Additional Traffic Sign Detection - A Comparative Study

Thomas Wenzel*, Steffen Brueggert† and Joachim Denzler‡

*Robert Bosch Car Multimedia GmbH, Hildesheim, Germany
†Robert Bosch GmbH, Hildesheim, Germany
‡Friedrich-Schiller-Universität Jena, Computer Vision Group, Jena, Germany

Abstract

Automated driving is a long term goal that currently generates a lot of interest and effort in the scientific community and the industry. A crucial step towards it is being able to read traffic signs along the roads. Unfortunately, state-of-the-art traffic sign detectors currently ignore the existence of additional traffic signs. Yet being able to recognize these is a requirement for the task of automated driving and automated map data updates, because they further determine the meaning or validity of main signs. In this paper we aim at the detection of these additional signs, a first step towards their recognition. We will have a careful look at suitable evaluation measures and then use these to compare our proposed MSER-based approach to a selection of five differing types of detectors from the literature. We achieved a substantial improvement of the state of the art with 90% successful detections with full sign content detection on a challenging dataset, while significantly reducing the number of false positives. We will present our database, which contains high-resolution images of German traffic signs suitable for optical character recognition. We rely on hand-labelled main signs to emphasize the focus on additional sign detection. Our results were confirmed on a validation set containing European additional signs.

I. Introduction

Traffic sign recognition has been of interest in the research community for a long time and significant progress has been made in this field. Recent results indicate, that available datasets do not pose a challenge anymore [1]. However, the problem of detecting and recognizing additional traffic signs has been widely disregarded in the scientific community. These signs are mounted closely to main signs and further specify their meaning or validity, e.g. by specifying vehicle types or time restrictions. In general, additional signs are rectangular-shaped, white (depending on the country of origin) and contain a variety of black pictograms, arbitrary text, or a combination of both. German panels are white, framed by a black boundary, and are mostly located below main signs. We observed up to four additional signs attached to a single main sign.

The problem of additional sign detection and recognition is unsolved so far. In practice a working main sign detector or annotated main sign ground truth is needed. When it is available, the variety of the additional sign content poses a challenge, and the smaller sizes compared to main traffic signs make their recognition very difficult under regular image acquisition scenarios with moderate camera resolutions of at most 2 megapixels. Additional sign detection is an inherently multi-scale detection task. It is challenging due to varying sign aspect ratios, relative sizes, numbers, and typical image acquisition problems like motion blur, varying illumination, occlusion, etc. Most of these problems are independent of the country-dependent additional sign layout. There is no public benchmark for additional sign detection available yet, requiring researchers to create their own databases. In contrast to previous publications our database is based on single high-resolution still images of German traffic signs captured from a moving car. Using

Fig. 1. Samples of successful additional sign detections with high localization accuracy from our proposed MSER-detector.

high-resolution data also allows us to evaluate roughly, how detection algorithms scale with resolution.
In this work we propose a novel detection algorithm using maximally stable extremal regions (MSER) [2] and multi-stage region-filtering. It is compared to four available detectors from the literature and a simple baseline, each using different cues and techniques. Detection approaches can be grouped into the following categories:

- Edge- or contour-based, where the contour of the rectangular sign plate or its black boundary is detected.
- Region-based, where the homogeneous white background is detected.
- Learned object detectors, e.g. from the domain of pedestrian detection.

Edge-based approaches are prone to detecting non-sign structures, region-based approaches are sensitive to low-contrast areas along the sign boundary, causing undesired region growth beyond sign areas, and object detectors suffer from poor localization due to the variations in aspect ratio. We investigated how to cope with these problems and included algorithms of each type into our evaluation. The results are presented in section V. Recognizing the sign content will be a future research topic.
Our contributions are twofold: Evaluation of five detection algorithms on a new database of high-resolution images using a suitable evaluation criterion, and proposition of a novel MSER-based detector combined with multi-stage region-filtering relying on thresholds learned from data, which outperforms the best available detectors.

II. Related work

Additional traffic sign detection has only been considered in a small number of publications. Most recently, Puthon [3] covered additional sign detection, recognition and aggregation based on her results presented earlier [4]. In [4], a detection algorithm employing region-growing from contrasted seeds was presented and compared, which defined the current state of the art. The databases contained German and French traffic signs taken from video-sequences, where the additional signs had an average width and height of 36px and 22px, respectively. Detection recall with an intersection over union above 0.5 was reported to be below 75%. Similar results were achieved by adapting the graph-based image segmentation approach proposed by Felsenszwalb et al. [5] to the task. Both approaches were combined with region filtering based on geometric constraints, but still suffer from high false positive (FP) rates with a precision below 25%.
Nienhueser et. al [6] employed a Hough transform-based voting scheme to detect rectangles by making use of geometric constraints from the image domain directly in Hough space. No evaluation results were presented, but [3] and [6] noted that very

precise ROIs are required for the algorithm to yield satisfactory results. Hamdoun et. al [7] presented a Canny-based rectangle detection without disclosing implementation details. Detection of 14 out of 18 additional signs in their database was claimed. In [3] and [4] it showed slightly worse results compared to the region-growing and graph-based approaches. We thus selected the detectors based on [4], [5], and [6] for our comparison. Wu et. al [8] made explicit use of temporal information from video sequences to detect (text on) sign plates by finding image regions satisfying a vertical plane condition. Detection rates of 89% were reported.
Additional sign detection is closely related to other problems: More abstractly, it is the problem of detecting homogeneous rectangular-shaped regions containing contrasting symbols. US speed limit and license plate detection are examples of similar problems. Approaches for solving license plate detection are summarized in [9] and range from geometry-based algorithms like line- and rectangle-detection to machine learning approaches like HoG-SVM. Other techniques are designed to perform a color-histogram based segmentation [10] or to detect high-frequency intensity changes, which is a typical property of scene text.

III. Database

The images used in this work were taken in regular driving situations with the camera app of a windshield-mounted Samsung Galaxy S4 smartphone. The dataset contains all typical situations from city to highway settings and weather conditions ranging from sunny to snowy. The majority of the images were taken during daytime, some at dusk or dawn, and none at night. We only rely on single images of each sign. All images were manually selected to contain at least one main sign with a humanly readable additional sign. Signs in the background, even unreadable ones, are also included. An example is shown in Figure 2. Different exposure compensation settings were used to control camera shutter speeds.
The image resolution is 4128×3096px. Since the sign-to-camera distances range from 5 to 150m (mean 31m), image sizes of main signs and their attached additional signs vary widely, as shown in Figure 3. These signs are about 2.5 times as wide as the signs in e.g. [4], are rotated by up to 15° and have aspect ratios between 0.6 and 5.2. Main signs are labelled manually to avoid database-bias and limitations of main sign types caused by using a pre-defined main sign detector. It also allows investigating additional sign detection performance for noisy main sign detection positions. Ground truth boxes are aligned with the true sign boundaries, allowing precise calculations of overlap measures. For validation purposes we included an additional dataset of 300 low-resolution French and Austrian signs with an average width of 35 pixels.
We also investigated which sign geometry is to be expected in the images. Measuring rolling shutter times and calibrating the cameras showed that lens distortion and the

Fig. 2. Typical full image frame example from the database with ground truth annotations of three main signs (green) and three additional signs (yellow).

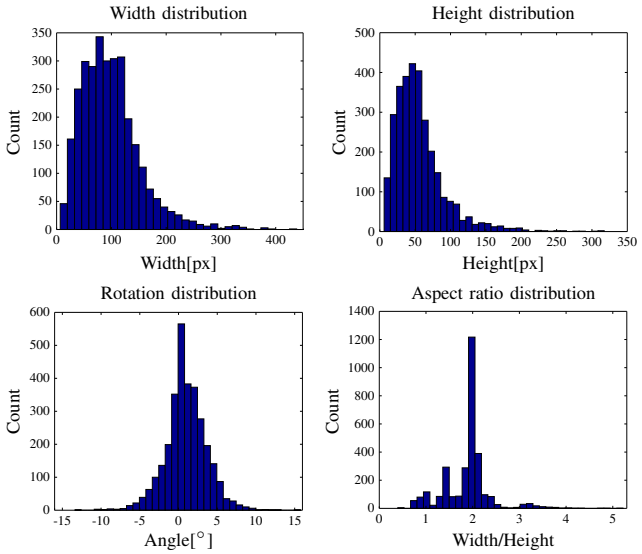

Fig. 3. Additional sign database statistics: Contains 3070 signs, mean sign width is 98px, height 55px, rotation mean and std. deviation is $0.97° \pm 2.61°$.

rolling shutter effect are negligible in our case due to the small relative sizes of additional signs compared to the image size. In combination with the large distances between camera and signs compared to very small depth variations of out-of-plane-rotated signs perspective effects are also negligible. Numerical verification confirmed that deviation from parallelism in manually labelled data has zero-mean and a standard deviation below $1°$. Thus we may assume a weak-perspective, i.e. parallelogram-shaped, image of additional traffic signs.

IV. Evaluation measures

The evaluation of a detection algorithm is a challenging task in practice. A standard measure is the PASCAL / intersection over union (IoU) measure. It is defined by dividing the intersection area of the ground truth box (B_{GT}) and the detection box (B_D) by their union area, generally considering a detection to be a true positive (TP), if $\text{IoU}(B_{GT}, B_D) > 0.5$.
Additional sign detection is an intermediate step, which aids a subsequent classification. Therefore the primary goal of a detector is to provide bounding boxes containing the full sign-content (e.g. text or pictograms), while avoiding the inclusion of clutter. Since content may be located arbitrarily close to the sign boundary, an optimal detection is the complete area inside the inner sign boundaries. We define the necessary condition for a successful detection as:

- All content elements have to be overlapped by a single detection by at least 95%, denoted by $COL(B_{GT}, B_D) \geq 0.95$.

In order to define a suitable sufficient condition, we additionally choose different IoU-thresholds. Our main evaluation measure is visualized in Figure 4 and defined as

$$\mathcal{C}_c := \text{IoU}(B_{GT}, B_D) > c \ \wedge \ COL(B_{GT}, B_D) > 0.95,$$

and for the sake of completeness we include

$$\text{IoU}_c := \text{IoU}(B_{GT}, B_D) > c, \qquad \text{with } c \in \{0.5, 0.8\}\,.$$

Any box B_D violating the considered evaluation measure is counted as FP. If more than one detection satisfies it, only the one with largest IoU is counted as TP, the others as FP.

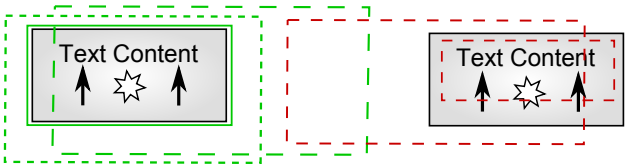

Fig. 4. Examples of allowed/disallowed overlap scenarios between ground truth and detections with respect to $\mathcal{C}_c$, larger line spacing corresponds to smaller c. Green boxes are counted as TP, red boxes as FP. Note that the latter are admissible according to $\text{IoU}_{0.5}$.

V. Detection

Assuming a given manually labelled main sign, we will first introduce how we select our ROIs for additional sign detection in this section. This process is shared by all implemented algorithms. Then we will present our MSER-based detector in detail and finally describe our implementations of detectors from the literature. For all detectors we convert the RGB input to grayscale first.

A. ROI generation

The first step in the detection of additional signs of a main sign M is determining the ROI for detection. Our approach is aimed at additional signs being located below a main sign, but it may be generalized to any principal orientation.
Our data showed that an incremental approach, where the ROI is dynamically extended downwards after successful detections is infeasible in practice. Instead, we choose the ROI below the main sign depending on the shape and size of M. The necessary extension factors are learned from a training set.

B. MSER-based detection

Our proposed algorithm mainly relies on the segmentation results obtained by the MSER flood-fill algorithm. It outputs regions, whose sizes remain stable during the flood-filling, and which may be mutually contained in each other. The detection process is divided into the following steps:

- Image contrast adjustment
- MSER detection
- Region filtering and post-processing

In the first step 1% of the ROI's grayscale values are saturated and the others are stretched to the full 8-bit range. This produced better results than normalizing the grayscale values. We then employ the OpenCV-implementation [11] of MSER for grayscale images, with a parameterization adjusted to outputting a large region-count. This process typically produces a set of several dozens of regions per ROI, which requires filtering to yield low FP rates.

1) Region filtering: Our proposed multi-stage approach is designed to prune the majority of the detected regions in a computationally efficient manner based on geometric constraints. In each stage a more complex bounding-box approximation, visualized in Figure 5, is calculated and a number of geometric quantities is checked against learned thresholds. In total, we perform 27 checks for each region, and immediately reject it, if a check fails.
Let R be one of those regions, representing a set of connected pixels at integer coordinates. Stage 0 only checks the ratio of region- to main sign-size. Stage 1 relies on the axis-aligned bounding-box of a region, B_a. Stage 2 determines a bounding rectangle based on fitting an ellipse E to R by estimating the covariance matrix of R using its normalized second central moments [12]. This yields its centroid $\mathbf{c}_e$, main axes lengths

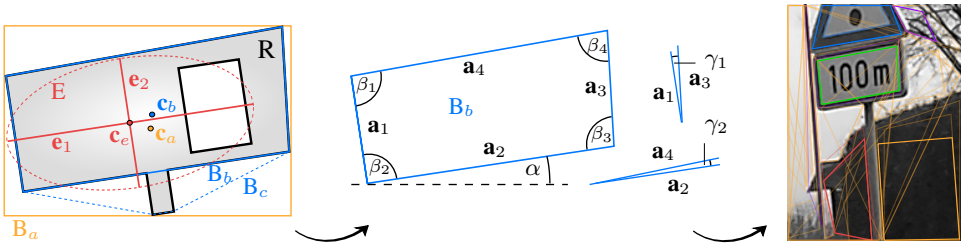

Fig. 5. Left, middle: Entities calculated for region filtering. Right: Filtered region examples, color denotes rejection stage. Purple: Stage 0. Orange: Stage 1, axis-aligned bounding box B_a. Red: Stage 2, main axes and centroid of the fitted ellipse E. Blue: Stage 3, convex hull B_c (dotted) and its approximation B_b (solid). Green: Accepted region. For improved visual clarity, all regions are represented by their stage 3-bounding box and most are drawn transparently.

$\|\mathbf{e}_1\|$ and $\|\mathbf{e}_2\|$ and rotation θ, which we omit due to sensitivity in the estimation. From centroid and axes we can determine an (axis-aligned) bounding box B_e. In stage 3 we finally compute the convex hull B_c of R and simplify it to a bounding box B_b, which only contains four remaining corners. We employ a variation of the greedy Visvalingam-Whyatt algorithm [13], where triangle heights are considered instead of triangle areas, which yielded better results. Both significantly outperformed the standard Douglas-Peucker algorithm [14].

While the ratio of $\|\mathbf{e}_1\|$ and $\|\mathbf{e}_2\|$ can be estimated accurately, their absolute lengths are estimated by assuming an e.g. elliptical point distribution in R. This results in inaccuracies when the assumption is violated. We compensate this by our threshold learning procedure described below.

In each of the last three stages we check the relative width w_i, height h_i, area relative to main sign area A_i, aspect ratio $r_i = \frac{w_i}{h_i}$, and centroid position $\mathbf{c}_i$ of B_i, where $i \in \{a, e, b\}$. Additional criteria are the mean grayscale value of the region and the norm of the relative distance error between $\mathbf{c}_e$ and $\mathbf{c}_a$. In the convex hull stage we also check the inner angles β_j of B_b and the angles between opposing edges γ_j, the relative shortest edge length, the ratio between shortest and longest edge of B_b and the error introduced in the simplification of the convex hull. The remaining regions are checked for mutual IoU, and if it is greater than 0.7, the smaller region is rejected. In a final post-processing step the obtained regions are stretched by 5% to always contain the complete sign content.

2) Threshold calculation: In order to determine the thresholds used in the above region filtering stage, we calculate all the above properties of all regions collected on a training set. Then for each property i the thresholds are chosen such that at most p_i percent of the TP regions are filtered out, according to a heuristic that maximizes the gain in FP filtering. By varying the p_i the amount of filtering can be adjusted, cf. section VI.

C. Region-growing from contrasted seeds

In [4], dark pixels surrounded by lighter pixels are determined by morphological reconstruction and a region-growing process is started from these locations. We added several improvements to the original proposal:

- Neighboring seeds are likely to yield regions with large mutual IoU. These regions are not filtered originally, we therefore delete the smaller of two regions with mutual IoU greater than 0.7, as described above.
- We use a subset of the criteria introduced in our region-filtering. Limiting it to stages 0 to 2 and outputting axis-aligned boxes showed the best results.

Parameters and thresholds are determined on a training set. These measures retain the original detection rates while improving FP rates dramatically. We name the original implementation RG and ours RG_+.

D. HoG-SVM

In this work we want to investigate, if standard object detectors can cope with the large aspect ratio variation in this detection problem. For the sake of simplicity we chose a standard HoG-SVM approach [15], since our experiments with the modern ACF-framework [16] yielded similar results. We evaluated three approaches for generating training patches:

1) Rescale all signs to a fixed aspect ratio and pad all patches equally.
2) Keep original sign aspect ratios and pad training patches as necessary.
3) Train multiple detectors for several aspect ratios.

As expected, our experiments confirmed that 1) works best, followed by 2). Approach 3) suffers from its increased need of training data and the need of a suitable normalization of detector outputs. We configured our detector as follows: ROIs were slightly extended to account for the training patch-padding, positive patches generated according to 1) with 10% padding in each direction, size 96×48px (based on Figure 3), 1.5k positive and 21k negative patches, randomly sampled from the training set. Hard negative mining slightly deteriorated the results, thus it was omitted. 8×8 fHoG-features, linear c-SVM, stronger weighting of positive instances. IoU-based non-maximal suppression with threshold 0.4 or 0.7 for higher recall, filtering of non-centered detections. Four scales, three aspect ratios, automatically selected on a validation set.

E. Hough transform-based

The Hough transform was considered in [6] and [7], but since no implementation details were provided in the latter, we implemented the approach described in [6], but used the classic Hough transform. Parameters for Canny filtering, angle thresholds and resolution, parallel line distances, and sizes of peaks in Hough space were determined on a training set.

F. Graph-based

Following the description in [3], we implemented the graph-based segmentation algorithm originally described in [5]. We added the contrast adjustment described above, used an 8-neighborhood, standard τ and optimized the ROI-size-dependent choice of k on a training set.

G. Position prior

For verification purposes we included a simple baseline: Based on the relative positions and sizes of additional signs in a training set we output two boxes at a fixed relative position and size within the ROI, leaving FP rejection to a subsequent classifier.

	Recall @ Precision ≥ 0.03				Avg.
	$\mathcal{C}_{0.5}$	$\mathcal{C}_{0.8}$	$\text{IoU}_{0.5}$	$\text{IoU}_{0.8}$	Runtime
Prior	0.552	0.112	0.661	0.115	**<1ms**
Hough	0.631	0.510	0.722	0.511	18ms
Graph	0.793	0.614	0.883	0.625	56ms
RG	0.879	0.746	0.902	0.751	65ms
HoG	**0.945**	0.314	**0.976**	0.320	19ms
MSER	0.942	**0.883**	0.967	**0.888**	11ms

TABLE I. MAXIMUM RECALL OBTAINED WITH EACH ALGORITHM ACCORDING TO DIFFERENT EVALUATION MEASURES. RUNTIMES DO NOT INCLUDE REGION-FILTERING. FOR MORE DETAILS SEE SECTION VI.

VI. Results

Our evaluation was conducted on a set of 3070 additional signs, 60% were used for training and validation, and 40% for testing. Due to their different nature and very low detection rates of all algorithms we excluded prism signs from the evaluation - detecting those signs will be a goal of future work.

We proceeded in two steps to evaluate the presented algorithms. In the first step, we maximized the recall of all detectors at a precision greater than 0.03 by disabling all types of region filtering and maximizing output of the SVM respectively. The evaluation acted as an oracle picking the optimal region. Results were obtained by repeating random sub-sampling validation 3 times, they are shown in Table I: RG, MSER, and HoG yield the best detection rates on our dataset. The position prior baseline proves that IoU_c is a significantly weaker evaluation measure than $\mathcal{C}_c$ and demonstrates that a dedicated detector is necessary for the task. The Hough-based approach suffers from the large ROI sizes and from resulting noisy Canny images, e.g. caused by sign posts or guard rails. It then detects spurious lines that do not correspond to sign boundaries. Adding a gradient direction-based filtering step only yielded a slight improvement. When the ROI-size is reduced, recall increases to 68%, but detection of multiple signs is hindered. We also worked on other edge-based approaches, but none of them yielded results similar to HoG or MSER. The graph-based approach is also outperformed: It suffers from its lack of parameters, a necessary adjustment towards noisy data results in oversegmentations of smooth data.

We selected RG, MSER and HoG for further investigation. Results were obtained by repeating random sub-sampling validation 10 times, they are displayed in Figure 6. MSER improves not only detection rates of RG_+, but also yields an improvement of the FP rate, which is dramatic compared to basic RG. This shows the effectiveness of our proposed region-filtering algorithm. Regarding detection and FP rates, HoG shows very similar results compared to MSER. Both are capable of satisfactory detection rates at less than 0.5 FPs per ROI. The center graph shows that RG yields the best localization performance, followed by MSER. Additionally, besides the Hough-based approach MSER is the only detector that allows estimating the sign rotation angle. The standard deviation of the estimation error is less than 2.25° on average, which is an improvement over assuming no rotation. The right graph shows results agreeing with the findings in [4] and suggests RG requires a higher sign resolution than provided by their dataset. HoG is the most robust against sign size variations, indicating its suitability for low-resolution data. MSER reduces the sign size required for best detection performance compared to RG. However, these results have to be considered with care, since they were obtained by interpolation. On the test set containing difficult low-resolution European signs MSER achieved 69% recall, RG_+ 67% and HoG 73%. HoG suffered from an unfavorable aspect ratio distribution on the dataset. On our dataset, MSER achieves the best overall

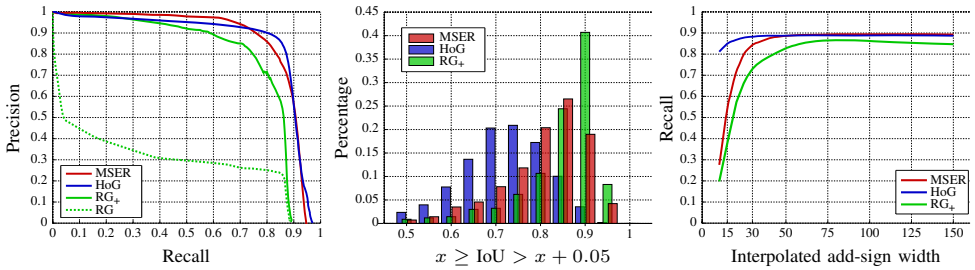

Fig. 6. Left: Precision/Recall-graph of the best-performing detectors. Results for both RG versions and MSER were obtained by varying all p_i in the threshold calculation. Average areas under the curves ($\pm$ standard deviation) are $\text{AUC}_{\text{MSER}} = 0.868 \pm 0.006$, $\text{AUC}_{\text{HoG}} = 0.863 \pm 0.012$, $\text{AUC}_{\text{RG}_+} = 0.786 \pm 0.007$, $\text{AUC}_{\text{RG}} = 0.298 \pm 0.008$. Middle: Distribution of IoU of true positive detections w.r.t. $\mathcal{C}_{0.5}$ at precision 0.5. Right: Scaling of detection algorithms with varying input data sizes at precision 0.5. Signs were bilinearly interpolated to a fixed width before presenting it to the detection algorithm.

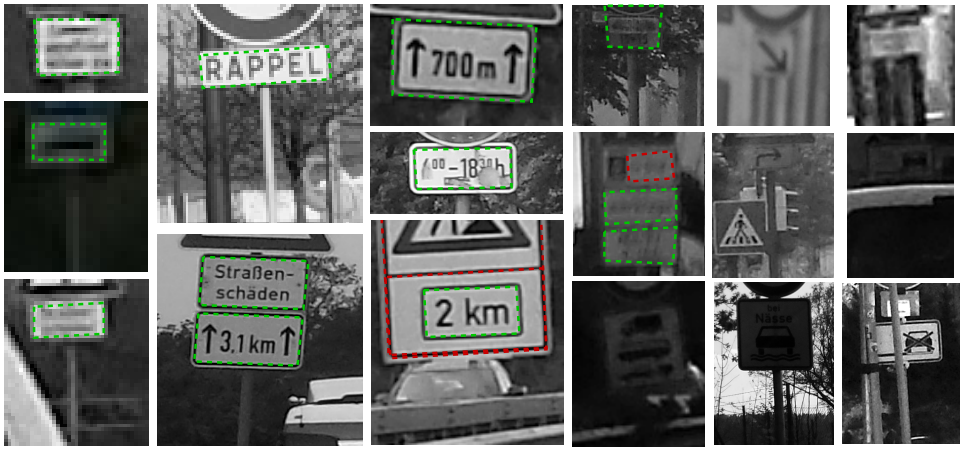

Fig. 7. Left columns: Successful detections of our proposed MSER detector. Right columns: Missed detections due to low contrast, dirt, occlusion, and motion blur. Green: Correct detections, red: false positives. Note that the FP in the fourth column is due to missed content and is admissible w.r.t. $\text{IoU}_{0.5}$.

results. It yields high detection rates paired with low FP rates, accurate localization and fast detection speeds. Additionally, both RG and MSER provide cues for a subsequent segmentation and recognition algorithm, in the form of detected seeds and contained dark regions within the detections, respectively.

A. *Runtime*

All algorithms were implemented in a combination of MATLAB and C++ and run on a single core of a modern laptop CPU. All core components were implemented in C++, and we used the SSE2-accelerated implementation of [17] for fHog-feature calculation and image resampling. Table I shows the results excluding region-filtering. Most of the implemented algorithms are close to real-time capable. When considering MSER runtimes including region filtering, contrast adjustment takes about 4ms on average, finding MSERs 7ms and region-filtering between 6 and 12ms, depending on the thresholding, summing up to 17 to 23ms per ROI. When the maximal ROI width is limited to at most 60 pixels (e.g. by interpolation), runtimes further decrease to a total of 10 ms/ROI.

VII. Conclusion and Future Work

In this work we proposed a novel additional sign detection algorithm and compared it to a variety of competitors on our new database of high-resolution smartphone images. We described our ROI generation procedure, which defines a region below main signs based on their sizes, shapes, and learned extension factors. We compared a baseline, three detectors from the literature, and a standard object detector, HoG-SVM, to our new MSER-based detection scheme, which relies on our effective multi-stage region-filtering using thresholds learned from data. Our results show that even on high-resolution data the detection task remains challenging. RG, MSER, and HoG-SVM yielded the best detection rates. When considering all criteria under investigation, i.e. AUC, FP rates, localization, and angle estimation accuracy, MSER yields the best overall results while running at 50 ROIs/s. Detection rates above 85% at less than 0.3 FPs per ROI were achieved. Additionally it may provide useful cues for a subsequent content recognition. The results presented in this work indicate that region-growing approaches apparently reached their limits in situations with low sign boundary contrast and lower-resolution data. Using machine-learning techniques may yield further improvements. Our research in this field will be continued before we tackle the problem of recognizing the sign content.

REFERENCES

[1] M. Mathias, R. Timofte, R. Benenson, and L. Van Gool, "Traffic sign recognition - how far are we from the solution?" in *ICJNN*, 2013.

[2] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image and vision computing*, vol. 22, no. 10, pp. 761–767, 2004.

[3] A.-S. Puthon, "Speed limit determination by real-time embedded visual and cartographical data fusion," Theses, Ecole Nationale Supérieure des Mines de Paris, Apr. 2013. [Online]. Available: https://pastel.archives-ouvertes.fr/pastel-00957392

[4] A.-S. Puthon, F. Moutarde, and F. Nashashibi, "Subsign detection with region-growing from contrasted seeds," in *Intelligent Transportation Systems (ITSC), 2012 15th International IEEE Conference on*. IEEE, 2012, pp. 969–974.

[5] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, 2004.

[6] D. Nienhuser, T. Gumpp, J. Zollner, and K. Natroshvili, "Fast and reliable recognition of supplementary traffic signs," in *Intelligent Vehicles Symposium (IV), 2010 IEEE*. IEEE, 2010, pp. 896–901.

[7] O. Hamdoun, A. Bargeton, F. Moutarde, B. Bradai, and L. L. Chanussot, "Detection and recognition of end-of-speed-limit and supplementary signs for improved european speed limit support," in *15th World Congress on Intelligent Transport Systems (ITS)*, 2008.

[8] W. Wu, X. Chen, and J. Yang, "Detection of text on road signs from video," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 6, no. 4, pp. 378–390, 2005.

[9] C. N. E. Anagnostopoulos, I. E. Anagnostopoulos, V. Loumos, and E. Kayafas, "A license plate-recognition algorithm for intelligent transportation system applications," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 7, no. 3, pp. 377–392, 2006.

[10] S.-L. Chang, L.-S. Chen, Y.-C. Chung, and S.-W. Chen, "Automatic license plate recognition," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 5, no. 1, pp. 42–53, 2004.

[11] G. Bradski, *Dr. Dobb's Journal of Software Tools*.

[12] B. Jähne, "Digital image processing."

[13] M. Visvalingam and J. Whyatt, "Line generalisation by repeated elimination of points," *The Cartographic Journal*, vol. 30, no. 1, pp. 46–51, 1993.

[14] D. H. Douglas and T. K. Peucker, "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature," *Cartographica: The International Journal for Geographic Information and Geovisualization*, vol. 10, no. 2, pp. 112–122, 1973.

[15] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1. IEEE, 2005, pp. 886–893.

[16] P. Dollár, R. Appel, S. Belongie, and P. Perona, "Fast feature pyramids for object detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 36, no. 8, pp. 1532–1545, 2014.

[17] P. Dollár, "Piotr's Computer Vision Matlab Toolbox (PMT)," http://vision.ucsd.edu/ pdollar/toolbox/doc/index.html.