Objective
To improve the accuracy of semantic segmentation by data augmentation approach to balance the label distribution using Generative Adversarial Networks (GANs).

Context
• The label distribution in most datasets are unbalanced.
• Manual pixel-level annotations of dataset have a huge cost.

Previous Work
• Photographic image generation by conditional GANs, e.g. Ting-Chun Wang et al. [1].
• Image-level label balance to improve the emotion classification accuracy [2].

Contributions
• Data augmentation pipeline by using GANs to generate supplementary data for semantic segmentation task.
• New method for image augmentation at pixel-level.

Feasibility Analysis
Figure 1. Average segmentation accuracy of top 5 ranked models on Cityscapes website and the proportion of each class in Cityscapes dataset.

Comparing those classes with low appearance frequency and those have low segmentation accuracy, we find out that two groups are highly overlapped. In other word, it is possible to balance the dataset and further improve segmentation accuracy by increasing the appearance frequency of some specific classes.

Methodology

Step 1: Train data generator
• Pix2pix HD [3] model is used to generate realistic images on condition of specific semantic label map.
• Minimax algorithm is used to train the generator G and discriminator D.

Step 2: Synthesize training-pair
• To begin with, separating each semantic label map in training set according to class of labels to form single-class label maps.
• Then, recombining these single-class label maps to form reconstructed label maps (Using three different reconstructed ways).
• Lastly, data generator transfer reconstructed label maps to corresponding photographic images.

Step 3: Segmentation validation
• Train semantic segmentation network with/without the generated supplementary data, we can improve the diversity of data and balance the data distribution.
• The results shown that by our method, the mean accuracy of a specific class can increase up to 5.5% and the average segmentation accuracy can increase 2.1%.

Concluding Remarks
• By adding generated label classes to original images as supplementary data, we can improve the diversity of data and balance the data distribution.
• The results shown that by our method, the mean accuracy of a specific class can increase up to 5.5% and the average segmentation accuracy can increase 2.1%.

Experimental Results
Figure 4. Detailed results comparison.

Table 1. Results of different generation method.

<table>
<thead>
<tr>
<th>Method</th>
<th>Baseline</th>
<th>Add single</th>
<th>Add multi</th>
<th>Recon</th>
</tr>
</thead>
<tbody>
<tr>
<td>mIoU</td>
<td>77.31</td>
<td>78.65</td>
<td>78.82</td>
<td>79.41</td>
</tr>
</tbody>
</table>

Table 2. Comparison with traditional data augmentation method and style transfer.

<table>
<thead>
<tr>
<th>Method</th>
<th>Tradition</th>
<th>Style Trans</th>
<th>Recon</th>
</tr>
</thead>
<tbody>
<tr>
<td>mIoU</td>
<td>77.31</td>
<td>78.82</td>
<td>79.41</td>
</tr>
</tbody>
</table>

Reference