



西安交通大学

XI'AN JIAOTONG UNIVERSITY

# Dual-Regularization Complementary Learning for Image Classification

*Lingjuan Ge, Mingming Gong, Yutian Lin, Bo Du*

*National Engineering Research Center for Multimedia Software,*

*Wuhan University*

ICME 2021

2022.4.29

# 论文概况

- 互补标签即指定图像不属于的类的标签，虽然其更容易获得且成本较低，但由于互补标签的固有的模糊性，其学习方法的性能远低于使用真实标签学习。
- 本文提出了一种提出了一种新的互补学习方法——双正则化互补学习(DRCL):
  - 同时训练两个深度神经网络，并强制它们对彼此的输出进行正则化；
  - 通过这种方式，两个网络可以相互学习，即使在有模糊互补标签的监督下，也倾向于产生一致的输出。

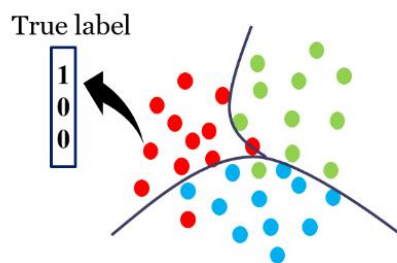
# 研究背景

现有的互补学习方法主要是通过修改常用的分类损失函数，从互补标记数据中恢复真实的分类风险。

尽管所有这些方法都在限制条件下提供了真实分类风险的一致估计，但在样本有限的数据集上，由于互补标签固有的模糊性，互补学习的表现可能比普通学习差得多。

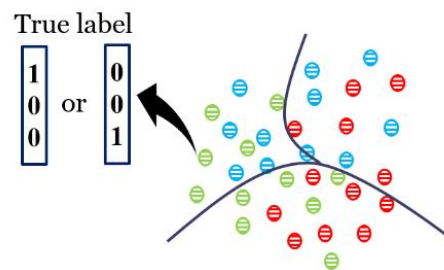


(a)  $Y = \text{'cat'}$   $\bar{Y} = \text{'not dog'}$  (b)  $Y = \text{'dog'}$   $\bar{Y} = \text{'not bird'}$  (c)  $Y = \text{'bird'}$   $\bar{Y} = \text{'not cat'}$



Class: ● :1 ● :2 ● :3

(a) Classifier trained with true labels

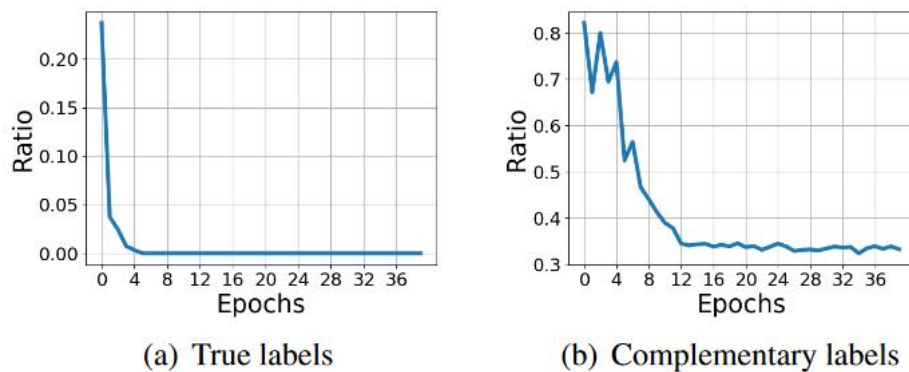


Class: ⚡ :not 1 ⚡ :not 2 ⚡ :not 3

(b) Classifier trained with complementary labels

## 研究背景

由于互补标签的模糊性，如果多次训练一个网络，它不太可能有与真实标签一致的输出。为了验证这一假设，使用真实标签和互补标签训练两个具有相同架构的独立网络。



**Fig. 3:** The ratio of inconsistent prediction of models trained by true labels and complementary labels.

由实验结果发现，由于互补标签的模糊性，两个独立训练的网络很难产生一致的输出，这意味着它们可能离真正的函数更远。

- 因此本文提出的双正则化互补学习强制两个网络满足必要的一致性条件，从而使两个网络更接近真实的分类函数。

# 研究方法

- 首先对两个神经网络分别独立地进行训练，使用根据转移矩阵修正的交叉熵损失：

$$\begin{aligned}\bar{\ell}(\bar{f}(X), \bar{Y}) &= \ell(\mathbf{Q}^T \mathbf{g}(X), \bar{Y}) \\ &= - \sum_{k=1}^K 1\{\bar{Y} = k\} \log(\mathbf{Q}^T \mathbf{g}(X))_k.\end{aligned}$$

具体而言，最小化下列经验风险：

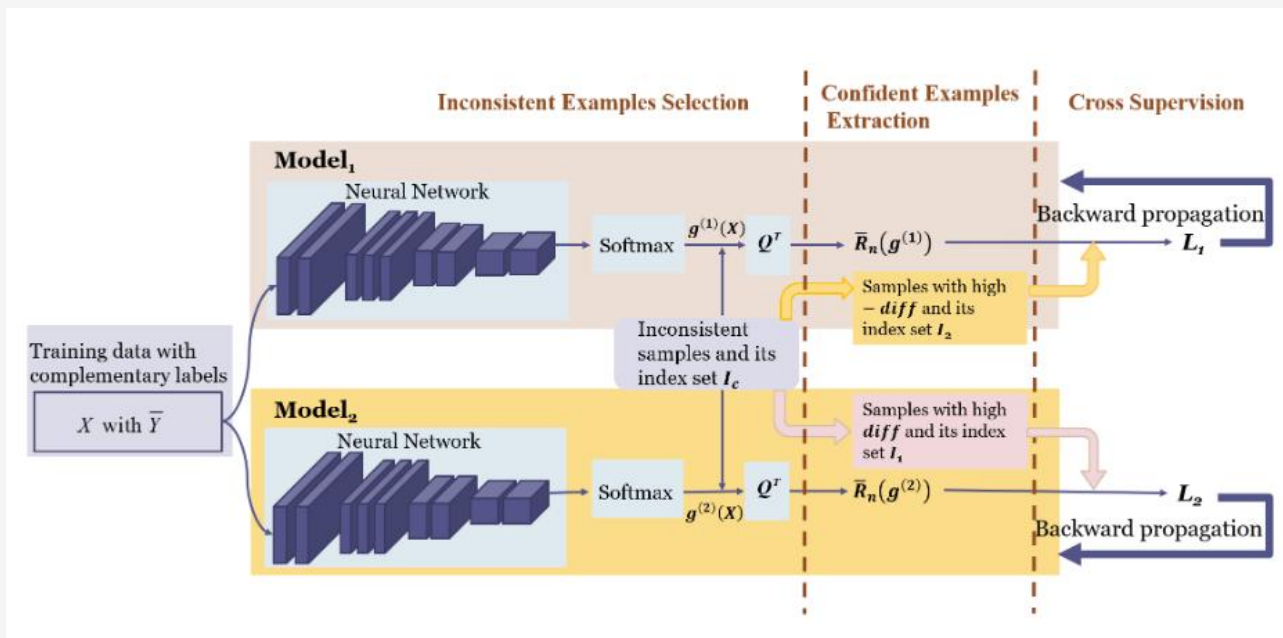
$$\begin{aligned}\bar{R}_n(g^{(1)}) &= 1/n * \sum_{i=1}^n l(\mathbf{Q}^T \mathbf{g}^{(1)}(\mathbf{x}_i), \bar{y}_i), \\ \bar{R}_n(g^{(2)}) &= 1/n * \sum_{i=1}^n l(\mathbf{Q}^T \mathbf{g}^{(2)}(\mathbf{x}_i), \bar{y}_i).\end{aligned}$$

- 之后再对两个网络输出的不一致进行修正，本文提出的DRCL方法具体分为三步：
  - inconsistent examples selection：挑选两网络中结果不一致的样本
  - confident examples extraction：在不一致的样本中选择高置信度样本
  - cross supervision：使用置信样本同时对两个网络进行增强



# 研究方法

- DRCL整体框架如下：



## 1. Inconsistent Examples Selection

挑选出两个网络预测结果不一致的样本集合：

$$I_c = \{i | i \in \{1, \dots, n\}, \hat{y}_i^{(1)} \neq \hat{y}_i^{(2)}\}.$$

$$\hat{y}_i^{(1)} = \arg \max_{k \in \{1, \dots, K\}} \mathbf{g}_k^{(1)}(\mathbf{x}_i)$$

$$\hat{y}_i^{(2)} = \arg \max_{k \in \{1, \dots, K\}} \mathbf{g}_k^{(2)}(\mathbf{x}_i)$$

## 2. Confident Examples Extraction

- 进一步选择一些在一个模型中预测置信度高而在另一个模型中预测置信度低的样本:

- 首先计算两个模型最终预测标签概率的差值  $diff(\mathbf{x}_i) = \mathbf{g}_{\hat{y}_i^{(1)}}^{(1)}(\mathbf{x}_i) - \mathbf{g}_{\hat{y}_i^{(2)}}^{(2)}(\mathbf{x}_i)$
- 对于模型1, 选择 $diff(x_i)$ 值最大的10个样本; 对于模型2, 根据 $-diff(x_i)$ 选择

(思想: 当两个网络有不一致的预测结果时, 要么其中一个网络预测正确, 要么两个网络都预测错误)

## 3. Cross Supervision

- 为了提升模型1、2的准确率, 使用提取出的置信样本来增强彼此的学习目标, 具体为:

$$L_1 = 1/(n + n_2) * \left( n * \bar{R}_n(\mathbf{g}^{(1)}) + \sum_{i \in I_2} l \left( \mathbf{g}^{(1)}(\mathbf{x}_i), \hat{y}_i^{(2)} \right) \right),$$

$$L_2 = 1/(n + n_1) * \left( n * \bar{R}_n(\mathbf{g}^{(2)}) + \sum_{i \in I_1} l \left( \mathbf{g}^{(2)}(\mathbf{x}_i), \hat{y}_i^{(1)} \right) \right).$$

# 实验结果

- 实验设置：
  - 数据集与模型：MNIST(LeNet5)、CIFAR-10(ResNet18)、CIFAR-100、Tiny ImageNet(ResNet34)
  - 实验时对两个网络使用不同的初始化方式，并将其中一个网络的原始图像旋转10度。
  - 对于转移矩阵 $Q$ ，使用ground-truth  $Q$ 和估计的 $Q$ 在三种设置下进行实验：uniform（除了真实类别之外的其余类别标签被选为互补标签的概率相同）、without0（其余类别标签均有可能被选中，但概率不同）、with0（部分类别标签被选中的概率为0）
- Baseline与对比方法：
  - 将使用互补数据训练的单个网络的预测结果作为baseline
  - 预训练之后，将本文的DRCL方法与co-training、co-teaching做对比（co-training只在另一个网络中选择高置信度的样本，而co-teaching则在另一个网络中选择损失小的样本。 ）



# 实验结果

**Table 1:** The accuracy on MNIST in the uniform setting with true  $Q$  ('TL' provides the upper bound of the comparison accuracy by training a single network with true labels, 'BL' represents the baseline of training a single complementary network, 'DRCL', 'CTR', 'CTE' represents models trained with our proposed method, co-training and co-teaching respectively. The another table keeps the same interpretation.)

Method	Percent of labeled data					
	1%	3%	5%	7%	9%	100%
TL	90.44	95.08	96.21	97.13	97.21	99.15
BL( $Model_1$ )	35.97 $\pm 5.75$	62.25 $\pm 6.98$	80.73 $\pm 5.39$	83.20 $\pm 6.75$	89.34 $\pm 3.10$	98.34 $\pm 0.04$
DRCL( $Model_1$ )	<b>40.67</b> $\pm 7.31$	<b>67.61</b> $\pm 12.96$	<b>89.79</b> $\pm 3.17$	<b>92.24</b> $\pm 3.00$	<b>94.70</b> $\pm 0.45$	<b>98.86</b> $\pm 0.04$
CTR( $Model_1$ )	39.95 $\pm 6.60$	61.18 $\pm 6.25$	84.56 $\pm 2.08$	85.75 $\pm 4.95$	90.74 $\pm 1.47$	98.58 $\pm 0.05$
CTE( $Model_1$ )	36.03 $\pm 5.79$	63.99 $\pm 7.92$	81.73 $\pm 5.33$	84.02 $\pm 6.72$	90.19 $\pm 2.60$	98.58 $\pm 0.05$
BL( $Model_2$ )	37.84 $\pm 7.42$	55.48 $\pm 9.16$	76.35 $\pm 6.79$	85.30 $\pm 4.75$	89.73 $\pm 2.46$	98.35 $\pm 0.05$
DRCL( $Model_2$ )	<b>40.71</b> $\pm 7.20$	<b>67.46</b> $\pm 13.12$	<b>89.80</b> $\pm 3.29$	<b>92.26</b> $\pm 2.96$	<b>94.65</b> $\pm 0.43$	<b>98.87</b> $\pm 0.05$
CTR( $Model_2$ )	40.37 $\pm 7.18$	59.18 $\pm 5.69$	81.59 $\pm 4.47$	87.48 $\pm 2.12$	91.15 $\pm 0.87$	98.58 $\pm 0.04$
CTE( $Model_2$ )	37.99 $\pm 7.73$	53.96 $\pm 16.65$	77.27 $\pm 6.68$	86.21 $\pm 4.36$	90.54 $\pm 1.72$	98.57 $\pm 0.03$

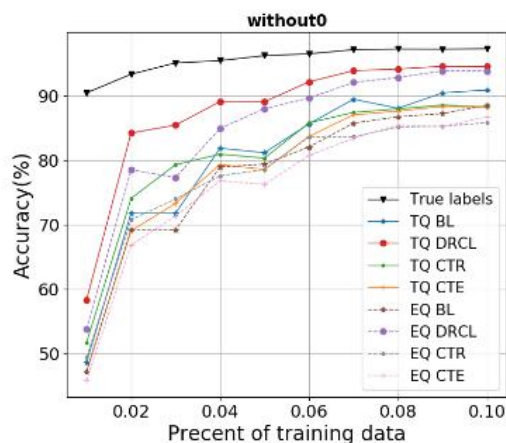
**Table 2:** The accuracy on CIFAR-10 in the uniform setting with true  $Q$ .

Method	Percent of labeled data					
	10%	30%	50%	70%	90%	100%
TL	69.35	83.95	87.34	89.87	90.93	91.60
BL( $Model_1$ )	25.52 $\pm 1.61$	44.04 $\pm 0.72$	52.55 $\pm 1.13$	61.31 $\pm 1.26$	69.40 $\pm 1.48$	72.06 $\pm 1.06$
DRCL( $Model_1$ )	<b>27.38</b> $\pm 1.16$	<b>47.67</b> $\pm 0.58$	<b>60.05</b> $\pm 1.43$	<b>70.07</b> $\pm 0.76$	<b>77.49</b> $\pm 0.59$	<b>79.99</b> $\pm 0.45$
CTR( $Model_1$ )	27.44 $\pm 1.32$	47.67 $\pm 0.75$	58.28 $\pm 1.05$	67.87 $\pm 0.87$	75.26 $\pm 1.08$	77.53 $\pm 0.56$
CTE( $Model_1$ )	26.44 $\pm 1.15$	46.18 $\pm 0.59$	56.48 $\pm 1.20$	65.99 $\pm 1.05$	74.10 $\pm 1.57$	76.35 $\pm 0.84$
BL( $Model_2$ )	25.98 $\pm 0.98$	43.77 $\pm 1.17$	51.98 $\pm 1.36$	61.94 $\pm 1.24$	69.15 $\pm 1.26$	71.97 $\pm 1.00$
DRCL( $Model_2$ )	<b>27.41</b> $\pm 1.15$	<b>47.67</b> $\pm 0.70$	<b>60.02</b> $\pm 1.36$	<b>70.07</b> $\pm 0.77$	<b>77.44</b> $\pm 0.55$	<b>79.90</b> $\pm 0.41$
CTR( $Model_2$ )	27.03 $\pm 1.00$	47.51 $\pm 0.96$	58.23 $\pm 1.02$	68.12 $\pm 0.75$	75.13 $\pm 0.80$	77.54 $\pm 0.48$
CTE( $Model_2$ )	26.00 $\pm 0.70$	46.21 $\pm 1.47$	56.02 $\pm 1.14$	66.52 $\pm 1.11$	73.77 $\pm 1.20$	76.56 $\pm 0.81$

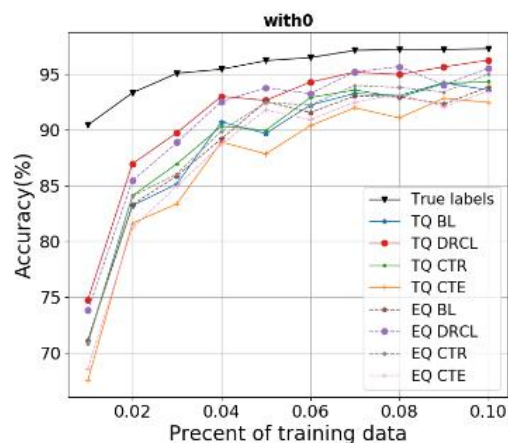
- percent of labeled data代表使用了部分训练集进行训练

# 实验结果

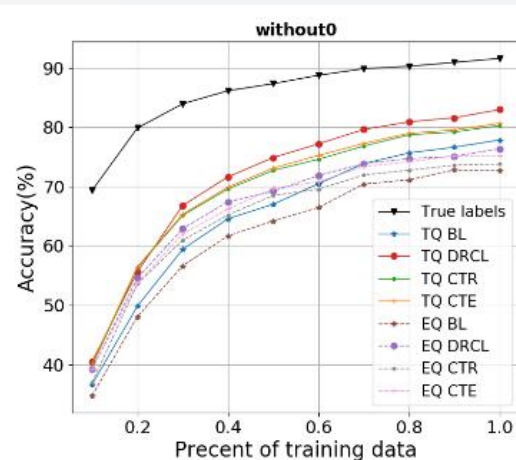
- 在without0和with0的设置下，使用真实的Q和估计的Q进行实验



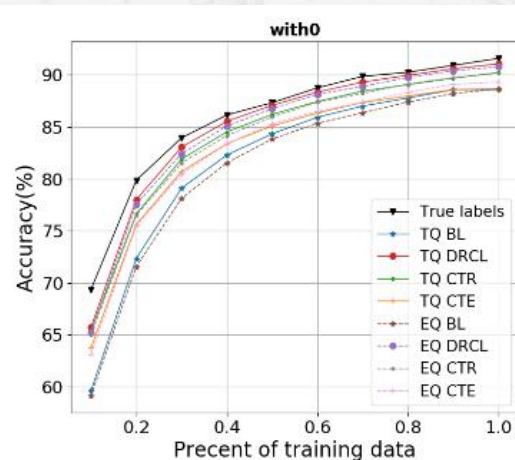
(a) MNIST in without0



(b) MNIST in with0



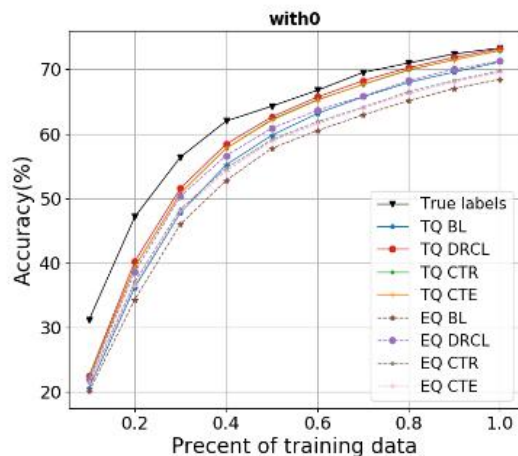
(a) CIFAR-10 in without0



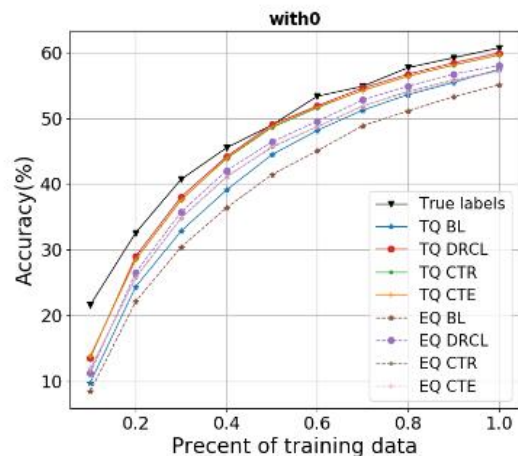
(b) CIFAR-10 in with0

# 实验结果

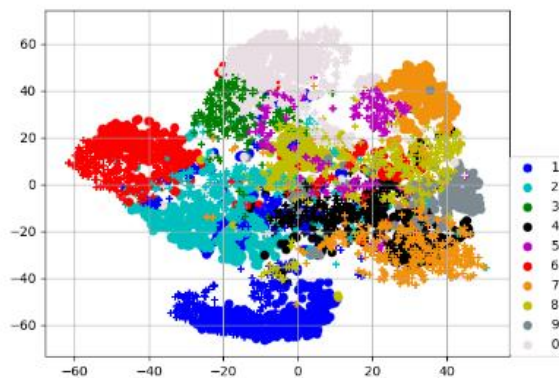
- 为了验证本文的一致性假设，使用4%MNIST训练数据在uniform的假设下分别训练两个网络以及使用DRCL训练将输出结果可视化，每个网络对同一类预测的聚合度较高，且两个网络预测的同一类的重叠部分增加，说明两个网络的一致性提高。



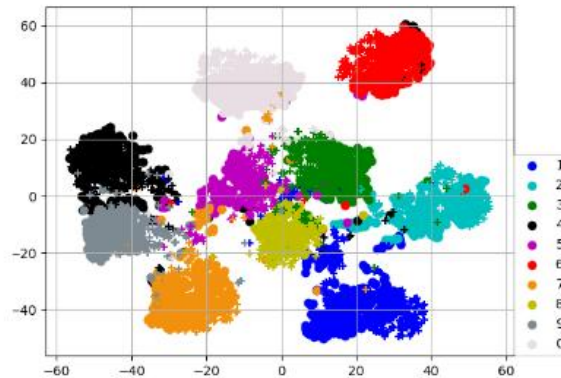
(a) CIFAR-100



(b) Tiny ImageNet



(a) BL



(b) DRCL



西安交通大学  
XI'AN JIAOTONG UNIVERSITY

# 感谢聆听

