Semi-Supervised Image Classification

Trends and Applications of Computer Vision

Alessandro Zinni, Filippo Momesso, Gabriele Padovani, Thomas De Min

Table of Contents

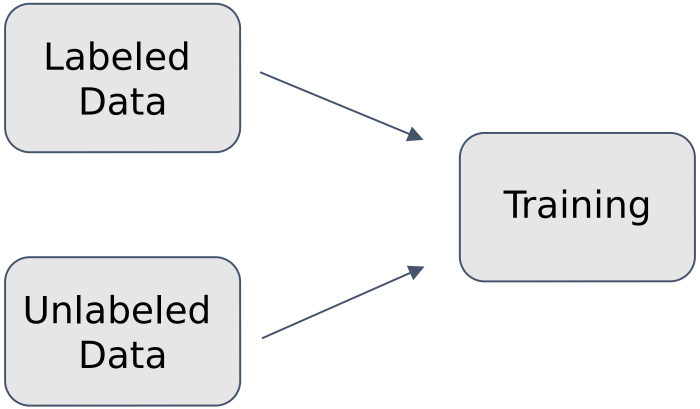
- Importance of the topic;
- Background on Semi-supervised image classification;
- CIFAR-100 failure cases;
- Cross-domain evaluation;
- Open Set Semi-supervised learning;
- Conclusions & Outlook.

Importance of the topic

- **Lots of data** required to train a Neural Network;
- Not always feasible to get large quantities of **annotated** examples;
- Unlabeled data is **cheaper** and does not require human annotations.

Importance of the Topic

- Exploit latent information contained in unlabeled data;
- Effective training even when **few labeled samples** are present.



Background

- The dataset is composed of **labeled** and **unlabeled** samples;
- Labeled images are usually much less wrt. unlabeled ones;
- **Objective**:
  - **Improve** over the baseline trained on only labeled subset;
  - **Exploit** the unlabeled subset to learn better representations.

$$\min_{\theta} \underbrace{\sum_{x \in X_L, y \in Y_L} \mathcal{L}_s(x, y, \theta)}_{\text{supervised loss}} + \alpha \underbrace{\sum_{x \in X_U} \mathcal{L}_u(x, \theta)}_{\text{unsupervised loss}} + \beta \underbrace{\sum_{x \in X} \mathcal{R}(x, \theta)}_{\text{regularization}}$$

Background - Common Methods

Pseudo Labeling:

- Treat high-confidence predictions on unlabeled samples as pseudo labels.

Consistency Regularization:

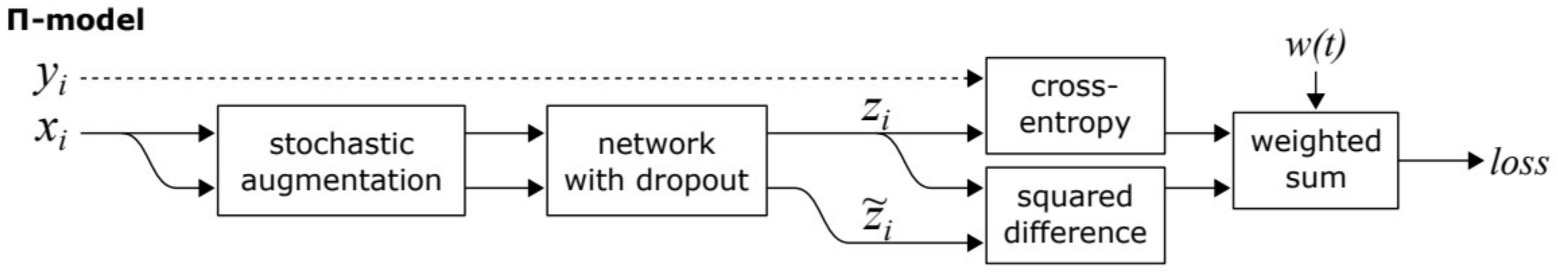
- Force similar outputs to different perturbations of the same sample.

Hybrid Methods:

- Combine the methods above.

Methods Recap - Π-Model

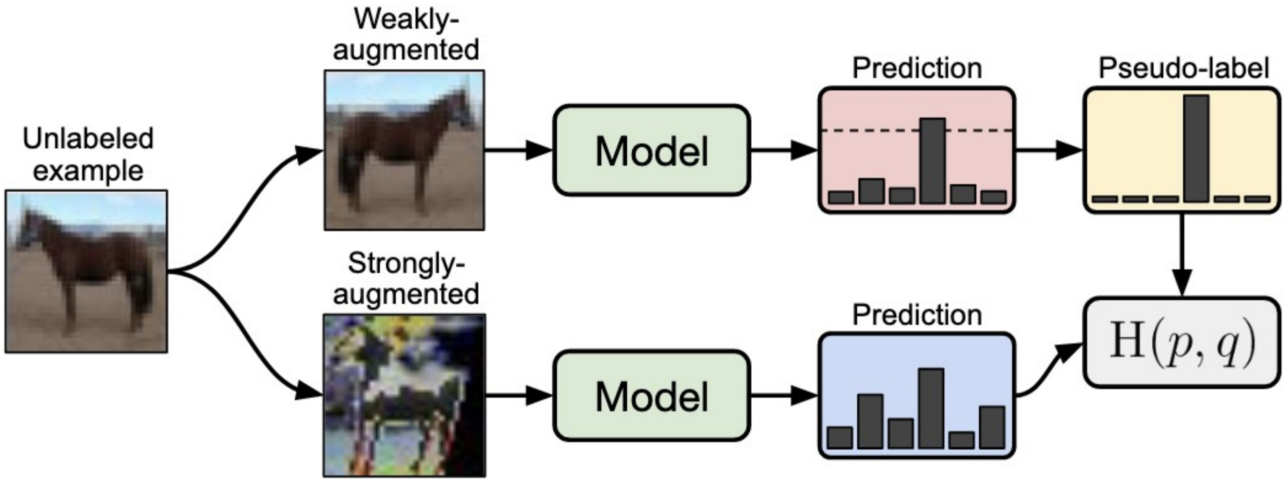
Basic consistency regularization method, poor performance related to other approaches.



Samuli Laine and Timo Aila. “Temporal ensembling for semi-supervised learning” (2016).

Methods Recap - FixMatch

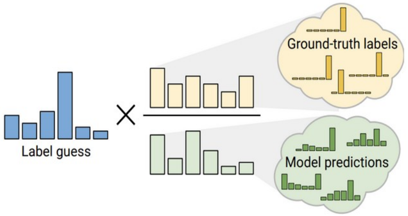
Hybrid method, consistency regularization and pseudo-labeling.



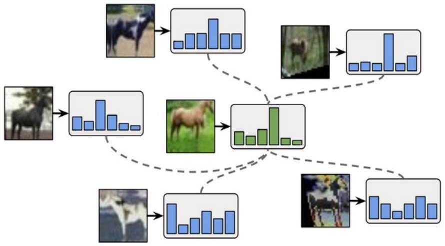
Kihyuk Sohn et al. FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence. 2020.

Methods Recap - ReMixMatch

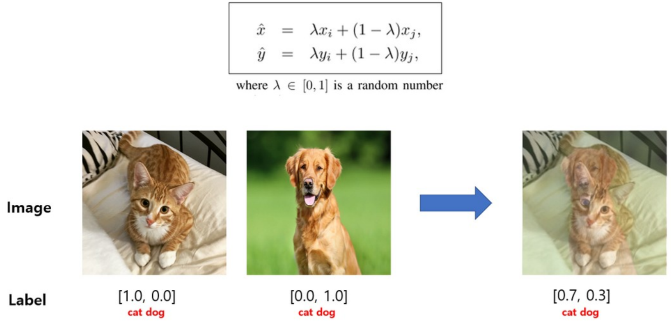
Hybrid method based on MixMatch.



+



+



Distribution alignment

Augmentation Anchoring

MixUp

David Berthelot et al. “ReMixMatch: Semi-Supervised Learning with Distribution Alignment and Augmentation Anchoring”. (2019)

UNIVERSITY OF TRENTO

USB

CONTRIBUTORS 10 | FORKS 56 | STARS 442 | ISSUES 6 OPEN

USB

Unified SSL Benchmark

USB: A Unified Semi-supervised learning Benchmark for CV, NLP, and Audio Classification

Paper · Benchmark · Demo · Docs · Issue · Blog · Blog (Chinese) · Video · Video (Chinese)

Wang et al. USB: A Unified Semi-supervised Learning Benchmark for Classification. (2022)

TABLE I
ACCURACIES ON CIFAR-100 WITH 200 AND 400 LABELS.

	Method	200 labels	400 labels
USB	Pseudo-Labeling	$66.84_{\pm 1.20}$	$74.71_{\pm 0.67}$
	Π-model	$63.76_{\pm 0.27}$	$73.51_{\pm 0.64}$
	Mean Teacher	$64.39_{\pm 0.38}$	$74.03_{\pm 0.37}$
	VAT	$68.39_{\pm 1.37}$	$78.71_{\pm 0.32}$
	MixMatch	$62.57_{\pm 0.58}$	$73.83_{\pm 0.24}$
	ReMixMatch	$\mathbf{79.15}_{\pm 1.42}$	$\mathbf{83.20}_{\pm 0.59}$
	FixMatch	$69.55_{\pm 0.65}$	$80.52_{\pm 0.93}$
	Dash	$69.81_{\pm 1.34}$	$81.10_{\pm 0.42}$
	CRMatch	$70.57_{\pm 1.11}$	$81.50_{\pm 0.26}$
	FlexMatch	$72.92_{\pm 0.90}$	$82.33_{\pm 0.66}$
Ours	Pseudo-Labeling	66.84	74.05
	Π-Model	64.06	74.33
	VAT	67.95	77.80
	ReMixMatch	72.83	**80.76**
	FixMatch	69.30	77.17
	FlexMatch	**73.32**	76.03
Lower Bound	*Supervised*	$64.37_{\pm 0.07}$	$73.92_{\pm 0.50}$
Upper Bound	*Fully Supervised*	$91.56_{\pm 0.07}$	

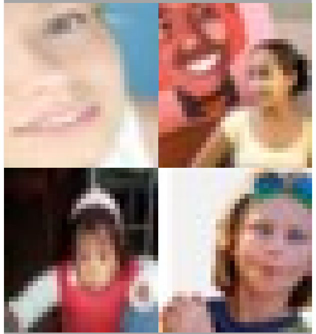
CIFAR-100 Results

- Comparison between USB results and ours;
- Some methods are worse due to insufficient computational power.

CIFAR-100 Failure Cases

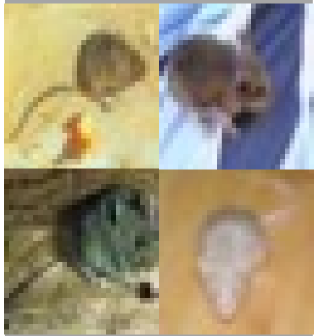
- **FixMatch** fine-tuned for 20 epochs on 400 labeled images and 50000 unlabeled ones:
  - 4 labels per class (0.8%);
- Accuracy **77.17 %**:
  - Lower Bound 73.92;
  - Upper Bound 91.56;

CIFAR-100 - Failure Cases



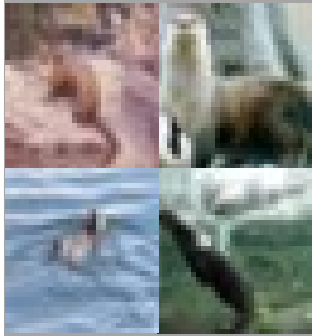
GT: Girl

Pred: Woman



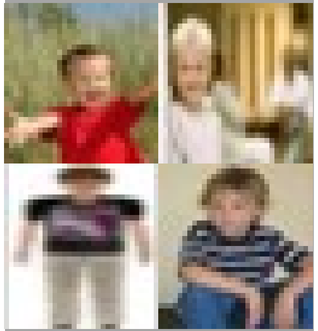
GT: Mouse

Pred: Hamster



GT: Otter

Pred: Seal



GT: Boy

Pred: Woman

Cross-Domain Evaluation

Rationale:

- Assumption in SSL is that labeled and unlabeled data come from the **same distribution** (otherwise results might **degrade**);
- The impact of the **domain shift** is proportional to the amount of unlabeled data;
- Performance may change depending on the approach.

Cross-Domain Evaluation - Setting

- **Dataset**: Adaptiope;
- Three domains:
  - Product;
  - Real World;
  - Synthetic;
- We use:
  - Labeled data -> “Product” domain;
  - Unlabeled data -> “Real World” domain.

(a) bicycle helmet (b) backpack (c) hot glue gun (d) diving fins (e) tyrannosaurus (f) acoustic guitar (g) boxing gloves (h) over-ear headphones

Tobias Ringwald, Rainer Stiefelhagen. Adaptiope: A Modern Benchmark for Unsupervised Domain Adaptation. (2021)

Cross-Domain Evaluation - Setting

- **Approaches**: Pi-Model, FixMatch and ReMixMatch.
- **Base Model**: ViT Small pre-trained on Imagenet-1k.
- **Training**: fine-tuning for 20 epochs on 615 labeled images and 12300 unlabeled images:
  - 5 labels per class (5%);

Cross-Domain Evaluation

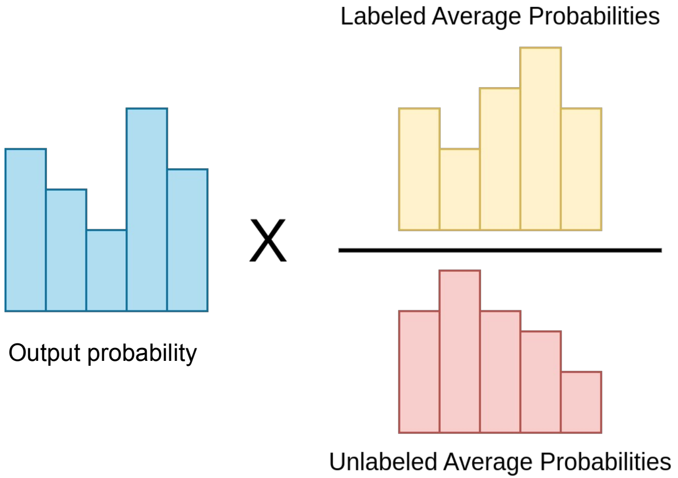
- **Pi-Model** suffers the domain shift:
  Probably **learns only** the unlabeled data distribution;
- **FixMatch** is not able to improve over the baseline by exploiting unlabeled data;

TABLE II
RESULTS IN CROSS-DOMAIN EVALUATION

Method	Accuracy
Π-Model	64.30
FixMatch	78.78
ReMixMatch	87.24
Supervised (LB)	78.17
Fully Supervised (UB)	93.01

Distribution Alignment

- **Objective**: Make unlabeled data useful to improve the performance;
- **Idea**: Align the output probability distribution of the unlabeled samples with the labeled ones.



Cross-Domain Evaluation

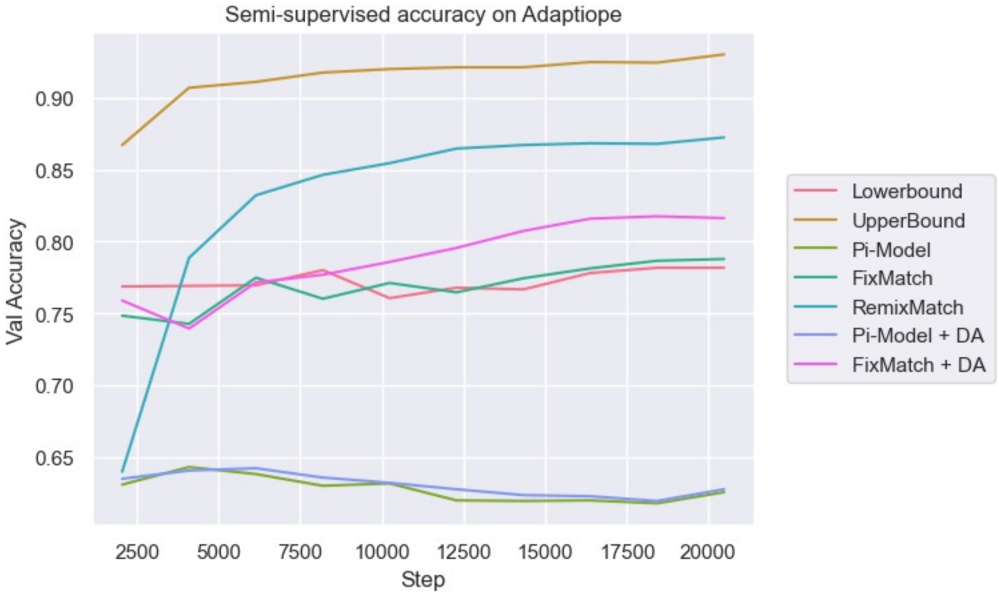
- **Distribution alignment** lead to improved performance only in Fixmatch;
- Pi-Model still does not manage to **generalize** to the test set.

TABLE II
RESULTS IN CROSS-DOMAIN EVALUATION

Method	Accuracy
Π-Model	64.30
FixMatch	78.78
ReMixMatch	87.24
Π-Model+DA	64.23
FixMatch+DA	81.17
Supervised (LB)	78.17
Fully Supervised (UB)	93.01

Cross-Domain Evaluation

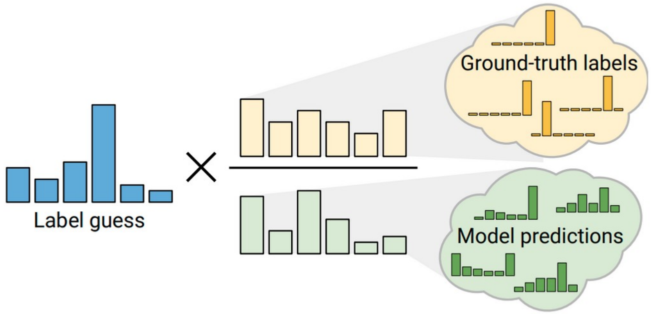
- ReMixMatch beats all the other methods **without any modification;**
- Correctly exploits the unlabeled data;
- **Not affected** by the domain shift.



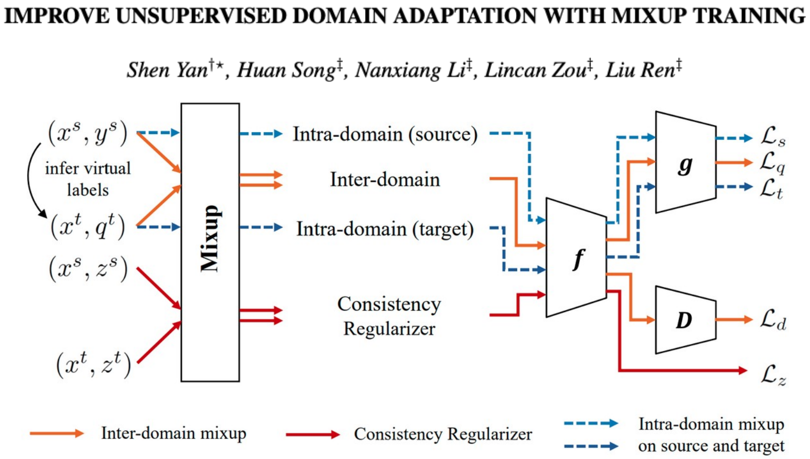
Why ReMixMatch?

Why does ReMixMatch not suffer from domain shift?

- Already uses distribution alignment to align pseudo-labels.



- Mixup augmentation produces augmented images belonging to both domains.



Adaptiope Failure Cases

GT: Hot Glue Gun

Pred: Power Drill

GT: Ring Binder

Pred: File Cabinet

GT: Roller Skates

Pred: Ice Skates

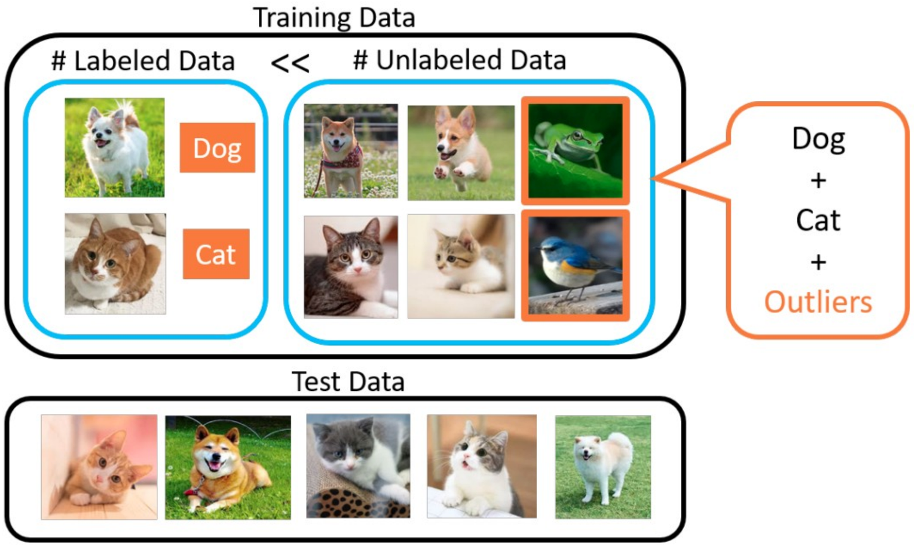
GT: VR Goggles

Pred: Over-Ear

Open Set Semi-Supervised Learning

Setting:
- Unlabeled dataset populated with samples out of distribution (**outliers**).

Objective:
- Make sure outliers do not affect the learning dynamics.



Saito, Kuniaki and Kim, Donghyun and Saenko, Kate. OpenMatch: Open-set Consistency Regularization for Semi-supervised Learning with Outliers. (2021)

Open Set Evaluation - Setting

- **Dataset**: CIFAR-100;
- **Unlabeled data**: from all classes;
- **Labeled data**: 80 classes out of 100, 4 labels per class;
- **Approach**: FixMatch.

Open Set Evaluation - Results

- **80 classes closed set:** 77.56% accuracy;
- **80 classes open set (+20 classes):** 78.41% accuracy;
- **Why these results?**
  - **Initial idea**: the threshold in fixmatch rules out outliers during unsupervised loss computation;
  - **Investigation**: 80% of the outliers where assigned to a class with more than 95% of confidence;
  - So we decided to inspect the predicted classes on the outliers.

Open Set Evaluation - Results

Actual	Predicted
Squirrel	Hamster (90%)
Worm	Snake (94%)
Whale	Dolphin (84%)
Tiger	Leopard (97%)
Willow Tree	Oak Tree (68%)

- The predicted classes are **very similar** to the actual ones;
- Hence **meaningful information** could have been extracted anyway;
- This is probably **dataset dependant** (trees: oak, maple, palm, pine, and willow).

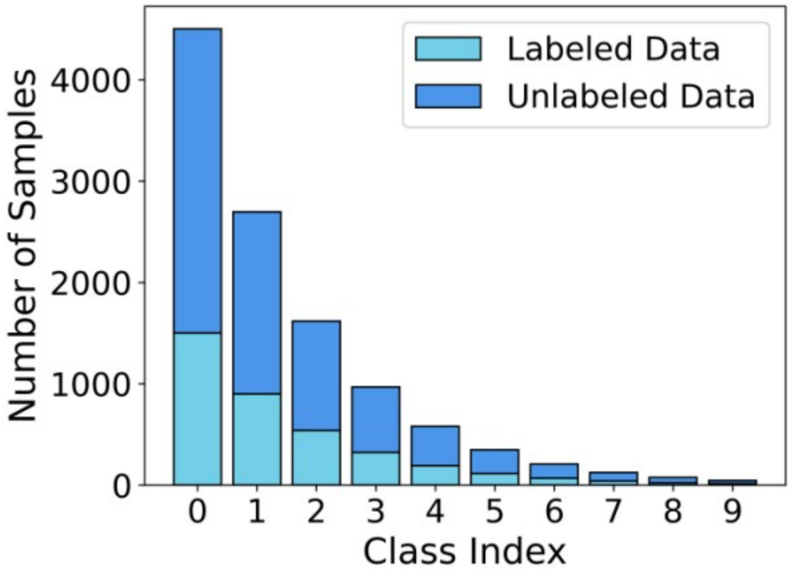
Other open challenges

Continual Learning in SSL setting

- Update model incrementally with no access to previous data.

Class imbalanced SSL

- Relaxes the assumption of balanced class distribution.



Conclusion

We deepened our understanding on Semi-Supervised Learning by:

- **Studying** several SSL approaches.
- **Experimenting** with Unified Semi-supervised Benchmark codebase.
- **Analyzing** current challenges on SSL, such as **Domain Shift** and **Open Set** SSL.

UNIVERSITY OF TRENTO

Thanks for your attention