Introduction

- Real-world datasets often exhibit a long-tailed class distribution.
- It is important to address this imbalance issue for robust real-world application.
- Otherwise classifiers tend to be biased towards the dominating classes and perform poorly on the tail classes.

<table>
<thead>
<tr>
<th>Transfer Learning</th>
<th>Re-weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within-class diversity</td>
<td>✓</td>
</tr>
<tr>
<td>Generalization</td>
<td>✓</td>
</tr>
<tr>
<td>Adaptive to data</td>
<td>✓</td>
</tr>
<tr>
<td>Lightweight</td>
<td>✓</td>
</tr>
</tbody>
</table>

### Class Volume

1. We define class volume $v_i$ to estimate the class boundary. $\sigma_{i,k}$ is the standard deviation of the $k$-th logit for samples in class $i$.
2. The re-weighting factor $\alpha_i$ can be viewed as a training sample’s power of influence to the decision boundary. The aggregated power of influence $\alpha_i n_j$ estimates the distance from a class to the decision boundary. The re-weighting factor is assigned such that the distance to the decision boundary is proportional to the radius of the class volume.
3. Re-weighting factors are then normalized using $\beta$ to preserve the expected loss: $E(\alpha \ell) = E(\ell)$

### Results

- The overall accuracy is improved.
- The classification bias is reduced.

Conclusion

- Class volume concept captures the bias introduced by both class imbalance and within-class diversity imbalance.
- Generalizable to all existing architectures.
- Adaptive to data without introducing any additional hyper-parameter.
- Lightweight. Does not add significant implementation effort or computation overhead.