

In this work our main objective is to resolve all previous existing issue and create a balanced system which will give a quality result in all parameters:

- Most of the time accuracy of emotion detection is very low so we will try to improve that
- Quality of emotion analysis is low so we will try to improve that
- Time complexity is a main issue so we will try to improve that
- There is need of balance algorithm which is able to manage time & quality.
- Real time video-based analysis

- Here we will use the deep learning concept and by using that first we make the training set, for trainings set we are using multiple previous existing emotions data sets.
- After this process first we will take input from the camera and generate the data sets, after this process we find the facial part on the video and based on the taring data sets we figure out the emotions and display those emotion in real time.

Initial stage we will take input as a video and after that convert that video in to the frame once we got the frame we will do the feature extraction analysis based on the previous training data and based on those data we will identify the different kind of emotions like:

1. Happy
2. Sad
3. Excited
4. Angry

Here we will design the algorithm in python only by using of different kind of library like:

- Opencv
- Tensor Flow
- Pillow
- Panda

V. RESULT ANALYSIS

As per our proposed approach we are able to achieve our targeted objectives, we are able to achieve the followings:

1. Improvement of 15-20% in accuracy
2. Improvement of 15-20% in time complexity
3. Improvement of 10-20% in quality

As compare to previous existing approach we are better in terms of the Accuracy, time complexity and quality wise.

VI. CONCLUSION

Human feeling examination is a difficult AI task with a wide scope of uses in human-PC communication, e-learning, medical services, publicizing and gaming. Feeling investigation is especially difficult as various information modalities, both visual and hear-able, assume a significant part in getting it. Given a video succession with a human subject, a portion of the significant prompts which help to comprehend the client's inclination are looks, developments and exercises. In this paper we proposed a swift algorithm which is able to detect the emotions in very quick time as compare to previous existing approaches.

REFERENCES

1. Aifanti, N., Papachristou, C., Delopoulos, A.: The MUG facial expression database. In: 11th International Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS 2010, pp. 1–4. IEEE (2010)
2. Allen, I.E., Seaman, J.: Staying the Course: Online Education in the United States. ERIC, Newburyport (2008)
3. Alsop, S., Watts, M.: Science education and affect. Int. J. Sci. Educ. 25(9), 1043– 1047 (2003)
4. Ark, W.S., Dryer, D.C., Lu, D.J.: The emotion mouse. In: HCI (1), pp. 818–823 (1999)
5. Bartlett, M.S., Littlewort, G., Fasel, I., Movellan, J.R.: Real time face detection and facial expression recognition: development and applications to human computer interaction. In: 2003 Conference on Computer Vision and Pattern Recognition Workshop, vol. 5, p. 53. IEEE (2003)
6. Baylor, A.L., Kim, Y.: Simulating instructional roles through pedagogical agents. Int. J. Artif. Intell. Educ. 15(2), 95–115 (2005)
7. Bettadapura, V.: Face expression recognition and analysis: the state of the art. arXiv preprint arXiv:1203.6722 (2012)
8. Borth, D., Chen, T., Ji, R., Chang, S.F.: SentiBank: large-scale ontology and classifiers for detecting sentiment and emotions in visual content. In: Proceedings of the 21st ACM International Conference on Multimedia, pp. 459–460 (2013)
9. Bower, B.L., Hardy, K.P.: From correspondence to cyberspace: changes and challenges in distance education. New Dir. Community Coll. 2004(128), 5–12 (2004)
10. Burkert, P., Trier, F., Afzal, M.Z., Dengel, A., Liwicki, M.: DeXpression: deep convolutional neural network for expression recognition. arXiv preprint arXiv:1509.05371 (2015) 330 W. Zhou et al.
11. Castellano, G., et al.: Towards empathic virtual and robotic tutors. In: Lane, H.C., Yacef, K., Mostow, J., Pavlik, P. (eds.) AIED 2013. LNCS (LNAI), vol. 7926, pp. 733–736. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-39112- 5 100
12. Craig, S.D., Gholson, B., Driscoll, D.M.: Animated pedagogical agents in multimedia educational environments: effects of agent properties, picture features and redundancy. J. Educ. Psychol. 94(2), 428 (2002)
13. Dimberg, U.: Facial reactions to facial expressions. Psychophysiology 19(6), 643– 647 (1982)
14. Dimberg, U., Thunberg, M., Elmehed, K.: Unconscious facial reactions to emotional facial expressions. Psychol. Sci. 11(1), 86–89 (2000)
15. D'Mello, S., Graesser, A.: Emotions during learning with autotutor. In: Adaptive Technologies for Training and Education, pp. 169–187 (2012)
16. Ekman, P.: Biological and cultural contributions to body and facial movement, pp. 34–84 (1977)
17. Ekman, P.: Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage, Revised edn. WW Norton & Company, New York (2009)
18. Ekman, P., Keltner, D.: Universal facial expressions of emotion. In: Segerstrale, U., Molnar, P. (eds.) Nonverbal Communication: Where Nature Meets Culture, pp. 27–46 (1997)
19. Fayek, H.M., Lech, M., Cavedon, L.: Evaluating deep learning architectures for Speech Emotion Recognition. Neural Netw. 92, 60–68 (2017)
20. Gourier, N., Hall, D., Crowley, J.L.: Estimating face orientation from robust detection of salient facial features. In: ICPR International Workshop on Visual Observation of Deictic Gestures. Citeseer (2004)
21. Gross, R., Matthews, I., Cohn, J., Kanade, T., Baker, S.: Multi-PIE. Image Vis. Comput. 28(5), 807–813 (2010)
22. Gunawardena, C.N., McIsaac, M.S.: Distance education. In: Handbook of Research on Educational Communications and Technology, pp. 361–401. Routledge (2013)
23. Happy, S., Patnaik, P., Routray, A., Guha, R.: The indian spontaneous expression database for emotion recognition. IEEE Trans. Affect. Comput. 8(1), 131–142 (2015)
24. Izard, C.E.: Innate and universal facial expressions: evidence from developmental and cross-cultural research (1994)
25. Cheng, J., Zhou, W., Lei, X., Adamo, N., Benes, B.: The effects of body gestures and gender on viewer's perception of animated pedagogical agent's emotions. In: Kurosu, M. (ed.) HCII 2020. LNCS, vol. 12182, pp. 169–186. Springer, Cham (2020)
26. Kahou, S.E., et al.: EmoNets: multimodal deep learning approaches for emotion recognition in video. J. Multimodal User Interfaces 10(2), 99–111 (2016). https:// doi.org/10.1007/s12193-015-0195-2

27. Kanade, T., Cohn, J.F., Tian, Y.: Comprehensive database for facial expression analysis. In: IEEE International Conference on Automatic Face and Gesture Recognition, pp. 46–53. IEEE (2000)

28. Kim, S., Georgiou, P.G., Lee, S., Narayanan, S.: Real-time emotion detection system using speech: multi-modal fusion of different timescale features. In: 2007 IEEE 9th Workshop on Multimedia Signal Processing, pp. 48–51. IEEE (2007)

29. Kim, Y., Baylor, A.L.: Pedagogical agents as social models to influence learner attitudes. Educ. Technol. 47(1), 23–28 (2007)

30. Kim, Y., Baylor, A.L., Shen, E.: Pedagogical agents as learning companions: the impact of agent emotion and gender. J. Comput. Assist. Learn. 23(3), 220–234 (2007) Learning-Based Emotion Recognition from Real-Time Videos 331

31. Kirouac, G., Dore, F.Y.: Accuracy of the judgment of facial expression of emotions as a function of sex and level of education. J. Nonverbal Behav. 9(1), 3–7 (1985). https://doi.org/10.1007/BF00987555

32. Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D.H., Hawk, S.T., Van Knippenberg, A.: Presentation and validation of the Radboud Faces Database. Cogn. Emot. 24(8), 1377–1388 (2010)

33. Le, Q.V., Zou, W.Y., Yeung, S.Y., Ng, A.Y.: Learning hierarchical invariant spatiotemporal features for action recognition with independent subspace analysis. In: CVPR 2011, pp. 3361–3368. IEEE (2011)

34. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature 521(7553), 436–444 (2015)

35. Lee, C.M., Narayanan, S.S.: Toward detecting emotions in spoken dialogs. IEEE Trans. Speech Audio Process. 13(2), 293–303 (2005)

36. Levi, G., Hassner, T.: Emotion recognition in the wild via convolutional neural networks and mapped binary patterns. In: Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, pp. 503–510 (2015)

37. Lisetti, C.L., Nasoz, F.: MAUI: a multimodal affective user interface. In: Proceedings of the Tenth ACM International Conference on Multimedia, pp. 161–170 (2002)

38. Lyons, M., Kamachi, M., Gyoba, J.: Japanese Female Facial Expression (JAFFE) Database, July 2017. https://figshare.com/articles/jaffe desc pdf/5245003

39. Martha, A.S.D., Santoso, H.B.: The design and impact of the pedagogical agent: a systematic literature review. J. Educ. Online 16(1), n1 (2019)

40. Miles, M.B., Saxl, E.R., Lieberman, A.: What skills do educational "change agents" need? An empirical view. Curric. Inq. 18(2), 157–193 (1988)

41. Mollahosseini, A., Hasani, B., Mahoor, M.H.: AffectNet: a database for facial expression, valence, and arousal computing in the wild. IEEE Trans. Affect. Comput. 10(1), 18–31 (2017)

42. Morency, L.P., et al.: SimSensei demonstration: a perceptive virtual human interviewer for healthcare applications. In: Twenty-Ninth AAAI Conference on Artificial Intelligence (2015)

43. Neri, L., et al.: Visuo-haptic simulations to improve students' understanding of friction concepts. In: IEEE Frontiers in Education, pp. 1–6. IEEE (2018)

44. Ng, H.W., Nguyen, V.D., Vonikakis, V., Winkler, S.: Deep learning for emotion recognition on small datasets using transfer learning. In: Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, pp. 443–449 (2015)

45. Pantic, M., Valstar, M., Rademaker, R., Maat, L.: Web-based database for facial expression analysis. In: 2005 IEEE International Conference on Multimedia and Expo, pp. 5–pp. IEEE (2005)

46. Pard`as, M., Bonafonte, A.: Facial animation parameters extraction and expression recognition using hidden Markov models. Sig. Process. Image Commun. 17(9), 675–688 (2002)

47. Payr, S.: The virtual university's faculty: an overview of educational agents. Appl. Artif. Intell. 17(1), 1–19 (2003)

48. Pekrun, R.: The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. Educ. Psychol. Rev. 18(4), 315–341 (2006). https://doi.org/10.1007/s10648-006-9029-9

49. Pekrun, R., Stephens, E.J.: Achievement emotions: a control-value approach. Soc. Pers. Psychol. Compass 4(4), 238–255 (2010)

50. Phipps, R., Merisotis, J., et al.: What's the difference? A review of contemporary research on the effectiveness of distance learning in higher education (1999) 332 W. Zhou et al.

51. Picard, R.W., Scheirer, J.: The Galvactivator: a glove that senses and communicates skin conductivity. In: Proceedings of the 9th International Conference on HCI (2001)

52. Porter, L.R.: Creating the Virtual Classroom: Distance Learning with the Internet. Wiley, Hoboken (1997)

53. Rowley, H.A., Baluja, S., Kanade, T.: Rotation invariant neural network-based face detection. In: Proceedings of the 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No. 98CB36231), pp. 38–44. IEEE (1998)

54. Russakovsky, O., et al.: ImageNet large scale visual recognition challenge. Int. J. Comput. Vis. 115(3), 211–252 (2015). https://doi.org/10.1007/s11263-015-0816-y

55. Russell, J.A.: Core affect and the psychological construction of emotion. Psychol. Rev. 110(1), 145 (2003)

56. Schneiderman, H., Kanade, T.: Probabilistic modeling of local appearance and spatial relationships for object recognition. In: Proceedings of the 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No. 98CB36231), pp. 45–51. IEEE (1998)

57. Schroeder, N.L., Adesope, O.O., Gilbert, R.B.: How effective are pedagogical agents for learning? A meta-analytic review. J. Educ. Comput. Res. 49(1), 1–39 (2013)

58. Tian, Y.I., Kanade, T., Cohn, J.F.: Recognizing action units for facial expression analysis. IEEE Trans. Pattern Anal. Mach. Intell. 23(2), 97–115 (2001)

59. Tie, Y., Guan, L.: A deformable 3-D facial expression model for dynamic human emotional state recognition. IEEE Trans. Circ. Syst. Video Technol. 23(1), 142–157 (2012)

60. Viola, P., Jones, M., et al.: Robust real-time object detection. Int. J. Comput. Vis. 4(34–47), 4 (2001)

61. Volery, T., Lord, D.: Critical success factors in online education. Int. J. Educ. Manag. 14(5), 216–223 (2000)

62. Wang, H., Chignell, M., Ishizuka, M.: Empathic tutoring software agents using real-time eye tracking. In: Proceedings of the 2006 Symposium on Eye Tracking Research & Applications, pp. 73–78 (2006)

63. Wilson, P.I., Fernandez, J.: Facial feature detection using Haar classifiers. J. Comput. Sci. Coll. 21(4), 127–133 (2006)

64. Yang, S., Luo, P., Loy, C.C., Tang, X.: From facial parts responses to face detection: a deep learning approach. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 3676–3684 (2015)