FACTA UNIVERSITATIS
Series: **Automatic Control and Robotics** Vol. 20, Nº 2, 2021, pp. 75 - 85
https://doi.org/10.22190/FUACR210319006R

Regular Paper

DEEP LEARNING-BASED OBSTACLE DETECTION AND DISTANCE ESTIMATION USING OBJECT BOUNDING BOX

UDC (159.923.2:(004.8+004.4))

Danijela Ristić-Durrant[1], Marten Franke[1], Kai Michels[1], Vlastimir Nikolić[2], Milan Banić[2], Miloš Simonović[2]

[1]University of Bremen, Institute of Automation, Germany
[2]University of Niš, Faculty of Mechanical Engineering, Republic of Serbia

Abstract. *In this paper, a system consisting of deep learning (DL)-based object detection followed by neural network based object distance estimation is considered. The accuracy of object distance estimation strongly depends on the size of the bounding box (BB) of the detected object extracted by the DL-based object detector. A method for improvement of the accuracy of object BB is proposed, which involves traditional computer vision-based edge segmentation of object BB image region. The proposed method is evaluated on the real-world images of railway scenes with obstacles on the rail tracks captured by thermal and RGB cameras. The evaluation results demonstrate the potential of traditional computer vision methods to complement state-of-the-art DL methods for accurate object detection and distance estimation.*

Key words: *Autonomous obstacle detection; deep learning-based object detection; object bounding box-based distance estimation; traditional computer vision edge-based segmentation*

1. INTRODUCTION

Environment perception with the purpose of obstacle detection for collision avoidance is crucial for the safety of a wide range of applications involving moving elements, ranging from robotic manipulators to manned and unmanned vehicles. As a result of developments in sensor technology and of Artificial Intelligence (AI), in recent years, there has been a rapid expansion in research and development of obstacle detection for road transport. Although railways are the other principal means of transport over land, research and development of obstacle detection in railways has to date lagged behind that

Received March 19, 2021 / Accepted June 09, 2021
Corresponding author: Danijela Ristić-Durrant
University of Bremen, Institute of Automation, Otto-Hahn-Allee 1, 28359 Bremen, Germany
E-mail: ristic@iat.uni-bremen.de

for road transport. Bearing in mind this advancement in obstacle detection for road vehicles over rail transport, and taking into account the similarities between these two means of land transport, it is understandable that the recently intensified research in obstacle detection for railways has built on research in the automotive field. That is, researchers and developers have applied and tested techniques that are known from the car driver assistance technologies, to develop obstacle detection systems for rail transport [1]. This includes using the same on-board sensors with predominant use of vision sensors, cameras of different types, and recent using of deep learning (DL)-based methods for object detection in cameras' images [2]. In this way, DL-based methods for obstacle detection in railways and automotive field currently share the same problems of lower performance of these methods in cameras' images of insufficiently good quality.

Currently, DL-based object detection frameworks offer an impressive performance on high-quality images. However, applying them to low-quality data causes significant performance drops. This performance drop is characterized with missing object detection and/or with imprecision in object detection. An example of low-quality image data is a thermal camera's image. Thermal cameras are desired on-board sensors for obstacle detection as they have potential for improved robustness in bad weather and nighttime compared with standard Red-Green-Blue (RGB) cameras. However, detecting objects such as persons, in thermal infrared imagery is a hard problem because image resolution and quality is typically far lower, especially for low-cost sensors. Also, the quality of thermal camera images are strongly influenced with environmental conditions and in the case of high environmental temperature the thermal camera images are of low contrast. Authors of [3] proposed a strategy to make use of elaborate Convolutional Neural Network (CNN)-based object detector frameworks which are pre-trained on visual RGB images. This strategy involves an appropriate traditional computer vision pre-processing method for transforming the thermal image data as close as possible to the RGB domain. This allows pre-trained RGB features to be effective on the novel, thermal image, domain.

However, even in the case of input images of good quality, state-of-the-art DL-based object detectors often suffer from the imperfect detection in real-world applications. Some of the possible reasons for this are: operating in complex scenes and insufficiently large datasets used for DL-models trainings. For the majority of DL-based object detectors their results are indicated by having the detected object encapsulated with a bounding box (BB) corresponding to the object's position and size. Due to the imperfections of DL-based methods, it is often difficult to obtain the precise BB as errors frequently occur producing oversized, partial, and false bounding boxes. In [4], a precise regression approach for correcting imprecise BB using deep reinforcement learning was proposed. An intelligent agent that moves BB to correct position and scales it to its correct size was trained by deep Q-learning and evaluated on several state-of-the-art multiple object tracking approaches. However, such method for DL-based bounding box refinement is expensive with respect to time and as such is not appropriate for real-world applications.

An accurately extracted object BB is of importance in a number of autonomous applications where an object BB is used as a basis for calculation of further object characteristics such as the object pose in 3D [5] and distance between the object and on-board camera [6]. In order to achieve correction of an object BB, extracted by a DL-based object detector, for the purpose of precise real-time object detection and distance estimation, in this paper a traditional computer vision-based post-processing of object BB is proposed. More precisely, the presented method was developed for the purpose of

reducing the error in object distance estimation using Artificial Neural Network (ANN) method named DisNet [6], which was developed within the project SMART-Smart Automation of Rail Transport [7].

This paper is organized as follows. In Section 2, the system for object detection and distance estimation using a DL-based object BB extraction is presented. A modification of the system with the purpose of refinement of object BB in order to improve the accuracy of object distance estimation is presented in Section 3. The experimental results are given in Section 4. The concluding remarks are given in the last Section 5.

2. System Overview

A block-diagram of the system for the DL-based object detection and distance estimation using object's BB is shown in Fig. 1. As illustrated, the system consists of two main parts. The first part is Convolutional Neural Network (CNN)-based object detection and the second part is multi hidden-layer neural network-based distance estimation.

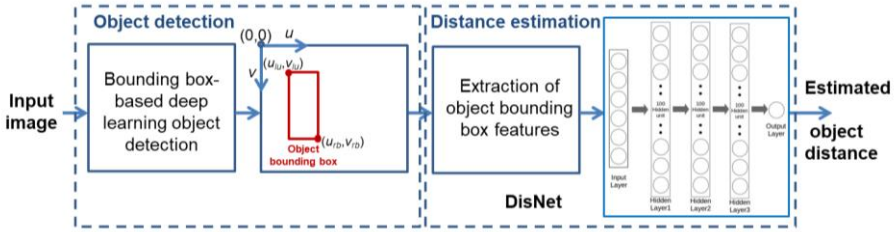

Fig. 1 DisNet-based object distance estimation system

The object detector can be any BB-based DL method which extracts the BB of an object detected in the input image as well as the object class. The results presented in this paper are based on the use of two of the state-of-the-art CNNs for the prediction of detected objects bounding boxes named YOLOv3 [8] and CenterNet [9]. One of the main advantages of both networks is their speed, making them appropriate for real-time applications such as obstacle detection in real-world operational railway applications. The YOLOv3 and CenterNet models used were originally trained with the Microsoft COCO dataset [10] of images of everyday scenes containing common objects in their natural context, consisting of 328000 images of 80 easily recognizable objects classes. In total, 2.5 million objects are labeled in the images of the dataset making it suitable for training DL models for object detection in different applications.

The distance estimator is a feedforward artificial neural network named DisNet. It was developed within the project SMART-Smart Automation of Rail Transport [7] and it consists of 3 hidden layers, each containing 100 hidden units. DisNet estimates distance between each detected object and the on-board camera based on the features of the object Bounding Box (BB) extracted by the DL-based object detector. The details of the development and training of DisNet networks are given in [6] and in the following an overview of this distance estimator is given.

The bounding boxes (BBs) of objects in a given input image can be directly extracted from the outputs of the pre-trained DL-based object detector model. Once a DL-based object detector has detected objects and extracted their BBs, the BB features to be used for subsequent object distance estimation by DisNet are calculated. These features are the height, width and diagonal of the BB, which are calculated respectively as follows:

$$B_h = \frac{\text{number of image rows}}{v_{lu} - v_{rb}}, \tag{1}$$

$$B_w = \frac{\text{number of image columns}}{u_{rb} - u_{lu}}, \tag{2}$$

$$B_d = \frac{\sqrt{\text{number of image rows}^2 + \text{number of image columns}^2}}{\sqrt{(v_{lu} - v_{rb})^2 + (u_{rb} - u_{lu})^2}}. \tag{3}$$

where (u_{lu}, v_{lu}) and (u_{rb}, v_{rb}) are respectively the image coordinates of the left upper corner and the right bottom corner of the object BB, as illustrated in Fig. 1. The above BB features are invariant to cameras image resolution so the DisNet trained model can be used with a variety of cameras independently of image resolution.

For each extracted object BB, a six-dimensional feature vector $\boldsymbol{v}$ is calculated:

$$\boldsymbol{v} = [B_h, B_w, B_d, C_h, C_w, C_d]^{\mathrm{T}}. \tag{4}$$

where C_h, C_w and C_d are the values of average height, width and depth of an object of the particular class. For example, for the class "human" C_h, C_w and C_d are respectively 175 cm, 55 cm and 30 cm, while for the class "car" these parameters are 160 cm, 180 cm and 400 cm respectively. The features C_h, C_w and C_d represent three-dimensional object features that complement information on object bounding boxes extracted from two-dimensional images and so they give more information to distinguish different objects. The feature vector (4) is the input to the DisNet feedforward neural network, which estimates the distance for each object. After training of DisNet, given an image including the BBs of the objects in it, the object-specific distance can be directly extracted from the outputs of the trained DisNet model as illustrated in Fig. 1. The dataset used to train DisNet was created by manually extracting the BBs of objects in the RGB images and assigning the ground truth distance to these labelled objects as described in [6].

As obvious from Fig.1, the accuracy of object distance estimation using the trained DisNet model strongly depends on the accuracy of DL-based object BB extraction as demonstrated by the following example illustrated by Fig. 2. The result of object (person) BB extraction in thermal camera image using the YOLO DL-network is shown in Fig. 2a, while Fig. 2b shows the result of object (person) BB extraction in RGB camera image using the CenterNet DL-network. The DisNet estimated distances from the cameras to the detected objects (persons) are given in Table 1. As obvious, the BB extracted for the person to the left in RGB image is of good quality as it correctly bounds the object (person) and consequently the DisNet-based object distance is estimated with high accuracy. However, the BB extracted for the person to the right in RGB image is of poor quality as it is of much larger size than the object (person) itself. The reason for such poor object detection result is an RGB image of low contrast due to the large zooming factor used with the goal of long-distance object detection. As a consequence of incorrect object BB extraction, the DisNet object distance estimation is inaccurate with a large estimation error of 41,39%. The same problem appears in the DisNet-based estimation of the object (person) detected in thermal camera image (Fig. 2a). Namely, the shown thermal camera

image is of a low contrast due to lower thermal camera performances influenced with high outside temperature of about 38°C during the field test performed in operational railway environment [11]. Due to such high environmental temperature, the temperature difference between the object (person) and environment was small so that object region is of low-contrast. As a consequence, the BB extracted by DL-based object detector is of poor quality as it is much larger than the object itself. As a consequence of this, DisNet using wrongly sized object BB estimated object distance as smaller than it really was.

(a) (b)

Fig. 2 DL-based object detection in railway scenes captured by thermal (a) and RGB camera (b).

Table 1 DisNet estimated object distance based on object BBs in images shown in Fig. 2

Object	Ground Truth	DisNet	Error
Person in thermal image	146 m	122,94 m	15,79%
Person to the left in RGB image	89,99 m	90.3 m	0,34%
Person to the right in RGB image	162 m	94,95 m	41,39%

In order to overcome the problem of incorrect object distance estimation because of the inaccurate object BB size, in this paper, a modification of DisNet-based object distance estimation system is proposed as illustrated in Fig. 3. The object detector part of the system is extended by a post-processing of an object bounding box extracted by a DL-based object detector.

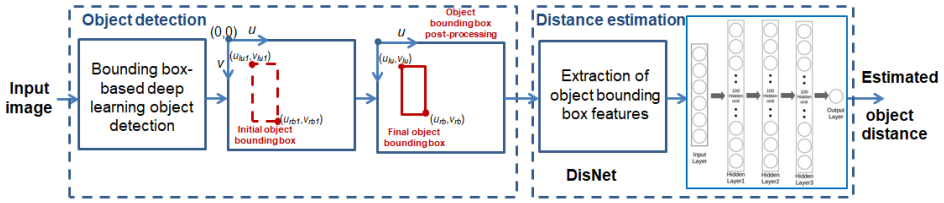

Fig. 3 Modified DisNet-based object distance estimation system with included object BB post-processing step

3. OBJECT BOUNDING BOX POST-PROCESSING: CLOSED-LOOP CONTROL OF EDGE-BASED SEGMENTATION

The main idea of the object detector extension by a BB post-processing step (Fig. 3) is to segment the original image region bounded with extracted BB so to check whether the bounded object touches the BB sides, that is to check whether BB properly bounds the detected object. It is proposed that traditional computer vision-based segmentation is applied to BB extracted by a DL-based object detector. More precisely, Canny edge detection [12], which is known as optimal traditional edge detector, is applied for BB edge-based segmentation. However, even though Canny edge detector is theoretically an optimal detector there is no guarantee that default Canny edge detector parameters, which are optimal for one application will give good results for another application. If the default optimal Canny edge detector parameters are applied to the object BB extracted from the low-cost thermal image (Fig. 4a) the segmentation result is poor, as shown in Fig. 4(d). Namely the result of edge-based segmentation is an over-segmented BB indicated that the segmentation (thresholding) parameter of Canny edge detector was "too low". In order to avoid "over-segmentation" so to avoid segmenting of not-object's edges, the thresholding parameter should be higher. The result of edge-based segmentation with a higher threshold is shown in Fig. 4b. As obvious, this edge segmentation result is poor as the resulting segmented BB image is an under-segmented indicating that used segmentation parameter was "too high". Neither an under- nor an over-segmented image is a good input for BB refinement as they have respectively too little that is too many edges making segmentation of object inside BB impossible.

(a) (b) (c) (d)

Fig. 4 Original object BB image (a). Under-segmented BB image (b). Optimally segmented BB image (c). Over-segmented BB image (d).

Evidently, it is crucial to define a segmentation (thresholding) parameter of the Canny edge detector which will yield optimal segmentation with as many edges of the object as possible and with as little noise edges (not-object edges) as possible as it is the case of the segmented BB image in Fig. 4(c). Such an optimal edge-based segmentation enables accurate segmentation of the object inside the original BB (extracted by DL-based object detector) so that BB can be refined to bounds the object exactly. The suggested cascade control structure, shown in Fig. 5, aims to calculate optimal segmentation parameter of Canny edge detector automatically.

The proposed method for automatic calculation of segmentation parameter follows the idea of implementation of closed-loop control structures for improvement of digital image processing [13]. The control actions in both control loops of the cascade structure in Fig. 5 are realized through an optimization procedure. The control objective of the inner closed-loop is to provide a binary segmented BB image that is free of not-object

edges (noise). It is accomplished by changing the segmentation (thresholding) parameter of the Canny edge detector so that the measured variable, two-dimensional entropy *S* of edge segmented pixels, is minimized. The control goal of outer closed-loop is to detect as many true segmented object's edges as possible. This is achieved by changing the parameter of a pre-processing contrast enhancement operation so that the contrast of input BB image is improved, allowing segmentation of more object edges. In this way, the second output variable, the number of segmented edge pixels, is maximized.

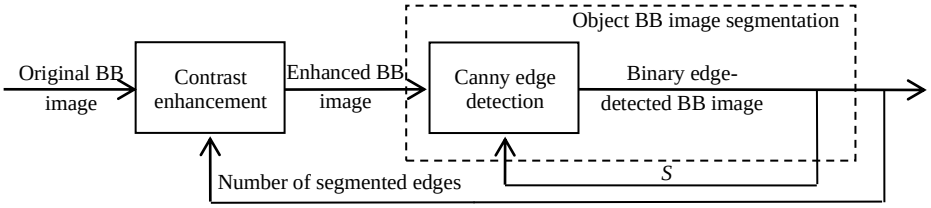

Fig. 5 Object BB post-processing

The measured variable of the inner closed-loop of above cascade control structure for optimal edge-based segmentation of object BB is a measure of quality of edge segmentation, so-called two-dimensional entropy *S* of edge segmented pixels. An edge-based segmented binary image is said to be of good quality if all pixels of object boundary are segmented. In other words, it is desired to have "unbroken" object boundaries (edges) in the segmented image. Also, it is desired that there are no pixels segmented as belonging to the object edges even though they actually do not belong to. In other words, it is desired that the segmented image is free of undesired edges-noise. Bearing in mind these requirements for the edge segmented image quality, it turns out that the measure of the connectivity of segmented pixels may be used as a suitable quality measure. A connectivity measure based on two-dimensional (2D) entropy of segmented pixels in a binary segmented image is introduced in [14]. It is defined as:

$$S = -\sum_{i=0}^{8} p_{(0,i)} \log_2 p_{(0,i)} \tag{5}$$

where $p_{(0,i)}$ is the relative frequency, that is, the estimate of the probability of occurrence of a pair $(0,i)$ representing the segmented pixel 0 surrounded with i segmented pixels in its 8-pixel neighborhood:

$$p_{(0,i)} = \frac{\text{number of segmented pixels surrounded with } i \text{ segmented pixels}}{\text{number of segmented pixels in the image}} \tag{6}$$

It can be demonstrated that the suggested entropy *S* has small values in case of well-connected segmented pixels and high value in case of scattered segmented pixels (broken edges and noisy edge image). For the purpose of this demonstration, synthetic segmented images shown in Fig. 6 are considered where black pixels represent segmented pixels.

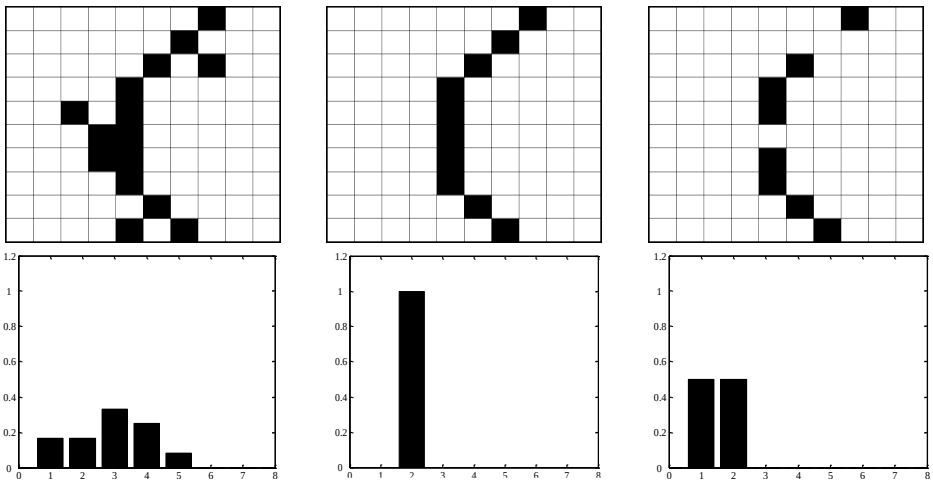

Fig. 6 Binary images of noisy (top left), ideal (top middle) and broken (top right) segmented edge with corresponding normalized histograms (respectively bottom left, bottom middle and bottom right) of distribution of pairs $(0,i)$.

Fig. 6 (top middle) shows an ideally segmented one pixel thick edge and the corresponding histogram is shown in Fig. 6 (bottom middle). This image, consisting of ideally connected pixels forming the segmented object edge, is referred to as being in a so-called well-ordered state. In real-word applications the well-ordered state is difficult (perhaps impossible) to achieve. Due to external disturbances which may include the image acquisition conditions, the camera characteristics and the imaged objects characteristics, the segmented object edges may appear "broken" as a consequence of insufficient contrast in the image to be segmented. Another deviation from the well-ordered state happens when pixels that do not belong to the edges are segmented as object pixels. A common reason for this is that other objects in the neighborhood of the object to be detected have the same characteristic (e.g. color) as the imaged object. These "undesirable" pixels in segmented image appear as noise. The noisy and broken detected edge, which can be referred to as being in a disordered state, together with the corresponding histograms of distributions of pairs $(0,i)$ are shown in Fig. 6 (top left), Fig. 6 (top right), Fig. 6 (bottom left) and Fig. 6 (bottom right) respectively.

Evidently, the histogram of connected pixels which represents an ideally segmented edge is narrower than the histograms of noisy and broken object edges. This is expected, as the number of different pairs $(0,i)$ in the case of an edge detected image of good quality is smaller than the number of different pairs $(0,i)$ in noisy and broken detected object edges. However, the relative frequency of a pair $(0,i)$ found in images of good quality is larger than the relative frequency of pair $(0,i)$ found in images of bad quality. Hence, the histogram of distribution of pairs $(0,i)$ in a segmented image of good quality can be considered as being obtained by shrinking the histograms of segmented images of bad quality. Let l and k be two bins of normalized histograms of distribution of pairs $(0,i)$. Since the estimate of the probability of occurrence of a pair $(0,i)$, $p(0,i)$, is positive and log function increases monotonically, the following holds:

$$\log_2(l+k) > \log_2 l \quad \text{and} \quad \log_2(l+k) > \log_2 k \tag{7}$$

$$(l+k)\log_2(l+k) = l\log_2(l+k) + k\log_2(l+k) > l\log_2 l + k\log_2 k \tag{8}$$

$$-(l+k)\log_2(l+k) < -l\log_2 l - k\log_2 k \tag{9}$$

It is evident from (9) that the sum of individual contributions of two histogram bins to the entropy is larger than the contribution of same bins when they are summed. Hence, the entropy of narrow histograms, like the histogram of edge segmented image of good quality, is smaller than the entropy of the stretched one, which is characteristic of edge segmented images of bad quality.

The results of 2D entropies *S* for considered images in Fig. 6, obtained while considering only segmented (black) pixels having full 8-neighborhood and so neglecting the image bounding pixels, which are given in Table 2, confirm the previous statement. This provides a basis for the use of the 2D entropy *S* as the measure of the connectivity of pixels, forming an object edge in a segmented image. In other words, obtained result provides a basis for the use of a 2D entropy *S* as the measure of the quality of an edge segmented image. The above discussion also demonstrates the possibility for using the two-dimensional entropy of segmented pixels *S* as a controlled variable for the closed-loop control of edge-based segmentation. The proposed inner closed-loop of cascade control structure for optimal segmentation of object BB image (Fig. 5) uses the two-dimensional entropy *S* as controlled variable and minimizes its value in an optimization process.

Table 2 2D entropy *S* of the pixels belonging to the object segmented edge

	Noisy	Ideal	Broken
Detected edge	2.1887	0	1

4. EXPERIMENTAL RESULTS

The presented method was evaluated on the images from railway scenes shown in Fig. 2. More precisely the presented method for closed-loop edge segmentation was applied to the "poor" BBs detected in thermal image and the RGB image shown in Fig. 2a and Fig. 2b respectively. The result of this post-processing of "poor" BBs were optimally segmented objects' edges as shown in Fig. 7b and Fig. 7e respectively. The resulted edge segmented BBs were scanned row by row starting from the left upper corner until the first segmented pixel was reached. The empty areas (areas without connected segmented pixels) were discarded so that the size of the BB was refined to bound the segmented object edge region properly. The initial object BBs and refined BBs are shown respectively in Fig. 7a and Fig. 7c, that is in Fig. 7d and Fig7f.

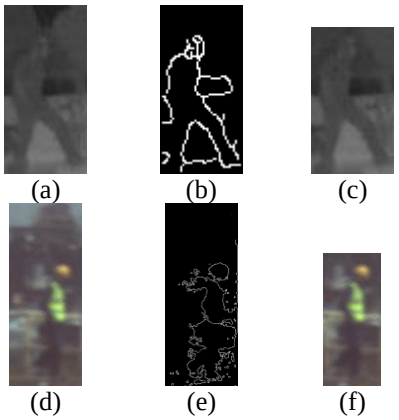

Fig. 7 Original objects' BBs (a) and (d). Optimally segmented BBs (b) and (e). Refined objects' BBs (c) and (f)

The DisNet estimated object distances using the refined BBs are given in Table 3. As evident from the comparison of original results given in Table 1, modification of the DisNet-based system by inclusion of post-processing of original object BB leads to significant improvement in object distance estimation.

Table 3 DisNet estimated object distance based on refined object BBs for the objects in images shown in Fig. 2

Object	Ground Truth	Modified DisNet	Error
Person in thermal image	146 m	148,52 m	1,72%
Person to the right in RGB image	162 m	175,87 m	8,56%

5. CONCLUSION

In this paper, a modification of the system for obstacle detection and distance estimation in railways, which was developed within the project SMART-Smart Automation of Rail Transport, is presented. The system consists of two main parts; the first part is Convolutional Neural Network (CNN)-based object detector, which outputs the bounding box of the detected object, and the second part is multi hidden-layer neural network-based distance estimator named DisNet. The accuracy of the DisNet estimated object distance strongly depends on the accuracy of the object BB extracted by a CNN-based object detector. In order to improve the accuracy of object BB and consequently to improve the accuracy of distance estimation, the extension of object detector by a post-processing of extracted object BB is proposed. The BB post-processing involves traditional computer vision-based edge segmentation. More precisely, the proposed method is based on the Canny edge segmentation with automatic calculation of Canny segmentation parameters by a closed-loop control of edge segmentation process. The proposed method is evaluated on the real-world images of railway scenes with obstacles on the rail tracks captured by thermal and RGB cameras. The evaluation results demonstrate the fact that traditional computer vision is beneficial in complementing state-of-the-art DL-based methods for object detection in real-world applications where input images are of not-good quality due to different external influences.

Acknowledgement: *This research has received funding from the Shift2Rail Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreements No 730836 and No. 881784.*

REFERENCES

[1] J. Weichselbaum, C. Zinner, O. Gebauer, W. Pree, "Accurate 3D-vision-based obstacle detection for an autonomous train", Computers in Industry, vol. 64, no. 9, pp. 1209–1220, 2013. https://doi.org/10.1016/j.compind.2013.03.015.

[2] Y. Xu, C. Gao, L. Yuan, S. Tang, G. Wei, "Real-time Obstacle Detection Over Rails Using Deep Convolutional Neural Network", in *Proceedings of 2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, Auckland, NZ, pp. 1007-1012, 2019. https://doi.org/10.1109/ITSC.2019.8917091.

[3] C. Herrmanna, M. Rufa, J. Beyerer, "CNN-based thermal infrared person detection by domain adaptation", in *Proceedings Volume 10643, Autonomous Systems: Sensors, Vehicles, Security, and the Internet of Everything*, 2018. https://doi.org/10.1117/12.2304400.

[4] Y. Jiang, H. Shin, H. Ko, "Precise Regression for Bounding Box Correction for Improved Tracking Based on Deep Reinforcement Learning", in *Proceedings of 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, AB, pp. 1643-1647, 2018. https://doi.org/10.1109/ICASSP.2018.8462063.

[5] J. Liu, S. He, "6D Object Pose Estimation Based on 2D Bounding Box, arXiv:1901.09366v1", [cs.CV] 27 Jan 2019, [Online]. Available: https://arxiv.org/abs/1901.09366.

[6] M. A. Haseeb, J. Guan, D. Ristić-Durrant, A. Gräser., "DisNet: A novel method for distance estimation from monocular camera", in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems - IROS 2018, 10th Workshop on Planning, Perception and Navigation for Intelligent Vehicles (PPNIV)*. [Online]. Available: https://project.inria.fr/ppniv18/program/.

[7] Shift2Rail project SMART-Smart Automation of Rail Transport. [Online]. Available: http://www.smartrail-automation-project.net/.

[8] J. Redmon, A. Farhadi, "YOLOv3: An Incremental Improvement", arXiv:1804.02767v1 [cs.CV], 2018. [Online]. Available: https://arxiv.org/abs/1804.02767v1.

[9] X. Zhou, D, Wang, P. Krähenbühl, "Objects as points", arXiv preprint arXiv:1904.07850, 2019. [Online]. Available: https://arxiv.org/abs/1904.07850v2.

[10] COCO dataset, [Online]. Available: https://arxiv.org/pdf/1405.0312.pdf.

[11] M. A. Haseeb, D. Ristić-Durrant, A. Gräser, M. Banić, D. Stamenković, *Multi-DisNet: Machine learning-based object distance estimation from multiple cameras*, D. Tzovaras et al. (Eds.): ICVS 2019, LNCS 11754, pp. 457–469, 2019, Springer Nature Switzerland AG 2019 https://doi.org/10.1007/978-3-030-34995-0_41.

[12] J. A. Canny, "Computational Approach To Edge Detection", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 8, no. 6, pp. 679–698, 1986. https://doi.org/10.1109/TPAMI.1986.4767851.

[13] D. Ristić, Feedback Structures in Image Processing. PhD Thesis, University of Bremen, 2007. ISBN 978-3-8322-6598-4.

[14] D. Ristić and A. Gräser, "Performance measure as feedback variable in image processing", EURASIP Journal on Applied Signal Processing, vol. 2006, https://doi.org/10.1155/ASP/2006/27848.