Rethinking the Value of Labels for Improving Class-Imbalanced Learning

Yuzhe Yang

Zhi Xu

NeurIPS 2020

http://www.mit.edu/~yuzhe/imbalanced-semi-self.html
Real-world data are often imbalanced (long-tailed)

- Image recognition / Object detection / Semantic segmentation...

Van Horn et al. 2018

Wang et al. 2017

Gupta et al. 2018
Real-world data are often imbalanced (long-tailed)

- Image recognition / Object detection / Semantic segmentation ...

- Critical applications

Autonomous driving

Medical diagnosis

Van Horn et al. 2018

Wang et al. 2017

Gupta et al. 2018
Dilemma: Value of imbalanced labels
Dilemma: Value of imbalanced labels

- Positive value
Dilemma: Value of imbalanced labels

- Positive value

Training data + imbalanced labels > Training data (alone)
Dilemma: Value of imbalanced labels

- Positive value
  - Training data + imbalanced labels > Training data (alone)

- Negative value
  - Imbalanced labels: Head classes vs. Tail classes
Dilemma: Value of imbalanced labels

- Positive value

- Negative value

“Label bias” driven by majority classes
Dilemma: Value of imbalanced labels

- Positive value

How to exploit the value of imbalanced labels?
Semi-supervision & Self-supervision help!
Semi-supervision & Self-supervision help!

- Positive viewpoint

![Diagram](image.png)

Training data + imbalanced labels
Semi-supervision & Self-supervision help!

- Positive viewpoint

![Diagram](image)
- Training data
- Imbalanced labels
- Unlabeled data
Semi-supervision & Self-supervision help!

- Positive viewpoint: Semi-supervised learning using imbalanced labels
Semi-supervision & Self-supervision help!

- Positive viewpoint: Semi-supervised learning using imbalanced labels

  - Training data
  - Imbalanced labels
  - Unlabeled data

- Negative viewpoint

  - Imbalanced labels
Semi-supervision & Self-supervision help!

- Positive viewpoint: Semi-supervised learning using imbalanced labels

- Negative viewpoint: Self-supervised pre-training in the first learning stage
Semi-supervision & Self-supervision help!

- Positive viewpoint: Semi-supervised learning using imbalanced labels
  
  ![Diagram](image)

  - Training data
  - Imbalanced labels
  - Unlabeled data

- Negative viewpoint: Self-supervised pre-training in the first learning stage
  
  ![Diagram](image)

  - Self-supervised pre-training
  - Compatible w/ any base training method

Potential imbalanced labels
Semi-supervised imbalanced learning

Training set

Test set

Standard CE

w/ unlabeled data
Semi-supervised imbalanced learning

Standard CE

Training set
Bad separation for tail classes

Test set
Mixed class boundary

w/ unlabeled data
Semi-supervised imbalanced learning

Training set:
- Standard CE
  - Bad separation for tail classes
- Tail class w/ unlabeled data

Test set:
- Mixed class boundary
- Clearer boundary

w/ unlabeled data
Self-supervised imbalanced learning

Standard CE

w/ self-supervised pre-training

Training set

Test set
Self-supervised imbalanced learning

Training set

Head class 2
Tail classes
Head class 1

Test set

Head class 1
Tail classes
Head class 2

Standard CE

w/ self-supervised pre-training
Self-supervised imbalanced learning

Training set

Test set

Standard CE

Head class 2

Tail classes

Head class 1

Leakage of tail classes

w/ self-supervised pre-training

Head class 1

Tail classes

Head class 2
Self-supervised imbalanced learning

Training set

Test set

Standard CE

Tail classes

Head class 2

Tail classes

Head class 1

w/ self-supervised pre-training

Head class 1

Tail classes

Head class 2

Leakage of tail classes

Less leakage
Consistent performance gains

- Semi-supervised imbalanced learning

<table>
<thead>
<tr>
<th></th>
<th>CIFAR-10-LT</th>
<th>SVHN-LT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard CE</td>
<td>70.36</td>
<td>80.02</td>
</tr>
<tr>
<td>w/ unlabeled data</td>
<td><strong>82.52 (+12.16)</strong></td>
<td><strong>86.98 (+6.96)</strong></td>
</tr>
</tbody>
</table>

- Self-supervised imbalanced learning

<table>
<thead>
<tr>
<th></th>
<th>CIFAR-10-LT</th>
<th>CIFAR-100-LT</th>
<th>ImageNet-LT</th>
<th>iNaturalist 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard CE</td>
<td>70.36</td>
<td>38.32</td>
<td>38.4</td>
<td>60.7</td>
</tr>
<tr>
<td>w/ SSP</td>
<td><strong>76.53 (+6.17)</strong></td>
<td><strong>43.06 (+4.74)</strong></td>
<td><strong>45.6 (+7.2)</strong></td>
<td><strong>64.4 (+3.7)</strong></td>
</tr>
</tbody>
</table>
Consistent performance gains

- Semi-supervised imbalanced learning

<table>
<thead>
<tr>
<th></th>
<th>CIFAR-10-LT</th>
<th>SVHN-LT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard CE</strong></td>
<td>70.36</td>
<td>80.02</td>
</tr>
<tr>
<td>w/ unlabeled data</td>
<td>82.52 (+12.16)</td>
<td>86.98 (+6.96)</td>
</tr>
</tbody>
</table>

Superior improvements across various datasets!
(more results in paper)

<table>
<thead>
<tr>
<th></th>
<th>CIFAR-10-LT</th>
<th>CIFAR-100-LT</th>
<th>ImageNet-LT</th>
<th>iNaturalist 2018</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard CE</strong></td>
<td>70.36</td>
<td>38.32</td>
<td>38.4</td>
<td>60.7</td>
</tr>
<tr>
<td>w/ SSP</td>
<td>76.53 (+6.17)</td>
<td>43.06 (+4.74)</td>
<td>45.6 (+7.2)</td>
<td>64.4 (+3.7)</td>
</tr>
</tbody>
</table>
Summary

- **Unlabeled data** helps imbalanced learning via a semi-supervised manner
- **Self-supervised pre-training** can substantially improve imbalanced performance
- Theoretical analysis + large-scale extensive experiments

Check out our code and models at...

- Code (relevant data + pretrained models): [https://github.com/YyzHarry/imbalanced-semi-self](https://github.com/YyzHarry/imbalanced-semi-self)