Boosting Semi-Supervised Learning by Exploiting All Unlabeled Data [CVPR 2023]

FixMatch's problem: suffer from the waste of complicated examples
Our contributions: Entropy Meaning Loss(EML), Adaptive Negative Learning(ANL), FullMatch

Preliminaries

The original consistency loss in SSL:

$$\sum_{i=1}^{B}(\left\|p_m(y|\omega(\mu^{(i)}) - p_m(y|\phi(\mu^{(i)}))\right\|_2^2)$$

The improved consistency loss function in FixMatch:

$$\frac{1}{B}\sum_{i=1}^{B}\mathbb{1}(max(Q^{(i)}) \geq \tau)H(\hat{Q^{(i)}}, P^{(i)})$$

$Q^{(i)} = p_m(y|\omega(\mu^{(i)}))$ and $P^{(i)} = p_m(y|\phi(\mu^{(i)}))$, $\hat{Q^{(i)}} = argmax(Q^{(i)})$ is the hard target.

Entropy Meaning Loss (EML)

We propose Entropy Meaning Loss to allocate more samples with pseudo-labels.

$Q^{(i)} = [q_1^{(i)}, ..., q_C^{(i)}]$: the prediction vector for the weakly-augmented version of sample i.

$S^{(i)} = [s_1^{(i)}, ..., s_C^{(i)}] \subseteq \{0,1\}^C$: a binary vector denoting the selected labels, $s_c^{(i)} = \mathbb{1}(q_c^{(i)} \geq \tau)$ and when it equals to 1, c is selected as a target class.

$U^{(i)} = [u_1^{(i)}, ..., u_C^{(i)}]$: $u_c^{(i)} = 1$ denotes class c is a non-target class and sample i is assigned a pseudo-label, $u_c^{(i)} = \mathbb{1}(max(Q^{(i)}) \geq \tau) \cdot \mathbb{1}(s_c^{(i)} = 0)$

$P^{(i)} = [p_1^{(i)}, ..., p_C^{(i)}]$:the prediction confidence vector on the strongly-augmented of sample i.

$p_{tc}^{(i)}$: the confidence of the target class, which will gradually converge to the given label.

For certain challenging examples, the competition between confusion classes and target class always leads to a failure in generating high-confidence prediction.
we impose an additional constraint on the rest of the categories (i.e., all non-target classes) to allow them to share the remaining confidence equally to avoid any class competition with the target class.
The label of the non-target classes:

$$y_c^{(i)} = \frac{1 - \mathbb{1}(u_c^{(i)} = 0) \cdot p_c^{(i)}}{\sum_c \mathbb{1}(u_c^{(i)} = 1)}$$

Our proposed Entropy Meaning Loss (EML) can be defined as:

$$\mathcal{L}_{eml} = -\frac{1}{BC}\sum_{i=1}^{B}\sum_{c=1}^{C} u_c^{(i)} \cdot [y_c^{(i)} log(p_c^{(i)}) \\ +(1-y_c^{(i)})(log(1-p_c^{(i)}))]$$

EML can cooperate with L_{us} to further promote the confidence of target class while constraining the distribution of the non-target classes.

Adaptive Negative Learning (ANL)

It is easy to produce ambiguous predictions on complicated scenarios, these examples are difficult to be assigned pseudo-label. Thus, we allocate an additional label with less noise to leverage these examples.

An ideal approach is to exploit an additional dataset to evaluate the top-k performance, thereby calculating a suitable k value so that the top-k error rate is close to zero. In this work, we present a scheme to approximately evaluate the top-k performance, referred to as Adaptive Negative Learning (ANL).

$$k = \arg\min_{\theta\in[2,C]} (Acc(P_t, \hat{Q}_t, \theta) = 100\%)$$

$\hat{Q}_t = \arg\max (Q, t)$ is $temp$ labels at step t

We assign negative pseudo-labels to categories ranked after top-k on the prediction distributions of the weakly-augmented version.

$$s_c^{(i)} = \mathbb{1}[q_c^{(i)} \geq \tau] + \mathbb{1}[Rank(q_c^{(i)}) > k]$$

where Rank is a category sorting function based on confidence scores in the descending order. The adaptive negative learning loss L_{anl} can be formulated as:

$$-\frac{1}{B}\sum_{i=1}^{B}\sum_{c=1}^{C} \mathbb{1}[Rank(q_c^{(i)}) > k]\, log(1-p_c^{(i)})$$

FullMatch

By integrating the Entropy Meaning Loss (EML) and Adaptive Negative Learning (ANL) into FixMatch, we propose an advanced SSL algorithm named FullMatch.

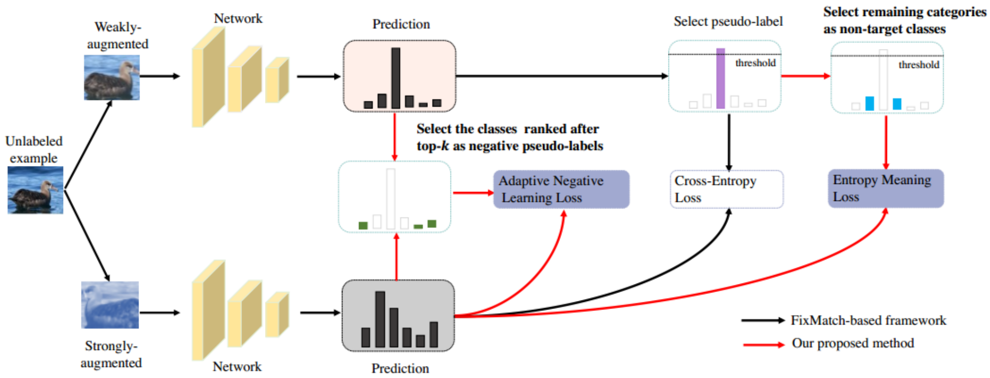

Figure 5. Overview of the proposed FullMatch. First, we allocate the negative pseudo-label (green bar) for all unlabeled data with the proposed Adaptive Negative Learning. Then, if the highest probability is above the predefined threshold (dotted line), we will assign the pseudo-label (purple bar) just like FixMatch, but we optimize further remaining non-target classes (blue bar) via the proposed Entropy Meaning Loss. The black line indicates the existing FixMatch-based methods, and the red line is our proposed method. (Best viewed in color).

$$\mathcal{L}_{sum} = \mathcal{L}_s + \mathcal{L}_{us} + \alpha \cdot \mathcal{L}_{anl} + \beta \cdot \mathcal{L}_{eml}$$