Structured Sparsity for Semi-Supervised Learning....

Ankan Bansal Joaquin Zepeda

1 Introduction

2 Approach

q_t is the output probability distribution over the classes. T is the batch size. $\mathcal{L}_{cls}$ is the cross-entropy loss.

2.1 Mean Entropy Penalty

$$\mathcal{L}_{MEL} = \frac{1}{T}\sum_{t=1}^{T} H(q_t) \tag{1}$$

where $H(.)$ represents the entropy.

2.2 Negative Batch Entropy Penalty

$$\mathcal{L}_{NBEL} = -H(\frac{1}{T}\sum_{t=1}^{T} q_t) \tag{2}$$

2.3 Locality Penalty

Class Activation Maps, C. Feature maps, F. Weights of the final layer, W. $C = A \times W$. The class with maximum probability is i. Calculate locality loss on C_i.

$$\mathcal{L}_{Loc} = \sum_{j=1}^{4} l_j(C_i) \tag{3}$$

where the sum is over the four groups (left-right, right-left, top-bottom, and bottom-top).

Consider one of these cases. Let, $C_i \in \mathbb{R}^{N\times N}$. We define the left-right group norms as $G_{lr} = [\left\|C_i^k\right\|_2 \mid C_i^k = C_i(1:k, 1:N), k = 1 \dots N]$. Then, $l_1(C_i) = \left\|G_{lr}\right\|_1$. Other group losses can be defined similarly.

Then the total loss is given as:

$$\mathcal{L} = \mathcal{L}_{cls} + \alpha\mathcal{L}_{MEL} + \beta\mathcal{L}_{NBEL} + \gamma\mathcal{L}_{Loc} \tag{4}$$

3 Results

References

1. Author, F.: Article title. Journal **2**(5), 99–110 (2016)