

A decorative graphic on the left side of the slide consisting of overlapping geometric shapes. It includes a blue parallelogram, a light green parallelogram, and a dark grey parallelogram, all arranged in a layered, diagonal fashion.

# 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:**

$\hat{x}^b = \text{Augment}(x^b)$

for  $k = 1$  to  $K$  do :

$\hat{u}^{b,k} = \text{Augment}(u^b$

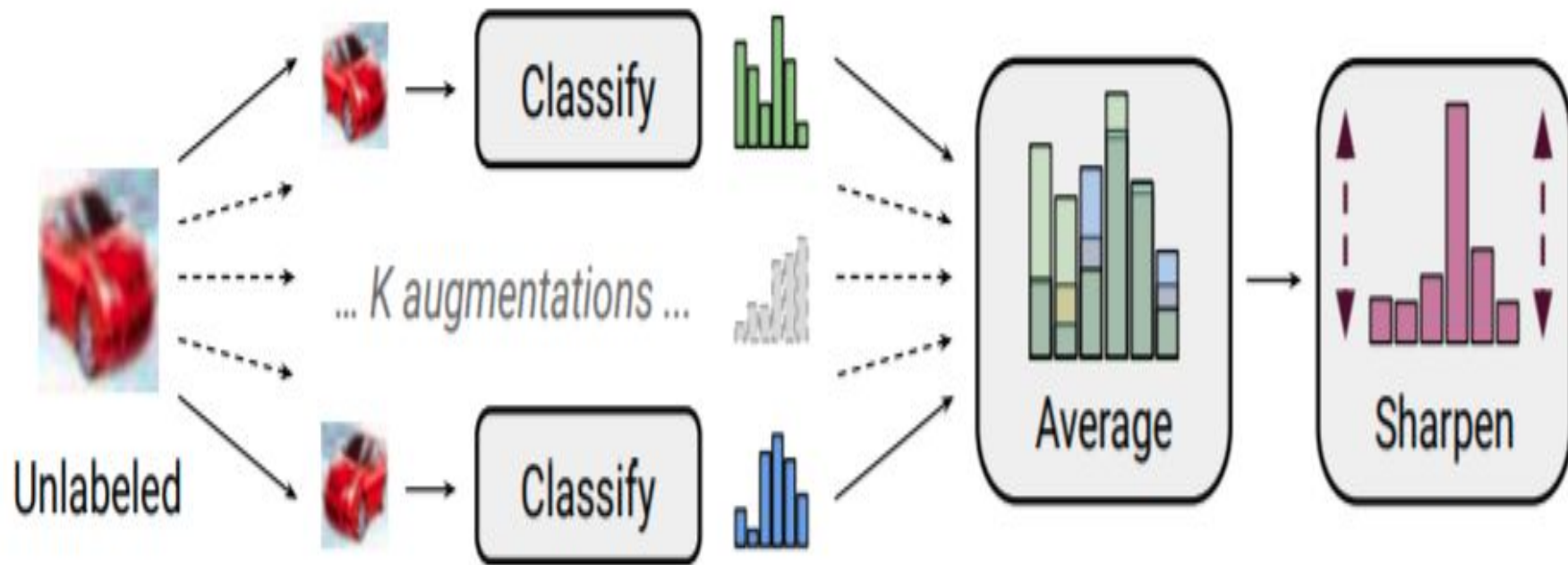


# Methodology

**Label guessing** : averaging the model predictions of  $K$  augmentations of  $u_b$ .

**Sharpening** : to reduce the entropy of the label distribution.

# Label guessing process





# Methodology

- Mixup:

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

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

$$\mathbf{x}_- = \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2$$

$$\mathbf{p}_- = \lambda \mathbf{p}_1 + (1 - \lambda) \mathbf{p}_2$$



## 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	20	64	12%
256	20	544	20%



Thanks..