**Introduction**

- In this paper, we investigate the robustness of traffic sign recognition algorithms under challenging conditions.
- Existing datasets are limited in terms of their size and challenging condition coverage, which motivated us to generate the Challenging Unreal and Real Environments for Traffic Sign Recognition (CURE-TSR) dataset. It includes more than two million traffic sign images that are based on real-world and simulator data.
- We benchmark the performance of existing solutions in real-world scenarios and analyze the performance variation with respect to challenging conditions.
- We show that challenging conditions can decrease the performance of baseline methods significantly.
- We also investigate the effect of data augmentation and show that utilization of virtual data along with real-world data enhances the average recognition performance.

**Dataset Generation and Visualization**

- In order to create realistic challenging scenarios, we generate challenging condition types and levels for entire scenes.
- The traffic signs are then cropped from such scenes.
- Each row in the figure below corresponds to a challenging condition and each column corresponds to a certain level of the condition.

**Dataset - Number of Images**

<table>
<thead>
<tr>
<th>Image Types</th>
<th>Training</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real Challenge-Free</td>
<td>7,292</td>
<td>3,334</td>
</tr>
<tr>
<td>Real Challenge</td>
<td>437,520</td>
<td>200,040</td>
</tr>
<tr>
<td>Virtual Challenge-Free</td>
<td>19,610</td>
<td>8,210</td>
</tr>
<tr>
<td>Virtual Challenge</td>
<td>1,078,550</td>
<td>451,550</td>
</tr>
</tbody>
</table>

**Dataset - Usage**

- To study traffic sign recognition under challenging conditions.
- Studying data augmentation techniques when traditional augmentation techniques fail.
- Studying regularization techniques when data is imperfect.
- Domain adaptation between virtual and real-world images.

**Dataset - Performance Benchmarks**

- Intensity-Softmax, RGB-Softmax, RGB-CNN, HoG-SVM
- Intensity-SVM, RGB-SVM, HoG-Softmax

![Performance Benchmarks](image)

**Data Augmentation**

**Traditional Strategies**

- Intensity-Softmax, RGB-Softmax, RGB-CNN, HoG-SVM
- Intensity-SVM, RGB-SVM, HoG-Softmax

**Real Challenging data**

**Virtual data**

**Contributions**

- We introduce the most comprehensive publicly-available traffic sign recognition dataset with controlled challenging conditions.
- We provide a detailed analysis of the benchmarked algorithms in terms of their recognition performance under challenging conditions thereby identifying the vulnerabilities of algorithms.
- We provide images that originate from captured sequences as well as synthesized sequences, that lead to a better understanding of the relationship between real-world and virtual data in terms of algorithmic performance. This understanding can be utilized to generate algorithmically invariant virtual datasets and minimize the need for real-world data collection that require significant resources.
- We use diverse data augmentation methods and show that utilization of limited virtual images along with real-world data can enhance the recognition performance even when the domain difference is not addressed.

**Dataset – Performance Benchmark Observations**

- Challenging conditions perturb original representation space to deceive classifier.
- Different challenges have different and characteristic degradation slopes.
- Decolorization, Darkening and Shadow have relatively consistent performance across challenge levels and algorithms.
- Lens blur, Codec error, Exposure and Gaussian blur show severe performance degradation.