MixMatch

A Holistic Approach to Semi-Supervised Learning

By : Soona Osman

INTRODUCTION

SSL objective : to make use of unlabeled data , by adding a loss term, which falls into one of three classes:

- consistency regularization: each unlabeled data point should be classified the same as its augmentation.
- entropy minimization : implement the clustering assumption by reducing the classes overlapping .
- generic regularization: imposing a constraint on a model to make it harder to memorize the training data and then generalize better to unseen data.

Mix Match combine all this together!

How??

- consistency regularization ---> By introducing data augmentation both in the labeled and unlabeled data.
- entropy minimization ---> By the use of label guessing and sharpening in the unlabeled data.
- generic regularization ---> Mixup.

Methodology

- **Data Augmentation**:

xˆb = Augment(xb)

for k = 1 to K do :

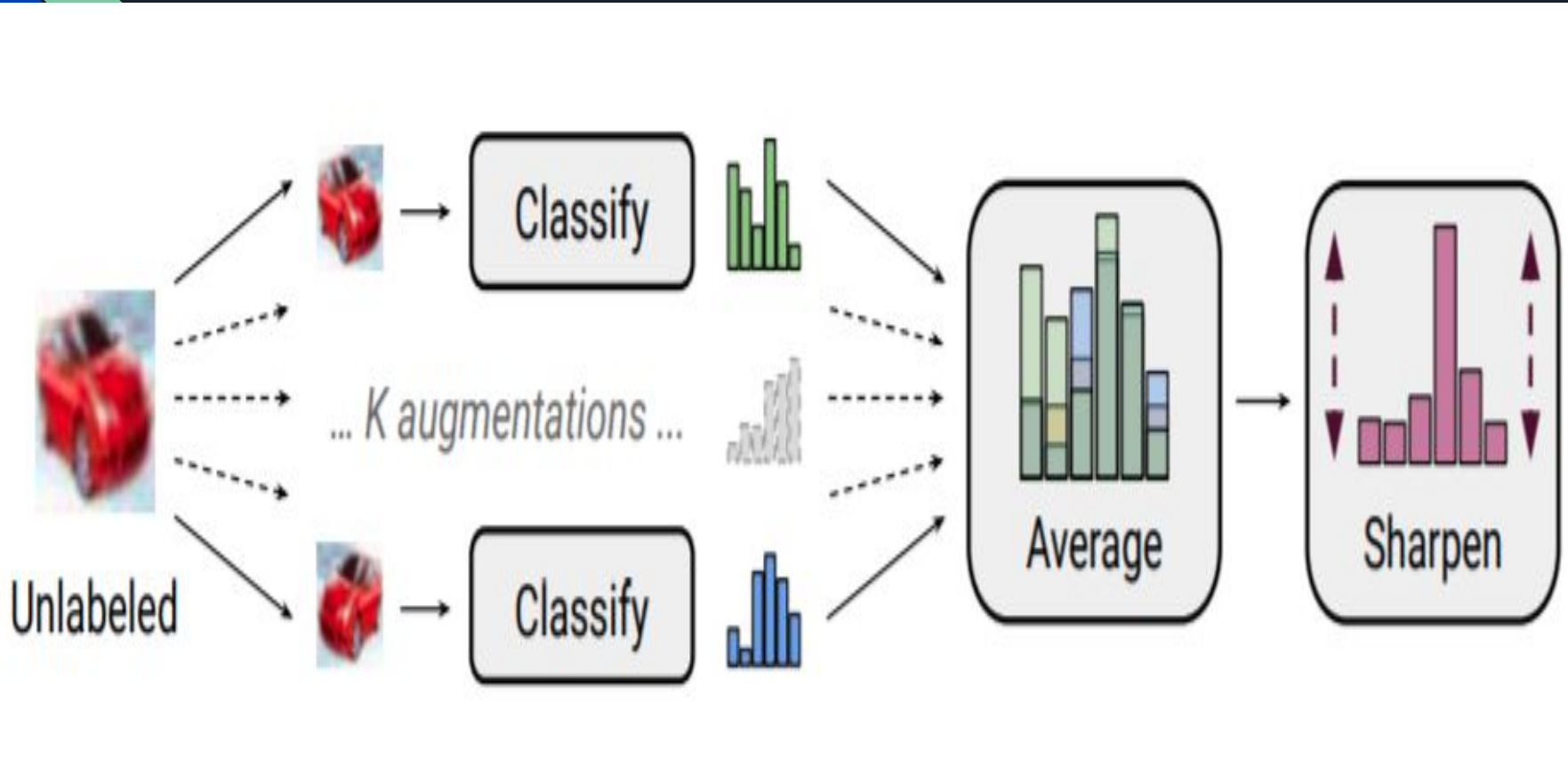
uˆb,k = Augment(ub)

Methodology

Label guessing : averaging the model predictions of K augmentations of ub.

Sharpening :to reduce the entropy of the label distribution.

Label guessing process



Methodology

- **Mixup**:

$$\lambda \sim \text{Beta}(\alpha, \alpha)$$

$$\lambda = \max(\lambda, 1 - \lambda)$$

$$x\_ = \lambda x1 + (1 - \lambda)x2$$

$$p\_ = \lambda p1 + (1 - \lambda)p2$$

Methodology

Loss Function :

Labeled loss: Cross entropy loss.

Unlabeled loss: L2 loss

Final loss = labeled + unlabeled

Experiments

Model:

Wide ResNet-28

Dataset:

CIFAR-10,CIFAR-100,SVHN and STL-10

Findings

num_label	Epoch	Iters	Accuracy
256	10	20	10%
256	10	64	12%
256	10	544	20%

Loss curve

Thanks..