A REVIEW OF HEAD POSE CLASSIFICATION WITH DIFFERENT METHODS

[1]Rashmi Jain, [2]Prof. Ela Kumar

[1]Mtech Research Scholar, [2]Professor and Dean student welfare
[1]Computer Science Engineering,
[1] Indira Gandhi Delhi Technical University for Women, Delhi, India

Abstract : Head pose classification is an area of interest for researchers as it is emerging as a popular field. A significant number of works have been done in this field. Detecting the head poses provides ease in various applications. It can enhance the process of face recognition and emotion analysis. This paper performed a survey and studied different approaches designed for Head Pose Estimation (HPE). It can be perceived from the study that neural networks are widely used networks to this end. Along with this, it has become a prominent technology to detect faces covered with masks due to Covid 19 disease. This survey showed the different HPE to understand all the latest approaches designed**.**

***IndexTerms* - Head Pose Estimation, Classification, Neural networks, Covid-19.**

I.INTRODUCTION

These days, the criteria for face recognition accuracy are vastly improved in many ways, particularly in security and online payments. However, the effects of the identification are often unfavorable for face photographs with wide head deflection angles. Thus, studies into the classification of head positions may provide improved facial picture identification. Furthermore, study into the head position can also be extended to contact between humans and machines, control of the driver's state, an examination of actions, and other realistic circumstances.

The final days of 2019 saw the identification of a group of patients affected with a novel corona viral disease (COVID-19, 2019) in Wuhan, China. COVID-19 infectious diseases have spread worldwide since then. COVID-19 has multiple effects on humans. The most frequent symptoms of infected patients comprise fever, fatigue, and dry cough [1]; further signs (i.e., dullness and discomfort, inflammation of the nose, runny nose, sore throat, and diarrhea) and others may be present [2]. COVID-19 revealed healthcare infrastructure vulnerabilities in many countries, and patient management deficiency in services of healthcare has generated fear. The lack of accuracy of clinical detection methods [3] is one of the critical factors behind COVID-19's rapid dissemination.

The head position and direction of the look will help understand the individual's purpose and actions. The theme of speech and visual identification and human-machine experiences has historically been examined [4]. Nevertheless, most current methods are obtained from a very near distance using the medium to high-resolution photographs under well-regulated conditions.

A variety of approaches have been recommended in recent decades. These techniques will usually, depending on whether they include face marks in pose estimation, be split into two main groups.

Methodologies that use facial landmarks approximate angles primarily on the basis that a geometric framework is essential for the posing angles amongst the facial points [5-7]

Techniques in [8-11] do not use landmarks to identify the issue of pose estimation. Initially, they use a set of facial pictures whose angles are defined for their position. The predictor is used to estimate the angles of their positions in new facial pictures. These techniques do not detect facial characteristics but derive appearance characteristics from facial images directly. Therefore it could be applied in a better way. However, some open problems with these techniques do remain. Thus this article illustrates the HPE technique and different works done in this field.

II.HEAD POSE ESTIMATION

Research in Computer Vision focuses on predicting a human head's pose in a frame. More precisely, the Euler angles of a human head are determined. The corners of the Euler are three: yaw, pitch, and roll. An image's pose can be classified into three broad categories; yaw, pitch, and roll. The yaw angles represent the horizontal orientation and the pitch vertical angle of a face image. The roll angles represent the image plane.

These three values characterize the rotation in 3D space of an object. By estimating values correctly, one can determine what path is faced by a human brain. The machine to work out the direction in which a human head faces offers several useful applications. For instance, the map can be used to map a 3D object similar to TikTok, Snapchat, and Instagram filters to a human head's location. Furthermore, it can be used to monitor whether or not a driver focuses on the path in self-driving vehicles.

Two main approaches to determine the head position are used. One approach involves estimating facial landmarks in an intermediate stage. The facial points are then converted to a 3D human head model. The mathematical formula for measuring the yaw, pitch, and roll value of the human head can be utilized by integrating the generated 3-D points of reference with camera details such as the focal lines, distortion, and optical center in image [12], except there are a few disadvantages to this strategy. According to Ruiz, Chong, and Rehg [13], this method's efficiency depends on projections of the facial hallmark, the representativeness of the 3D head model, and the mathematical model used to complete the calculation.

Another method is predicting specifically the lagoon, pitch, and roll values without first estimating face characteristics. This method is expected to be more stable, quicker, and reliable if the facial landmark calculation is skipped [13]. This straightforward approach to the head position calculation is also given by a team of the University of Cambridge [14]. Further the classification of head pose is performed using different techniques is explained in the following section.

III. RELATED WORK

J. Sun and S. Lu proposed an effective and robust approach to track the driver's concentration in [15]. An object detection algorithm was applied on the single scale anchors, and the simulation is performed by using angle classification Experiments on YawDD result that their method can effectively execute identification tasks and prediction tasks under the real driving condition of varying luminosity.

In [16], a robust method is demostrated in a image to approximate head pose angles. A scalable gradient booster is then used to train and identify the incorporated characteristics. Experiments on the pointing 04 datasets have been tested, and it reveals that the technique proposed is more sophisticated than other approaches with lower head angle defects in pitch and yaw at 6.16 ° and 7.17 °.

A Head Pose Classification system HGL is developed in [17] by applying the picture and line portrait color texture analysis. The suggested HGL solution integrates the H-channel of the HSV color space face and grey image and trains CNN to remove classification features. The MAFA Data Collection review indicates that the suggested approach has achieved a higher result (Front Precision 93.64 percent, side accuracy: 87.17 percent, compared with algorithms focused on facial landmark recognition and convolutionary neural network.

The research in [18] suggested a system based on the line portrait generation algorithm. Color FERET and BIWI database tests demonstrated that the planned method performed better than the head pose classification algorithm, depending on facial recognition and is highly robust.

Robust and efficient algorithms for classification through deep neural networking (DNNs) in the head pose database are presented in [19]. The DNNs will learn from raw pictures and process large-scale image data. The DNN approach suggested is quickly converged and has a better capacity to understand. The outcome signifies that the approach proposed is more reliable and stable than current approaches.

To resolve particular stressors, authors in [20] looked at differences in head location characteristics. A proper experiment was done to cover various types of stress effects of neutral and stressful environments. Functions consistent with head gestures and poses were measured and evaluated computationally. Facial landmarks were aligned to direction with Active Appearance Templates (AAM). A 3D head-posing model was calculated with the 2D AAM facial landmarks. Results delineate that particular tension factors in translation and rotational factors improve mobility speed.

An observational experiments on standard in-the-wild data sets that display state-of-the-art outcomes are explained in [21]. Besides, they validate the approach using a dataset typically used for depth measurement and start narrowing the distance using specialized methods for depth measurement.

Researchers in [22] tried to create a deep-learning model that identifies the pan angle of the head by applying an appropriate convolutionary neural network for low-resolution face images without preliminary image processing. The single customized model consisting of a few convolutionary layers and dropout scheme demonstrated improved precision in the face-angle forecast in contrast with the transfer learning on the basis of pre-trained model.

A perilous activity identification system with depth information is suggested in [23] that rely on the orientation and head location. This device attempts to determine the viewing line from the head angle and the hazardous actions to detect that the direction of travel is very different from the direction of the head. The likelihood of behavior, graded from head and motion into three stages, verified the identification precision. Experimental findings demonstrated the precision of hazardous behavior identification in this method.

Conditional Random Forests is utilized in [24] to gather low-level speech transformation patterns. Heterogeneous derivative characteristics are tested on pairs of images (e.g., motions the features or texture variations). Besides, PCRF collections may also be based on head pose estimates for dynamic FER multi-view. Thus, this method tends to be the inevitable expansion of Random Forests to spatial and temporal patterns, probably from several perspectives. Experiments on the most common data sets revealed that the process contributed to substantial changes in state-of-the-art methods in many different contexts, including a new, publicly accessible multi-view video corpus.

A head-service identification system with 2 stages is anticipated in [25] using the FCN, which produces scale-conscious proposals, and CNN, divided into two groups for each proposal, i.e., head- and context. The test findings demonstrated that the object recall rate and medium accuracy are increased with scale-aware suggestions received from FCN.

D. Li et al. (2019) [26] gave a generative opponent network that can practice posture-invariant and expression-discriminatory representations to mitigate the impact of poses. In particular, they assumed that an expression, an identity component, a head-posing feature, and a remaining part of the face picture could be broken down. A feature classification loss and l-1 pixel loss are used to ensure reconstructed image accuracy and create a more narrow view. Two multi-use data sets suggest that the quantitative and qualitative evaluation indicated that the suggested algorithm is fantastic with existing techniques.

Two radio signals in a monostatic arrangement are resented in [27] at broadly different frequencies. By integrating information on Doppler movement, speed, and position at each frequency, the number of features was increased have. A K-Nearest Neighbor classification algorithm is used to train the extracted functional package, increasing car consciousness, and making it less invasive than cameras. The training findings from four participants show that classification accuracy is 77.4% at 1.8 GHz, 87.4% at 30 GHz, and multi-frequency configuration increases precision to 92%.

According to study of [28], 2D recognition has established facial recognition to detect 3D face pictures that generally require examining the face. Because of this issue's low awareness in traditional artificial neural networks (ANN), the Convolution Neural Network (CNN) was the most potential classification for determining pose estimates for a 3-dimensional face image. Distortion and disorientation effects of the target are assumed to be reduced by integration and effectively reducing the necessary parameters. However, this decrease in the recognition rate dramatically for buried face pictures shows that CNN needs to improve working with noisy conditions. The conclusion advised that the CNN system can estimate a 3D face frame with a high recognition rate.

A vision-based multi-task control system is show in [29] that simultaneously analyses the driver's headache, eyes and mouth, and somnolence levels. The feasibility of the suggested structure is demonstrated by experimental findings at both the frame and sequence stages.

A new methodology is recommended in [30] for recognizing head locations in inadequate lighting and low-quality video photos in crowded public spaces. A new Head Pose Desk is formulated by indexing each pixel in a head picture to indicate head picture models at various locations using resemblance distance charts. These distance characteristics maps are used to train multi-class Vector Classification Support machines. In demanding lighting and vision conditions, our approach is tested by using underground scene data set. Model's findings increase the head position evaluation's precision, with over 80 percent, compared to 32 percent from the best current models.

Overall, it is perceived from the literature survey that head pose classification techniques are widely emerging methods and research in this field has become very popular. Researchers have used different methodologies that included NN, CNN, DNN, k-nearest neighbor, RF classification. These approaches offer various advantages in terms of better performance but in depth analysis also revealed that there are some limitations as well that makes this area open for research.

The drawbacks includes that neural networks such as DNN, CNN utilized for classification of the images can be enhanced by introducing RNN to overcome the limitations currently used networks. It would help in diminishing the complexity by reducing the training process. Further, novel feature extraction approaches can also be implemented to make progress in the performance HPE. Moreover, the novel application of head pose classification method is the detection of the face poses covered with masks as Covid-19 pandemic has affected the people worldwide.

IV.CONCLUSION

This paper studied the different works designed and developed in the field of HPE. The head pose classification problem is considered in every work and neural network is considered as an optimal solution to such problems. The analysis of study also showed that it could be beneficial to detect faces covered with the mask. Recent work is developed to this end in which CNN is utilized. Though the efficacy of the approaches is seen in the study, it is perceived that there is a scope of enhancing the technique used to detect faces during the COVID-19 pandemic by commencing Artificial Intelligence Techniques

V.ACKNOWLEDGMENT

I would like to show my gratitude to Prof Ela Kumar for sharing her knowledge and observation. I am very thankful for her guidance in the entire research..

REFERENCES

1. S. Kooraki, M. Hosseiny, L. Myers, and A. J. J. o. t. A. c. o. r. Gholamrezanezhad 2020, Coronavirus (COVID-19) outbreak: what the department of radiology should know.
2. Y. Wang, Y. Wang, Y. Chen, and Q. J. Qin 2019, Unique epidemiological and clinical features of the emerging 2019 novel coronavirus pneumonia (COVID-19) implicate special control measures, vol. 92, no. 6, pp. 568-576.
3. Ai, Tao & Yang, Zhenlu & Hou, Hongyan & Zhan, Chenao & Chen, Chong & Lv, Wenzhi & Tao, Qian & Sun, Ziyong & Xia, Liming. (2020). Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. Radiology. 296. Pp.200642.
4. E. Murphy-Chutorian and M.M. Trivedi 2009, Head pose estimation in computer vision: A survey.PAMI, vol. 31 no.4.
5. Y. Cai, M. Yang, and Z. Li 2015, Robust head pose estimation using a 3D morphable model, Mathematical Problems in Engineering, pp. 1-10.
6. J. Aghajanian and S. J. D. Prince 2009, Face pose estimation in uncontrolled environments, in Proceedings of the 20th British Machine Vision Conference (BMVC '09), vol. 1, pp. 1–11, September 2009.
7. X. Zhu and D. Ramanan 2012, Face detection, pose estimation, and landmark localization in the wild, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2879–2886, IEEE, Providence, RI, USA
8. B. Ma, R. Huang, and L. Qin 2015, VoD: a novel image representation for head yaw estimation, Neurocomputing, vol. 148, pp. 455–466, 2015.
9. R. Zhu, G. Sang, Y. Cai, et al. 2013, Head pose estimation with improved random regression forests, in Proceedings of the 8th Chinese Conference on Biometric Recognition (CCBR '13), pp. 457–465, Jinan, China.
10. M. Fenzi, L. Leal-Taixe, B. Rosenhahn, and J. Ostermann, 2013, Class generative models based on feature regression for pose estimation of object categories, in Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition, pp. 755–762, Portland, Ore, USA, June 2013
11. X. Geng and Y. Xia 2014, Head pose estimation based on multivariate label distribution, in Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '14), pp. 1837–1842, IEEE, Columbus, Ohio, USA
12. Mallick, S. (2016, September 26). Head Pose Estimation using OpenCV and Dlib. Learn OpenCV. Retrieved from https://www.learnopencv.com/head-pose-estimation-using-opencv-and-dlib
13. Ruiz, N., Chong, E., Rehg, M., J. 2018, Fine-Grained Head Pose Estimation Without Keypoints. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pp. 2074–2083.
14. Yang, H., Mou, W., Zhang, Y., Patras, I., Gunes, H., & Robinson, P. (2015). Face Alignment Assisted by Head Pose Estimation.
15. J. Sun and S. Lu 2020, An Improved Single Shot Multibox for Video-Rate Head Pose Prediction, in IEEE Sensors Journal, vol. 20, no. 20, pp. 12326-12333, 15 Oct.15, 2020,
16. M. T. Vo, T. Nguyen and T. Le 2020, Robust Head Pose Estimation Using Extreme Gradient Boosting Machine on Stacked Autoencoders Neural Network, in IEEE Access, vol. 8, pp. 3687-3694.
17. S. Li et al. 2020, Multi-angle Head Pose Classification when Wearing the Mask for Face Recognition under the COVID-19 Coronavirus Epidemic, 2020 International Conference on High-Performance Big Data and Intelligent Systems (HPBD&IS), Shenzhen, China, pp. 1-5
18. Li, S., Sun, L., Ning, X., Shi, Y., & Dong, X. 2019, Head Pose Classification Based on Line Portrait, International Conference on High-Performance Big Data and Intelligent Systems (HPBD&IS).
19. Lu, Y., Yi, S., Hou, N., Zhu, J., & Ma, T. 2016, Deep neural networks for head pose classification. 12th World Congress on Intelligent Control and Automation (WCICA).
20. G. Giannakakis, D. Manousos, V. Chaniotakis and M. Tsiknakis 2018, Evaluation of head pose features for stress detection and classification, 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), Las Vegas, NV, 2018, pp. 406-409,
21. N. Ruiz, E. Chong, and J. M. Rehg, 2018, Fine-Grained Head Pose Estimation Without Keypoints, IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, , pp. 2155-215509,
22. S. Lee, H. Kim and M. Ishikawa 2020, Deep Learning Approach to Face Pose Estimation for High-Speed Camera Network System, 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Fukuoka, Japan, pp. 084-088,

23. K. Miyoshi, H. Nomiya and T. Houchin 2018, Detection of Dangerous Behavior by Estimation of Head Pose and Moving Direction, 2018 5th International Conference on Computational Science/ Intelligence and Applied Informatics (CSII), Yonago, 2018, pp. 121-126,
24. A Dapogny, K. Bailly and S. Dubuisson 2019, Dynamic Pose-Robust Facial Expression Recognition by Multi-View Pairwise Conditional Random Forests, in IEEE Transactions on Affective Computing, vol. 10, no. 2, pp. 167-181, 1 April-June 2019
25. S. D. Khan, Y. Ali, B. Zafar and A. Noorwali 2020, Robust Head Detection in Complex Videos Using Two-Stage Deep Convolution Framework, in IEEE Access, vol. 8, pp. 98679-98692.
26. D. Li, Z. Li, R. Luo, J. Deng and S. Sun 2019, Multi-Pose Facial Expression Recognition Based on Generative Adversarial Network, in IEEE Access, vol. 7, pp. 143980-143989, 2019.
27. M. Raja, A. Hughes, Y. Xu, P. Zarei, D. G. Michelson and S. Sigg 2019, Wireless Multi-frequency Feature Set to Simplify Human 3-D Pose Estimation, in IEEE Antennas and Wireless Propagation Letters, vol. 18, no. 5, pp. 876-880.
28. B. Kamanditya, R. P. Kuswara, M. A. Nugroho, and B. Kusumoputro 2018, Convolution Neural Network for Pose Estimation of Noisy Three-Dimensional Face Images, 2018 IEEE 5th ICETAS, Bangkok, Thailand, 2018, pp. 1-5.
29. Celona, L., Mammana, L., Bianco, S., & Schettini, R. 2018, A Multi-Task CNN Framework for Driver Face Monitoring, 2018 IEEE 8th International Conference on Consumer Electronics - Berlin (ICCE-Berlin).
30. Javier Orozco, Shaogang Gong, Tao Xiang, Head Pose Classification in Crowded Scenes pp: 1-11, Queen Mary Vision Laboratory School of EECS.