Estimating the repeatability and reproducibility of an AI-embedded measuring device: application to road markings

Térence Bordet[1*], *Maxime* Redondin[1], *Stefan* Bornhofen[2], *Sébastien* Denaës[1] *Aymeric* Histace[2]

[1]CORE Center by COLAS, Data Science Digital Road Inspection and Material, 4 rue Jean Mermoz, F-78114, Magny-Les-Hameaux
[2] ETIS UMR 8051, Cergy Paris University, ENSEA, CNRS, F-95000, Cergy, France

Abstract. The road markings inspection using on-board cameras is gaining popularity, particularly for road inspection, despite most research focusing on autonomous vehicles. Traditional inspection methods rely on costly equipment like RetroTek-D, resulting in infrequent inspections. Recent advancements in artificial intelligence offer a more cost-effective solution, enabling the use of standard cameras with AI models for continuous road marking detection. While metrics such as F1-score evaluate AI performance, they do not ensure effectiveness on new sites or account for uncertainty in AI applications. Once deployed, AI systems may not adapt to new data, necessitating retraining for any updates. This paper proposes a method to estimate the repeatability and reproducibility of AI-based measurement devices, drawing inspiration from ISO 5725 standards, to enhance their reliability in road marking detection.

1 Introduction

A road marking material is a painting on the pavement with injected glass microbeads to ensure good nighttime visibility [1]. Beads reflect the light emitted by vehicle's headlights back to the driver. This phenomenon is called retroreflection and is measured by retroreflectometer such as RetroTek-D. The quantity of retroreflected light is expressed in mcd/m²/lx. This process checks if a given line respects a minimal service and current marking durability models are based on this method [2]. However, this operation is expensive and do not allow for more regular inspections than once a year.

Since the 90's [3], the main alternative has been an on-board camera inspection. This solution is less expensive but is not fully industrialized. Main obstacles are the need for a reliable algorithm that can detect lines and define a measure of their deterioration. The first is resolved since the 2010s, thanks to the development of driving aids partially based on the detection of road markings to define traffic lanes [4], [5]. The second is still an issue but two main solutions are proposed: to photograph a marking and given to assessors [6] or an AI [7] for them to evaluate the deterioration; to estimate the contrast between pavement and marking [8]. However, these solutions are still not industrialized and are dependent of environmental conditions.

Our final objective is to propose a full framework to ensure the proper use of marking inspection assisted by an on-board camera under the widest range of environmental conditions possible. The first step is to define the visibility geometry of the road, the type of deployable camera and the AI family useful for marking detection for road inspection purposes only. Following the current standards, the geometry defined is EN 1436, which establishes the markings visibility model for a driver. For the couple of cameras and AI, we need a method for detecting road markings.

This point is a widely researched field thanks to the development of autonomous vehicles. A good amount of traditional image processing approaches was and is still used to detect road markings [4], but in the last ten years, deep learning methods have revolutionized the field [5], to a point where road markings detection is today almost exclusively working with AI.

Measuring the robustness of an AI is not fully established. Main performance scores announced, such as the F1-score, are calculated from its test dataset. It is composed by chosen images supposedly representative of roads. Concretely, if an AI is trained on images taken in Sri Lanka [9], then is this AI still reliable on French roads? If not, is there an AI model capable of adapting to French roads based on a Sri Lanka dataset? Even assuming that the performance score is unbiased and certain, there is no way of knowing whether a randomly selected road will show this performance. Concretely, even if an AI shows a 90% probability of correctly detecting a line, this doesn't mean that 90% of lines on any given road will be detected. For a road maintenance company such as COLAS, it is important to have a procedure at the end of any inspection campaign to measure a valid performance score for this campaign.

To answer these questions, this paper introduces a combination of AI analysis approaches and metrological control of measurements. Our method uses the principles of the round-robin test to determine the repeatability and reproducibility of our AI-embedded measuring device. To use the round-robin test, we readapt important metrological elements to adapt them to our experiment, and we collect our own data on a specific road track, that will later be sampled using the Cochran sampling method and analyzed using an analyze of variance (ANOVA).

* Corresponding author: terence.bordet@colas.com

2 Related Work

2.1 Metrological standard and AI objectives

The method searches to respect ISO 5725-1:2023 [10].

Repeatability is the fidelity under repeatability conditions. Repeatability conditions are conditions where independent test results are obtained by the same method on identical test or measurement individuals on the same test or measurement facility, by the same operator, using the same equipment and over a short time interval.[10].

Reproducibility is the fidelity under reproducibility conditions. Reproducibility conditions are conditions where independent test results are obtained by the same method on identical test or measurement individuals on different test or measurement facilities with different operators and using different equipment.[10]

Although many types exist, AI models only categorize lines as continuous or discontinuous at best and can't specify the line type. Only continuous lines or discontinuous lines are considered as lines and should be detected by AI models.

2.2 Selected artificial intelligence models

An AI model efficiency is deeply tied to the dataset it is trained with. An appropriate dataset is required to build any AI-based application. A same model can behave differently depending on the dataset it was trained on. In this article, we define an AI as a couple of a model and a dataset.

YOLO is introduced in 2015 by Redmon *et al* [11]. It is an innovative single-staged object detection deep convolutional neural network (CNN). It became popular because of its high speed and detection accuracy, while being open source. YOLO divide the image into a grid of same sized cells. Each grid cell predicts several bounding boxes and a set of class probabilities, the latter regardless of the number of predicted boxes. The grid cell in which the center of an object fall is responsible for detecting that object. Doing detection and classification in one go, while reframing object detection as a regression problem enable YOLO to reach enough speed to be a real-time object detector. Today, YOLO is a collection of algorithms evolving and getting better at each iteration. Each iteration enhances the previous one, using lot of improvements like better backbone, better loss functions or better architectural design. Yolov8 can be used has a solid solution to detect road markings [5], [12] In this article, we use YOLOv8.

U-Net is introduced in 2015 by Ronneberger *et al* [13]. It is a two-sided, multi-layer fully CNN model aiming to segment biomedical images. The first side is the encoding part, known as contracting path. It is a succession of double-convolution and max pooling, used to extract features and contextual information while decreasing the spatial resolution. Each convolution results on this side will be used layer as skip connection. The second side is the decoding part, known as expansive path. It consists of a succession of double-convolution and up-convolution, reducing the number of features at each up-convolution. Skip connections are used to help the decoding part locate the features in the image. The final part of U-Net is a simple 1x1 convolution. The resulting output is a mask composed of a pixel-level classes annotation. In the biomedical field, it can be a mask locating cancerous tumors, as an example. U-Net as demonstrated such performance that it is used as a general segmentation algorithm, beyond the biomedical field. On the road, it has been used to segment road markings [5], [14]. In this article, we use U-Net to segment line-type road markings.

SCNN is introduced in 2017 by X. Pan *et al* [15]. It is a specifically designed CNN structure to more effectively learn spatial relationships of lane markings. Unlike traditional CNN, which are built using convolutional operations layer by layer, the SCNN propose a slice-by-slice convolutional method inside the feature map. The feature map is a 3D tensor of dimension C x H x W. Spatial operation separate this tensor into slices. There are four spatial operations: down, up, right, left. For example, the down spatial operation split the tensor into H slices, and each slice is sent into a convolution layer of C kernels, where the result is stored into the next slice. By doing these spatial operations, spatial coherence is reinforced and allow message to pass between pixels. SCNN demonstrate solid results in lane marking detection. In this article, we use it to detect line-type road markings.

Hough transform is introduced in 1962 by Hough, is a feature extraction method used to find simple shape in image, such as straight lines, circles or ellipses. While not being an AI model, the Hough transform is a very popular computer vision and image processing method, especially in the field of autonomous vehicle and road marking detection [4], [16]. Because line-type road markings are often straight or slightly incurved in most situations, a simple Hough transform may be enough to enable the accurate detection of road lanes detection. With good pre-processing techniques, the Hough transform can detect even the border of the road to more precisely detect lane even when there is no road marking, while still being way less computationally heavy than any AI alternative. In fact, a simple Hough transform can solve the lane detection problem in 90% of the highway cases [4], [17]. The remaining 10% are due to various issues like road illumination, shadows, rain, fog or glare for example [4]. In this article, we use a simple Hough transform to detect line-type road markings and observe how it behaves in various non-highway situations, compared to AI models.

2.3 Datasets

Training an AI model require a large amount of data, especially on the field of road marking detection due to the vast amount of different scenarios possible and the large number of different objects a driver or autonomous vehicle may encounter. Many road scene datasets have been created in the past, and for this experiment we chose three of them.

Mapillary [18] is a large dataset of 25,000 high-resolution street-level images dataset. It contains 100 instance categories and 124 semantic objects categories,

covering different road instances across six continents. Images contains every possible condition, including different daytime, weather conditions, cameras and viewpoint. It is a very complete streetscape dataset. An example of an AI trained by Mapillary is [19].

Tusimple [20] is a dataset of 6,408 road images of US highways. It annotates lane-marking under various lighting conditions and different levels of occlusion. The resolution of Tusimple images is 1280x720. An example of an AI trained with Tusimple is [21].

CULane [22] is a dataset of 133,235 streetscape images located in Beijing and including various lightning and weather conditions. In addition, CULane reference eight lane scenarios: normal, crowded, night, no line, shadow, arrow, dazzle light, curve and crossroad. X.Pan is the reference example [15].

3 Method

Our goal is to determine a F1-score which corresponds more to the reality, while estimating the repeatability and the reproducibility of our AI-embedded measuring device, in our case, a camera. To achieve this, we create variation in our measuring system, in our measuring device and in our way of applying the method. Our strategy is to define different AI, different cameras and cameras resolution, different vehicles, different operators, choosing a single road circuit for the data acquisition and operate on it multiple times at different time of the day.

Then, when the data acquisition is complete and processed by our defined AI, we sample the processed frames to extract a fraction of frames representative of the global results. We use the Cochran sampling method [23] for creating the sample. Those results are then reviewed manually by our team.

Finally, with the newly reviewed results for every possible combination of AI, camera, operator and time of the day, we can compute the new more representative F1-score. With the newly computed F1-scores, we run an ANOVA method to extract all the information we want. We also use those newly computed F1-scores across our different groups to do a round-robin test, estimating the repeatability and reproducibility of our method. While not technically being an AI, it is the standard non-deep-learning method, and we use it to compare its behavior to real AI.

3.1 Road markings inspection strategy

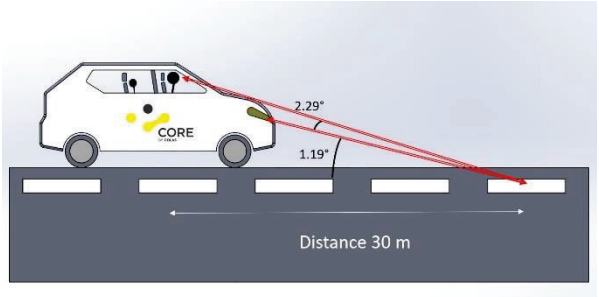


Figure 1 – EN 1436 standard

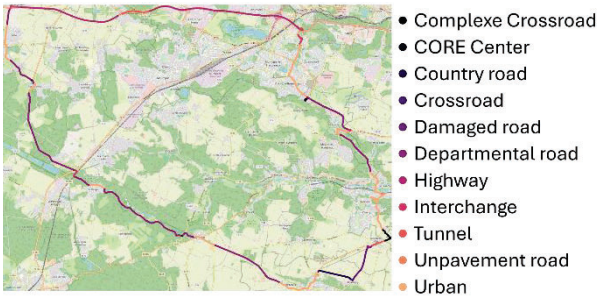


Figure 2 – Selected Road track for our study

An AI is defined as a couple between a model and a dataset. We define four AI: Yolov8 trained (by ourselves) with the Mapillary dataset; U-Net (by ourselves) trained with Tusimple; SCNN trained with CULane [24] ; a Hough Transform with no dataset [25].

We define a camera by its resolution: 1280x720 (720p), 1920x1080 (1080p) and 3840x2160 (4K). The objective is to check if the resolution could influence the performance of an AI. Cameras are mounted inside of the vehicle. The height and angle of the camera is inspired by the EN 1436 [26] geometry (Figure 1).

We chose three operators (drivers) and three vehicles. Before driving, operators remount the camera inside their vehicle. Each operator follows a road track, and their departure is spaced by a two-minute interval. Operators drives at different daytime of a same day: morning, noon and afternoon. We drove on dry day condition according to EN 1436. The road track covers standard French road scenarios: different markings material; road with no markings, different pavement, highway, urban (Figure 2)… We made a total of 9 videos of 100 minutes and 6000 images. Selected AIs processed each video and return the number of lines detected by image. It resulted in a total of 216,000 processed images to manually validate. This task is too much time-consuming and we need a smaller representative frame explained on the following section.

3.2 F1-score estimate and Cochran sampling

In our context, the F1-score (1) enables the quantity of correct detected lines (True Positives) against the quantity of false alarms (False Positives) and missed lines (False Negatives).

$$\boldsymbol{F1} = \frac{TP}{TP + \frac{FP + FN}{2}} \quad (1)$$

Currently, the F1-score is determined on the test dataset to confirm the training process. Because this dataset may not be representative of given work site, the communicated F1-score may not be adapted. In our method, we estimate a F1-score for each video acquired.

Each video has the same or different operators, environment, measuring system, sample or time-lapse to estimate in second time the reproducibility and repeatability. However, we need to validate the predicted values on each processed image. To avoid this, the solution proposed here and summarized in Figure 3, is to perform a double Cochran sampling.

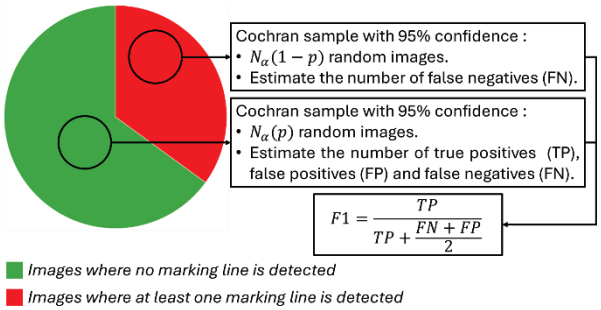


Figure 3 - Principle of Cochran sampling to control AI results

$$N_\alpha(p) = \frac{t_\alpha^2 \cdot p \cdot (1-p)}{(1-\alpha)^2} \quad (2)$$

Introduced by Cochran [23] and adapted here to the detection of marking lines, the minimum size of a sample is set by equation (2) where α=0.95 is the desired confidence level, t_α=1.96 is the solution to F(x)=α with F is the cumulative density function of the centered-reduced normal distribution and p∈[0,1] is the empirical probability that at least one marking line is detected by the AI on an random image. The choice of a 95% confidence implies an acceptable risk of 5% to obtain a sampling not representative of the real AI performance. A sample size is between 0 and 384 images.

On a random image, one AI detects no line or at least one. To analyze those two classes of situation, two samples are produced. Sample $N_\alpha(p)$ represents images where AI has detected at least one line. The number of TP, FP are counted, number of FN is partially counted. Sample $N_\alpha(1\text{-}p)$ represents images where AI has detected no line. The number of FN is fully counted (Figure 3).

An F1-score is calculated for each AI and each video. According to the latest reviews, we require a minimum of 75% to accept an AI and validate an inspection campaign [27], [5]. Different F1-scores must be compared, and an ANOVA approach is introduced to estimate the difference according to the different camera and weather conditions.

3.3 ANOVA

Analysis of variance (ANOVA) is a statistical technique used to evaluate differences between the means of several groups. In our case, a group is either a specific AI, camera or daytime. The ANOVA is based on the comparison of variances: it examines the variability of the data within each group and the variability between groups. By breaking down the total variance observed into two main components, the between-group variance and the within-group variance, the ANOVA enables us to determine whether the differences observed in the means are not linked to chance, meaning the difference are significant and due to something. The results of an ANOVA are often presented as a p-value. A p-value below a predefined threshold (in our case 0.05) indicates that there are significant differences between the groups. If the ANOVA reveals significant differences, post hoc analyses can be performed to identify which groups are different from each other, like the Honestly Significant Difference (HSD) test, created by Tukey and aiming to know which pair of groups are different by calculating a critical value based on the range distribution studied, then used to compare the differences between the means of each pair. ANOVA also enable us to study interactions, defined as the existing differences between the groups of one factor and all the groups of another factor. If the difference is not the same, there is an interaction. In our case, our factors are the AIs, cameras and daytimes. In the next section, we study the significant differences and the interactions of our factors and groups.

4 Results

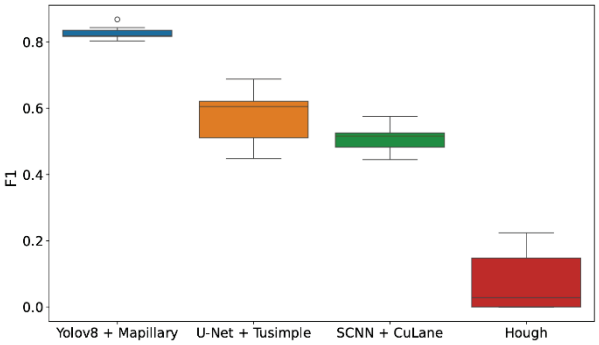
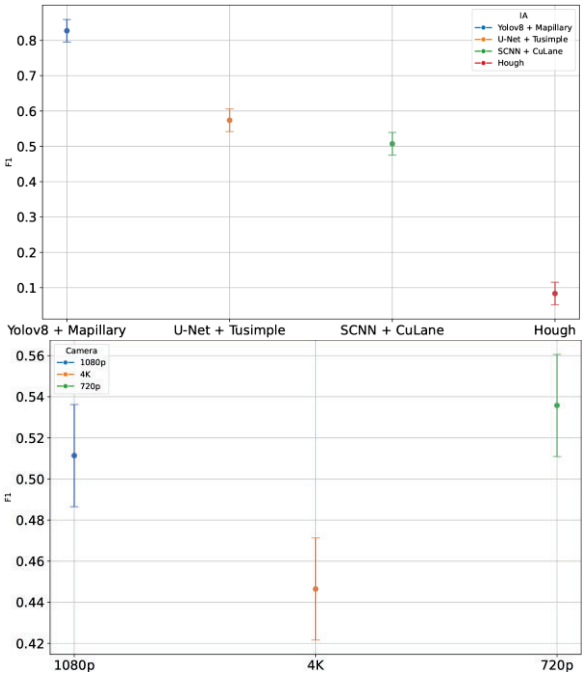


Figure 4 - Estimated F1-score for each model

Figure 4 shows F1-score per model, camera and daytime. Results are heterogeneous: Hough Transform shows an average of 8% with values being spread between 0% and 22% ; Yolov8 shows an average of 83% with values being spread between 80% and 87% ; U-Net is slightly superior to the SCNN, with respectively an average F1-score of 57% against 51% ; SCNN seems more constant and stable than U-Net, with an interquartile range (IQR) of 5% versus 13%. The IQR of Hough is 18%, and Yolov8 one is 3%. These raw results shows that Yolov8 is the most efficient.



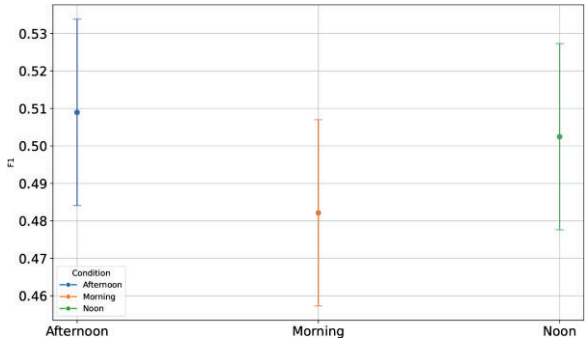
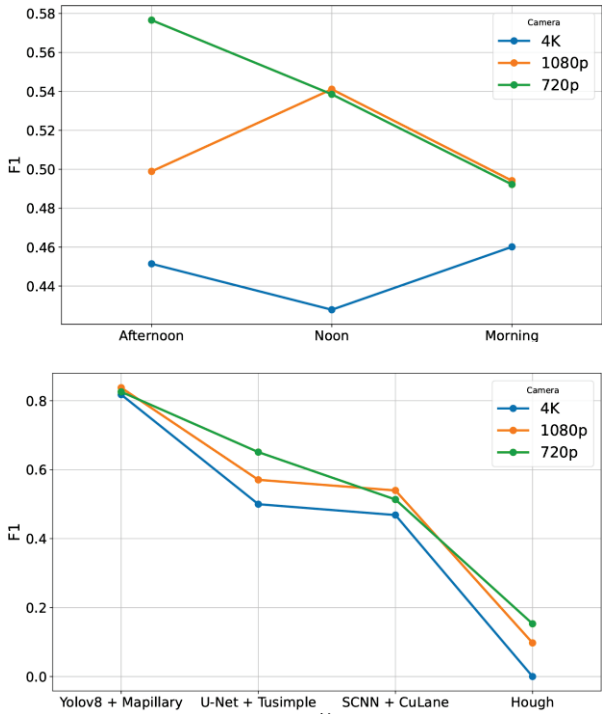


Figure 5 - HSD results for Ais (up), cameras (middle) and daytimes (down)

Figure 5 shows the HSD results between each group of a given factor. For AI, there is a significant difference between each AI. Especially between Yolov8+Mapillary and Hough. While the difference between the SCNN+CULane and UNet+Tusimple seems to be very small, it is still considered as significant. It means that the choice of an AI is fundamental for the methods, and they cannot be interchanged. For cameras, there is a significant difference between 720p and 4K, and 1080p and 4K. However, there is no significant difference between 720p and 1080p. Here, the 4K is the less performant camera resolution. In our case, because 720p and 1080p does not have significant difference, they can be interchanged. Finally, for daytimes, there is no significant difference between morning, noon, or afternoon. We can operate on the whole day without having an impact on our performances.

4.1 Interactions



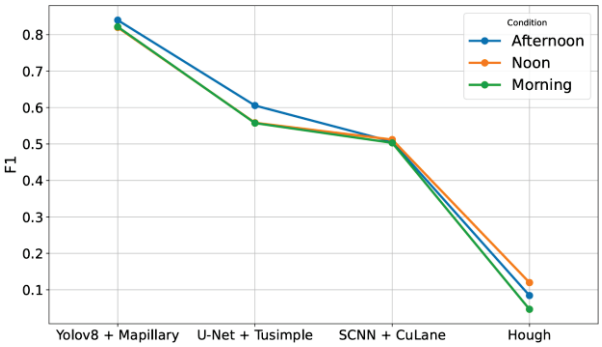


Figure 6 – Interactions results for Daytimes-Cameras (up), Ais-Cameras (middle) and Ais-Daytimes (down)

Figure 6 shows that there is no interaction between any of our factors. It means that no factor influences the others.

4.2 Repeatability and reproducibility

Table 1 - Round-robin test repeatability and reproducibility results

	Repeatability	**Reproducibility**
Variance	0.002	0.06
Standard deviation	0.05	0.25
Score 95%	**0.15**	**0.74**
Variation coefficient	9.26	46.23

Using the raw F1-score results we determined earlier on our road track, we can estimate the repeatability and reproducibility of our measuring method and AI-embedded measuring device (Table 1 - Round-robin test repeatability and reproducibility results). With 95% confidence, we can estimate our repeatability as 0.15 and our reproducibility at 0.74. It means that for a same measure on same conditions and a short time-lapse, two samples having a difference inferior to 15% are considered as the same sample. With varying factors, two samples having a difference inferior to 74% is considered as the same sample. These intervals are huge, but explainable (see Section 5).

5 Conclusion

This paper proposes a method to estimate the repeatability and reproducibility of an AI-embedded measuring device applied to road markings detection. The current way F1-score is determined, using a test dataset, is biased as the latter cannot be exhaustive. By using the principles of the round-robin test and adapting important metrological elements to adapt them to our experiment, we were able to compute a more accurate F1-score and estimate its representative repeatability and reproducibility. We used three AI models (Yolov8, UNet, SCNN) and one more classic image processing approach (Hough transform), and obtained average F1-score of 83% for Yolov8, 57% for UNet, 51% for the SCNN and 8% for the Hough transform. It is important

to remind that our objective is to detect line markings and not lanes. SCNN and Hough are more suitable for detecting lanes, explaining those results compared to general literature. We then estimated the repeatability and reproducibility, scoring respectively 0.15 and 0.74. These results are huge, and it is explained by the fact that the different AIs (model+dataset) we used are too different from one another. The raw F1-score determined by our method was sufficient to separate these AIs. However, we demonstrated that it is possible to run round-robin tests for AI-embedded devices. We did obtain repeatability and reproducibility results.

Next step now is to use more similar AI models to Yolov8 to obtain coherent results of repeatability and reproducibility, because it performed extremely well in this paper. We also plan on taking this method on the next level to estimate the accuracy and the uncertainty of our AI-embedded measuring device.

References

[1] D. Babić, T. Burghardt, D. B.-I. J. for T. and, and undefined 2015, "Application and characteristics of waterborne road marking paint," *ijtte.comD Babić, TE Burghardt, D BabićInternational Journal for Traffic and Transport Engineering, 2015•ijtte.com*, vol. 5, no. 2, pp. 150–169, 2015, doi: 10.7708/ijtte.2015.5(2).06.

[2] R. Maxime, B. Laurent, D. D.-I. J. of, and undefined 2021, "Em approach for weibull analysis in a strongly censored data context-application to road markings," *ijpe-online.comR Maxime, B Laurent, D DimitriInternational Journal of Performability Engineering, 2021•ijpe-online.com*, Accessed: Feb. 26, 2025. [Online]. Available: https://www.ijpe-online.com/EN/Y2021/V17/I4/333

[3] P. Charbonnier, F. Diebolt, … Y. G.-… of C. on, and undefined 1997, "Road markings recognition using image processing," *ieeexplore.ieee.orgP Charbonnier, F Diebolt, Y Guillard, F PeyretProceedings of Conference on Intelligent Transportation Systems, 1997•ieeexplore.ieee.org*, Accessed: Feb. 26, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/660595/

[4] A. Bar Hillel *et al.*, "Recent progress in road and lane detection: a survey", doi: 10.1007/s00138-011-0404-2.

[5] Y. Zhang, Z. Lu, X. Zhang, … J. X.-I. T. on, and undefined 2021, "Deep learning in lane marking detection: A survey," *ieeexplore.ieee.orgY Zhang, Z Lu, X Zhang, JH Xue, Q LiaoIEEE Transactions on Intelligent Transportation Systems, 2021•ieeexplore.ieee.org*, Accessed: Jun. 18, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9398517/

[6] Dumont Éric, "Évaluation du degré d'usure des marquages routiers par traitement d'images," Dec. 2001.

[7] J. Wu, W. Liu, Y. M.- Sensors, and undefined 2024, "Street View Image-Based Road Marking Inspection System Using Computer Vision and Deep Learning Techniques," *mdpi.comJ Wu, W Liu, Y MaruyamaSensors, 2024•mdpi.com*, Accessed: Feb. 19, 2025. [Online]. Available: https://www.mdpi.com/1424-8220/24/23/7724

[8] N. Hautiére, R. Labayrade, and D. Aubert, "Real-time disparity contrast combination for onboard estimation of the visibility distance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 2, pp. 201–211, Jun. 2006, doi: 10.1109/TITS.2006.874682.

[9] O. Jayasinghe, … S. H.-P. of the, and undefined 2022, "Ceymo: See more on roads-a novel benchmark dataset for road marking detection," *openaccess.thecvf.com*, Accessed: Aug. 02, 2024. [Online]. Available: http://openaccess.thecvf.com/content/WACV2022/html/Jayasinghe_CeyMo_See_More_on_Roads_-_A_Novel_Benchmark_Dataset_WACV_2022_paper.html

[10] "ISO 5725-1:2023 - Exactitude (justesse et fidélité) des résultats et méthodes de mesure — Partie 1: Principes généraux et définitions." Accessed: Feb. 20, 2025. [Online]. Available: https://www.iso.org/fr/standard/69418.html

[11] J. Redmon, S. Divvala, R. G.-… and pattern recognition, and undefined 2016, "You only look once: Unified, real-time object detection," *cv-foundation.org*, Accessed: Aug. 02, 2024. [Online]. Available: https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html

[12] C. Dewi, R. C. Chen, Y. C. Zhuang, X. Jiang, and H. Yu, "Recognizing Road Surface Traffic Signs Based on Yolo Models Considering Image Flips," *Big Data and Cognitive Computing 2023, Vol. 7, Page 54*, vol. 7, no. 1, p. 54, Mar. 2023, doi: 10.3390/BDCC7010054.

[13] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9351, pp. 234–241, 2015, doi: 10.1007/978-3-319-24574-4_28.

[14] M. Ghafoorian, C. Nugteren, N. Baka, O. Booij, and M. Hofmann, "EL-GAN: Embedding Loss Driven Generative Adversarial Networks for Lane Detection," 2018.

[15] X. Pan, J. Shi, P. Luo, X. Wang, X. T.-P. of the AAAI, and undefined 2018, "Spatial as deep: Spatial cnn for traffic scene understanding," *ojs.aaai.orgX Pan, J Shi, P Luo, X Wang, X TangProceedings of the AAAI conference on artificial intelligence, 2018•ojs.aaai.org*, Accessed: Jul. 30, 2024. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/12301

[16] R. Mori, … K. K.-S. 2004 A., and undefined 2004, "Hough-based robust lane boundary detection for the omni-directional camera," *ieeexplore.ieee.orgR Mori, K Kobayashi, K WatanabeSICE 2004 Annual Conference, 2004•ieeexplore.ieee.org*, Accessed: Jan. 23, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/1491793/

[17] A. Borkar, M. Hayes, M. T. Smith, and S. Pankanti, "A layered approach to robust lane detection at night," *2009 IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems, CIVVS 2009 - Proceedings*, pp. 51–57, 2009, doi: 10.1109/CIVVS.2009.4938723.

[18] "Mapillary Datasets." Accessed: Feb. 24, 2025. [Online]. Available: https://www.mapillary.com/datasets?locale=fr_FR

[19] Z. Yang, C. Zhao, H. Maeda, and Y. Sekimoto, "Development of a Large-Scale Roadside Facility Detection Model Based on the Mapillary Dataset," *Sensors 2022, Vol. 22, Page 9992*, vol. 22, no. 24, p. 9992, Dec. 2022, doi: 10.3390/S22249992.

[20] "TuSimple Dataset." Accessed: Feb. 24, 2025. [Online]. Available: https://paperswithcode.com/dataset/tusimple

[21] D. Neven Bert De Brabandere Stamatios Georgoulis Marc Proesmans Luc Van Gool ESAT-PSI and K. Leuven, "Towards End-to-End Lane Detection: an Instance Segmentation Approach".

[22] "CULane Dataset." Accessed: Feb. 24, 2025. [Online]. Available: https://xingangpan.github.io/projects/CULane.html

[23] R. R.-P. of the E. M. Society and undefined 1963, "w. G. Cochran, Sampling Techniques (John Wiley & Sons, 1963), ix+ 413 pp., 72s.," *cambridge.org*, doi: 10.1017/S0013091500025724.

[24] "GitHub - harryhan618/SCNN_Pytorch: Pytorch implementation of 'Spatial As Deep: Spatial CNN for Traffic Scene Understanding.'" Accessed: Feb. 24, 2025. [Online]. Available: https://github.com/harryhan618/SCNN_Pytorch

[25] "GitHub - chuanenlin/lane-detector: https://towardsdatascience.com/tutorial-build-a-lane-detector-679fd8953132." Accessed: Feb. 26, 2025. [Online]. Available: https://github.com/chuanenlin/lane-detector

[26] "Norme NF EN 1436." Accessed: Jan. 28, 2025. [Online]. Available: https://www.boutique.afnor.org/fr-fr/norme/nf-en-1436/produits-de-marquage-routier-performances-des-marquages-appliques-sur-la-ro/fa179243/81040

[27] D. Liang, Y. Guo, S. Zhang, … T. M.-J. of C., and undefined 2020, "Lane detection: A survey with new results," *SpringerD Liang, YC Guo, SK Zhang, TJ Mu, X HuangJournal of Computer Science and Technology, 2020•Springer*, vol. 35, no. 3, pp. 493–505, May 2020, doi: 10.1007/s11390-020-0476-4.