IEEE VIP Cup 2017
Video and Image Processing Cup

This competition is sponsored by the *IEEE Signal Processing Society (SPS) MMSP TC*.

Traffic Sign Detection under Challenging Conditions

Outline

1 General Information

1.1 Competition Overview

Robust and reliable traffic sign detection is necessary to bring autonomous vehicles onto our roads. State of the art traffic sign detection algorithms in the literature successfully perform the task over existing databases that mostly lack challenging road and environmental conditions. This competition focuses on detecting such traffic signs under challenging conditions.

To facilitate such task and competition, we provide video dataset that contains a variety of road conditions. In such video sequences, we vary the type and the level of the challenging conditions including but not limited to a range of lighting conditions, blur, haze, rain and snow levels. The goal of this challenge is to implement traffic sign detection algorithms that can robustly perform under such challenging environmental conditions.

1.2 Competition Organization

The competition is open to any team that is eligible to participate and make a submission, which is due ***July 1, 2017*** and the finalists of the competition are announced by ***August 1, 2017***. Finalists will be selected by the organizers of the competition and will be invited to ICIP 2017 based on the overall performance whose details are provided in Section 2. A judging panel will evaluate the three finalists to determine the ultimate winners at ICIP 2017, which will be held ***September 17-20, 2017***. In addition to algorithmic performances, demonstration and presentation performances will also affect the final ranking. Presentations will be followed by questions of judging committee and general audience.

1.3 Database Overview

We provide a video dataset whose sequences include processed versions of captured traffic videos and synthesized videos. Examples of the simulated environmental conditions are extreme lighting conditions, blur, haze, heavy rain and snow. These conditions span a wide range of challenging levels from mild to severe. The participants are expected to design a traffic sign detection algorithm that robustly works for the challenging conditions in the provided dataset.

1.4 Detailed Description of Competition

Current transportation technologies shape our industry and landscape. Transportation-related goods and services account for more than $1 trillion of U.S Gross Domestic Product [1]. High-ways, roads, and parking lots consume a significant portion of our cities and dictates their designs. The technological enhancements shape the practices in transportation systems and autonomous driving is a good example that has the potential to transform current systems. Robust understanding of a scene is vital to have autonomous cars in our roads without subjective intervention and traffic sign detection is one of the most critical challenges in scene understanding for autonomous vehicles. Because of its significance, the researchers have conducted various studies to solve the traffic sign detection issue in the literature [2-12].

The authors in [2] provide an overview of existing traffic sign detection methods by categorizing the main components of these approaches into three classes as detection, classification, and temporal integration. They also propose a shape-based method and validate the detection performance through offline and online experiments. Even though the authors claim that challenging environmental conditions affect the detection accuracy, they do not analyze the relationship between these conditions and their effects. In [3], the authors conduct a survey of the traffic sign detection literature with a focus on driver assistance systems. The public sign databases described in this survey include *German Traffic Sign Recognition Benchmark (GTSRB)* [4, 5], *Belgium Traffic Sign Data Set (BelgiumTS or KUL)* [6], *Swedish Traffic Signs (STS) Data Set* [7], *Traffic Sign Database* by *Petkov et al.* [8], and *Stereopolis Database* [9]. In addition to *GTSRB* and *BelgiumTS*, *German Traffic Sign Detection Benchmark* [10] and *Belgium Traffic Sign Classification* (BTSC) [11] database are also utilized in the literature [12]. However, these traffic sign detection and recognition databases do not explicitly address the relationship between challenging environmental conditions and algorithmic performance.

In order to fully realize self-driving cars, we need to utilize algorithms that can successfully operate under real-life challenging conditions. Self-driving cars are expected to recognize traffic signs in day and night under various visibility conditions. Speed of the car, lighting conditions, haze, rain and snow can lead to significant challenges for detection and recognition algorithms. To serve this competition and to analyze the effect of challenging conditions on traffic sign detection performance, we provide a video dataset, which includes processed versions of captured traffic videos and synthesized videos. We simulate challenging conditions including but not limited to extreme lighting conditions, blur, haze, heavy rain and snow and the level of these conditions vary from mild to severe. We provide examples corresponding to motion blur, rain, and extreme lighting conditions in *Fig. 1*, *Fig. 2*, and *Fig. 3*.

The most common approach to collect driving data is to mount cameras on a car, drive the car around to record videos, and then process the videos to build a database. However, there are several drawbacks in this approach. First, it consumes considerable time. Second, and more importantly, it is impossible to control environmental factors and capture the same scene under varying environmental conditions. In other words, if we want to record the same scene at different times to make a controlled experiment, we need to keep everything identical at all times of recording except for the environmental condition that we want to vary. Apparently, it is not possible to control all environmental and surrounding conditions. For instance, we cannot guarantee following exactly the same route every time. Moreover, there are several other factors that are difficult to control including the intensity of a weather phenomenon such as rain or the intensity of the lighting condition.

Fig. 1. Sample images that are processed to simulate motion blur.

Fig. 2. Sample images that are processed to simulate rain.

Fig. 3. Sample images that are processed to simulate extreme lighting conditions.

Fig. 4. Sample images with different distortion levels in synthesized videos.

To complement the limitations of captured sequences, we generate synthesized videos using a professional game development tool. To simulate the driving experience, we created a simple car-driving game including a car object and a mounted camera. To avoid manually controlling the car, we generated a path that is followed by the car object. Therefore, we generate training and testing sequences by recording the game. In *Fig. 1-4*, we provide example images in which the levels of distortion are systematically controlled and in *Fig. 5*, we show examples of different types of challenges.

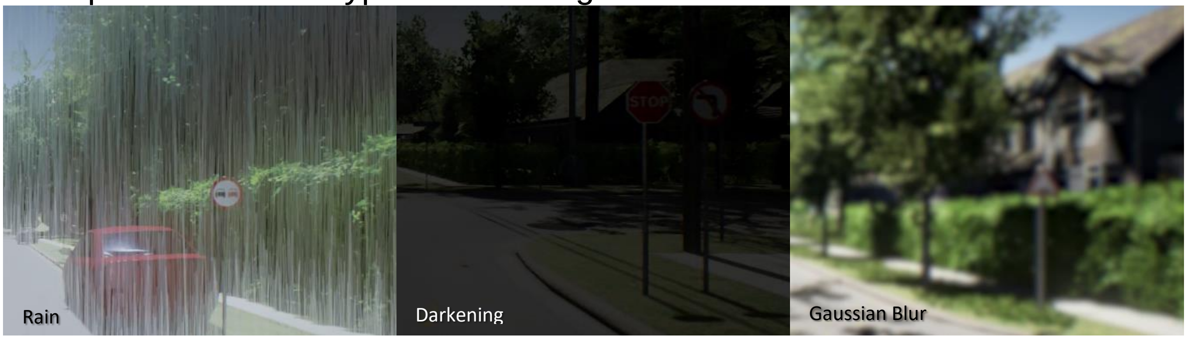

Fig. 5. Sample images with different weather conditions.

In this competition, we ask participating teams to design and implement traffic sign detection algorithms that can robustly perform under challenging environmental conditions. Proposed algorithms are expected to recognize traffic signs in both real-world videos and synthesized sequences without utilizing upcoming frames for current prediction. We provide a detailed description of the requirements and the database in Section 2.

1.5 Formation of Competition Teams

Each team is to be composed of one faculty member (as the supervisor), at most one graduate student (as a tutor/mentor), and at least three but no more than 10 undergraduate students. At least three of the undergraduate team members must be either IEEE SP members or student members.

1.6 Submission from Participating Teams and Evaluation

All the submissions must be received by ***July 1, 2017***. Each submission should include a report, in the form of an IEEE conference paper, up to 6 pages, on the technical details of the methods used and programs developed. The organizing committee will evaluate the submissions. By ***August 1, 2017***, the best three teams will be identified to participate in a session for the final competition, to be held at ICIP 2017.

1.7 Final Competition at ICIP 2017

A maximum of three members from each of the three teams selected will be paid to attend ICIP 2017 for the final competition. More members are also welcome to attend, but they will not be provided with any travel grant. In addition, those team members who will not be presenting a paper at the conference will be offered a complimentary ICIP registration. A Judging Panel (composed of members of the TC concerned and other TC Chairs) will be set up to pick the ultimate winners at the conference. The teams need to present the technical details regarding how they solve the challenging problem, and demonstrate their results at a session. In addition to algorithmic performances, demonstration and presentation performances will also affect the final ranking. Presentations will be followed by the questions of judging committee and general audience.

1.8 Important Dates

Call for Competition	February 15, 2017
Data are Available Online	March 15, 2017
Submission Deadline	July 1, 2017
Announcement of Best Three Teams	August 1, 2017
Conference	September 17-20 2017

1.9 Budget

Each team member is offered up to $1,200 for continental travel or $1,700 for intercontinental travel, and at most three people from each team will be supported. The participants can claim their travel expenses on a reimbursement basis. Furthermore, the selected teams will be invited to join the Conference Banquet as well as the Student Career Luncheon so that they can meet and talk to SPS leaders. The prizes offered to the three teams for each of the categories in the final competition are suggested to be as follows:

The champion: $5,000
The first runner-up: $2,500
The second runner-up: $1,500

1.10 Online Resources

The official announcement of the VIP cup on the IEEE SPS website can be found at:
http://signalprocessingsociety.org/get-involved/video-image-processing-cup.

Team Registration form for the VIP cup can be accessed at:
https://www2.securecms.com/VIPCup/VIPRegistration.asp.

All updates, news, dataset, and information on the competition throughout the course of the competition can be found at: https://ghassanalregib.com/vip-cup/.

1.11 Organizing Team

The organization of 2017 VIP Cup is supported by the members of the IEEE SPS Multimedia and Signal Processing (MMSP) Technical Committee in collaboration with team members from Georgia Institute of Technology's School of Electrical and Computer Engineering:

- Ghassan AlRegib
- Can Temel
- Tariq Alshawi
- Min-Hung (Steve) Chen

VIP Cup is overseen by the Student Service Committee of the IEEE Signal Processing Society (Patrizio Campisi, Chair).

2 Detailed Information about Competition and Database

2.1 Video Sequence Information

The video sequences in the database are grouped into two classes: real data and unreal data. Real data correspond to processed versions of sequences acquired from real world. Unreal data corresponds to synthesized sequences generated in a virtual environment. There are *49* real sequences and *49* unreal sequences that do not include any specific challenge. We separated the sequences into *70%* and *%30* splits. Therefore, we have *34* training videos and *15* test videos in both real and unreal sequences that are challenge-free. There are 300 frames in each video sequence. There are *49* challenge-free real video sequences processed with *12* different types of effects and *5* different challenge levels, which result in 2,989 (49*12*5+49) video sequences. Moreover, there are *49* synthesized video sequences processed with *11* different types of effects and *5* different challenge levels, which leads to 2,744 (49*11*5+49) video sequences. In total, there are 5,733 video sequences.

2.2 Video File Name Format

The name format of the provided video sequences is as follows:

"`sequenceType_sequenceNumber_challengeSourceType_challengeType_challengeLevel.txt`".

<u>`sequenceType`:</u>
01 - Real data
02 - Unreal data

<u>`sequenceNumber`:</u>
A number in between [01 - 49]

<u>`challengeSourceType`:</u>
00 - No challenge source (which means no challenge)
01 - After affect

`challengeType`:
00 - No challenge
01 - Decolorization
02 - Lens blur
03 - Codec error
04 - Darkening
05 - Dirty lens
06 - Exposure
07 - Gaussian blur
08 - Noise
09 - Rain
10 - Shadow
11 - Snow
12 - Haze

`challengeLevel`:
A number in between [01-05] where 01 is the least severe and 05 is the most severe challenge.

2.3 Test and Training Sequences

We split the video sequences into 70% training set and 30% test set. The sequence numbers corresponding to test set are given below:

[01_04_x_x_x, 01_05_x_x_x, 01_06_x_x_x, 01_07_x_x_x, 01_08_x_x_x, 01_18_x_x_x, 01_19_x_x_x, 01_21_x_x_x, 01_24_x_x_x, 01_26_x_x_x, 01_31_x_x_x, 01_38_x_x_x, 01_39_x_x_x, 01_41_x_x_x, 01_47_x_x_x, 02_02_x_x_x, 02_04_x_x_x, 02_06_x_x_x, 02_09_x_x_x, 02_12_x_x_x, 02_13_x_x_x, 02_16_x_x_x, 02_17_x_x_x, 02_18_x_x_x, 02_20_x_x_x, 02_22_x_x_x, 02_28_x_x_x, 02_31_x_x_x, 02_32_x_x_x, 02_36_x_x_x]

The videos with all other sequence numbers are in the training set.

Note that “x” above refers to the variations listed earlier.

2.4 Annotation Format

The naming format of the provided annotation files is given as follows:

"`sequenceType_sequenceNumber.txt`".

Challenge source type, challenge type, and challenge level do not affect the annotations. Therefore, the video sequences that start with the same sequence type and the sequence number have the same annotations.

`sequenceType`:
01 - Real data
02 - Unreal data

`sequenceNumber`:
A number in between [01 - 49]

The format of each line in the annotation file (txt) should be: "frameNumber_signType_llx_lly_lrx_lry_ulx_uly_urx_ury".

`frameNumber`:
A number in between [001-300]

`signType`:
01 - speed_limit
02 - goods_vehicles
03 - no_overtaking
04 - no_stopping
05 - no_parking
06 - stop
07 - bicycle
08 - hump
09 - no_left
10 - no_right
11 - priority_to
12 - no_entry
13 - yield
14 - parking

llx:	lower left edge,	x coordinate of bounding box
lly:	lower left edge,	y coordinate of bounding box
lrx:	lower right edge,	x coordinate of bounding box
lry:	lower right edge,	y coordinate of bounding box
ulx:	upper left edge,	x coordinate of bounding box
uly:	upper left edge,	y coordinate of bounding box
urx:	upper right edge,	x coordinate of bounding box
ury:	upper right edge,	y coordinate of bounding box

The *top left* corner of the images corresponds to the origin of the coordinate system.

2.5 Allowed Coding Languages and Platforms

- Matlab
- Python
- C++

Please provide detailed description if you utilize any libraries and toolboxes along with specific instructions.

2.6 Evaluation Criteria

Algorithms will be ranked based on their precision-recall curve. Brief descriptions of true positive, false positive, true negative, and false negative are provided below.

True Positive (TP) = sign type is correctly identified with at least %50 overlap with the ground truth annotation

False Positive (FP) = a non-sign region is identified as a sign or a sign type is identified incorrectly or the overlap between estimated and ground truth bounding box is less than 50%

True Negative (TN) = a non-sign region is correctly identified as a non-sign region

False Negative (FN) = sign is identified as a non-sign region

2.7 Submission Package

Each participating team must submit the following:

I. A report in the form of an IEEE conference paper up to 6 pages, on the technical details of the methods used, programs developed, and results. Check the IEEE SPS website for a template.

II. Estimated detection files for each test sequence. These files should be in a zip folder named as `detections.zip`, which should include all the estimated detection files adheres to the following naming convention: "`sequenceType_sequenceNumber_challengeSourceType_challengeType_challengeLevel.txt`". These files should follow the format described in Section 2.2 and Section 2.4.

III. All codes with detailed comments and readme files that are required to produce files in `detections.zip`.

Note: The submission package should be carefully prepared so that it can be run on any PC by following the instructions in the submission.

2.8 Important Rules and Reminders

- Teams should be formed of three to ten undergraduate students, at most one graduate student (as a tutor/mentor), and one faculty member (as the supervisor). At least three of the undergraduate team members must be either IEEE SP members or student members. Participant teams that do not satisfy these formation requirements will be disqualified.
- Any algorithm that utilizes future frames for previous prediction will be disqualified.
- Any algorithm that utilizes testing labels in the final evaluation will be disqualified.
- Any algorithm that utilizes testing sequences or labels in the training will be disqualified.
- The submissions should include detailed instructions and necessary codes to replicate the results. Otherwise, the participants can be disqualified.
- Reproducing results including training, testing, or any other processes should not exceed a reasonable amount of time that allow us to evaluate all submissions within the given time window.
- We plan to provide tips and resources along the competition, which will be announced in https://ghassanalregib.com/vip-cup/. Please, stay tuned!

3 References

[1] Bureau of Transportation Statistics, "Economic Impact on Transportation," http://www.rita.dot.gov/bts/programs/freight_transportation/html/transportation.html.
[2] D. Gavrila, "Traffic sign recognition revisited," in Mustererkennung 1999, 21. DAGM-Symposium, London, UK, UK, 1999, pp. 86–93, Springer-Verlag.
[3] A. Mogelmose, M. M. Trivedi, and T. B. Moeslund, "Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey," IEEE Transactions on Intelligent Transportation Systems, vol. 13, no. 4, pp. 1484–1497, Dec 2012.
[4] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "The german traffic sign recognition benchmark: A multi-class classification competition," in Neural Networks (IJCNN), The 2011 International Joint Conference on, July 2011, pp. 1453–1460.
[5] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "Man versus computer: Benchmarking machine learning algorithms for traffic sign recognition," Neural Networks, vol. 32, pp. 323 – 332, 2012, Selected Papers from fIJCNNg 2011.
[6] R. Timofte, K. Zimmermann, and L. V. Gool, "Multiview traffic sign detection, recognition, and 3d localisation," in Applications of Computer Vision (WACV), 2009.
Workshop on, Dec 2009, pp. 1–8.
[7] F. Larsson and M. ' Felsberg, "Using fourier descriptors and spatial models for traffic sign recognition," in Proceedings of the 17th Scandinavian Conference on Image Analysis, Berlin, Heidelberg, 2011, SCIA'11, pp. 238–249, Springer-Verlag.
[8] C. Grigorescu and N. Petkov, "Distance sets for shape filters and shape recognition," IEEE Transactions on Image Processing, vol. 12, no. 10, pp. 1274–1286, Oct 2003.
[9] R. Belaroussi, P. Foucher, J. P. Tarel, B. Soheilian, P. Charbonnier, and N. Paparoditis, "Road sign detection in images: A case study," in Pattern Recognition (ICPR), 2010 20th International Conference on, Aug 2010, pp. 484–488.
[10] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, and C. Igel, "Detection of traffic signs in real-world images: The german traffic sign detection benchmark," in Neural Networks (IJCNN), The 2013 International Joint Conference on, Aug 2013, pp. 1–8.
[11] R. Timofle and L. Van Gool, "Sparse representation based projections," in BMVC, 2011.
[12] M. Mathias, R. Timofte, R. Benenson, and L. Van Gool, "Traffic sign recognition - how far are we from the solution?" in Neural Networks (IJCNN), The 2013 International Joint Conference on, Aug 2013, pp. 1–8.