2021 5th International Conference on Computer, Communication and Signal Processing (ICCCSP - 2021)

Driver's Face Detection in Poor Illumination for ADAS Applications

Saksham Srivastava
Department of Electronics and Communication Engineering
Amrita School of Engineering
Amrita Vishwa Vidyapeetham
Coimbatore, India
cb.en.p2ael19019@cb.students.amrita.edu

Adarsh S
Department of Electronics and Communication Engineering
Amrita School of Engineering
Amrita Vishwa Vidyapeetham
Coimbatore, India
s_adarsh@cb.amrita.edu

Binoy B Nair
Department of Electronics and Communication Engineering
Amrita School of Engineering
Amrita Vishwa Vidyapeetham
Coimbatore, India
b_binoy@cb.amrita.edu

K I Ramachandran
Center for Computational Engineering and Networking
Amrita School of Engineering
Amrita Vishwa Vidyapeetham
Coimbatore, India
ki_ram@cb.amrita.edu

2021 5th International Conference on Computer, Communication and Signal Processing (ICCCSP) | 978-1-6654-3277-1/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICCCSP52374.2021.9465533

Abstract— **The detection of the driver's face is an important part of the active safety system for advanced driver assistance systems (ADAS) and this task becomes challenging and complex when normal visual cameras are used in poor illumination. This paper incorporates the adaptive attenuation quantification retinex (AAQR) method to improve the details of low and poorly illuminated images and videos with the Viola-Jones face detection algorithm. This paper also makes a comparative study of the implementation of the proposed solution and analyzes its performance on three different datasets: EBDD, 3MDAD IR2, and our self-generated validation dataset. The results have shown that when AAQR is combined with Viola-Jones, the face detection has been improved in the range of 11.18% to 24.21% for different illumination categories of EBDD, 16.71% for 3MDAD IR2 (night), and 46.79% for validation datasets.**

Keywords—Poor illumination, Adaptive attenuation quantification retinex, Face detection, Viola-Jones, Visual information fidelity, Advanced driver assistance systems.

I. INTRODUCTION

The statistics and reports show that most road accidents happen during nighttime. According to The Royal Society for the Prevention of Accidents[1], the chances of road accident increases by 40% during dark hours. In 2015 the Great Britain reported 41,825 injuries in accidents during the period of 7pm to 6am[1]. While we drive only one-quarter in the night as compared to day, the 50% of deaths happen at night. The Transport research wing of the Ministry of Road Transport and Highways, India reported[2], [3] a total of 188,126 traffic accidents in 2018 and 183,452 traffic accidents in 2019 from 6pm to 6am in India, contributing 40.2% in 2018 and 41% in 2019 in the total accidents occurred throughout the day. Poor visibility, fatigue, distracted and impaired driving are the primary causes of road accidents in the night. And most of the work done[4], [5] revolves around detecting the driver's face for fatigue and distraction through vision-based sensors[6], [7] only during the day when the lighting and illumination conditions are good as compared to nighttime when visibility of the driver's face to be detected by the camera falls due to lack of proper illumination and light. The detection of the driver's face is an important criterion to conclude whether a driver is attentive, inattentive, or drowsy. The infrared system has also been used to mitigate the driver's face visibility issue in poor illumination[8]. The sensors other than normal visual cameras to detect faces are expensive.

This paper suggests an effective and cost-efficient approach to minimize road fatalities by integrating normal visual camera and image enhancement methods to enhance the details of the captured images with a face detection algorithm to detect the driver's face in poor illumination.

The structure of the paper is as follows, section II discusses various works related to improving the poorly illuminated images, face detection in poor illumination, and face detection methodologies through different approaches. Section III covers our proposed methodology to detect the face in different illumination levels and lighting conditions along with evaluation methods. Section IV illustrates the experimental setup, section V contains results and their analysis, and section VI contains the conclusion and the future scope.

II. LITERATURE SURVEY

R. Ki[9] has used the infrared sensors to read the facial temperature of the driver to detect the driver's fatigue but the performance of infrared techniques reduces when lighting condition changes during traffic[10]. Thus illumination methods are much better and Retinex[11] is one of the ways to achieve this. It imitates human vision methods.

W. Kim[12], proposes the adoption of maximum values of diffused intensity of each pixel to improve the quality of poorly illuminated images.

The Low-Light Image Enhancement Method (LIME)[12] uses maximum values among all the diffused intensities positioned at each pixel and the Multi-Scale Retinex with Color Restoration method (MSRCR)[13] uses Gaussian filters to divide illumination and reflectance to improve the image.

J. Shen et al [14], studied and compared different methods: LIME[12], MSRCR[13], Multi scale Fusion (MF)[15], Simultaneous Reflectance and Illumination Estimation (SRIE)[16], Naturalness Preserved Enhancement (NPE)[17], and Optimized de-hazing (DONG)[18] to detect the face in low illumination and proposed a new Retinex[11] method and termed as Adaptive Attenuation Quantification Retinex (AAQR)[14] which improves the details of an image by improving the brightness and contrast level.

The performance analysis of image enhancement methods done in[14] concludes that AAQR is better than other methods when all the factors: execution time, visual information fidelity (VIF)[19], and confidence values are taken into account. It uses a convolutional neural network to detect and classify faces.

The Viola-Jones face detection method[20] to differentiate faces and non-faces is still popular and because of easy to implement steps, it is still being used today. But this algorithm is sensitive to lighting conditions and performs well for the frontal face only.

The convolutional neural network has been used to extract facial features and a support vector machine (SVM) is used as a classifier to distinguish between facial and non-facial images[21].

The improvised Viola-Jones method through neural network increases the accuracy of face detection to 90% from 78.6%[22].

III. PROPOSED SOLUTION AND EVALUATION METHODS

A. Proposed Solution

The proposed solution to improve the active safety system for advanced driver assistance systems has two steps; firstly it enhances the image and then it applies a face detection algorithm to detect the face. Fig.2 shows the flow of solution.

1. Image Enhancement

   The driver's face image obtained in poor and low illumination has low values of contrast and luminosity that is enhanced through adaptive attenuation quantification retinex (AAQR)[14] method so that the face detection can happen in the next step through the face detection algorithm. It is a Retinex[11] method and it involves the three steps[14] which are explained below:

   a. Attenuation Restriction

      The pixels of the digital images hold all the information. The brightness of an image is the mean pixel values while the stability of illumination is the variance so the range of brightness varies between mean and variance. The attenuation coefficient is the logarithmic ratio of mean and variance values.

   b. Attenuation Prediction

      The logarithmic retinex model of the image is obtained by subtracting the logarithmic low pass filtered image from the logarithmic original image. The median and average values of retinex images are used to find image inclination and attenuation range which are used in the next step.

   c. Adaptive Quantification

      The low values of quantization reduce noises but it also causes the loss of information but on the other hand, a high value keeps the original information but introduces noises so an optimized quantization value is needed. So the quantization range is obtained through requantification obtained from the attenuation range calculated in the previous step. And hence the enhanced image is obtained.

2. Face detection

   The convolutional neural network (CNN) and Viola-Jones[20] algorithm have their advantages and disadvantages. The convolutional neural network (CNN) is faster but takes more computational memory space and on the other hand, Viola-Jones takes less memory space and is easy to implement and when merged with other methods like image enhancement, can give improvised results.

   The Viola-Jones framework is used to detect the driver's face. It is a machine learning algorithm and can find faces at a fast rate. Following steps[20] are involved in this algorithm to detect the driver's face:

   a. Haar Features

      It uses Haar features instead of RGB pixel values and hence it is fast and more effective. A mathematical operation called convolution is performed on different windows and sub-windows of images. Basically, 3 kinds of features as shown in fig.1 are used which are listed below:

      - 2 rectangle feature
      - 3 rectangle feature
      - 4 rectangle feature

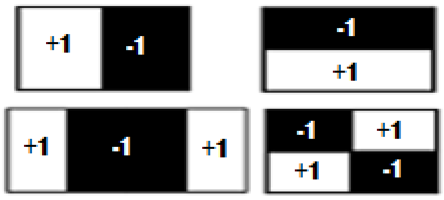

Fig.1. Haar Features

      The difference is used to identify the various subparts of an image. For example: in a human face, the cheek area is fairer or less dark than the area below the eye.

   b. Integral Image

      The Haar features are computed quickly with this new image representation. The few operations per pixels of the image are used to calculate the integral image.

   c. Adaptive boosting

      It is also known as Adaboost[23], a machine learning algorithm which gives a strong classifier on combining weak classifiers. The more weights are given to the wrongly classified data so that we can get more probability which can be used for classification.

   d. Cascade classifier

      The cascading is a continuous stream of stages which determines whether the given sub-window is a face or not and if it is not a face then it is rejected instantly and does not move up in further stages.

B. Evaluation Methods

1. Face detection percentage

We have defined the accuracy of face detection as face detection percentage given by:

$$\frac{Total\ number\ of\ frames\ of\ video\ in\ which\ face\ is\ detected}{Total\ number\ of\ frames\ of\ video\ in\ which\ face\ is\ present} * 100 \qquad (1)$$

2. Visual information fidelity (VIF)

The visual information fidelity index[19] is an image quality evaluation parameter that tells how much information is lost when another image is produced from the original image through some process. The values lie between 0 and 1. It is calculated as:

$$\frac{Amount\ of\ information\ in\ newly\ transformed\ image}{Amount\ of\ information\ in\ the\ original\ image} \qquad (2)$$

3. Throughput

The throughput for the AAQR[14] process is defined as:

$$\frac{1}{Computation\ time\ to\ enhance\ the\ details\ of\ one\ frame\ through\ AAQR} \qquad (3)$$

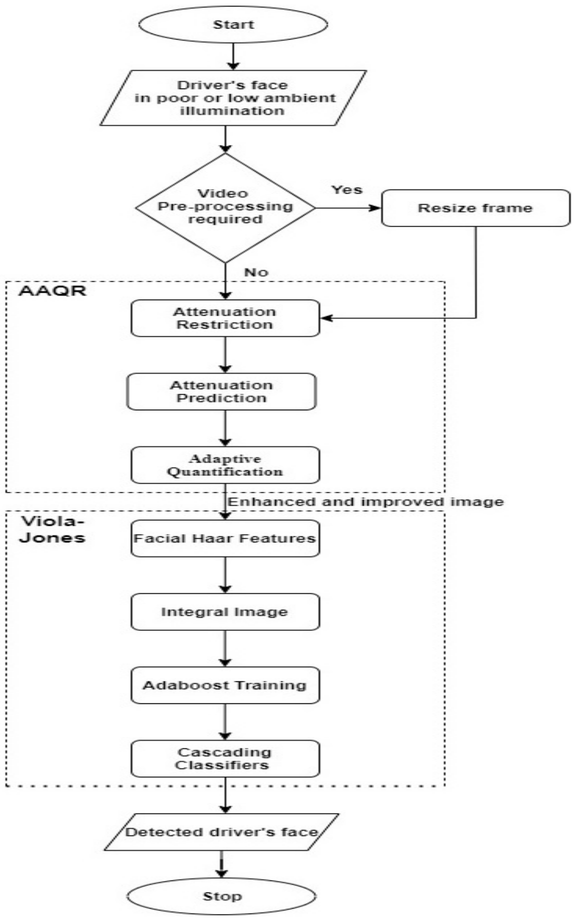

Fig.2. Flow diagram of the proposed solution

IV. EXPERIMENTAL SETUP

A. Dataset

Following datasets with details mentioned in TABLE I are used:

i. EEE BUET Distracted Driving (EBDD)[24]

The EBDD video database consists of 59 videos of the driver under six illuminations categories captured from front view through Sony Cyber shot 14.1 megapixels camera. The dataset has frame dimensions of 854*480 pixels having a frame rate of 30fps. The duration of 52 videos varies between 26 seconds to 44 seconds and the remaining 7 videos have duration of 120 seconds.

ii. 3MDAD : Multimodal Multiview and Multispectral Driver Action Dataset[25]

This dataset has images belonging to day and nighttime captured through front and side view using Microsoft Kinect camera. We have used the 3MDAD (night) IR 2 images for our study.

iii. Validation[26]

This is a self-captured 15 videos of drivers sitting in a vehicle in the nighttime under low or poor illumination. The normal front camera of a mobile phone of 13MP is used to capture the face of the driver. The duration of videos varies between 25 seconds and 40 seconds.

TABLE I. DETAILS OF THE DATASET

Data set Name	Frame speed (fps)	Frame dimens-ions (pixels)	Illumination	View-ing Angle
EBDD	30	854*480	Rainy	Front
			Low	
			Low+Sunny	
			Normal+Low	
			Normal+Sunny	
			Normal	
3MDAD IR2	NA	640*480	Night	Front
Valid-ation	30	1440*720	Night	Front

B. Setup

The studies are done in MATLAB R2019b and Python 3.7.7. The Intel Core i5 clocked at 3.2 GHz and 8GB RAM are used.

C. Pre-processing of images and videos

The details of pre-processing are mentioned below:

EBDD: The videos are resized to 640*480 pixels and the frame rate is kept unchanged to 30fps.

3MDAD IR2 (Night): This dataset consists of 1656 infrared images in tagged image file format (.tiff) and it is converted into 16 videos of MPEG-4 Part 14 (.mp4) format with frame speed of 15fps. The video duration varies between 4 seconds to 11 seconds.

Fig.3. The first row shows the different frames of the driver's face under different illumination and lighting conditions. The second row shows the improved enhanced images after AAQR is performed. The third row shows the detected face.

Validation: The videos are resized to 640*480 dimensions with frame rate kept unchanged to 30fps.

V. RESULTS AND ANALYSIS

Fig. 3 shows the outcomes of the proposed method of different stages for different illumination categories of the dataset. The second column has infrared images as input and other columns have normal visual camera images as input. So the AAQR performs well for both infrared and visual cameras. The brightness and luminous values are enhanced through this method.

Fig. 4, depicts that the face detection has been increased to 24.21%, 17.23%, 11.18%, 17.01%, 22.06% and, 19.75% respectively for rainy, normal+sunny, normal, low+sunny, low, and normal+low illumination levels of the EBDD dataset. There has not been much improvement in the face detection percentage when the illumination levels are normal as there is enough amount of brightness and luminosity in the image even before the AAQR is applied and Viola-Jones performs well when illumination and lighting conditions are good. The face detection percentage for normal+sunny illumination is less than expected because the sunlight has already made the frame brighter and when AAQR is implemented, the brightness increases further and resulting in saturation. That's why the proposed method doesn't improve the detection much in this case. This is similar to the situation when sunlight is falling directly on the face and making the face shine but making it difficult for the camera to capture good quality pictures. The face detection for the 3MDAD IR2 (night) dataset is improved by 16.71%. There is a significant improvement of 46.79% in face detection for the self-generated validation dataset.

The face detection percentage also depends on the viewing angle. Although all the three datasets that are taken into account have front as viewing angle but mounting angle and orientation of the cameras are different. The camera placed directly in the front of the driver as used in the validation dataset gives a better result as Viola-Jones is effective to detect the frontal faces.

TABLE II shows the mean execution time and throughput to perform AAQR technique.

TABLE II. EXECUTION TIME AND THROUGHPUT FOR AAQR METHOD

Dataset	Execution time (seconds)	Frame dimensions (pixels)	Throughput (frames/second)
EBDD	0.24	640*480	4.17
3MDAD IR 2 (Night-Front)	0.21	640*480	4.76
Validation	0.26	640*480	3.84

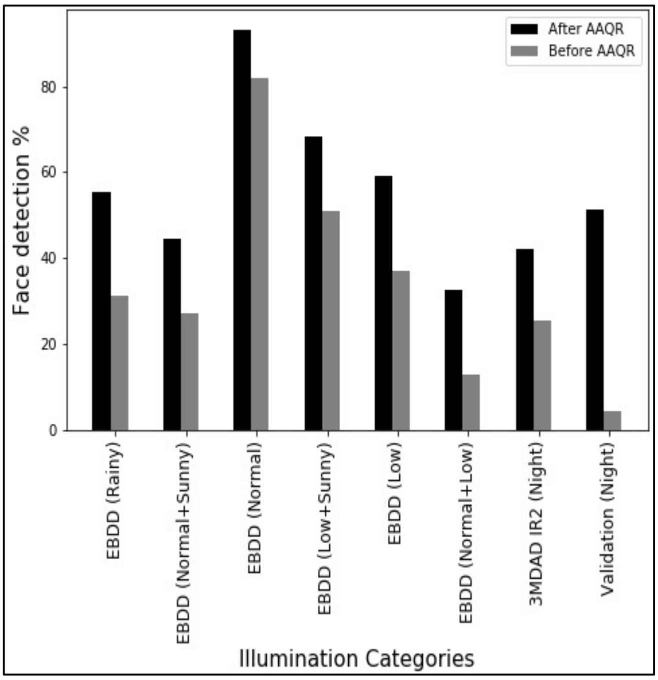

Fig.4. Improvement in face detection percentage for different illumination categories for different datasets.

TABLE III shows the visual information fidelity values (VIF) for datasets. The mean VIF values for the EBDD dataset is 0.53. The visual field information index value is above 0.50 which means at least half of the information is preserved after the transformation.

TABLE III. COMPARISON OF VIF VALUES

Dataset Name	Frame dimensions (pixels)	Illumination	VIF
EBDD	640*480	Rainy	0.52
		Low	0.53
		Normal+Low	0.51
		Normal+Sunny	0.52
		Normal	0.63
		Low+Sunny	0.51
Validation	640*480	Night	0.76

VI. CONCLUSION AND FUTURE SCOPE

This paper incorporates the adaptive attenuation quantification retinex (AAQR) method with the Viola-Jones face detection algorithm to improve the overall face detection. The Viola-Jones face detection algorithm is sensitive to poor illumination and fails when the lighting conditions are poor. The results of our proposed method enhance the details of the poorly illuminated driver's face and show the improvement in face detection percentage for the images captured through both normal visual and infrared cameras. The neural network needs more computational memory space but Viola-Jones takes less memory space and it is easy to implement, so the Viola-Jones with AAQR can be used as proposed in our solution to detect the driver face in poor lighting and low illumination conditions. The detected face in poor illumination or low lighting condition can be then analyzed for distraction or drowsiness in an advanced driver assistance systems (ADAS) for the active safety system.

The merger of the Kanade–Lucas–Tomasi (KLT) feature tracker algorithm with the proposed solution can be studied in the future to improve the driver's face detection.

REFERENCES

[1] N. Factsheet, "Road Safety Factsheet Driving at Night Factsheet," Birmingham, UK, 2017. [Online]. Available: https://www.rospa.com/rospaweb/docs/advice-services/road-safety/drivers/driving-at-night.pdf.

[2] Transport Research wing, Ministry of Road Transport and Highways, India, "Road Accidents in India-2018", New Delhi, India, 2019. [Online]. Available: https://morth.nic.in/sites/default/files/Road_Accident.pdf.

[3] Transport Research wing, Ministry of Road Transport and Highways, India, "Road Accidents in India-2019", New Delhi, India, 2020. [Online]. Available: https://www.morth.nic.in/sites/default/files/RA_Uploading.pdf.

[4] P. M. Manjula, S. Adarsh, and K. I. Ramachandran, "Driver Inattention Monitoring System Based on the Orientation of the Face Using Convolutional Neural Network," *11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Kharagpur, India, 2020, pp.1-7, doi: 10.1109/ICCCNT49239.2020.9225600.

[5] S. Misal and B. B. Nair, "A Machine Learning Based Approach to Driver Drowsiness Detection," in *Communications in Computer and Information Science*, 2019, vol. 835, pp. 150–159, doi: 10.1007/978-981-13-5992-7_13.

[6] A. Fernández, R. Usamentiaga, J. L. Carús, and R. Casado, "Driver distraction using visual-based sensors and algorithms," *Sensors (Switzerland)*, vol. 16, no. 11, pp. 1–44, 2016, doi: 10.3390/s16111805.

[7] V. Sanjay Kumar, S. N. Ashish, I. V. Gowtham, S. P. Ashwin Balaji, and E. Prabhu, "Smart driver assistance system using raspberry pi and sensor networks," *Microprocess. Microsyst.*, vol. 79, no. January, p. 103275, 2020, doi: 10.1016/j.micpro.2020.103275.

[8] J. Kang, D. V Anderson, S. Member, M. H. Hayes, and L. Fellow, "Face recognition for vehicle personalization with near-IR frame differencing and pose clustering," 2015 *IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, USA, 2015, pp. 455-456, doi: 10.1109/ICCE.2015.7066484.

[9] R. Ki, "Driver Fatigue detection using Infrared sensor and Adaptive Neuro-Fuzzy Inference System," *International Journal of Pure and Applied Mathematics*, 2018, vol. 118, no. 20, pp. 4209–4214.

[10] D. W. Hansen and Q. Ji, "In the Eye of the Beholder: A Survey of Models for Eyes and Gaze," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 3, pp. 478–500, 2010, doi: 10.1109/TPAMI.2009.30.

[11] H. L. Edwin and J. M. John, "Lightness and Retinex Theory," *J. Opt. Soc. Am.*, vol. 61, no. 1, pp. 1–11, 1971.

[12] W. Kim, R. Lee, M. Park, and S. H. Lee, "Low-Light Image Enhancement Based on Maximal Diffusion Values," *IEEE Access*, vol. 7, pp. 129150–129163, 2019, doi: 10.1109/ACCESS.2019.2940452.

[13] B. Jiang, G. A. Woodell, and D. J. Jobson, "Novel multi-scale retinex with color restoration on graphics processing unit," *J. Real-Time Image Process.*, vol. 10, no. 2, pp. 239–253, 2015, doi: 10.1007/s11554-014-0399-9.

[14] J. Shen *et al.*, "Nighttime Driving Safety Improvement via Image Enhancement for Driver Face Detection," *IEEE Access*, vol. 6, pp. 45625–45634, 2018, doi: 10.1109/ACCESS.2018.2864629.

[15] X. Fu, D. Zeng, Y. Huang, Y. Liao, X. Ding, and J. Paisley, "A fusion-based enhancing method for weakly illuminated images," *Signal Processing*, vol. 129, pp. 82–96, 2016, doi: 10.1016/j.sigpro.2016.05.031.

[16] X. Fu, D. Zeng, Y. Huang, X. P. Zhang, and X. Ding, "A Weighted Variational Model for Simultaneous Reflectance and Illumination Estimation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, vol. 2016-Decem, pp. 2782–2790, doi: 10.1109/CVPR.2016.304.

[17] S. Wang, J. Zheng, H. M. Hu, and B. Li, "Naturalness preserved enhancement algorithm for non-uniform illumination images," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3538–3548, 2013, doi: 10.1109/TIP.2013.2261309.

[18] X. Dong *et al.*, "Fast efficient algorithm for enhancement of low lighting video," in *Proceedings - IEEE International Conference on Multimedia and Expo*, 2011, vol. 3, pp. 3–8, doi: 10.1109/ICME.2011.6012107.

[19] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2004, vol. 3, no. 2, pp. 430–444, doi: 10.1109/icassp.2004.1326643.

[20] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001, vol. 1, doi: 10.1109/cvpr.2001.990517.

[21] S. Guo, S. Chen, and Y. Li, "Face recognition based on convolutional neural network & support vector machine," in *2016 IEEE International Conference on Information and Automation, IEEE ICIA 2016*, 2017, no. August, pp. 1787–1792, doi: 10.1109/ICInfA.2016.7832107.

[22] M. N. Chaudhari, M. Deshmukh, G. Ramrakhiani, and R. Parvatikar, "Face Detection Using Viola Jones Algorithm and Neural Networks," 2018 *Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, Pune, India, 2018, pp. 1-6, doi: 10.1109/ICCUBEA.2018.8697768.

[23] Y. Freund and R. E. Schapire, "A desicion-theoretic generalization of on-line learning and an application to boosting BT - Computational learning theory," *Comput. Learn. theory*, vol. 904, no. Chapter 2, pp. 23–37, 2005.

[24] T. Billah, S. M. M. Rahman, M. O. Ahmad, and M. N. S. Swamy, "Recognizing Distractions for Assistive Driving by Tracking Body Parts," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 4, pp. 1048-1062, April 2019, doi: 10.1109/TCSVT.2018.2818407.

[25] I. Jegham, A. Ben, I. Alouani, and M. Ali, "Signal Processing : Image Communication A novel public dataset for multimodal multiview and multispectral driver distraction analysis : 3MDAD," in *Signal Processing: Image Communication*, 2020, vol. 88, no. July, p. 115966, doi: 10.1016/j.image.2020.115960.

[26] "Validation Dataset," 2021. https://www.drive.google.com/drive/folders/1e9I4Ci4h80FR5drcwYkI6NJzwP7N7Dvm?usp=sharing.