ISSN:2229-6107

INTERNATIONAL JOURNAL OF PURE AND APPLIED SCIENCE & TECHNOLOGY

E-mail :
editor.ijpast@gmail.com
editor@ijpast.in

www.ijpast.in

INTERNATIONAL JOURNAL OF PURE AND APPLIED SCIENCE & TECHNOLOGY
IJPAST
ISSN 2229-6107
www.ijpast.in
Vol 14 Issuse 1 Jan 2024

Uncovering Anomalous Driving behaviour patterns through Deep Learning Fusions

Dr.M.Vanitha[1],D. Prathima[2],D. Haripriya[3] ,G. Manusha[4]

ABSTRACT

Video-based abnormal driving behavior detection is becoming more and more popular for the time being, as it is highly important in ensuring safeties of drivers and passengers in the vehicle, and it is an essential step in realizing automatic driving at the current stage. Thanks to recent developments in deep learning techniques, this challenging detection task can be largely facilitated via the prominent generalization capability of sophisticated deep learning models as well as large volumes of video clips which are indispensable for thoroughly training these data-driven deep learning models. deep learning fusion techniques are emphasized, and three novel deep learning-based fusion models inspired by the recently proposed and popular densely connected convolutional network (DenseNet) are introduced, to fulfill the video-based abnormal driving behavior detection task for the first time.

I. INTRODUCTION

It is widely acknowledged that, high-resolution videos are more and more commonly seen within a great number of visual applications at the current stage. For instance, in video surveillance, multiple high-resolution cameras are necessary to be placed at different locations. They work together to identify and track the moving target making the later high-level analyses based on the moving target (e.g., behavior or even potential intention) more feasible. In emotional computation, high-resolution cameras need to be utilized to capture both obvious and fine changes of emotions of the target person in real-time which has significant impacts in security issues nowadays It is easy to perceive from the above descriptions that, acquiring and storing a large volume of high-resolution videos are often not difficult to be realized for the time being. However, the main challenge resides in how to efficiently and effectively make correct high-level decisions based on those low-level video clips of large volumes. In this study, high-resolution videos of drivers recorded within vehicles are emphasized. The high-level decision here is to correctly detect abnormal driving behavior (i.e., patterns) of drivers.Automatic abnormal driving behavior detection is generally accepted as the first issue in realizing the popular fully autonomous driving task.It is certain that,fortheautonomous driving task, safety issues are undoubtedly first priorities. It is widely known that, behavior of

[1]Professor, Department of CSE,Malla Reddy Engineering College for Women,Hyderabad, TS, India,vanitha.official@gmail.com
[2,3,4]UG Students, Department of CSE, MallaReddy Engineering College for Women, Hyderabad, TS, India

INTERNATIONAL JOURNAL OF PURE AND APPLIED SCIENCE & TECHNOLOGY
IJPAST
ISSN 2229-6107
www.ijpast.in
Vol 14 Issuse 1 Jan 2024

drivers need to be well restricted in order to avoid any potential accident. Therefore, multiple high-resolution cameras equipped with in the driver's vehicle can be utilized to monitor the driver's status in real time. Generally speaking, videos captured by high-resolution cameras also need to be processed immediately, in order to determine whether the current status of the driver is normal or not. It can be acknowledged from the above descriptions that, both the effectiveness (i.e., the detection accuracy) and the efficiency (i.e., the detection speed) of abnormal driving behavior detection are highly demanded. Also, high-speed wireless transmissions are necessary to realize the swift and reliable transmission of high-quality videos, which further facilitates the above automatic abnormal driving behavior detection task [9]– [23].In order to detect abnormal behavior of drivers, an official and precise definition of abnormal driving is often necessary. According to the International Organization for Standardization (ISO), abnormal driving is defined as the phenomenon that a driver's ability to drive is impaired due to her / his own focus on other activities unrelated to normal driving.

II. LITERATURE SURVEY

1. Fast deep neural networks with knowledge guided training and predicted regions of interests for realtime video object detection, Jianhe Yuan,Wenming Cao,Zhihai He,Zhi Zhang, It has been recognized that deeper and wider neural networks are continuously advancing the state-of-the-art performance of various computer vision and machine learning tasks. However, they often require large sets of labeled data for effective training and suffer from extremely high computational complexity, preventing them from being deployed in real-time systems, for example vehicle object detection from vehicle cameras for assisted driving. In this paper, we aim to develop a fast deep neural network for real-time video object detection by exploring the ideas of knowledge-guided training and predicted regions of interest. Specifically, we will develop a new framework for training deep neural networks on datasets with limited labeled samples using cross-network knowledge projection which is able to improve the network performance while reducing the overall computational complexity significantly. A large pre-trained teacher network is

ISSN 2229-6107
www.ijpast.in
Vol 14 Issuse 1 Jan 2024

used to observe samples from the training data. A projection matrix is learned to project this teacher-level knowledge and its visual representations from an intermediate layer of the teacher network to an intermediate layer of a thinner and faster student network to guide and regulate the training process. To further speed up the network, we propose to train a low-complexity object detection using traditional machine learning methods, such as support vector machine. Using this low-complexity object detector, we identify the regions of interest that contain the target objects with high confidence. We obtain a mathematical formula to estimate the regions of interest to save the computation for each convolution layer. Our experimental results on vehicle detection from videos demonstrated that the proposed method is able to speed up the network by up to 16 times while maintaining the object detection performance.

2. Cascaded regional spatio-temporal feature-routing networks for video object detection, Hui shuai, Qingshan Liu, Kaihua Zhang,This paper presents a cascaded regional spatiotemporal feature-routing networks for video object detection. Region proposal networks in faster region-based convolutional neural network (CNN) generate spatial proposals, whereas neglecting the temporal property of the videos. We incorporate the correlation filter tracking on the convolutional feature maps to explore an efficient and effective spatiotemporal region proposal generation method. To gradually refine the bounding boxes of proposals, three region classification and regression networks are cascaded. Feature maps from different layers in CNNs extract hierarchical information of the input, so we propose a router function which selects feature maps according to the scale of proposals. In addition, object co-occurrence inference is exploited to suppress conflicting false positives, which leads to a semantically coherent interpretation on the video. Extensive experiments on the Pascal VOC 2007 dataset and the ImageNet VID dataset show that the proposed method achieves the state-of-the-art performance for detecting unconstrained objects in cluttered scenes.

3. A neuromorphic person re-identification framework for video surveillance, Aparajita Nanda,Pankaj Kumar Sa,Suman Kumar

ISSN 2229-6107
www.ijpast.in
Vol 14 Issue 1 Jan 2024

Choudhury,Sambit Bakshi, This paper presents a neuromorphic person re-identification (NPReId) framework to establish the correspondence among individuals observed across two disjoint camera views. The proposed framework comprises three modules (observation, cognition, and contemplation), inspired by the form-and-color-and-depth (FACADE) theory model of object recognition system. In the observation module, a semantic partitioning scheme is introduced to segment a pedestrian into several logical parts, and an exhaustive set of experiments have been carried out to select the best possible complementary feature cues. In the cognition module, an unsupervised procedure is suggested to partition the gallery candidates into multiple consensus clusters with high intra-cluster and low inter-cluster similarity. A supervised classifier is then deployed to learn the relationship between each gallery candidate and its associated cluster, which is subsequently used to identify a set of inlier consensus clusters. This module also includes weighing of contribution of each feature channel toward defining a consensus cluster. Finally, in the contemplation module, the contributory weights are employed in a correlation-based similarity measure to find the corresponding match within the inlier set. The proposed framework is compared with several state-of-the-art methods on three challenging data sets: VIPeR, iLIDS-VID, and CUHK01. The experimental results, with respect to recognition rates, demonstrate that the proposed framework can obtain superior performance as compared with the counterparts. The proposed framework, along with its low-rank bound property, further establishes its suitability in practical scenarios through yielding high cluster hit rate with low database penetration.

4. Semi-coupled dictionary learning with relaxation label space transformation for video-based person re-identification, Lingchuan Sun,Zhuqing Jiang,Hongchao Song,Qishuo Lu, Video-based person reidentification (re-id) is a challenging problem due to much discrepancy between different videos by person pose, illumination, viewpoint change, background clutter, and occlusion within each camera and across different cameras. However, most existing video-based person re-id methods usually focus on dealing with the discrepancy between different cameras and do not fully consider the correlation between different cameras. In this paper,

ISSN 2229-6107
www.ijpast.in
Vol 14 Issuse 1 Jan 2024

we propose a semicoupled dictionary learning with relaxation label space transformation approach to capture the intrinsic relationship of the same person under different cameras. First, to reduce the discrepancy between different views, we transform the original feature spaces into the common feature space by local Fisher discriminant analysis. Two dictionaries are learned from this common feature space. Second, a relaxation label space is introduced to associate the same person under different views. In this label space, the distance between different persons can be enlarged as much as possible, such that label information has stronger discriminative capability. A single dictionary is learned from the relaxation label space. Finally, in order to further enhance the correlation of the same person between different cameras, we use a pair of transformation matrices which map the coding coefficients learned from the common feature space to the coding coefficients learned from the relaxation label space, respectively. Extensive experimental results on two public iLIDS Video re-IDentification and Person Re-ID 2011 video-based person re-id datasets demonstrate the effectiveness of the proposed method.

III. EXISTING SYSTEM:

- abnormal driving detection and deep learning techniques, which are closely related to this study, are emphasized. Recent developments in the two aspects are briefly reviewed, with pros and cons been discussed.
- can be summarized based on literatures of automatic abnormal driving behavior detection that, there are often three commonly used detection schemes.
- The first one is based on the detection of human physiological signals (i.e., electrooculogram, electro-encephalogram, respiratory, blood flow changes, etc.) using diverse kinds of sensors [27], [28]. The second one is based on facial details [29] (i.e., changes in eye movement, mouth movement, head movement, hand features, etc.).
- deep learning techniques receive vast popularity when powerful computational hardware and large-scale data become more and more available nowadays.
- Generally speaking, most contemporary deep learning models can be categorized into two types, i.e., deep generative learning models and deep discriminant

ISSN 2229-6107
www.ijpast.in
Vol 14 Issuse 1 Jan 2024

learning models. To be specific, deep generative learning models mainly aim to replicate ‘‘fake-but-realistic’’ data based on real data, and popular deep generative learning models include but not limited to VAE (i.e., variational auto-encoder) [34], GAN (i.e., generative adversarial network) [35], GLOW (i.e., generative flow)

DISADVANTAGES OF EXISTING SYSTEM:

- High-speed wireless transmissions are necessary to realize the swift and reliable transmission of highquality videos, which further facilitates the above automatic abnormal driving behavior detection task.

IV. PROPOSED SYSTEM:

- The proposed deep learning-based fusion models in automatically detecting abnormal driving behavior of this study, the Kaggle state farm distracted driver detection database
- demonstrates the trend of accuracies increasing with respect of training epochs in all compared deep learning models. First, it can be noticed that, accuracies of all deep learning models keep on increasing and then become stable when their training epochs further increase, which is a significant indicator of the thorough training and convergence of all deep learningmodels.Second,threedeepl earning-basedfusionmodels, DenseNet, as well as ResNet outperform other conventional CNN-based models (i.e., CNN, Wide CNN, Group CNN) as revealed in Figure 9. For comparisons between three deep learning-based fusion models and DenseNet, it is interesting to notice that, the former reaches the stable stage faster (i.e., less epochs) than DenseNet, and significant robustness can be obtained from new deep learning-based fusion models.

ADVANTAGES OF PROPOSED SYSTEM:

- The main advantage of affecting drivers’ normal drivings cannot be neglected, either. Furthermore, physiological signals of human beings vary greatly due to the physiological difference in each individual person and her / his

INTERNATIONAL JOURNAL OF PURE AND APPLIED SCIENCE & TECHNOLOGY
IJPAST
ISSN 2229-6107
www.ijpast.in
Vol 14 Issuse 1 Jan 2024

environmental conditions.

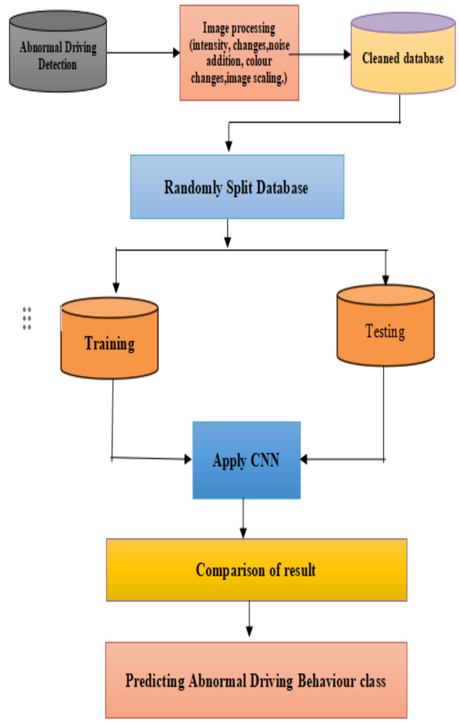


Fig1 : System architecture

V. IMPLEMENTATION

1.MODULES DESCRIPTION:

1) **Generate & Load AWGRD Model:** Using this module AWGRD train model will be generated from input images download from Kaggle state farm distracted driver detection database. This database contains 22424 images and model is built by using all those images.
2) **Upload Video:** using this module we can upload video to this application and then start playing video using Python OPENCV library.
3) **Start Behaviour Monitoring:** Using this module we will extract each frame from video and then resize image according to AWGRD Model. AWGRD Model will be applied on this frame to predict behaviour of driving person. All behaviours will be displayed on playing video.
4)

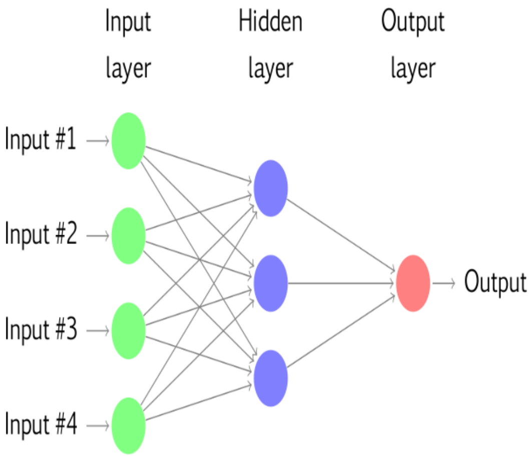
VI. ALGORITHM:

CONVOLUTIONAL NEURAL NETWORK

To demonstrate how to build a convolutional neural network based image classifier, we shall build a 6 layer neural network that will identify and separate one image from other. This network that we shall build is a very small network that we can run on a CPU as well. Traditional neural networks that are very good at doing image classification have many more parameters and take a lot of time if trained on normal CPU. However, our objective is to show how to build a real-world convolutional neural network using TENSORFLOW. Neural Networks are essentially mathematical

INTERNATIONAL JOURNAL OF PURE AND APPLIED SCIENCE & TECHNOLOGY
IJPAST
ISSN 2229-6107
www.ijpast.in
Vol 14 Issue 1 Jan 2024

models to solve an optimization problem. They are made of neurons, the basic computation unit of neural networks. A neuron takes an input (say x), do some computation on it (say: multiply it with a variable w and adds another variable b) to produce a value (say; z= wx+b). This value is passed to a non-linear function called activation function (f) to produce the final output(activation) of a neuron. There are many kinds of activation functions.



To predict image class multiple layers operate on each other to get best match layer and this process continues till no more improvement left.**FUNCTIONAL REQUIREMENTS**

Functional requirements are represented or stated in the form of input to be given to the system, the operation performed and the output expected. System should collect the data from any resources. All the

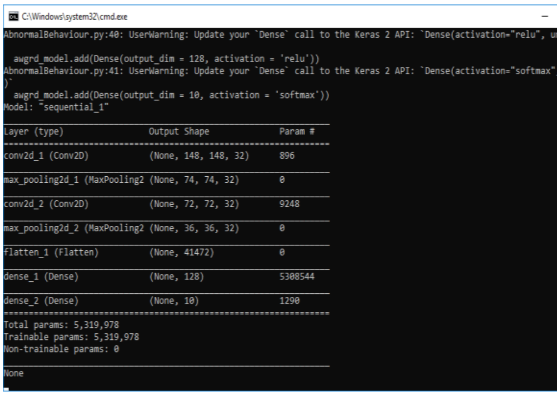
One of the popular activation function is Sigmoid. The neuron which uses sigmoid function as an activation function will be called sigmoid neuron. Depending on the activation functions, neurons are named and there are many kinds of them like RELU, TanH. If you stack neurons in a single line, it's called a layer; which is the next building block of neural networks. See below image with layers

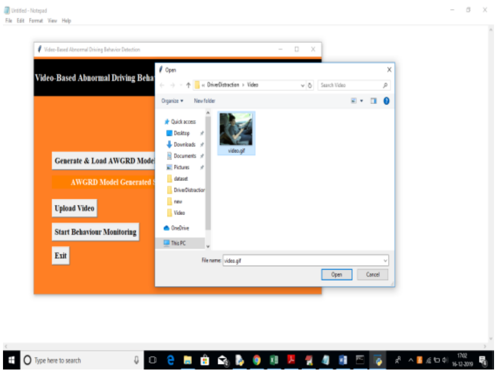
collected data should be processed for proper use, some analysis should be done for understanding the data properly.

1.Upload Video dataset
2.Data Preprocessing
3.Training And Testing
4.Modiling
5.Predicting

1.Upload Video Dataset: Collect and store diverse driving videos in a centralized repository with associated metadata for subsequent analysis.

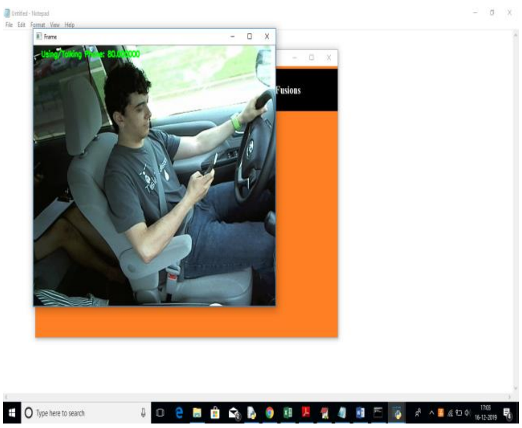
ISSN 2229-6107
www.ijpast.in
Vol 14 Issuse 1 Jan 2024





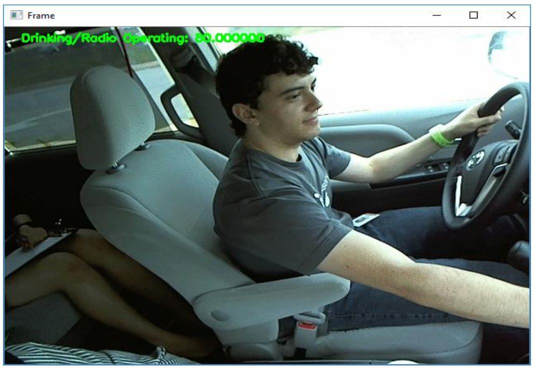
2.Data Preprocessing: Extract frames, normalize video quality, perform feature extraction, and reduce noise for effective video analysis.

3.Training and Testing: Select, train, and evaluate deep learning models using split datasets for anomaly detection in driving behavior.



4.Modeling: Develop fusion models integrating multimodal data, design architectures, and fine-tune parameters for enhanced anomaly identification.

5.Predicting: Implement real-time detection systems and conduct offline analyses using trained models to flag abnormal driving behavior patterns.



VIII. CONCLUSION

The video-based abnormal driving behavior detection study is highly important nowadays, as it is a reliable and automatic manner to ensure safeties of drivers. Also, it receives vast popularity as it is an essential step to realize fully automatic driving (i.e., particularly in Level-3 and Level-4 stages according to the ‘‘autonomous driving’’ definition provided by the US Department of Transportation’s National Highway Traffic Safety Administration). In this study, three novel deep learning-based fusion models are introduced for the first time, to fulfill the video-based abnormal driving behavior detection task. Technically, these new models are

ISSN 2229-6107
www.ijpast.in
Vol 14 Issue 1 Jan 2024

inspired by the popular DenseNet, which was proposed in recent years. For WGD, it emphasizes on important issues of designs of modern deep learning models, including the depth, the width, and the cardinality. The width and the cardinality of WGD significantly increase, therein. For WGRD and AWGRD, they are more sophisticated as the important idea of residual networks with superpositions of previous layers is incorporated. This idea is highly valuable in the video-based abnormal driving behavior detection task, as temporary and spatial latent information can be comprehensively described with the help of superpositions of previous layers. Extensive experiments based on the standard Kaggle state farm distracted driver detection dataset as well as rigorous comparisons with several other popular deep learning models suggest the superiority of newly proposed deep learning-based fusion models in both effectiveness and efficiency.

IX.BIBLIOGRAPHY

[1] W. Cao, J. Yuan, Z. He, Z. Zhang, and Z. He, ''Fast deep neural networks with knowledge guided training and predicted regions of interests for realtime video object detection,'' IEEE Access, vol. 6, pp. 8990–8999, 2018.

[2] H. Shuai, Q. Liu, K. Zhang, J. Yang, and J. Deng, ''Cascaded regional spatio-temporal feature-routing networks for video object detection,'' IEEE Access, vol. 6, pp. 3096–3106, 2018.

[3] A. Nanda, P. K. Sa, S. K. Choudhury, S. Bakshi, and B. Majhi, ''A neuromorphic person re-identification framework for video surveillance,'' IEEE Access, vol. 5, pp. 6471–6482, 2017.

[4] L. Sun, Z. Jiang, H. Song, Q. Lu, and A. Men, ''Semi-coupled dictionary learning with relaxation label space transformation for video-based person re-identification,'' IEEE Access, vol. 6, pp. 12587–12597, 2018.

[5] Y. Wu, Y. Sui, and G. Wang, ''Vision-based real-time aerial object localization and tracking for UAV sensing system,'' IEEE Access, vol. 5, pp. 23969–23978, 2017.

[6] S.-H. Lee, M.-Y. Kim, and S.-H. Bae, ''Learning discriminative appearance models for online multi-object tracking with appearance discriminability measures,'' IEEE Access, vol. 6, pp. 67316–67328, 2018.

[7] M. S. Hossain and G. Muhammad, ''An emotion recognition system for

ISSN 2229-6107
www.ijpast.in
Vol 14 Issuse 1 Jan 2024

mobile applications,'' IEEE Access, vol. 5, pp. 2281–2287, 2017.

[8] Z. Pan, X. Yi, and L. Chen, ''Motion and disparity vectors early determination for texture video in 3D-HEVC,'' Multimedia Tools Appl., to be published. doi: 10.1007/s11042-018-6830-7.

[9] J. Wang, Z. Zhang, B. Li, S. Lee, and R. S. Sherratt, ''An enhanced fall detection system for elderly person monitoring using consumer home networks,'' IEEE Trans. Consum. Electron., vol. 60, no. 1, pp. 23–29, Feb. 2014.

[10] Z. Zhang, X. Guo, and Y. Lin, ''Trust management method of D2D communication based on RF fingerprint identification,'' IEEE Access, vol. 6, pp. 66082–66087, 2018.

[11] J. Wang, Y. Cao, B. Li, H.-J. Kim, and S. Lee, ''Particle swarm optimization based clustering algorithm with mobile sink for WSNs,'' Future Gener. Comput. Syst., vol. 76, pp. 452–457, Nov. 2017. [12] Z. Xue, J. Wang, G. Ding, Q. Wu, Y. Lin, and T. A. Tsiftsis, ''Deviceto-device communications underlying UAV-supported social networking,'' IEEE Access, vol. 6, pp. 34488–34502, 2018.

[13] J. Wang, J. Cao, R. S. Sherratt, and J. H. Park, ''An improved ant colony optimization-based approach with mobile sink for wireless sensor networks,'' J. Supercomput., vol. 74, no. 12, pp. 6633–6645, Dec. 2018.

[14] Y. Tu, Y. Lin, J. Wang, and J.-U. Kim, ''Semi-supervised learning with generative adversarial networks on digital signal modulation classification,'' Comput. Mater. Continua, vol. 55, no. 2, pp. 243–254, 2018.

[15] J. Wang, Y. Gao, X. Yin, F. Li, and H.-J. Kim, ''An enhanced PEGASIS algorithm with mobile sink support for wireless sensor networks,'' Wireless Commun. Mobile Comput., vol. 2018, Dec. 2018, Art. no. 9472075.

[16] J. Sun et al., ''A multi-focus image fusion algorithm in 5G communications,'' Multimedia Tools Appl., vol. 3, pp. 1–20, Feb. 2018.

[17] J. Wang, C. Ju, Y. Gao, A. K. Sangaiah, and G.-J. Kim, ''A PSO based energy efficient coverage control algorithm for wireless sensor networks,'' Comput. Mater. Continua, vol. 56, no. 3, pp. 433–446, Sep. 2018.

[18] Y. Lin, X. Zhu, Z. Zheng, Z. Dou, and R. Zhou, ''The individual identification method of wireless device based on dimensionality reduction and machine learning,'' J. Supercomput., vol. 5, pp. 1–18, Dec. 2017.

[19] E. B. Tirkolaee, A. A. R. Hosseinabadi, M. Soltani, A. K.

ISSN 2229-6107
www.ijpast.in
Vol 14 Issue 1 Jan 2024

Sangaiah, and J. Wang, ''A hybrid genetic algorithm for multi-trip green capacitated arc routing problem in the scope of urban services,'' Sustainability, vol. 10, no. 5, p. 1366, 2018.

[20] W. Qidi, L. Yibing, L. Yun, and Y. Xiaodong, ''The nonlocal sparse reconstruction algorithm by similarity measurement with shearlet feature vector,'' Math. Problems Eng., vol. 2014, Mar. 2014, Art. no. 586014.

[21] Q. Wu, Y. Li, and Y. Lin, ''The application of nonlocal total variation in image denoising for mobile transmission,'' Multimedia Tools Appl., vol. 76, no. 16, pp. 17179–17191, Aug. 2016.

[22] D. Zhang, D. Meng, and J. Han, ''Co-saliency detection via a selfpaced multiple-instance learning framework,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 5, pp. 865–878, May 2017.

[23] D. Zeng, Y. Dai, F. Li, R. S. Sherratt, and J. Wang, ''Adversarial learning for distant supervised relation extraction,'' Comput. Mater. Continua, vol. 55, no. 1, pp. 121–136, 2018.

[24] R. Olson, R. Hanowski, J. Hickman, and J. Bocanegra, ''Driver distraction in commercial vehicle operations,'' U.S. Dept. Transp., Rep. Federal Motor Carrier Saf. Admin., Washinton, DC, USA, Tech. Rep. FMCSA-RRT-09- 042, 2009.

[25] G. Huang, Z. Liu, L. Van der Maaten, and K. Q. Weinberger, ''Densely connected convolutional networks,'' in Proc. CVPR, Jul. 2017, pp. 4700–4708.

[26] K. He, X. Zhang, S. Ren, and J. Sun, ''Deep residual learning for image recognition,'' in Proc. CVPR, Jun. 2016, pp. 770–778.

[27] A. Saeed, S. Trajanovski, M. Van Keulen, and J. Van Erp, ''Deep physiological arousal detection in a driving simulator using wearable sensors,'' in Proc. ICDMW, Nov. 2017, pp. 486–493.

[28] Z. Pan, H. Qin, X. Yi, Y. Zheng, and A. Khan, ''Low complexity versatile video coding for traffic surveillance system,'' Int. J. Sensor Netw., vol. 30, no. 2, pp. 116–125, 2019.

[29] M. Jeong, B. C. Ko, S. Kwak, and J.-Y. Nam, ''Driver facial landmark detection in real driving situations,'' IEEE Trans. Circuits Syst. Video Technol., vol. 28, no. 10, pp. 2753–2767, Oct. 2018.

[30] V. Balasubramanian and R. Bhardwaj, ''Grip and electrophysiological sensor-based estimation of muscle fatigue while holding steering wheel in different positions,'' IEEE Sensors J., vol. 19, no. 5, pp. 1951–1960, Mar. 2019.