https://doi.org/10.2352/lim.2024.5.1.3

The Influence of Read Noise on an Automatic License Plate Recognition System

Nikola Plavac, Seyed Ali Amirshahi, Marius Pedersen, Sophie Triantaphillidou; Norwegian University of Science and Technology (NTNU), Gjøvik, Norway

Abstract

This study aims to investigate how a specific type of distortion in imaging pipelines, such as read noise, affects the performance of an automatic license plate recognition algorithm. We first evaluated a pretrained three-stage license plate recognition algorithm using undistorted license plate images. Subsequently, we applied 15 different levels of read noise using a well-known imaging pipeline simulation tool and assessed the recognition performance on the distorted images. Our analysis reveals that recognition accuracy decreases as read noise becomes more prevalent in the imaging pipeline. However, we observed that, contrary to expectations, a small amount of noise can increase vehicle detection accuracy, particularly in the case of the YOLO-based vehicle detection module. The part of the automatic license plate recognition system that is mostly prone to errors and is mostly affected by read noise is the optical character recognition module. The results highlight the importance of considering imaging pipeline distortions when designing and deploying automatic license plate recognition systems.

Introduction

For more than four decades, Automatic License Plate Recognition (ALPR) systems have played a crucial role in various applications, including toll collection, traffic management, and law enforcement. The first prototype of an ALPR system was introduced in 1979 by the UK's Police Scientific Development Branch [1]. Traditional ALPR systems relied on classical image processing methods, which incorporate different features, such as global image information, texture, color, character features, or combinations of these [2]. However, recent advancements in deep learning have revolutionized ALPR systems, with many now employing machine learning [1], neural networks [3], and different deep learning techniques [4, 5].

Typically, ALPR systems consist of three main components: a license plate detection module, a character segmentation module, and a character recognition module [6]. The license plate detection module identifies the region in an image containing the license plate, while the character segmentation and recognition modules process and interpret the characters on the plate. Despite all the advancements, challenges persist, particularly in scenarios involving adverse weather conditions, variable illumination, and distortions within the imaging pipeline [7, 8]. Understanding the impact of such distortions on ALPR performance is essential for developing robust and reliable systems.

Read noise is a type of electronic noise generated during the process of analog-to-digital signal conversion. When the camera's sensor reads the electrical charge generated by photons impinging on the sensor's pixels and converts that charge into digital data. This type of noise is independent of the signal level and can be modeled using a Gaussian distribution [9].

To investigate the influence of specific camera distortions, such as read noise, on ALPR performance, imaging pipeline simulations serve as invaluable tools. These simulations enable researchers and engineers to observe the effects of various design decisions and components on system performance without the need for costly real-world experimentation. Design considerations for digital imaging systems have been addressed since as early as 1977 [10]. The Image Systems Evaluation Toolkit (ISET) [11], introduced by Farrell et al. in 2003, facilitates a comprehensive simulation of the imaging pipeline, allowing for a detailed analysis of system behavior. At about the same time, Kerekes et al. [12] proposed a full-spectrum spectral imaging system analytical model, which was another end-to-end system model, but designed for applications in remote sensing. While limiting his study to the optical part of the pipeline and color processing, Florin [13] proposed the simulation of a digital camera pipeline.

In this study, we utilize the ISET imaging pipeline simulation to explore how read noise affects the accuracy of a pretrained three-stage license plate recognition algorithm. By systematically introducing varying levels of read noise and evaluating recognition performance, we aim to provide insights into the robustness of an ALPR system in the presence of this specific distortion.

Database and Model

The UFPR-ALPR dataset, adapted for this study, consists of images captured by different cameras mounted inside a car in Paraná, Brazil [14]. The car tracked 150 different vehicles for 30 frames each, resulting in a total of 4,500 images. Among tracked vehicles, 30 are motorcycles with double-line license plates, while the remaining 120 vehicles include cars, vans, buses, or trucks with single-line license plates. To align with the capabilities of our selected license plate recognition algorithm and streamline the baseline dataset, motorcycle images were excluded and only the first frame of each remaining vehicle was retained, resulting in a baseline dataset of 120 images (Figure 1). All images in the dataset are captured during daytime, with no weather-related distortions such as snow, rain, fog, or similar.

Figure 1. *Examples of the images from UFPR-ALPR dataset [14].*

The three-stage license plate detection and recognition algorithm proposed by Silva et al. [15] serves as the foundation for our study. From the results presented in [15], it can be observed that the proposed ALPR system performs similarly to other state-of-the-art systems under normal conditions and performs better than them under challenging capture conditions, such as distortions that can occur due to oblique camera views. The algorithm comprises the following stages: vehicle detection module, license plate detection and rectification module, and optical character recognition (OCR) module. Due to its very good performance in terms of speed and accuracy, the YOLOv2 [16] network is employed for vehicle detection, without architectural modifications. The authors proposed a warped planar object detection network (WPOD-NET) for the license plate detection and rectification module. This network searches for license plates in the image and regresses one affine transformation (six unknown parameters) per license plate detection. The predicted affine transformation is used to perform the rectification of the license plate, resembling the frontal camera view. The optical character recognition module, based on a modified YOLO [17] network, handles character segmentation and recognition on the rectified license plate. The pretrained algorithm from [15] is accessed and utilized without any alterations.

Imaging pipeline distortions

In this study, we utilized the ISET toolbox developed by [11] to simulate the imaging pipeline. The pipeline consists of four distinct stages: scene, optics, sensor, and image processing.

The scene stage ideally requires a radiometric description of the illumination in the scene, often represented as a hyperspectral cube. However, the complexity of the scene representation can be reduced by using only an RGB image. In this case, a pseudo-spectral representation of the scene is created using the spectral primaries of a standard display (e.g. sRGB), and the display's white point is assumed to be the display illuminant (e.g. D65).

The optics component of the pipeline converts the scene radiance data into irradiance data, representing the sum of rays that impinge on a sensor surface ($photons/s/nm/m^2$). Optics components essentially gather the rays from each point of the scene and focus them onto the sensor. The modeled optics components can diverge in complexity, from the simplest case where an ideal lens with a circular aperture is used to rather complex lens configurations and lens distortion models. Since the optical configuration was not within the scope of this study, we used an ideal lens with a user-specified focal length (7×10^{-3} mm) and an F number (0.1), assuming a camera-to-scene distance of 3 m.

The sensor stage simulates the transformation of the irradiance on the sensor surface into an electrical signal. For each spatial pixel of the imaging sensor, a voltage value will be the final output of this component. The key configurations included the use of a Bayer RGB sensor, a voltage swing of 1.15 V, a pixel conversion gain of $1.28 * 10^{-4} \frac{V}{electrons}$, a pixel fill factor of 45%, a pixel size of $2\mu m$ and a pixel dark voltage of $1 * 10^{-5} \frac{V}{s}$. Pixel read noise was systematically varied from 0.2 to 14.2 mV in increments of 1 mV (Figure 2).

The image processing module covers the conversion from the digitized voltage values generated by the two-dimensional sensor array to a three-dimensional (RGB) image that can be rendered on a specific display. This is accomplished through four essential steps: interpolating missing RGB sensor values (demosaicing) and transforming sensor RGB values into an internal color space for encoding and display (color-balancing, color-rendering and optionally color conversion) [11]. The adaptive Laplacian demosaicing algorithm was employed to minimize demosaicing artifacts.

Figure 2. *Example of the distorted images created from image A. Number next to D indicates the distortion level. Every other level of the distortion is shown.*

In total, 1800 distorted images were generated, representing 15 different levels of read noise, to evaluate the performance of the ALPR system [15] in contrast to the one obtained for 120 undistorted images. This experimental design aimed to explore the impact of read noise on ALPR performance and assess the network's sensitivity to sensor-type distortions in the imaging pipeline.

It is notable that creating a pseudo-spectral representation from RGB images is not an ideal solution. However, to the best of our knowledge, no existing license plate recognition dataset containing spectral scene representation is available. Additionally, one may observe that the read noise is added on top of already existing noise in images. This is certainly another limitation of using existing datasets, as it is not possible to characterize the pre-existing noise in images.

Results

In this section, we analyze the performance of the ALPR system [15] on undistorted and distorted images, assessing the success of vehicle detection, license plate detection and OCR components, as well as the success of the entire system. The recognition success of the system is defined as the correct recognition of all characters on the license plate. The same criteria are used for the OCR module of the system, while in case of the vehicle detection and license plate detection modules, the success of the modules is defined as correctly setting the bounding box around the vehicle and the license plate, respectively.

Upon evaluation of the ALPR system [15] on undistorted images, an initial accuracy rate of 44.17% was observed. Investigation into the causes of such low accuracy rate revealed frequent confusion between certain letters and numbers, such as 'I' and '1', 'O' and '0', and 'S' and '5'. Inspired by the work

Table 1: Results of each component of the ALPR system [15] before and after introducing the HC rules on the undistorted images.

	Fails /120	Acc. (%)	Fails HC /120	Acc. HC (%)
YOLO VD	9	92.50	9	92.50
WPOD-NET	10	90.83	10	90.83
OCR module	48	60.00	16	86.67
System	67	44.17	35	70.83

Table 2: Results of each component of the ALPR system [15] on distorted images, indicating changes in accuracy rates compared to the baseline configuration (Table 1).

	Fails /1800	Acc. (%)	Change (%)
YOLO VD	121	93.28	0.78
WPOD-NET	243	86.50	-4.33
OCR module	373	79.28	-7.39
System	737	59.06	-11.77

of Montazzolli and Jung [18], we use heuristic correction rules based on the standard pattern of Brazilian license plates. Namely, Brazilian license plates consist of three letters followed by four digits. Each output label is then divided into two parts, and the following Heuristic Correction (HC) rules were applied:

- First three positions: $0 \rightarrow$ O, $4 \rightarrow$ A, $8 \rightarrow$ B, $2 \rightarrow$ Z, $5 \rightarrow$ S, $1 \rightarrow$ I;
- From fourth position to the end: I $\rightarrow 1$, Q $\rightarrow 0$, G $\rightarrow 6$.

After applying these rules, the performance of the system significantly improved to 70.83%. The performance of each component of the ALPR system [15] was assessed individually, including the YOLOv2-based vehicle detection network (YOLO VD), the license plate detection and rectification network (WPOD-NET), and the optical character recognition module (OCR module). Results (Table 1) show an increase in the accuracy of the ALPR system [15] after using the HC rules. The vehicle detection and license plate recognition modules exhibited very high performance on the undistorted images, with accuracy rates exceeding 90%. However, the optical character recognition module, while showing a significant improvement after the introduction of heuristic correction rules, remained the most error-prone component, with a baseline accuracy of 86.67%. Notably, failures in the vehicle detection and license plate detection modules were primarily observed for non-standard vehicles such as buses, trucks, and vans.

The analysis was further extended to distorted images, with a focus on the effect of read noise distortion on recognition accuracy. It is important to note that all further analyses will incorporate the HC rules, since it was shown that applying this approach significantly increases the performance of the system. Results (Table 2) show a change in the accuracy rate of each component of the ALPR system [15] when applied to the distorted images compared to when the undistorted images were used. The vehicle detection network slightly improved its accuracy when read noise distortion was applied, while the license plate recognition network and the OCR module exhibited a decrease in performance. Overall system accuracy decreased from 70.83% to 59.06% if all levels of distortion are observed jointly.

In our experiments, the undistorted images correspond to distortion level zero, and the accuracy on level zero corresponds to the accuracy of the system when the undistorted images were used. The remaining 15 distortion levels correspond to read noise introduced at the sensor level in the imaging pipeline simulations, ranging from 0.2 to 14.2 mv with 1 mV increment. The results show a decrease in the accuracy of the ALPR system [15] as the applied read noise level increases (Figure 3). Compared to the baseline accuracy of 70.83%, the accuracy of the system decreased to 55% at the distortion level 15 which corresponds to the read noise of 14.2 mV, although the decrease is not perfectly monotonic. Most often, the initial correct recognition of the license plate failed when the very first level of read noise (0.2 mV) was added (Figure 4). For higher levels of read noise applied, recognition failed for significantly fewer samples. This finding highlights the sensitivity of the system to low levels of read noise.

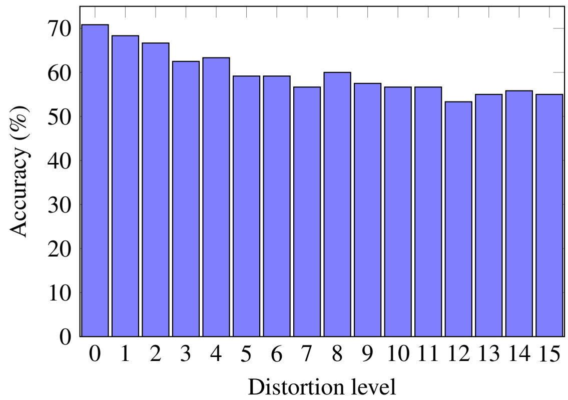

Figure 3. *Accuracy of the ALPR system [15] for each distortion level. Level zero represents the accuracy obtained for undistorted images.*

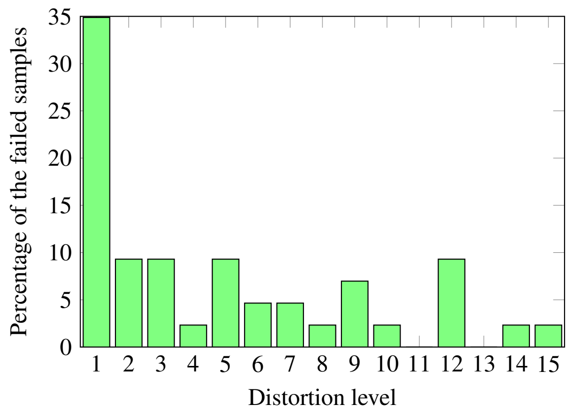

Figure 4. *The graph shows the distortion level in which the recognition failed for the first time. Percentages are computed based on the total number of images for which recognition failed at some of the distortion levels.*

Further analysis explored the persistence of recognition failures at different distortion levels (Figure 5). Ideally, the system should change it's prediction (for example recognition fails) at a certain level of distortion, and retain the same behavior for next levels of the distortion. In 28% percent of the cases where the system changed its initial prediction, this change occurred only once, i.e. the system continued to behave in the same way with adding further distortions. However, the system changed its prediction two, three, or even five times for a significant percentage of the cases ($>$ 10%), which means that if the system failed to recognize a license plate after adding a certain distortion level, it will not necessarily fail for the next distortion levels added.

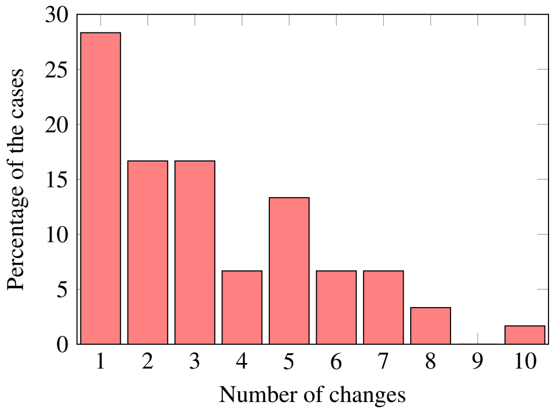

Figure 5. *Frequency of the systems recognition change.*

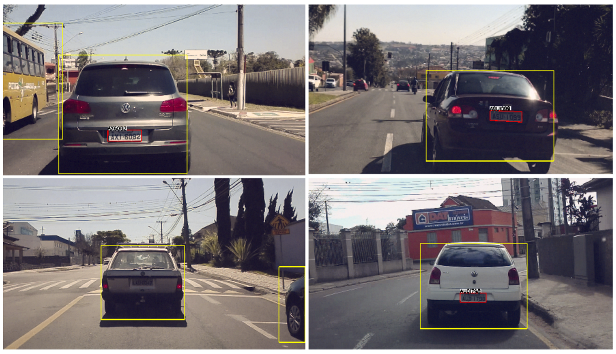

Figure 6. *Examples of vehicles for which license plate recognition failed after the lowest level of read noise was added.*

Discussion

Reflecting on the results presented, it becomes evident that the introduction of read noise has a notable impact on the recognition accuracy in the ALPR system [15]. Particularly noteworthy is the observation that even a low level of read noise, as minimal as 0.2 mV, leads to the first recognition failure in most cases. Upon closer examination of examples such as those depicted in Figure 6, it is apparent that recognition failures often occur when license plates are partially or completely obscured by shadows. This phenomenon warrants further investigation, perhaps through the exploration of differences in contrast between successful recognition cases and those affected by shadowing. Such insights could inform the development of more robust algorithms capable of handling challenging lighting conditions or adding the contrast enhancement step before the OCR module.

In contrast, intriguing observations emerge from cases where the introduction of read noise results in successful recognition, as evidenced by the examples in Figure 7. Notably, the YOLOv2-based vehicle detection module appears to benefit from small amounts of read noise, effectively detecting vehicles that were previously undetected. Although the existing literature mentions the potential benefits of noise or blur in enhancing object detection networks [19, 20], this phenomenon's application to license plate detection and recognition systems remains unexplored. Further investigation of the underlying mechanisms driving this phenomenon could offer valuable insights for optimizing the performance of the ALPR system.

Despite the improvements achieved through the HC rules, it is evident that the OCR module remains the least accurate and robust component of the ALPR system. While the HC rules significantly enhance performance, they are not universal solutions and will not generalize well to license plates from different regions. However, it is important to note that there are inherent visual similarities between characters '1' and 'I', as well as '0' and 'O' on license plate images in the dataset.

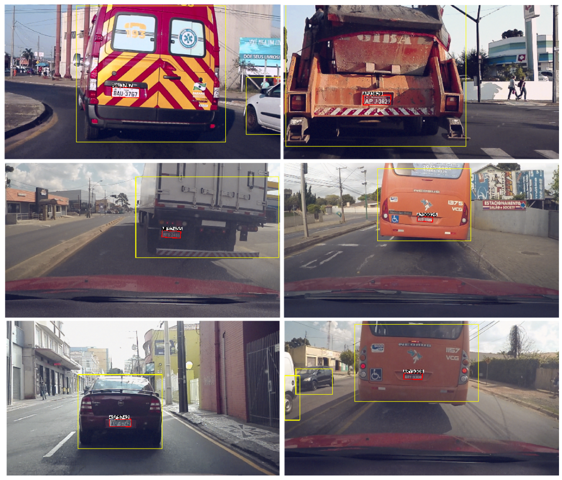

Figure 7. *Examples of vehicles for which recognition became successful after introducing read noise.*

In designing the experiment, our aim was to replicate real-world conditions by simulating read noise levels. Although we were unable to find reliable sources for typical read noise levels in the literature, online sources suggest upper limits of 3-5 mV. The upper limits in our experiment significantly exceeded these limits, but the simulation results (Figure 4) indicate that relatively low levels of read noise (0.2 to 4.2 mV) significantly affect the recognition accuracy.

Conclusion

In this study, we investigated the impact of read noise on the performance of an ALPR system [15]. Through comprehensive analysis and experimentation, we uncovered valuable insights into the behavior of ALPR systems under varying levels of distortion. Our findings highlight several key observations. The introduction of read noise distortion has a discernible effect on recognition accuracy, with even minimal levels of noise leading to a notable decrease in performance. This underscores the importance of considering imaging pipeline distortions when designing and deploying ALPR systems. Furthermore, our study revealed intriguing phenomena, such as noise-induced recognition, where small amounts of read noise can unexpectedly improve the performance of certain components within the ALPR system, such as the vehicle detection module. Further exploration of these phenomena could yield novel approaches to improving the robustness and efficiency of ALPR systems. We identified the OCR module as a critical component with significant room for improvement. Although using heuristic correction rules provided performance enhancements, they will not generalize well to diverse license plate formats, highlighting the need for more sophisticated recognition algorithms capable of handling various character variations. Overall, our study provides valuable insights into the effects of read noise distortion and highlights areas for further research and development. For example, enhancing contrast levels in the license plate area could benefit ALPR, as well as investigating the causes of the increased YOLOv2-based vehicle detection module when a small level of read noise is added. By addressing the challenges identified in the study and leveraging emerging technologies, we can work towards building more robust and reliable ALPR systems capable of meeting the demands of diverse real-world applications.

Acknowledgments

Seyed Ali Amirshahi is supported by the project "VQ4MedicS: Video Quality Assessment and Enhancement for Pre-Hospital Medical Services" (grant number 329034) from the Research Council of Norway. Marius Pedersen and Sophie Triantaphillidou are supported by the Research Council of Norway through the "Quality and Content" project (grant number 324663).

References

[1] R. K. Prajapati, Y. Bhardwaj, R. K. Jain and K. K. Hiran, A Review Paper on Automatic Number Plate Recognition Using Machine Learning: An In-Depth Analysis of Machine Learning Techniques in Automatic Number Plate Recognition: Opportunities and Limitations, 2023 Int. Conf. on Computational Intelligence, Communication Technology and Networking (CICTN), pp. 527-532 (2023).

[2] S. Du, M. Ibrahim, M. Shehata and W. Badawy, Automatic License Plate Recognition (ALPR): A State-of-the-Art Review, IEEE Transactions on Circuits and Systems for Video Technology, vol. 23, no. 2, pp. 311-325 (2013).

[3] M. M. Khan, M. U. Ilyas, I. R. Khan, S. M. Alshomrani and S. Rahardja, Lincese Plate Recognition Methods Employing Neural Networks, IEEE Access, vol. 11, pp. 73613-73646 (2023).

[4] C. Lin and Y. Li, A License Plate Recognition System for Severe Tilt Angles Using Mask R-CNN, IEEE 2019 International Conference on Advanced Mechatronic Systems (ICAMechS), Japan, pp. 229-234 (2019).

[5] H. Shi and D. Zhao, License Plate Recognition System Based on Improved YOLOv5 and GRU, IEEE Access, vol. 10, pp. 10429-10439 (2023).

[6] I. Gorovyi, Efficient Two-Step Approach for Automatic Number Plate Detection, International Journal of Electronics and Telecommunications, vol. 61, no. 4, pp. 351-356 (2015).

[7] H. Padmasiri, J. Shashirangana, D. Meedeniya, O. Rana and C. Perera, Automated License Plate Recognition for Resource-Constrained Environments, Sensors, vol. 22, no. 1434 (2022).

[8] H. Li, P. Wang and C. Shen, Toward End-to-End Car License Plate Detection and Recognition With Deep Neural Networks, IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 3, pp. 1126-1136 (2019).

[9] K. Wei, Y. Fu, Y. Zheng and J. Yang, Physics-Based Noise Modeling for Extreme Low-Light Photography, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 11, pp. 8520-8537 (2023).

[10] J. M. Booth and J. B. Schroeder, Design Considerations for Digital Image Processing Systems, *Computer*, vol. 10, no. 8, pp. 15-20 (1977).

[11] J. E. Farrell, F. Xiao, P. B. Catrysse, B. and A. Wandell, A simulation tool for evaluating digital camera image quality, Proc. SPIE 5294, Image Quality and System Performance, vol. 5294, (2003).

[12] J. P. Kerekes and J. E. Baum, Full-spectrum spectral imaging system analytical model, IEEE Trans. on Geoscience and Remote Sensing, vol. 43, no. 3, pp. 571-580 (2005).

[13] T. Florin, Simulation of a Digital Camera Pipeline, 2007 International Symposium on Signals, Circuits and Systems, (2007), doi: https://doi.org/10.1109/ISSCS.2007.4292789

[14] R. Laroca, E. Severo, L. A. Zanlorensi, L. S. Oliveira, G. R. Gonçalves, W. R. Schwartz and D. Menotti, A Robust Real-Time Automatic License Plate Recognition Based on the YOLO Detector, in 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1-10 (2018).

[15] S.M. Silva and C. R. Jung, License Plate Detection and Recognition in Unconstrained Scenarios, 2018 European Conference on Computer Vision (ECCV), pp. 580-596 (2018).

[16] J. Redmon and A. Farhadi, YOLO9000: Better, faster, stronger, in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6517-6525 (2017).

[17] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, You only look once: Unified, real-time object detection, in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788 (2016).

[18] S. Montazzolli and C. Jung, Real-Time Brazilian License Plate Detection and Recognition Using Deep Convolutional Neural Networks, 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Niteroi, Brazil, pp. 55-62 (2017).

[19] J. A. Rodríguez-Rodríguez, E. López-Rubio, J. A. Ángel-Ruiz and M. A. Molina-Cabello, The Impact of Noise and Brightness on Object Detection Methods, Sensors, vol. 24, no. 821 (2024).

[20] I. Vasiljevic, A. Chakrabarti and G. Shakhnarovich, Examining the Impact of Blur on Recognition by Convolutional Networks, in ArXiv, https://arxiv.org/abs/1611.05760 (2016).

Author Biography

Nikola Plavac received his BSc in Biomedical Engineering from the University of Novi Sad, Serbia (2022). He is currently completing his MSc in Erasmus+ Master Program Computational Colour and Spectral Imaging. Additionally, Nikola is an integrated PhD candidate at the Department of Computer Science, Norwegian University of Science and Technology (NTNU).

Seyed Ali Amirshahi received his BSc in 2008 from Amirkabir University of Technology (Tehran Polytechnic) in Electrical Engineering. In 2010 he graduated from the Master Erasmus Mundus program in Color in Informatics and MEdia Technology (CIMET). He finished his PhD at the Computer Vision Group at the Friedrich Schiller University of Jena (Germany) in 2015 followed by a postdoctoral positions at the International Computer Science Institute (ICSI) (Berkeley, California). He then joined the Department of Computer Science at NTNU in Gjøvik first as a postdoctoral fellow and starting from late 2019 as an Associate Professor.

Marius Pedersen received his BSc in Computer Engineering in 2006, and MiT in Media Technology in 2007, both from Gjøvik University College, Norway. He completed a PhD program in color imaging in 2011 from the University of Oslo, Norway, sponsored by Océ. He is professor at the Department of Computer Science at NTNU in Gjøvik, Norway. He has since 2012 led the Colourlab, a research group at the Department of Computer Science.

Sophie Triantaphillidou received her BSc in 1995 and PhD in 2001 in Imaging Science from the University of Westminster in London, UK, where she subsequently taught and led research in imaging and colour sciences for 20 years. She is currently Adjunct Professor at the Department of Computer Science at NTNU in Gjøvik, Norway.