

Objects Motion Detection in Domain-adapted Assisted Driving

Francesco Rundo¹^a, Roberto Leotta² and Sebastiano Battiato²^b

¹*STMicroelectronics, ADG Central R&D, Catania, Italy*

²*Department of Mathematics and Computer Science, University of Catania, Catania, Italy*

Keywords: ADAS, Automotive, Deep Learning, Road Classification, Intelligent Suspension.

Abstract: The modern Advanced Driver Assistance Systems (ADAS) contributed to reduce road accidents due to the driver's inexperience or unexpected scenarios. ADAS technologies allow the intelligent monitoring of the driving scenario. Recently, estimation of the visual saliency i.e. the part of the visual scene in which the driver put high visual attention has received significant research interests. This work makes further contributions to video saliency investigation for automotive applications. The difficulty to collect robust labeled data as well as the several features of the driving scenarios require the usage of such domain adaptation methods. A new approach to Gradient-Reversal domain adaptation in deep architectures is proposed. More in detail, the proposed pipeline enables an intelligent identification and segmentation of the motion salient objects in different driving scenarios and domains. The performed test results confirmed the effectiveness of the overall proposed pipeline.

1 INTRODUCTION

The term Advanced Driver Assistance Systems (ADAS) includes different type of intelligent solutions including systems providing driver assistance, advice and warnings, self autonomous driving and so on (Okuda et al., 2014). In this context the car assisting-information systems or LiDAR/RADAR based applications can be included (Spelt and Tufano, 1998). Recent ADAS technology enhancement includes Intelligent Speed Adaptation systems, collision warning systems, car driver drowsiness monitoring and pedestrian tracking systems (Ogitsu and Mizoguchi, 2015; Wang et al., 2019; Ganin and Lempitsky, 2015). The ADAS warnings system may be auditory, visual or haptic, covering such level of such standard automotive (Zhan et al., 2020). Deep Learning solutions have significantly improved the ability of algorithms to address several issues in automotive and ADAS fields.

Often, the problem of lacking labeled data, can impact the performance of such artificial intelligence based solutions. To address this relevant issue, ad-hoc intelligent domain adaptation approaches have been implemented and published in scientific literature database (Ganin and Lempitsky, 2015). The pro-

posed pipeline embeds innovative domain adaptation approach based on the usage of the Gradient Reversal algorithm. More in detail, the authors propose an overall ADAS system embedding a physio-based car driver drowsiness tracking system combined with a domain-adapted intelligent risk assessment of the associated driving scenario. Specifically, by means of the designed innovative domain adaptation method, the proposed pipeline will be able to detect and track the driving motion objects, providing an associated overall driving risk assessment. About objects motion estimation, different solutions have been proposed. A summary about scientific state of the arts is reported.

In (Zheng et al., 2018) the authors proposed an approach based on odometry for object motion estimation to be extended to automotive market. The reported performance confirmed the effectiveness of the implemented pipeline. In (Barjenbruch et al., 2015) the authors implemented an interesting motion detection pipeline based on the usage of doppler effect over radar technologies. Even the investigated approach showed very interesting results, the drawback to need the radar equipment was highlighted. In (Hee Lee et al., 2013) the authors proposed a visual ego-motion estimation algorithm for a self-driving car equipped with a commercial multi-camera system. The results obtained over a large dataset confirmed the robustness of the proposed architecture (Hee Lee et al., 2013).

^a <https://orcid.org/0000-0003-1766-3065>

^b <https://orcid.org/0000-0001-6127-2470>

Further supervised and unsupervised based deep solutions were widely investigated with aim to address the issue of the efficient driving object tracking and motion estimation (Grigorescu et al., 2019; Alletto et al., 2018; Wang et al., 2021). Considerable interest has been found by the unsupervised and semi-supervised domain adaptation techniques.

In (Singh et al., 2021) the authors proposed a semi-supervised domain adaptation approach which leveraged limited labeled target samples with unlabeled data to manage the distribution shift across the source and target domains. The proposed approach contributed significantly in bridging the domain gap as confirmed by the experimental results reported in (Singh et al., 2021). The survey reported in (Carré et al., 2018) showed the considerable advantages that domain adaptation techniques entail for ADAS and automotive applications. About driver attention monitoring systems, the authors of the pipeline herein proposed have deeply investigated that issue (Vinciguerra et al., 2018; Conoci et al., 2018; Rundo et al., 2018a; Rundo et al., 2018b; Trenta et al., 2019; Rundo et al., 2019a; Rundo et al., 2020a; Rundo et al., 2020b). More in detail, the performed scientific investigation has confirmed that the car driver physiological signals, especially the Photoplethysmography (PPG), can be efficiently used to real-time monitoring of the subject drowsiness (Rundo et al., 2020b; Rundo et al., 2019b; Lee et al., 2019). The proposed full solution can be contextualized in the intelligent driving assistance approaches in different driving scenarios and domains.

2 DOMAIN ADAPTATION FOR MOTION ASSESSMENT

As introduced, the core of the proposed approach regards a domain-adapted driving risk assessment system for ad-hoc object motion tracking. In Fig. 1 the overall diagram of the proposed pipeline.

The system reported in Fig. 1 allow a robust driving risk level assessment through an intelligent processing of the sampled driving visual frames. The sampled driving scene frames will be processed by ad-hoc Semantic Segmentation Fully Convolutional Neural Network embedding a Gradient Reversal layer (SS-FCN-GRL) (Ganin and Lempitsky, 2015). Through a semantic segmentation of the captured driving visual frames, the saliency map of the analyzed source driving scene will be reconstructed. This saliency map will be fed as input data of the downstream driving safety assessment sub-system to retrieve the correlated motion dynamic.

The proposed SS-FCN-GRL architecture will be described in detail. The designed encoder block (Encoding) composed by 5 layers is able to process the visual features of the captured driving frames. The first two blocks embed (for each block) two separable convolution layers with 3×3 kernel filters followed by a batch normalization, ReLU layer and a downstream 2×2 max-pooling layer. The remaining three blocks include two separable convolution layers with 3×3 kernel filter followed by a batch normalization, another convolutional layer with 3×3 kernel, batch normalization and ReLU with a downstream 2×2 max-pooling layer. The Decoder stage of the proposed pipeline is composed as per encoder structure i.e. up-sampling the encoded visual features through ad-hoc decoding processing. The decoder is composed by five blocks including 2D convolutional layers with 3×3 kernel, batch normalization layers, ReLU. Classical skip-connections through convolutional block were embedded in the backbone. In the decoder side we have interpolated such up-sampling blocks (with bi-cubic algorithm) to adjust the size of the generated feature maps.

To improve the domain adaptation capability, the authors have embedded the mentioned Gradient Reversal Layer block (Ganin and Lempitsky, 2015) as per Fig. 1. The designed he model works with input samples $x \in X$, where X is the input space while y (label data) from the label space Y . Let's defined a classification problems where Y is a finite set ($Y = \{1, 2, \dots, L\}$), handling any output label space. We further assumed that there exist two distributions $S(x, y)$ and $T(x, y) X \otimes Y$, which will be referred to as the source distribution and the target distribution (or the source domain and the target domain). Both distributions were assumed as unknown and specifically S is "shifted" from T through a not specified domain remapping. Our objective function is to predict labels y given the input x for the target distribution. At training time, we supposed to collect large training samples $\{x_1, x_2, \dots, x_N\}$ from both the source and the target domains distributed according to the defined statistical distributions $S(x)$ and $T(x)$.

We denote with d_i the binary variable (domain label) for the i -th example, which means that ($x_i \sim S(x)$ if $d_i = 0$) or ($x_i \sim T(x)$ if $d_i = 1$). We now define a custom deep feed-forward architecture which for each input x predicts its label $y \in Y$ and its domain label $d \in \{0, 1\}$. The authors assumed that the input x is re/mapped through the function G_f (a feature extractor) to a D -dimensional feature vector $f \in R^D$. The proposed feature mapping includes feed-forward layers and we denote the vector of parameters of all layers in this mapping as θ_f , i.e. $f = G_f(x; \theta_f)$. Then,

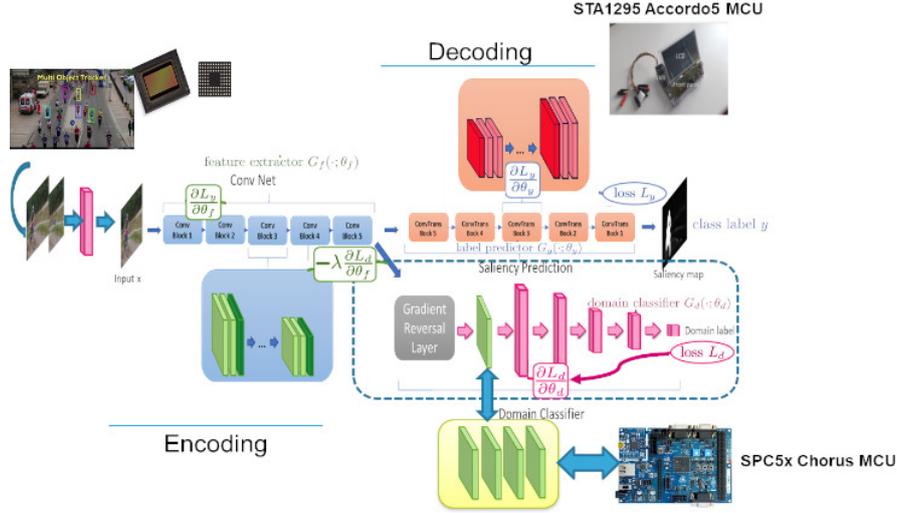


Figure 1: The proposed domain adapted motion assessment pipeline.

the feature vector f is mapped by a mapping G_y (label predictor) to the label y , and we denote the parameters of this mapping with θ_y . Finally, the same feature vector f is mapped to the domain label d by a mapping G_d (domain classifier) with the parameters θ_d (Fig. 1).

During the learning session, the implemented deep system tries to minimize the label prediction loss on the labeled part (i.e. the source part) of the training set. At the same time, the parameters of both the feature extractor and the label predictor are thus optimized in order to minimize the empirical loss for the source domain unlabeled samples. In the following section, a mathematical formalization of the proposed GRL approach:

$$\begin{aligned}
 E(\theta_f, \theta_y, \theta_d) &= \sum_{\substack{i=1 \dots N \\ d_i=0}} L_y(G_y(G_f(x_i; \theta_f); \theta_y), y_i) - \\
 &\lambda \sum_{i=1 \dots N} L_d(G_d(G_f(x_i; \theta_f); \theta_d), y_i) = \\
 &= \sum_{\substack{i=1 \dots N \\ d_i=0}} L_y^i(\theta_f, \theta_y) - \lambda \sum_{i=1 \dots N} L_d^i(\theta_f, \theta_d)
 \end{aligned} \quad (1)$$

where $L_y(\cdot, \cdot)$ is the loss for the label prediction while $L_d(\cdot, \cdot)$ is the loss for the domain classification. The terms L_y^i and L_d^i denote the corresponding loss functions evaluated at the i -th training input frames. Based on our proposed architecture, we are seeking the parameters $\hat{\theta}_f, \hat{\theta}_y, \hat{\theta}_d$ that find a saddle point of the Eq. 1:

$$(\hat{\theta}_f, \hat{\theta}_y) = \underset{\theta_f, \theta_y}{\operatorname{argmin}} E(\theta_f, \theta_y, \hat{\theta}_d) \quad (2)$$

$$\hat{\theta}_d = \underset{\theta_d}{\operatorname{argmax}} E(\hat{\theta}_f, \hat{\theta}_y, \theta_d) \quad (3)$$

At the saddle point, the parameters θ_d of the domain classifier minimizes the domain classification loss while the parameters θ_y of the label predictor minimizes the label prediction loss. The feature mapping parameters θ_f minimizes the label prediction loss, while maximizing the domain classification loss (i.e. the features are domain-invariant). The learning rate λ modulates the two objective dynamics. The authors whose designed the GRL approach have showed that the classical Stochastic Gradient Descent (SGD) learning is able to find the needed saddle point (Ganin and Lempitsky, 2015).

After that SGD-based learning, the label predictor $y(x) = G_y(G_f(x; \theta_f); \theta_y)$ can be used to predict labels for samples from the target domain (and clearly the source domain). The output of the so designed SS-FCN-GRL is the feature saliency map of the sampled driving frame. Specifically, a no-zero saliency map will be generated in case of motion objects while null saliency map will be generated in case of slow-motion or static objects in the sampled driving scenario frames.

Through the action of the GRL the proposed solutions are able to train the deep network both with label data and target ones (different driving scenario frames) suitable to retrieve a robust domain-invariant motion-objects saliency map, thus characterizing the risk of driving accordingly.

In Fig. 2 the authors have reported some instances of the so generated saliency maps for moving and slow-motion objects in the analyzed driving scenario. The proposed SS-FCN-GRL architecture has been validated and tested on the DHF1K dataset (Min and Corso, 2019). The proposed solution has showed acceptable performance on DHF1K dataset (Rundo

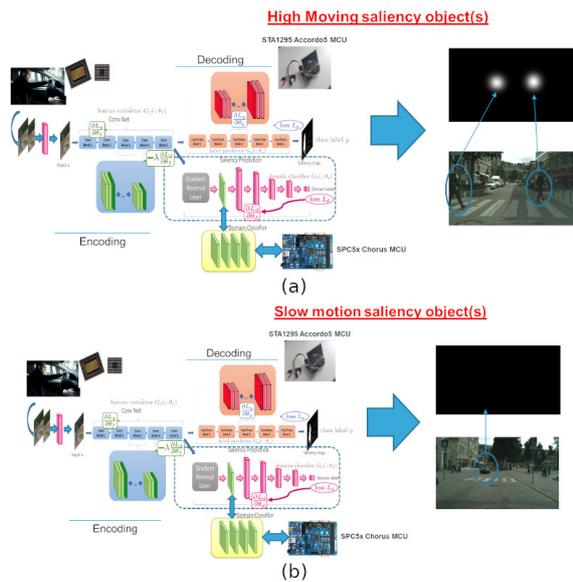


Figure 2: Intelligent Driving Objects Tracking and Motion Detection: (a) High moving saliency objects tracking; (b) Slow motion objects tracking and detection (Null saliency map).

et al., 2019b) (Area Under the Curve: 0.892; Similarity: 0.415; Correlation Coefficient: 0.480; Normalized Scanpath Saliency: 2.598) with respect to similar intelligent backbones. A careful porting of the proposed deep solution as per Fig. 1 is running over ad-hoc hardware with graphic/AI accelerations and provided by STMicroelectronics (MCUs framework based STA1295A Accordo5 and SPC5x Chorus devices) (Rundo et al., 2021).

3 THE PHYSIO-BASED CAR DRIVER DROWSINESS MONITORING

As introduced, the proposed ADAS solution embeds a robust intelligent car driver drowsiness monitoring system. Specifically, we proposed a car-driver attention level monitoring based on the usage of the driver’s Photoplethysmographic (PPG) signal.

The PPG signal can be considered as a less-invasive physio-signal suitable to monitor cardiovascular dynamics of a human subject. Both heart pulse and respiratory rate as well as vascular and cardiac disorders may be monitored by means of ad-hoc analysis of the PPG dynamic (Ganin and Lempitsky, 2015). Through the PPG signal the authors were able to perform less-invasive measure of the subject blood volume changes. A classical PPG waveform embeds a pulsatile (‘AC’) physiological signal which

is correlated to cardiac-synchronous changes in the blood volume superimposed with a slowly varying (‘DC’) component containing lower frequency sub-signals correlated to respiration and other physiological parameters. The change in volume caused by the periodic heart pressure pulses can be tracked by illuminating the skin of the subject and then by measuring the amount of light either transmitted or back-scattered by means of ad-hoc combined detector (Ogitsu and Mizoguchi, 2015; Wang et al., 2019). More detail about PPG patter formation in (Okuda et al., 2014; Spelt and Tufano, 1998; Panagiotopoulos and Dimitrakopoulos, 2019; Ogitsu and Mizoguchi, 2015; Wang et al., 2019; Ganin and Lempitsky, 2015; Zhan et al., 2020). For the proposed pipeline, the authors have used the PPG sampling embedding a Silicon Photomultiplier (SiPM) device provided by STMicroelectronics (Vinciguerra et al., 2018; Conoci et al., 2018; Rundo et al., 2018a).

The proposed PPG sensing probes includes a large area n-on-p Silicon Photomultipliers (SiPMs) fabricated at STMicroelectronics (Conoci et al., 2018; Rundo et al., 2018a). $4.0 \times 4.5 \text{ mm}^2$ and 4871 square microcells with $60 \mu\text{m}$ pitch. The devices have a geometrical fill factor of 67.4% and are packaged in a surface mount housing (SMD) with about $5.1 \times 5.1 \text{ mm}^2$ total area (Conoci et al., 2018). We propose the usage of Pixelteq dichroic bandpass filter with a pass band centered at about 840 nm with a Full Width at Half Maximum (FWHM) of 70 nm and an optical transmission higher than 90 – 95% in the pass band range was glued on the SMD package by using a Loctite 352TM adhesive. With the dichroic filter at 3V-OV the SiPM has a maximum detection efficiency of about 30% at 565 nm and a PDE of about 27.5% at 830 nm (central wavelength in the filter pass band). We have applied a dichroic filter to reduce the absorption of environmental light of more than 60% when the detector works in the linear range in Geiger mode above its breakdown voltage ($\sim 27 \text{ V}$).

As described, the so designed PPG probe embeds a set of OSRAM LT M673 LEDs in SMD package emitting at 830 nm and based on InGaN technology (Conoci et al., 2018). The used LEDs devices have an area of $2.3 \times 1.5 \text{ mm}^2$, viewing angle of 120° , spectral bandwidth of 33 nm and lower power emission (mW) in the standard operation range. The authors designed an embedded motherboard populated by a 4 V portable battery, a power management circuits, a conditioning circuit for output SiPMs signals, several USB connectors for PPG probes and related SMA output connectors (Conoci et al., 2018; Rundo et al., 2018b). We designed to embed several PPG sensing probes on the car steering.

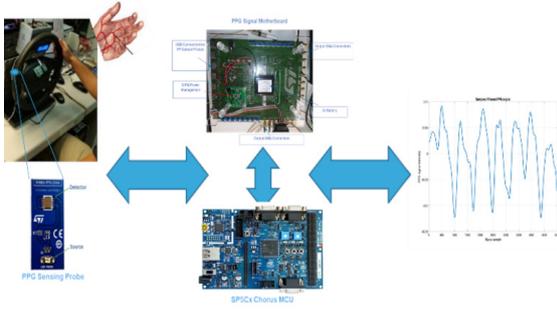


Figure 3: The designed PPG sensing platform.

In Fig. 3 we report an overall scheme of the proposed PPG sensing framework. The filtering and stabilization of the collected raw PPG signal will be performed by the developed algorithms running as firmware in the SPC5x Chorus MCU (Conoci et al., 2018; Rundo et al., 2018b; Trenta et al., 2019; Rundo et al., 2019a; Rundo et al., 2020a; Rundo et al., 2020b; Rundo et al., 2019b).

The designed hyper-filtering approach (Rundo et al., 2018b; Rundo et al., 2019b) will be applied to the collected steady-state PPG raw data in order to retrieve such discriminative features to be correlated to the driver attention level.

More in detail, the idea inside the hyper-filtering approach was inspired by hyper-spectral method usually applied to 2D data (Rundo et al., 2019b). Basically, the authors investigated the discrimination level of the features retrieved by the "hyper-filtering" of the source car driver PPG signal. More in detail, instead of applying a single filter setup (low pass and high pass) having a well-defined cut-off frequency, we have analyzed a range of dynamic frequencies in which the PPG signal shows useful information. Considering that the useful frequency range is included in the 0.5 – 10Hz, we have investigated the performance of an hyper-filtered PPG-based classification system in which the signal frequency spectrum (0.5 – 10Hz) was divided into several sub-bands. We have configured two spectral layers of hyper-filtering layer. A first layer changes the frequencies in the low-pass filter maintaining instead the cut-off frequency of the high-pass filter (Hyper low-pass filtering layer) and vice versa a layer that changes the cut-off frequencies of the high-pass filter while maintaining fixed the frequency setup of the low pass filter ((Hyper high-pass filtering layer). Due to an efficient noise-modulations in the bandwidth, we adopted the Butterworth filter types in both layers of Hyper filtering (Rundo et al., 2020a; Rundo et al., 2020b; Rundo et al., 2019b). Through a Reinforcement Learning algorithm (Rundo et al., 2020b; Rundo et al., 2019b) we optimized the setup of hyper-filtering layer to be applied to the col-

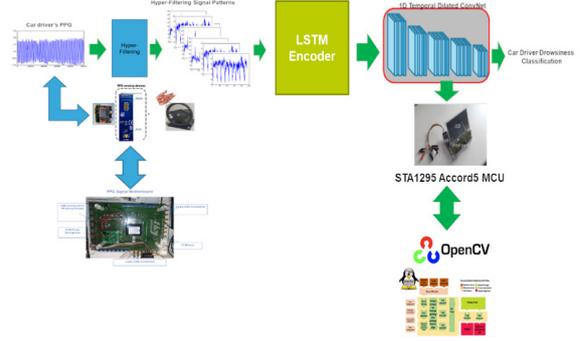


Figure 4: The proposed downstream PPG deep 1D-CNN classifier.

lected car driver PPG signal. This setup is reported in the following Table 1 e Table 2.

Table 1: Hyper Low-pass filtering setup (in Hz).

F	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11
HP	0.5	/	/	/	/	/	/	/	/	/	/
LP	0.0	1.2	3.3	3.5	3.6	3.8	4.0	4.2	5.0	5.1	6.1

Table 2: Hyper High-pass filtering setup (in Hz).

F	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11
HP	0.0	1.2	2.3	2.6	3.1	3.5	4.0	4.3	5.0	5.5	6.2
LP	0.6	/	/	/	/	/	/	/	/	/	/

At this point, once the hyper-filtering configuration has been completed, the car driver PPG raw signal that is gradually sampled will be processed accordingly to the so configured hyper-filtered framework. For each sample of the single PPG waveform, a dataset of hyper filtered signals will be generated, each having a temporal dynamics represented by the time-dynamic of that signal-sample. Formally, if we indicate with $W_i^{PPG}(t, k)$ the single segmented PPG compliant waveform of each hyper-filtered PPG time-serie, we proceed computing for each sample $s(t, k)$ of the waveform a signal-pattern depending on how that signal samples $s(t, k)$ changes in intensity in each of the previously generated hyper-filtered signals.

Through the above detailed hyper-filtered pipeline, a large dataset of hyper-filtered signals will be generated and collected. The above RL algorithms confirmed that a well optimization is reached by 11 sub-bands of hyper-filtering frequencies. The so generated signal-patterns will be fed into the Deep Learning block as described in the Fig. 4.

The signal-patterns generated by the previous hyper-filtering PPG processing pipeline (Rundo et al., 2019b) will be preliminary encoded by a Long-Short-Term Memory (LSTM) network. The LSTM composed by 6 layers of vanilla-unit backbone (Trenta et al., 2019), is able to encode the hyper-filtered signal-patterns in order to provide a folding-

embedding of the input signals. These features will be fed as input to the deep 1D downstream classifier (Rundo et al., 2018b; Trenta et al., 2019). The Deep 1D Temporal Dilated Convolutional Neural Network (1D-CNN) with residual block (Rundo et al., 2020b). Specifically, we have implemented a 1D-CNN embedding 36 blocks with a downstream softmax layer. The output of the deep network is a classification of the input hyper-filtered PPG patterns i.e. a classification of drowsy (0.0, 0.5) or wakeful (0.51, 1.0) driver. Each of the 1D-CNN block consists of a dilated convolution layer having 3×3 kernel filters, a spatial dropout layer, another dilated convolution layer, ReLU layer and a final spatial drop. The dilation size start from 2 and increase (power of 2) for each block till the max value of 32. A softmax layer completes the proposed pipeline. The so designed Deep Learning framework is able to estimate and monitor the car driver drowsiness level. As reported in Fig. 4, the proposed deep classifier is running over the STA1295A Accordo5 MCU with ad-hoc Graphics accelerator and Linux YOCTO and OpenCV based software framework (Rundo et al., 2020b; Rundo et al., 2019b).

4 EXPERIMENTAL RESULTS AND CONCLUSION

We tested the proposed pipeline, firstly validating each of the implemented sub-systems and then arranging a combined testing scenario. Specifically, we have considered the following risk assessment in relation to the tracking of the salient moving objects: detected no-zero map for salient moving objects (medium/high driving risk) against a scenario with detected zero-map slow moving salient objects with associated null generated saliency map (low risk driving scenario).

Therefore, the proposed full pipeline provide an overall driving risk assessment comparing the saliency-motion-based risk evaluation with the PPG physio-based drowsiness monitoring retrieving if that attention level is adequate or not.

More in detail, if high or medium risk level is detected, the proposed driving monitoring system will check if the designed 1D-TCNN detects a corresponding "wakeful driver" classification. Otherwise, acoustic alert-signal will be generated. In the scenario in which the driver's PPG signal is not available for some reasons, the authors have developed a Visual-to-PPG replacement algorithm (Trenta et al., 2019). About the physio-based car driver drowsiness assessment, we have validated the proposed pipeline by collecting several PPG measurements of different sub-

Table 3: Car Driver Drowsiness monitoring performance.

Method	Driver Drowsiness Monitoring	
	Drowsy Driver	Wakeful Driver
Proposed	99.76%	99.89%
1D-Temporal CNN w/o LSTM	98.71%	99.03%
(Rundo et al., 2019b)	96.50%	98.40%

jects in different scenarios (Drowsy driver vs Wakeful driver) under authorization of the Ethical Committee CT1 authorization Nr. 113 / 2018 / PO. The PPG sampling sessions have been supervised by experienced physicians. We have collected data from 70 patients with different ages, sex, and so on (Rundo et al., 2019a). We have used the herein introduced PPG hardware setup with sampling frequency of 1 kHz. For each condition (Drowsy vs Wakeful subject) we have collected 5 minutes of PPG signals. All acquired PPG time-series have been arranged as follow: 70% for the training and validation phase of the Deep learning framework while the remaining 30% have been used for testing. We have used different driving scenarios: some labeled driving scenarios and some unlabeled target domain scenarios for testing the proposed GRL approach. The results reported in Table 3 confirmed that the physio-sensing system for car driver drowsiness monitoring outperformed similar pipelines in terms of accuracy.

The SS-FCN-GRL and the 1D-TCNN have been trained with a classic SGD algorithm with dropout factor of 0.75 and initial learning rate of 0.001. The LSTM layer was trained with an initial learning rate of 0.002. The Table 3 shows the performance of the proposed pipeline compared with similar pipeline based on deep learning (Rundo et al., 2019b) both in labeled and target domain adapted driving scenarios. We consider such interval of about 8/12 seconds of PPG sampling in order to show the near real-time performance of the proposed pipeline. Finally, we have tested the combined full system. Specifically, we have validated the proposed low/high risk assessment of the analyzed driving scenarios. As highlighted by the experimental results reported in Table 4, the architecture that exports the domain adaptation GRL layer shows high performance in risk assessment compared to the benchmark architectures. The use of the GRL significantly improves the characterization of the motion of the tracked objects and therefore the consequent risk assessment (accuracy on average of 96% against 91% of the system without GRL).

The collected performance confirmed the robustness and the effectiveness of the proposed overall approach.

Table 4: Proposed Intelligent Saliency-Motion Driving risk assessment system.

Method	Intelligent Domain Adapted Risk Assessment Performance	
	Low Risk (Static Salient Objects)	High Risk (Salient Moving Objects)
Proposed	96.78%	96.66%
Proposed w/o GRL	91.87%	91.09%
Classic SS-FCN With Attention (Rundo et al., 2021)	91.65%	90.90%
Classic SS-FCN DenseNet Backbone	89.30%	90.11%
Classic SS-FCN ResNet-101 backbone (Min and Corso, 2019)	87.79%	90.01%

Future works aim on embedding such features of deep LSTM with Attention to further improve the performance of overall pipeline (Rundo, 2019).

ACKNOWLEDGEMENTS

The authors thank the physiologists belonging to the Department of Biomedical and Biotechnological Sciences (BIOMETEC) of the University of Catania, who collaborated in this work in the context of the clinical study Ethical Committee CT1 authorization n.113 / 2018 / PO. This research was funded by the National Funded Program 2014-2020 under grant agreement n. 1733, (ADAS + Project). The reported information is covered by the following registered patents: IT Patent Nr. 102017000120714, 24 October 2017; IT Patent Nr. 102019000005868, 16 April 2018; IT Patent Nr. 102019000000133, 07 January 2019.

REFERENCES

- Alletto, S., Abati, D., Calderara, S., Cucchiara, R., and Rigazio, L. (2018). Self-supervised optical flow estimation by projective bootstrap. *IEEE Transactions on Intelligent Transportation Systems*, 20(9):3294–3302.
- Barjenbruch, M., Kellner, D., Klappstein, J., Dickmann, J., and Dietmayer, K. (2015). Joint spatial-and doppler-based ego-motion estimation for automotive radars. In *2015 IEEE Intelligent Vehicles Symposium (IV)*, pages 839–844. IEEE.
- Carré, M., Exposito, E., and Ibanez-Guzman, J. (2018). Challenges for the self-safety in autonomous vehicles. In *2018 13th Annual Conference on System of Systems Engineering (SoSE)*, pages 181–188. IEEE.
- Conoci, S., Rundo, F., Fallica, G., Lena, D., Buraioli, I., and Demarchi, D. (2018). Live demonstration of portable systems based on silicon sensors for the monitoring of physiological parameters of driver drowsiness and pulse wave velocity. In *2018 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 1–3. IEEE.
- Ganin, Y. and Lempitsky, V. (2015). Unsupervised domain adaptation by backpropagation. In *International conference on machine learning*, pages 1180–1189. PMLR.
- Grigorescu, S. M., Trasnea, B., Marina, L., Vasilcoi, A., and Cocias, T. (2019). Neurotrajectory: A neuroevolutionary approach to local state trajectory learning for autonomous vehicles. *IEEE Robotics and Automation Letters*, 4(4):3441–3448.
- Hee Lee, G., Faundorfer, F., and Pollefeys, M. (2013). Motion estimation for self-driving cars with a generalized camera. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2746–2753.
- Lee, H., Lee, J., and Shin, M. (2019). Using wearable ecg/ppg sensors for driver drowsiness detection based on distinguishable pattern of recurrence plots. *Electronics*, 8(2):192.
- Min, K. and Corso, J. J. (2019). Tased-net: Temporally-aggregating spatial encoder-decoder network for video saliency detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2394–2403.
- Ogitsu, T. and Mizoguchi, H. (2015). A study on driver training on advanced driver assistance systems by using a driving simulator. In *2015 International Conference on Connected Vehicles and Expo (ICCVE)*, pages 352–353. IEEE.
- Okuda, R., Kajiwara, Y., and Terashima, K. (2014). A survey of technical trend of adas and autonomous driving. In *Technical Papers of 2014 International Symposium on VLSI Design, Automation and Test*, pages 1–4. IEEE.
- Panagiotopoulos, I. and Dimitrakopoulos, G. (2019). Cognitive infotainment systems for intelligent vehicles. In *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)*, pages 1–8. IEEE.
- Rundo, F. (2019). Deep lstm with reinforcement learning layer for financial trend prediction in fx high frequency trading systems. *Applied Sciences*, 9(20):4460.
- Rundo, F., Conoci, S., Battiato, S., Trenta, F., and Spampinato, C. (2020a). Innovative saliency based deep driving scene understanding system for automatic safety assessment in next-generation cars. In *2020 AEIT International Conference of Electrical and Electronic Technologies for Automotive (AEIT AUTOMOTIVE)*, pages 1–6. IEEE.
- Rundo, F., Conoci, S., Ortis, A., and Battiato, S. (2018a). An advanced bio-inspired photoplethysmography (ppg) and ecg pattern recognition system for medical assessment. *Sensors*, 18(2):405.
- Rundo, F., Leotta, R., and Battiato, S. (2021). Real-time deep neuro-vision embedded processing system for saliency-based car driving safety monitoring. In *2021 4th International Conference on Circuits, Systems and Simulation (ICCS)*, pages 218–224. IEEE.

- Rundo, F., Petralia, S., Fallica, G., and Conoci, S. (2018b). A nonlinear pattern recognition pipeline for ppg/ecg medical assessments. In *Convegno Nazionale Sensori*, pages 473–480. Springer.
- Rundo, F., Rinella, S., Massimino, S., Coco, M., Fallica, G., Parenti, R., Conoci, S., and Perciavalle, V. (2019a). An innovative deep learning algorithm for drowsiness detection from eeg signal. *Computation*, 7(1):13.
- Rundo, F., Spampinato, C., Battiato, S., Trenta, F., and Conoci, S. (2020b). Advanced 1d temporal deep dilated convolutional embedded perceptual system for fast car-driver drowsiness monitoring. In *2020 AEIT International Conference of Electrical and Electronic Technologies for Automotive (AEIT AUTOMOTIVE)*, pages 1–6. IEEE.
- Rundo, F., Spampinato, C., and Conoci, S. (2019b). Ad-hoc shallow neural network to learn hyper filtered photoplethysmographic (ppg) signal for efficient car-driver drowsiness monitoring. *Electronics*, 8(8):890.
- Singh, A., Doraiswamy, N., Takamuku, S., Bhalerao, M., Dutta, T., Biswas, S., Chepuri, A., Vengatesan, B., and Natori, N. (2021). Improving semi-supervised domain adaptation using effective target selection and semantics. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2709–2718.
- Spelt, P. and Tufano, D. (1998). An in-vehicle information system for its information management. In *17th DASC. AIAA/IEEE/SAE. Digital Avionics Systems Conference. Proceedings (Cat. No. 98CH36267)*, volume 2, pages 131–1. IEEE.
- Trenta, F., Conoci, S., Rundo, F., and Battiato, S. (2019). Advanced motion-tracking system with multi-layers deep learning framework for innovative car-driver drowsiness monitoring. In *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*, pages 1–5. IEEE.
- Vinciguerra, V., Ambra, E., Maddiona, L., Romeo, M., Mazzillo, M., Rundo, F., Fallica, G., di Pompeo, F., Chiarelli, A. M., Zappasodi, F., et al. (2018). Ppg/ecg multisite combo system based on sipm technology. In *Convegno Nazionale Sensori*, pages 353–360. Springer.
- Wang, C., Sun, Q., Guo, Y., Fu, R., and Yuan, W. (2019). Improving the user acceptability of advanced driver assistance systems based on different driving styles: A case study of lane change warning systems. *IEEE Transactions on Intelligent Transportation Systems*, 21(10):4196–4208.
- Wang, K., Jiasheng, N., and Yanqiang, L. (2021). A robust lidar state estimation and map building approach for urban road. In *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, pages 502–506. IEEE.
- Zhan, H., Wan, D., and Huang, Z. (2020). On the responsible subjects of self-driving cars under the sae system: An improvement scheme. In *2020 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1–5. IEEE.
- Zheng, F., Tang, H., and Liu, Y.-H. (2018). Odometry-vision-based ground vehicle motion estimation with se (2)-constrained se (3) poses. *IEEE transactions on cybernetics*, 49(7):2652–2663.