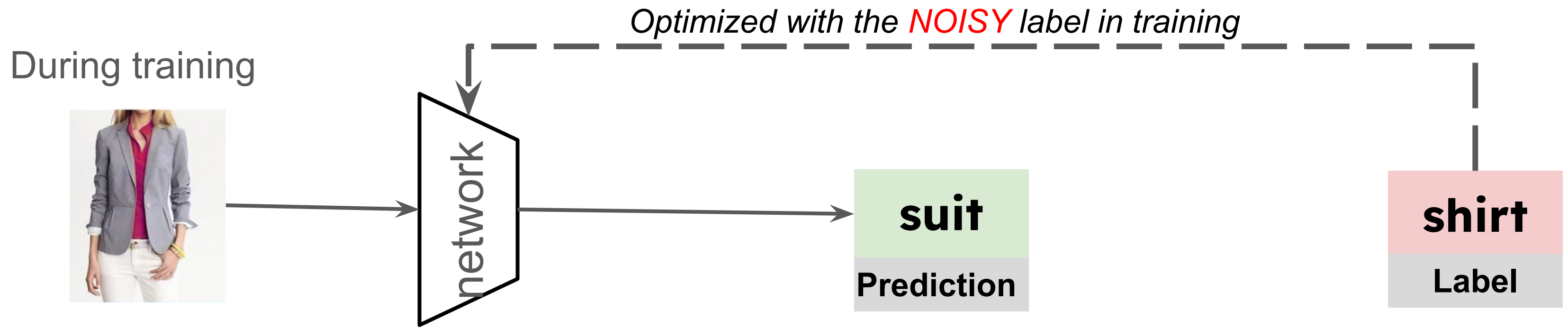
LNL+K: Enhancing Learning with Noisy Labels Through Noise Source Knowledge Integration

Siqi Wang, Bryan A. Plummer

BOSTON UNIVERSITY

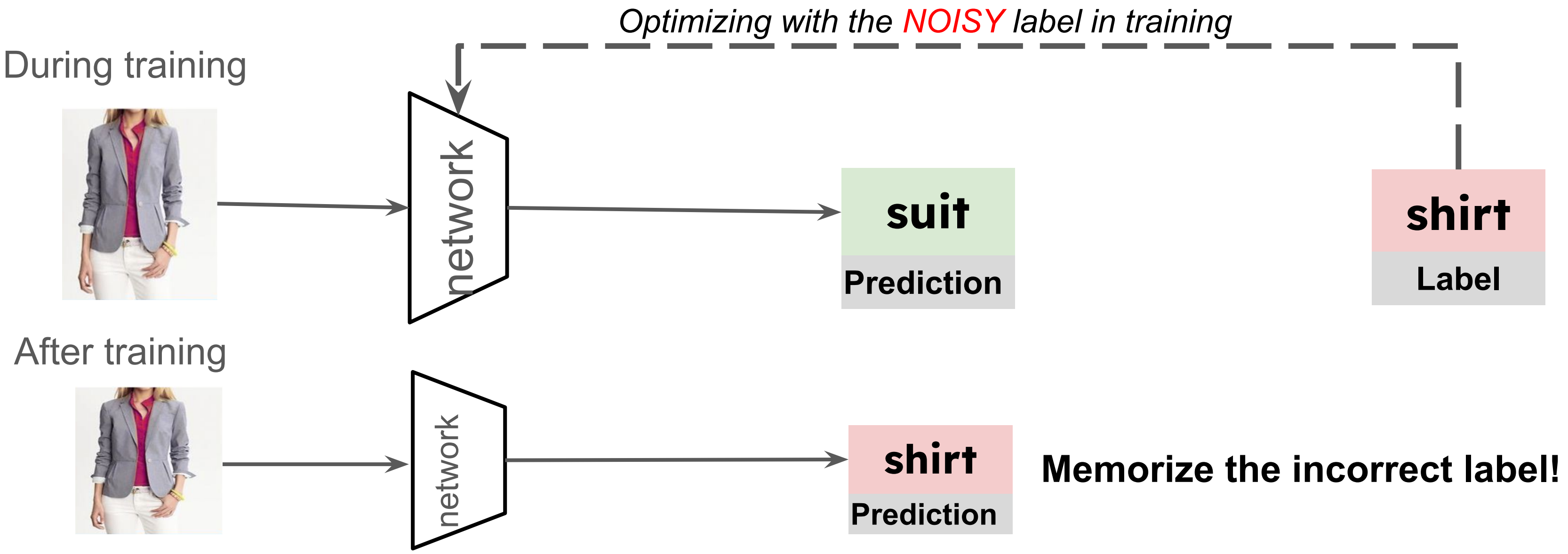
Learning with Noisy Labels (LNL)

When labels are noisy in the supervised learning...Model overfits to noisy labels!



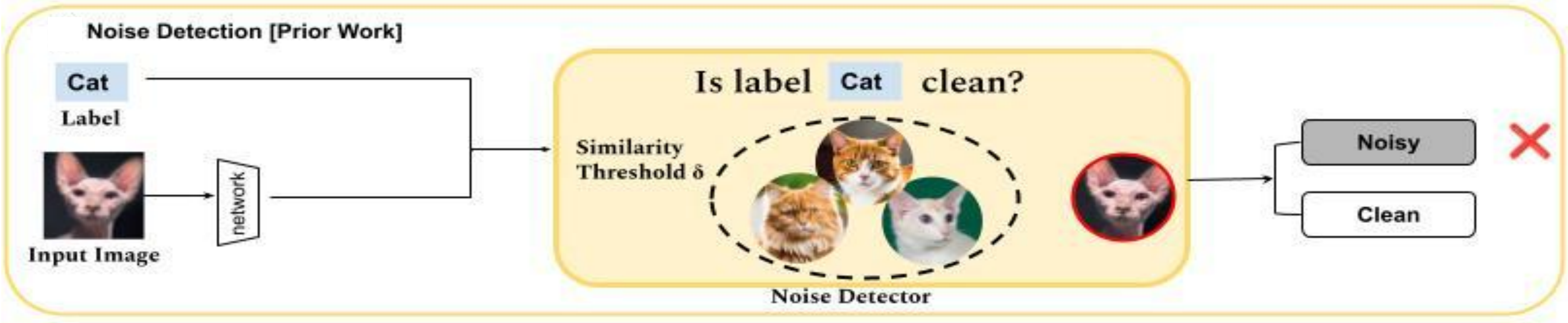
Learning with Noisy Labels (LNL)

When labels are noisy in the supervised learning…Model overfits to noisy labels!



LNL aims to train a high-performing model using noisy training data.

Motivation



Noise Detection [Prior Work]
Cat
Label
Input Image
network
Is label Cat clean?
Similarity
Threshold δ
Noise Detector
Noisy
Clean

Motivation

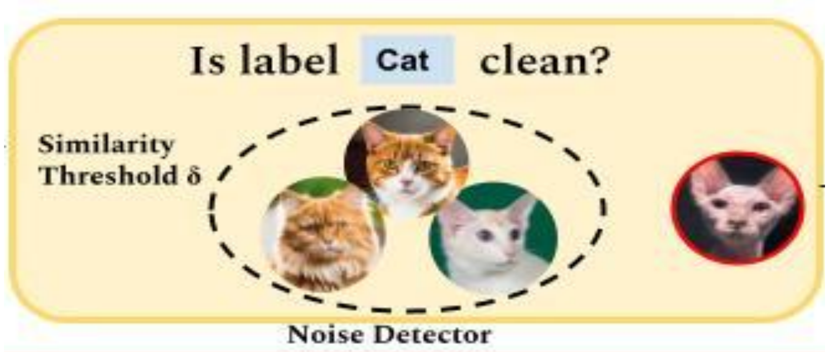
Challenges of Noise Detection 🤔

Similar Category

Similar Background

High Noise Ratio

Cell painting images are labeled with treatments, but many treatments have little to no effect. This results in a noise ratio of over **50%**.



Motivation

Challenges of Noise Detection 🤔

Similar Category

Similar Background

High Noise Ratio

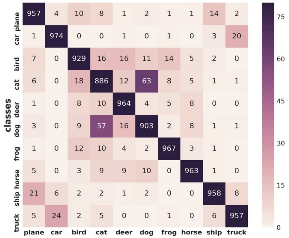
Cell painting images are labeled with treatments, but many treatments have little to no effect. This results in a noise ratio of over **50%**.

Available Noise Source Knowledge 💡

Dataset Meta-Data

"ANIMAL-10N dataset contains 5 pairs of confusing animals with a total of 55,000 images. **The 5 pairs are as following**: …"

Confusion Matrix from Related work



Domain Knowledge

Most weak treatment cells visually resemble the control class.

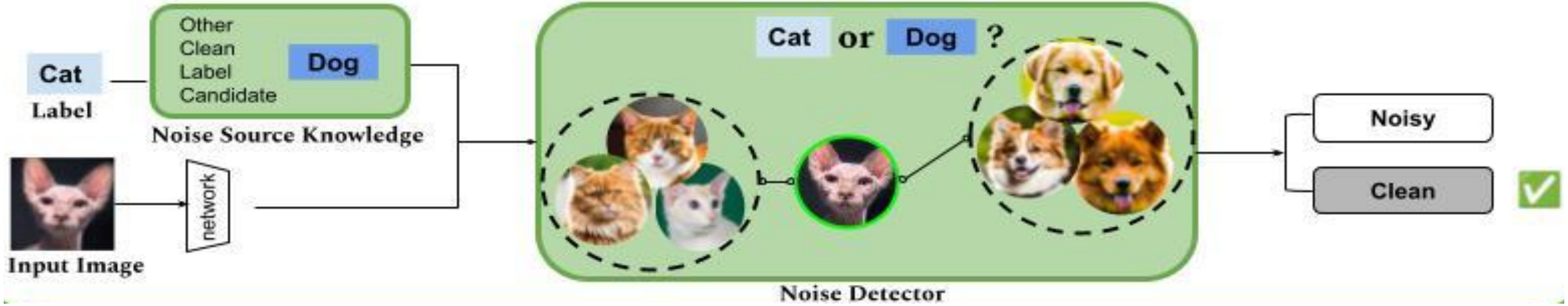
Motivation

Q1: How can we integrate noise source knowledge into existing LNL methods?
- Method

Q2: How helpful is noise source knowledge?
- Experiments

Q3: What are insights for designing new LNL methods?
- Discussions

Method



$$y_i = c \leftrightarrow \tilde{y_i} = c \wedge p(c|x_i) > \max(\{p(c_n|x_i)|c_n \in D_{c-ns}\}).$$

A Unified Framework

$(x_i, \tilde{y_i})$ i-th sample in the dataset

$\{\tilde{y_i}\}_{i=1}^{n}$ noisy labels

$\{y_i\}_{i=1}^{n}$ true labels

D_{c-ns} Noise source classes for class *c*

$p(c|x_i)$ Probability of sample x_i with clean label c

Experiments (Dominant Noise)

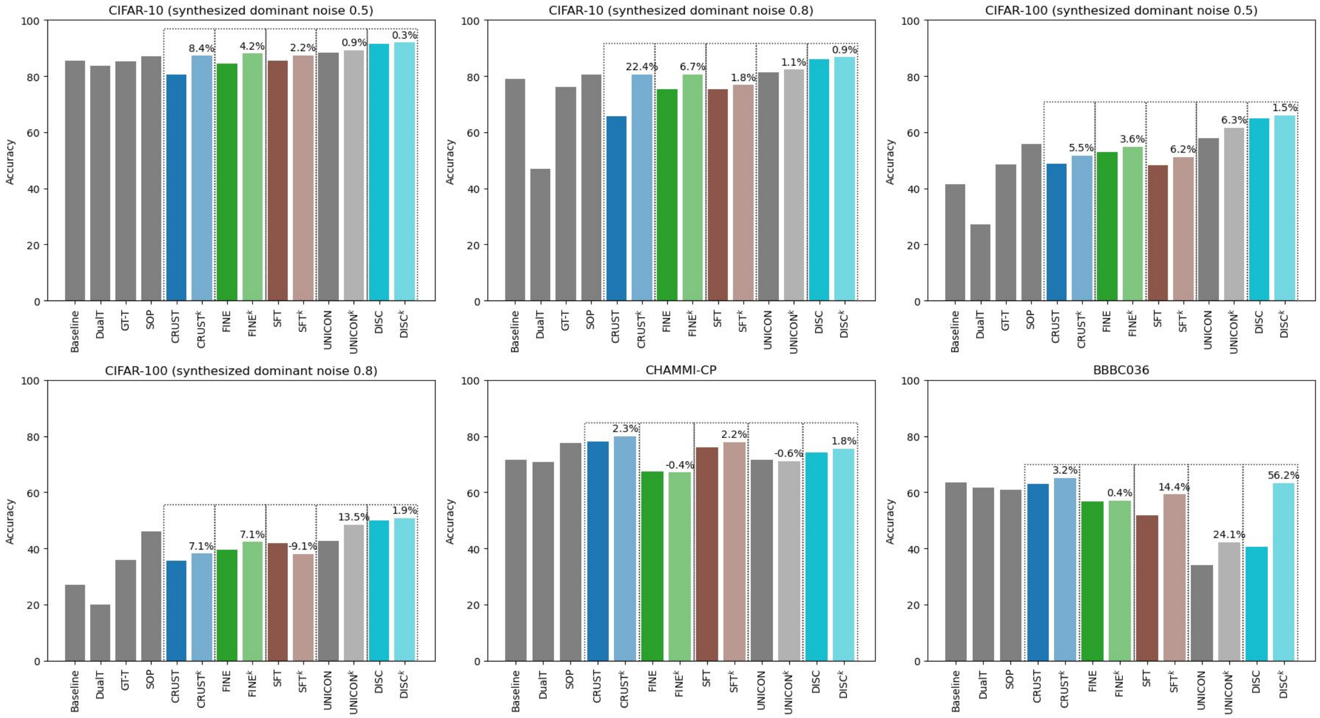
Class A

Class B

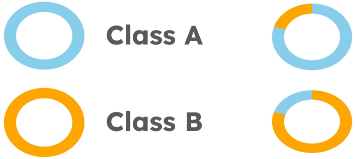
Class A with Dominant Noise

	CIFAR-10		CIFAR-100		CHAMMI-CP	BBBC036
Noise ratio	0.5	0.8	0.5	0.8		
Baseline	85.46±0.25	78.99±0.07	41.41±1.47	27.03±0.12	71.54±0.45	63.49±0.62
DualT	83.70±0.04	46.96±0.07	27.04±0.07	19.94±0.04	70.73±0.17	61.54±0.61
GT-T	85.24±0.06	76.03±0.04	48.39±0.21	35.96±0.04	-	-
SOP	86.94±0.37	80.65±0.71	55.78±0.68	45.94±0.62	77.55±0.23	60.94±0.38
CRUST	80.46±0.17	65.79±0.62	48.87±0.31	35.56±1.38	78.02±0.31	63.06±0.65
CRUST^{+k}	87.19±0.08	80.54±0.30	51.56±0.31	38.07±2.05	**79.81**±0.56	**65.07**±0.71
FINE	84.43±0.38	75.45±0.74	52.87±0.98	39.45±0.25	67.27±0.82	56.80±0.87
FINE^{+k}	88.00±0.11	80.52±0.28	54.77±1.68	42.25±0.27	67.02±0.73	57.01±0.40
SFT	85.43±0.13	75.43±0.12	48.21±1.21	41.76±1.34	76.08±0.25	51.71±0.82
SFT^{+k}	87.31±0.15	76.78±0.38	51.21±1.14	37.96±0.05	77.75±0.42	59.18±1.33
UNICON	88.43±0.14	81.37±0.43	57.92±0.43	42.70±0.50	71.45±0.03	33.98±1.03
UNICON^{+k}	89.21±0.42	82.27±0.29	61.55±0.13	48.47±0.40	71.04±0.14	42.17±0.31
DISC	91.58±0.21	85.89±0.16	64.97±0.17	49.79±0.20	74.04±0.11	40.55±0.18
DISC^{+k}	**91.88**±0.15	**86.70**±0.03	**65.96**±0.15	**50.74**±0.11	75.38±0.30	63.32±0.49

Experiments (Dominant Noise)



Experiments (Asymmetric Noise)



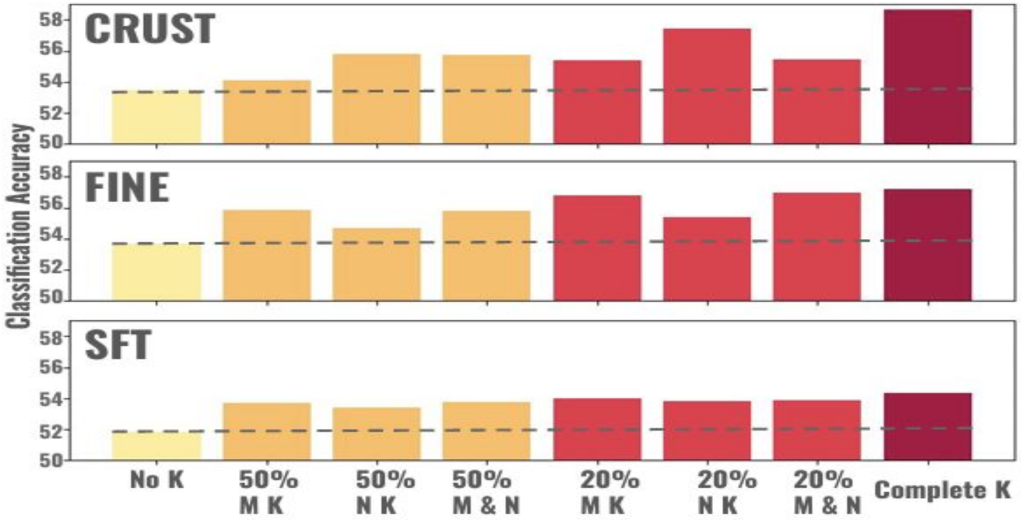
	CIFAR-10		CIFAR-100		Animal-10N
Noise ratio	0.2	0.4	0.2	0.4	
Baseline	86.12±0.42	77.18±0.30	62.96±0.12	59.07±0.08	80.32±0.20
DualT	92.24±0.10	66.23±0.03	53.61±1.49	52.03±1.92	81.14±0.28
GT-T	92.51±0.03	89.68±0.13	73.88±0.04	66.61±0.03	-
SOP	92.85±0.49	89.93±0.25	72.60±0.70	70.58±0.30	83.93±0.35
CRUST	91.94±0.05	89.40±0.03	60.75±1.87	59.79±0.89	81.88±0.13
CRUST^{+k}	89.47±0.17	84.96±0.91	62.44±0.84	61.07±0.16	81.74±0.08
FINE	89.07±0.03	85.51±0.18	65.42±0.11	65.11±0.11	81.15±0.11
FINE^{+k}	90.87±0.04	89.15±0.26	73.59±0.12	72.87±0.11	82.27±0.10
SFT	92.67±0.04	89.77±0.14	74.41±0.05	69.51±0.06	82.24±0.10
SFT^{+k}	93.19±0.08	90.55±0.06	74.29±0.14	70.94±0.13	82.88±0.18
UNICON	92.42±0.04	91.51±0.12	75.95±0.04	73.08±0.07	87.76±0.06
UNICON^{+k}	92.60±0.07	91.35±0.24	76.87±0.24	73.97±0.11	**88.28**±0.29
DISC	94.82±0.04	93.24±0.04	76.02±0.15	74.36±0.16	86.44±0.14
DISC^{+k}	**95.40**±0.08	**94.05**±0.07	**77.13**±0.05	**75.50**±0.08	86.90±0.10

Clothing1M

Baseline	DivideMix*	ELR*	CORES^2*	SOP*	UNICON	UNICON^{+k} (ours)	DISC	DISC^{+k} (ours)
69.45	74.76	72.87	73.24	73.50	74.56	**75.13**	73.30	73.87

Discussions

Incomplete or Noisy Knowledge



Discussions

Combining noise estimation and noise detection algorithms

Asym Noise Ratio		CIFAR-10 0.2	CIFAR-10 0.4	CIFAR-100 0.2	CIFAR-100 0.4	CHAMMI-CP	Animal-10N
FINE		89.07	85.51	65.42	65.11	67.27	81.15
DualT	+FINE^{+k}	89.89	88.87	66.36	62.80	**70.70**	81.84
FINE^{+k}		**90.87**	**89.15**	**73.59**	**72.87**	67.02	**82.27**

Thanks for Watching!

Siqi Wang, Bryan A. Plummer

siqiwang@bu.edu