Class-Balanced Loss

Based on Cui et al., CVPR 2019

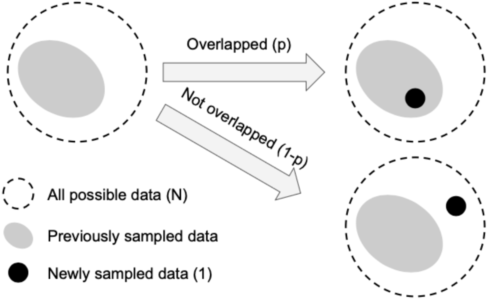
Cui Y, Jia M, Lin T Y, et al. Class-balanced loss based on effective number of samples[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 9268-9277.

1. Motivation & Recursive Derivation

Problem: In long-tailed datasets, head classes dominate training. We need a smoother re-weighting than $1/n_i$.

Idea: Effective Number of Samples

Each sample covers a small volume (1) in feature space. As samples overlap, the *effective* coverage E_n grows slower than n.



1. Motivation & Recursive Derivation

Variables:

N : total volume of possible data,

E_n : expected covered volume after n samples,

$p = E_{n-1}/N$.

Expected coverage:

$$
\begin{aligned}
E_n &= pE_{n-1} + (1-p)(E_{n-1}+1) \\
&= E_{n-1} + (1-p) \\
&= E_{n-1} + \left(1 - \frac{E_{n-1}}{N}\right) \\
&= 1 + \underbrace{\left(1 - \frac{1}{N}\right)}_{\beta} E_{n-1}
\end{aligned}
$$

$$\boxed{E_n = 1 + \beta E_{n-1}}, \quad \beta = \frac{N-1}{N}$$

2. Closed-Form & Class-Balanced Loss

Solving the recursion:

$$E_n = 1 + \beta + \beta^2 + \cdots + \beta^{n-1} = \frac{1-\beta^n}{1-\beta}$$

Interpretation:

- $\beta \to 1$: little overlap $\Rightarrow E_n \approx n$
- β small: heavy overlap $\Rightarrow E_n$ saturates

Class-Balanced Weight

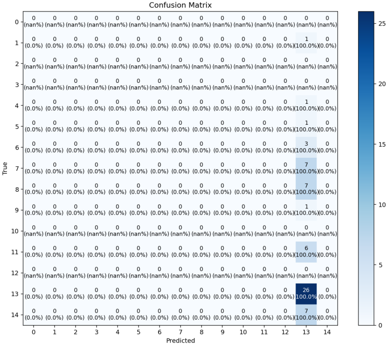
For class i with n_i samples:

$$w_i = \frac{1-\beta}{1-\beta^{n_i}}$$

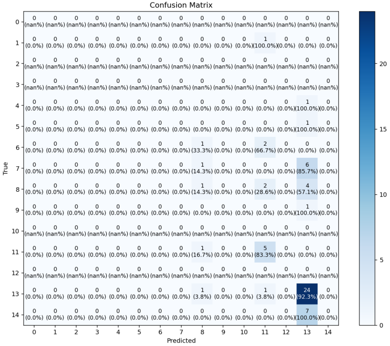
$$\text{CB-CE}(p, y) = \frac{1-\beta}{1-\beta^{n_y}} \left[-\log p_y\right]$$

$\Rightarrow \beta$ smoothly interpolates between
no weighting ($\beta = 0$) and inverse-frequency weighting ($\beta \to 1$).

3. experiment result



(a) Before



(b) After