Class-Imbalanced Semi-Supervised Learning with Inverse Auxiliary Classifier

BMVC 2023

Tiansong Jiang, Sheng Wan*, Chen Gong*
School of Computer Science and Engineering, Nanjing University of Science and Technology.

Introduction & Motivation

We introduce IAC (Inverse Auxiliary Classifier), which is a novel and effective plug-in module for existing CISSL algorithms to emphasize the importance of minority classes during the feature learning process.

The proposed IDA (Inverse Distribution Alignment) loss is utilized to enrich the limited supervision, facilitating IAC in effectively learning from the minority-class data.

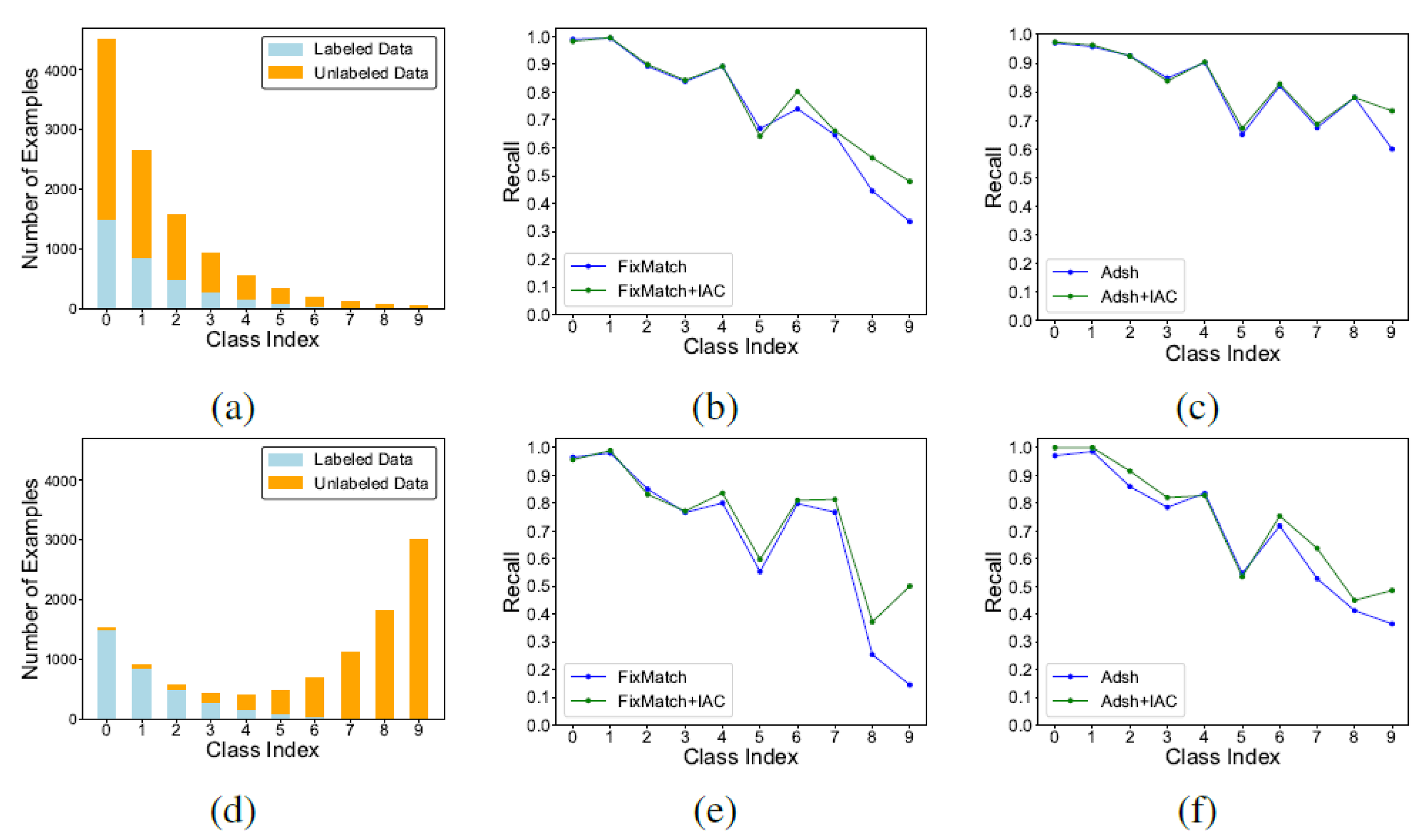

Figure 1: Learning with consistent and inconsistent class distributions between labeled and unlabeled examples. (a) Consistent class distributions between labeled and unlabeled examples. (b)-(c) Recall of the generated pseudo-labels and predicted labels under consistent class distributions, respectively. (d) Inconsistent class distributions between labeled and unlabeled examples. (e)-(f) Recall of the generated pseudo-labels and predicted labels under inconsistent class distributions, respectively.

Method

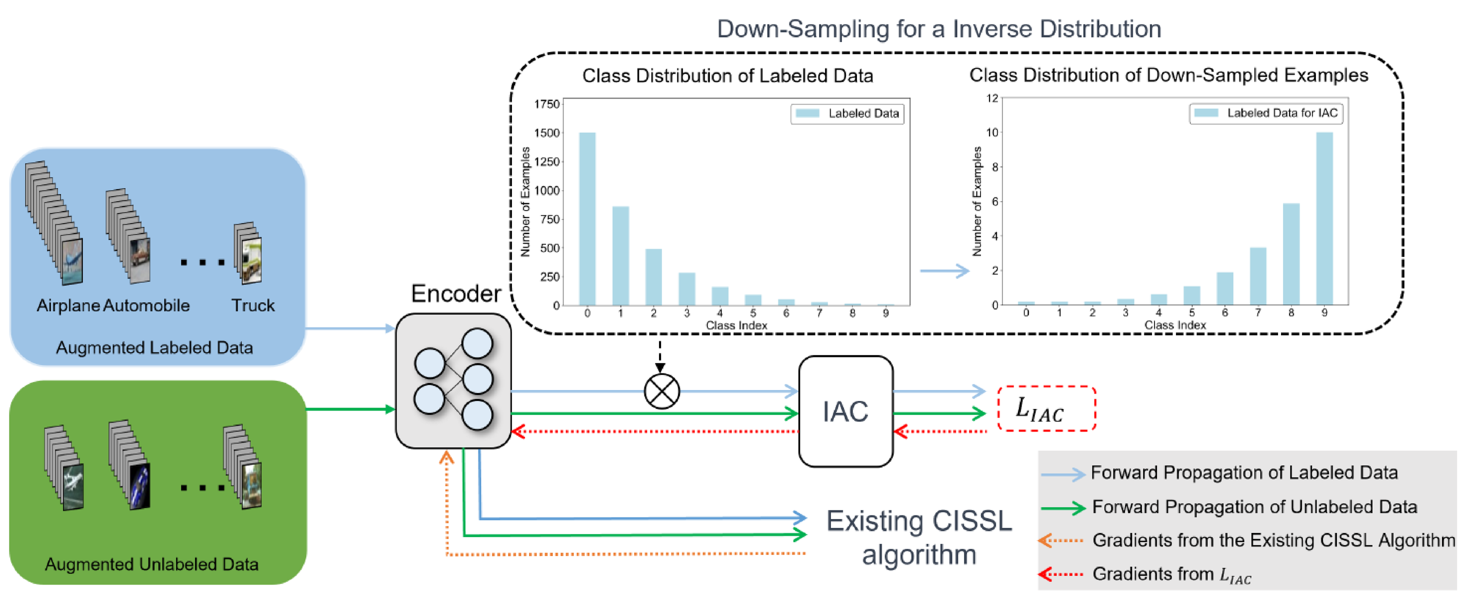

Figure 2: The overall framework of the CISSL method that incorporates our proposed IAC.

$$L_s = \frac{1}{B}\sum_{b=1}^{B} M(\mathbf{x}_b)H(p(y_b), q_b)$$

$$M(\mathbf{x}_b) = Ber(max(\frac{\delta}{N_{y_b}}, (\frac{N_K}{N_{y_b}})^{\rho}))$$

$$L_u = \frac{1}{\mu B}\sum_{b=1}^{\mu B} \mathbb{1}(max(q_b) \geq \tau)H(p_i(y \mid A(\mathbf{u}_b)), \hat{q}_b)$$

$$L_{IDA} = H(\tilde{p}(y), \frac{\sum_{b=1}^{\mu B} q_b}{B})$$

$$L_{IAC} = L_s + \lambda_1 L_u + \lambda_2 L_{IDA}$$

$$L = L_{CISSL} + L_{IAC}$$

Experiments

Algorithm	CIFAR-10							
	$N_1 = 500, M_1 = 4000$					$N_1 = 1500, M_1 = 3000$		
	$\gamma_l = 150$	$\gamma_l = 100$				$\gamma_l = 150$	$\gamma_l = 100$	
	$\gamma_u = 150$	$\gamma_u = 150$	$\gamma_u = 100$	$\gamma_u = 1$	$\gamma_u = \frac{1}{100}$	$\gamma_u = 150$	$\gamma_u = 100$	$\gamma_u = \frac{1}{100}$
Supervised	43.66/36.60	46.47/41.46	46.47/41.46	46.47/41.46	46.47/41.46	60.37/57.09	63.39/61.14	63.39/61.14
FixMatch	64.10/59.14	71.47/70.04	74.17/73.43	77.66/73.30	60.05/54.78	73.20/72.30	77.66/76.21	71.57/69.78
w/IAC	**67.81/64.92**	**73.26/72.42**	**74.72/74.02**	**78.54/78.27**	**65.26/61.68**	**74.44/73.80**	**78.10/77.75**	**75.27/74.51**
CReST	72.28/70.79	73.66/72.48	78.72/78.45	92.86/92.20	71.80/68.35	74.00/73.21	79.22/78.48	86.80/86.61
w/IAC	**73.03/71.62**	**75.71/72.97**	**79.02/78.12**	**92.89/92.26**	**79.47/76.96**	**75.22/73.38**	**79.77/79.65**	**87.61/87.44**
Adsh	67.47/64.26	73.03/72.23	76.34/75.95	82.22/82.57	66.91/66.39	74.00/73.09	78.15/77.70	70.09/69.33
w/IAC	**70.39/68.71**	**74.75/74.00**	**77.23/76.88**	**85.51/85.76**	**71.17/71.15**	**74.35/73.60**	**79.40/79.14**	**74.09/74.11**
DARP	68.99/**67.05**	74.10/73.22	75.34/74.81	71.69/68.51	61.44/56.19	**74.34**/72.41	78.27/77.90	69.31/67.34
w/IAC	**69.01**/66.50	**74.59/73.85**	**75.85/75.32**	**84.09/83.96**	**64.87/61.05**	74.01/**73.18**	**78.92/78.62**	**75.31/74.81**
DASO	64.84/62.03	67.71/66.34	69.87/69.09	71.57/70.75	69.77/69.44	72.68/**72.12**	76.90/76.63	78.03/78.14
w/IAC	**67.13/65.67**	**68.02/66.65**	**71.19/70.64**	**74.55/74.37**	**78.80/78.96**	**72.70**/71.99	**77.09/76.86**	**81.65/81.78**
ABC	76.41/75.78	80.63/80.49	81.88/81.79	87.07/87.09	79.43/79.22	80.73/80.67	84.25/84.23	83.76/83.54
w/IAC	**78.54/78.24**	**80.69/80.53**	**81.93/81.82**	**88.32/87.49**	**81.94/81.57**	**82.05/82.03**	**84.49/84.48**	**84.65/84.59**
CoSSL	76.95/76.98	80.68/80.79	82.55/82.52	85.98/85.95	72.27/72.04	**83.13/83.30**	84.98/85.03	74.12/73.22
w/IAC	**78.79/78.94**	**80.81/80.99**	**82.85/82.91**	**86.89/86.88**	**76.23/76.07**	83.11/83.20	**85.37/85.41**	**76.64/75.77**

Table 1: Comparison results (ACC/F1) on CIFAR-10.

Algorithm	SVHN		CIFAR-100				STL-10	
	$N_1 = 500, M_1 = 4000$		$N_1 = 150, M_1 = 300$				$N_1 = 150, M = 100k$	
	$\gamma_l = 100$		$\gamma_l = 10$		$\gamma_l = 15$		$\gamma_l = 10$	$\gamma_l = 20$
	$\gamma_u = 100$	$\gamma_u = \frac{1}{100}$	$\gamma_u = 10$	$\gamma_u = \frac{1}{10}$	$\gamma_u = 15$	$\gamma_u = \frac{1}{15}$	$\gamma_u : N/A$	$\gamma_u : N/A$
Supervised	82.29/79.73	82.29/79.73	48.23/46.61	48.23/46.61	45.92/43.80	45.92/43.80	46.42/44.62	40.04/35.76
FixMatch	91.93/90.73	89.72/88.20	57.96/56.47	56.91/56.06	54.50/52.34	53.78/52.38	65.61/63.97	54.56/48.89
w/IAC	**91.95/90.67**	**92.66/92.34**	**58.58/57.22**	**59.33/58.66**	**55.74/53.80**	**56.40/55.44**	**67.01/64.94**	**55.85/50.95**
CReST	93.12/**91.56**	92.13/**91.72**	57.09/55.55	59.17/58.33	54.54/52.08	56.38/55.38	67.01/61.94	59.23/54.46
w/IAC	**93.27**/90.78	**92.47**/90.98	**58.57/57.23**	**61.06/60.26**	**54.90/52.14**	**58.57/57.55**	**67.73/62.15**	**60.26/54.57**
Adsh	92.32/91.29	87.60/87.08	58.20/57.21	55.22/54.74	54.14/52.57	51.80/50.92	69.68/70.03	65.50/64.68
w/IAC	**92.34/91.36**	**91.48/91.10**	**58.33/57.33**	**57.25/56.95**	**54.53/52.83**	**54.39/53.83**	**70.48/70.35**	**66.33/64.94**
DARP	91.81,90.60	91.95/90.99	57.88/56.69	57.82/56.91	54.49/52.61	54.69/53.43	63.94/61.85	55.26/51.81
w/IAC	**91.84/90.74**	**93.27/92.94**	**58.55/57.43**	**59.13/58.45**	**55.16/53.13**	**56.94/55.93**	**66.83/65.63**	**57.63/54.28**
DASO	88.59/87.19	89.54/88.72	58.67/56.51	59.31/58.22	55.11/52.37	56.18/54.72	69.38/68.54	58.08/53.66
w/IAC	**89.01/87.76**	**93.00/92.67**	**58.77/57.50**	**59.70/59.06**	**55.56/53.48**	**56.89/55.87**	**70.55/69.62**	**62.54/60.43**
ABC	93.75/**93.12**	92.74/**92.39**	59.83/58.64	59.88/59.02	**56.87/55.47**	57.25/56.36	70.83/69.92	65.69/63.87
w/IAC	**93.76**/93.01	**92.76**/92.27	**60.15/59.14**	**60.51/59.81**	56.73/55.35	**57.73/56.72**	**71.99/71.16**	**67.64/66.81**
CoSSL	92.68/91.69	90.27/88.03	58.58/57.58	58.11/57.45	56.21/55.31	55.56/54.53	70.75/70.15	64.85/64.09
w/IAC	**92.97/92.29**	**92.46/91.57**	**59.45/58.53**	**59.67/59.10**	**56.43/55.62**	**56.33/55.51**	**71.14/70.54**	**65.18/64.34**

Table 2: Comparison results (ACC/F1) on SVHN, CIFAR-100 and STL-10.

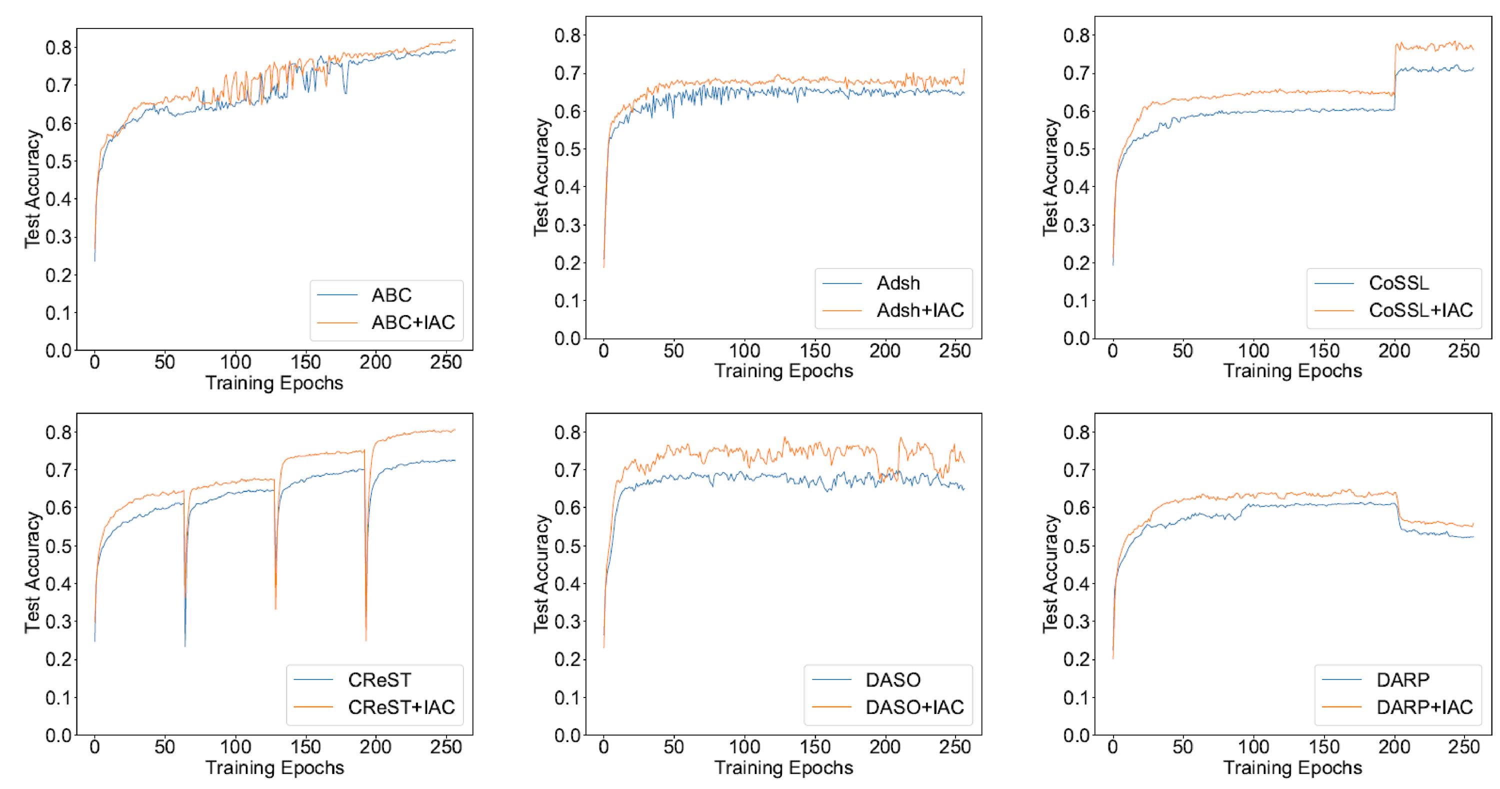

Figure 3: Training curves for the test accuracy on CIFAR-10, where $N_1 = 500$, $M_1 = 4000$, $\gamma_l = 100$ and $\gamma_u = \frac{1}{100}$.

Algorithm	CIFAR-10	
	$N_1 = 500, M_1 = 4000$	
	$\gamma_l = 100$	
	$\gamma u = 100$	$\gamma u = \frac{1}{100}$
CoSSL	82.55	72.27
CoSSL+IAC	**82.79**	**76.23**
Without L_u	81.80	71.63
Without L_{IDA}	82.65	74.63
$\lambda_1 = 2.0$	82.41	**77.53**
$\lambda_2 = 1.0$	82.47	75.37
$\rho = 1.0$	82.58	72.16
$\rho = 2.5$	82.78	75.83
Without δ	82.74	75.96
$\delta = 1.0$	82.76	70.34
Without $\delta, \rho = 2.5$	82.10	73.21

Table 3: The ablation study for our method on CoSSL on CIFAR-10