Entropy is not Enough for Test-Time Adaptation: From the Perspective of Disentangled Factors

Jonghyun Lee*, Dahuin Jung, Saehyung Lee, Junsung Park,
Juhyeon Shin, Uiwon Hwang, Sungroh Yoon

Seoul National University

DSAIL Data Science & AI Laboratory

ICLR

*speaker

Test-Time Adaptation

- Requirements
  - The pre-trained model $\mathcal{M}_s$ trained on the source (training) data $\mathcal{D}_s$
  - Target (test) data $\mathcal{D}_t = \{x_t\}$
- Goal
  - Using $\mathcal{M}_s$ and the stream of $\mathcal{D}_t$, obtain the best performance on target domain
- Constraints
  - No source (train) data $\mathcal{D}_s$
  - Efficiency (memory, runtime)

Test-Time Adaptation (cont'd)

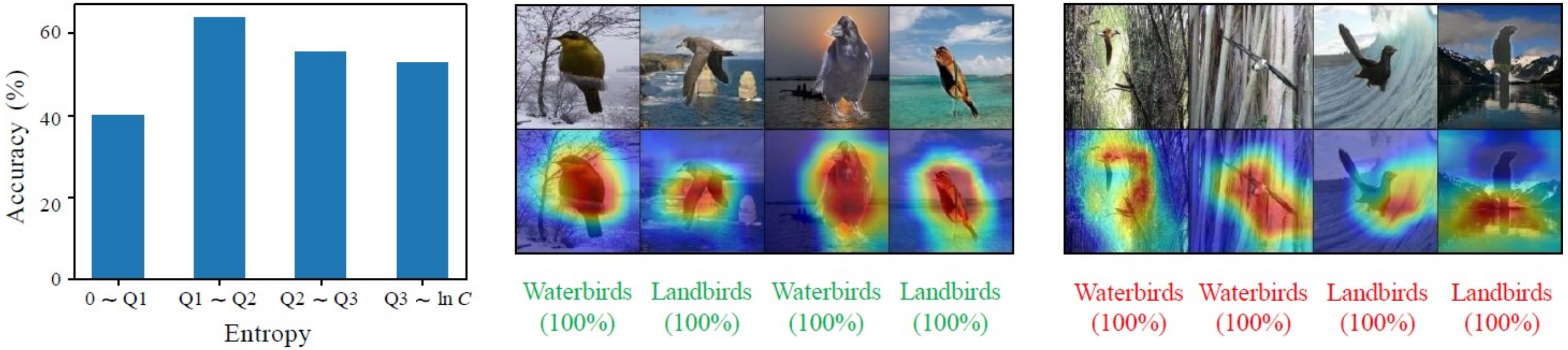
- TTA cannot access whole target data $\mathcal{D}_t$ before adaptation.
  - Impossible to estimate the target distribution
    - Prone to inaccurate prediction in the early stage
    - Error accumulation!
  - Need to adapt using samples that have a lower likelihood of causing error
    - Sample filtering and reweighting with confidence metric!

Previous work

- TENT (ICLR 2021 spotlight)
  - Entropy minimization loss
  - No filtering
- EATA (ICML 2022)
  - TENT + filtering + regularization
    - Reliable (entropy) filtering + redundant filtering
    - Fisher regularization
- SAR (ICLR 2023 oral)
  - Reliable (entropy) filtering + sharpness-aware entropy minimization loss

Previous work (cont'd)

- Observations



(a) Entropy level vs accuracy (b) Grad-CAM of correct samples (c) Grad-CAM of wrong samples

- The lowest entropy interval shows the lowest accuracy
  - Unreliable!
- Samples in 0~Q1
  - Correct samples: focus on bird (target) well
  - Wrong samples: relatively bad

Entropy is not enough for TTA!

- Entropy cannot reflect 'which part' of the image assigns low entropy.
  - If birds (object)? ➔ good!
  - If spurious features (background)? ➔ bad…

- Let's take 'disentangled factors' notation!

Entropy is not enough for TTA! (cont'd)

- Disentangled factors
  - Disentangled latent vector $\mathbf{v}(\mathbf{x}) = (\mathrm{v}_0(\mathbf{x}), \mathrm{v}_1(\mathbf{x}), \dots, \mathrm{v}_{d_v}(\mathbf{x})) \in \mathbb{R}^{d_v}$
    - Each element is independent
    - $\mathbf{v}(\mathbf{x})$ can perfectly reconstruct an input image $\mathbf{x}$
    - $\mathrm{v}_i(\mathbf{x}) \in [0,1]$: the i-th factor of $\mathbf{x}$
  - In binary classification ($\mathrm{y} \in \{-1, +1\}$)
    - Each factor $\mathrm{v}_i(\mathbf{x})$ has a correlation with a true label y
    - Under distribution shift, the correlation can be changed!
    - Define two correlations
      - $\mathrm{corr}_i^{\mathrm{train}} = \mathrm{corr}(\mathrm{y}^{\mathrm{train}}, \mathrm{v}_i^{\mathrm{train}})$, $\mathrm{corr}_i^{\mathrm{test}} = \mathrm{corr}(\mathrm{y}^{\mathrm{test}}, \mathrm{v}_i^{\mathrm{test}})$
    - Then we could divide $\mathbf{v}(\mathbf{x})$ into four partitions

$$\mathbf{v}_{pp} = \{\mathrm{v}_i | \mathrm{corr}_i^{\mathrm{train}} > 0, \mathrm{corr}_i^{\mathrm{test}} > 0\}, \quad \mathbf{v}_{pn} = \{\mathrm{v}_i | \mathrm{corr}_i^{\mathrm{train}} > 0, \mathrm{corr}_i^{\mathrm{test}} \leq 0\},$$
$$\mathbf{v}_{np} = \{\mathrm{v}_i | \mathrm{corr}_i^{\mathrm{train}} \leq 0, \mathrm{corr}_i^{\mathrm{test}} > 0\}, \quad \mathbf{v}_{nn} = \{\mathrm{v}_i | \mathrm{corr}_i^{\mathrm{train}} \leq 0, \mathrm{corr}_i^{\mathrm{test}} \leq 0\}.$$

Entropy is not enough for TTA! (cont'd)

- Assume the pretrained model $\mathcal{M}_\theta$ as a linear classifier.

Proposition 1. *Let us consider a pre-trained linear classifier $\mathcal{M}_{\boldsymbol{\theta}}$ that uses the latent disentangled factors $\mathbf{v}(\mathbf{x})$ of sample $\mathbf{x}$ as input. We define a **harmful** sample as one that reduces the difference in the mean logits between classes when used for adaptation. A sample $\mathbf{x} \in \mathcal{X}^{\text{test}}$ is a **harmful** sample for adaptation using entropy minimization loss if it satisfies the following condition:*

$$\hat{\mathrm{y}}\mathbf{v}(\mathbf{x}) \cdot \left(\mathbb{E}_{\mathbf{x}^{\text{test}} \sim \mathcal{X}^{\text{test}}_{+1}}[\mathbf{v}(\mathbf{x}^{\text{test}})] - \mathbb{E}_{\mathbf{x}^{\text{test}} \sim \mathcal{X}^{\text{test}}_{-1}}[\mathbf{v}(\mathbf{x}^{\text{test}})]\right) < 0, \tag{5}$$

where $\mathcal{X}^{\text{test}}_y = \{\mathbf{x} | (\mathbf{x}, \mathrm{y}) \in \mathcal{D}^{\text{test}}, \mathrm{y} = y\}$, and $y \in \{1, -1\}$.

  - With Proposition 1, we can explain why the samples with low entropy can be ***harmful***.

Entropy is not enough for TTA! (cont'd)

- The partitions of optimal parameters $\boldsymbol{\theta}^*$ for the training data

$$\mathbf{v}_{pp} = \{\mathrm{v}_i | \mathrm{corr}_i^{\mathrm{train}} > 0, \mathrm{corr}_i^{\mathrm{test}} > 0\}, \quad \mathbf{v}_{pn} = \{\mathrm{v}_i | \mathrm{corr}_i^{\mathrm{train}} > 0, \mathrm{corr}_i^{\mathrm{test}} \leq 0\}, \qquad \boldsymbol{\theta}_{pp}^*, \boldsymbol{\theta}_{pn}^* > 0$$
$$\mathbf{v}_{np} = \{\mathrm{v}_i | \mathrm{corr}_i^{\mathrm{train}} \leq 0, \mathrm{corr}_i^{\mathrm{test}} > 0\}, \quad \mathbf{v}_{nn} = \{\mathrm{v}_i | \mathrm{corr}_i^{\mathrm{train}} \leq 0, \mathrm{corr}_i^{\mathrm{test}} \leq 0\}. \qquad \boldsymbol{\theta}_{np}^*, \boldsymbol{\theta}_{nn}^* \leq 0$$

- In the early stages of adaptation, $\mathbf{x}$ with a high-confidence pseudo-label of $\hat{\mathrm{y}} = +1$ satisfies

$$\mathrm{a}_{\boldsymbol{\theta}}(\mathbf{x}) = \boldsymbol{\theta}_{pp} \cdot \mathbf{v}_{pp} + \boldsymbol{\theta}_{pn} \cdot \mathbf{v}_{pn} + \boldsymbol{\theta}_{np} \cdot \mathbf{v}_{np} + \boldsymbol{\theta}_{nn} \cdot \mathbf{v}_{nn} \gg 0,$$
$$|\boldsymbol{\theta}_{pp} \cdot \mathbf{v}_{pp} + \boldsymbol{\theta}_{pn} \cdot \mathbf{v}_{pn}| \gg |\boldsymbol{\theta}_{np} \cdot \mathbf{v}_{np} + \boldsymbol{\theta}_{nn} \cdot \mathbf{v}_{nn}|.$$

- Two dominant factors
  - $\mathbf{v}_{pp}$: Commonly Positively-coRrelated with label (CPR) factors
  - $\mathbf{v}_{pn}$: TRAin-time only Positively-correlated with label (TRAP) factors

Entropy is not enough for TTA! (cont'd)

- By definition, the expectations of CPR factors and TRAP factors are as follows:

$$\mathbb{E}_{\mathbf{x}^{\text{test}} \sim \mathcal{X}_{+1}^{\text{test}}}[\mathbf{v}_{pp}(\mathbf{x}^{\text{test}})] > \mathbb{E}_{\mathbf{x}^{\text{test}} \sim \mathcal{X}_{-1}^{\text{test}}}[\mathbf{v}_{pp}(\mathbf{x}^{\text{test}})],$$
$$\mathbb{E}_{\mathbf{x}^{\text{test}} \sim \mathcal{X}_{+1}^{\text{test}}}[\mathbf{v}_{pn}(\mathbf{x}^{\text{test}})] \leq \mathbb{E}_{\mathbf{x}^{\text{test}} \sim \mathcal{X}_{-1}^{\text{test}}}[\mathbf{v}_{pn}(\mathbf{x}^{\text{test}})].$$

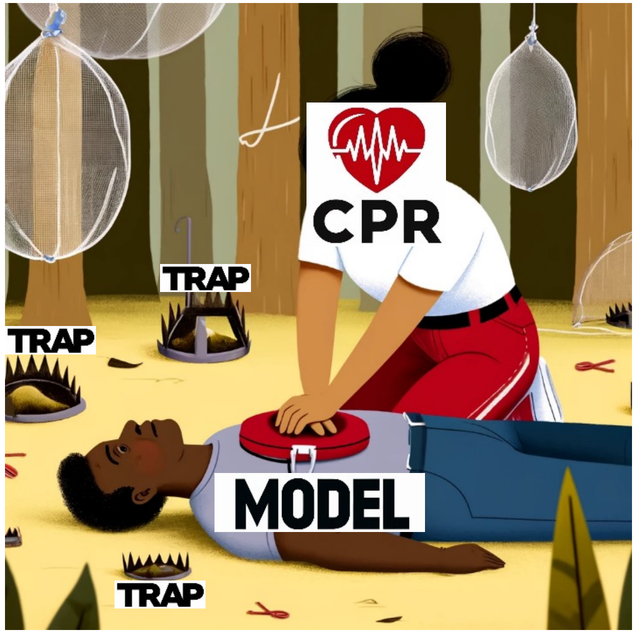
- Reformulation of Eq. (5) of Proposition 1.

$$\begin{aligned}
&\hat{\mathbf{y}}\mathbf{v}(\mathbf{x})\cdot(\mathbb{E}_{\mathbf{x}^{\text{test}} \sim \mathcal{X}_{+1}^{\text{test}}}[\mathbf{v}(\mathbf{x}^{\text{test}})] - \mathbb{E}_{\mathbf{x}^{\text{test}} \sim \mathcal{X}_{-1}^{\text{test}}}[\mathbf{v}(\mathbf{x}^{\text{test}})])\\
&\approx \underbrace{\mathbf{v}_{pp}(\mathbf{x})\cdot(\mathbb{E}_{\mathbf{x}^{\text{test}} \sim \mathcal{X}_{+1}^{\text{test}}}[\mathbf{v}_{pp}(\mathbf{x}^{\text{test}})] - \mathbb{E}_{\mathbf{x}^{\text{test}} \sim \mathcal{X}_{-1}^{\text{test}}}[\mathbf{v}_{pp}(\mathbf{x}^{\text{test}})])}_{(7.\text{a})}\\
&+ \underbrace{\mathbf{v}_{pn}(\mathbf{x})\cdot(\mathbb{E}_{\mathbf{x}^{\text{test}} \sim \mathcal{X}_{+1}^{\text{test}}}[\mathbf{v}_{pn}(\mathbf{x}^{\text{test}})] - \mathbb{E}_{\mathbf{x}^{\text{test}} \sim \mathcal{X}_{-1}^{\text{test}}}[\mathbf{v}_{pn}(\mathbf{x}^{\text{test}})])}_{(7.\text{b})} < 0
\end{aligned}$$

  - Eq. (7.a) (related to CPR factors) becomes positive, and Eq. (7.b) (related to TRAP factors) becomes negative
    - $|(7.\text{a})| \ll |(7.b)|$ → a ***harmful*** sample
      - When TRAP factors affect more than CPR factors, the sample becomes ***harmful***

Entropy is not enough for TTA! (cont'd)

- If $|(7.\mathrm{a})| \ll |(7.b)|$ and $|\boldsymbol{\theta}_{pp} \cdot \mathbf{v}_{pp} + \boldsymbol{\theta}_{pn} \cdot \mathbf{v}_{pn}| \gg |\boldsymbol{\theta}_{np} \cdot \mathbf{v}_{np} + \boldsymbol{\theta}_{nn} \cdot \mathbf{v}_{nn}|$
  → $\mathbf{x}$ becomes a ***harmful*** sample with low entropy (high confidence)
  - Low entropy filtering cannot discern good & bad samples
- Then how to adapt?
  - Utilize the CPR factors and avoid the TRAP factors!



All sources are created by DALL·E

Methodology

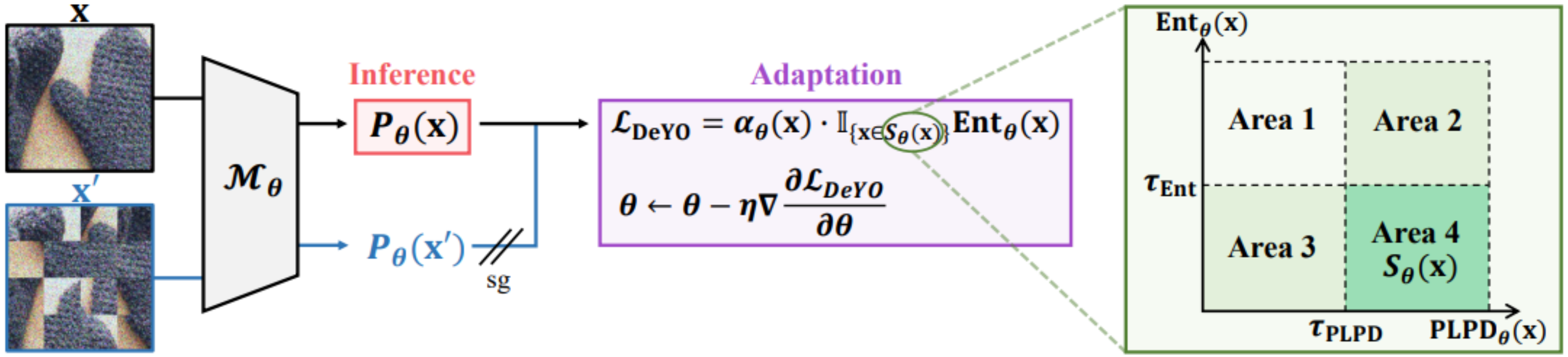
- Destroy Your Object (DeYO)
  - Utilize the factor that aligns with g.t. class (CPR factor: $\mathbf{v}_{pp}$) under **any** test distribution
    - Classification task: the shape of object!
  - We can simply apply patch shuffling to destroy the shape information, preserving the patch-level local features.
  - If a prediction becomes uncertain when the shape of object is destroyed,
    → The model considers the shape of the object as the dominant factor when classifying the sample.
  - Pseudo-Label Probability Difference (PLPD)
    - Measures the extent to which the probability of pseudo-label decreases after applying the patch shuffling
  - Utilize samples with low entropy & high PLPD!

$$S_{\boldsymbol{\theta}}(\mathbf{x}) = \{\mathbf{x} | \mathrm{Ent}_{\boldsymbol{\theta}}(\mathbf{x}) < \tau_{\mathrm{Ent}},\ \mathrm{PLPD}_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{x}') > \tau_{\mathrm{PLPD}}\},\ \text{where}$$

$$\mathrm{PLPD}_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{x}') = (\mathbf{p}_{\boldsymbol{\theta}}(\mathbf{x}) - \mathbf{p}_{\boldsymbol{\theta}}(\mathbf{x}'))_{\hat{\mathrm{y}}},$$

Methodology (cont'd)

- Overall procedure of DeYO



$$S_{\theta}(\mathbf{x}) = \{\mathbf{x} | \mathrm{Ent}_{\theta}(\mathbf{x}) < \tau_{\mathrm{Ent}}, \boxed{\mathrm{PLPD}_{\theta}(\mathbf{x}, \mathbf{x}') > \tau_{\mathrm{PLPD}}}\}, \text{ where}$$

$$\mathrm{PLPD}_{\theta}(\mathbf{x}, \mathbf{x}') = (\mathbf{p}_{\theta}(\mathbf{x}) - \mathbf{p}_{\theta}(\mathbf{x}'))_{\hat{y}},$$

$$\alpha_{\theta}(\mathbf{x}) = \frac{1}{\exp\{(\mathrm{Ent}_{\theta}(\mathbf{x}) - \mathrm{Ent}_0)\}} + \boxed{\frac{1}{\exp\{-\mathrm{PLPD}_{\theta}(\mathbf{x}, \mathbf{x}')\}}},$$

$$\mathcal{L}_{\mathrm{DeYO}}(\mathbf{x}; \boldsymbol{\theta}) = \alpha_{\theta}(\mathbf{x}) \cdot \mathbb{I}_{\{\mathbf{x} \in S_{\theta}(\mathbf{x})\}} \mathrm{Ent}_{\theta}(\mathbf{x}),$$

Experiments

- TTA on an i.i.d. sampling (mild) scenario (ResNet-50-BN)

Table 1: Comparisons with baselines on ImageNet-C at severity level 5 under a mild scenario regarding accuracy (%). The **bold** value signifies the top-performing result.

	Noise			Blur				Weather				Digital				
Mild	Gauss.	Shot	Impul.	Defoc.	Glass	Motion	Zoom	Snow	Frost	Fog	Brit.	Contr.	Elastic	Pixel	JPEG	Avg.
ResNet-50-BN	2.2	2.9	1.8	17.9	9.8	14.8	22.5	16.9	23.3	24.4	58.9	5.4	16.9	20.7	31.7	18.0
• MEMO	7.5	8.8	8.9	19.8	13.0	20.7	27.7	25.3	28.7	32.2	61.0	11.0	23.8	33.0	37.6	23.9
• Tent	29.2	31.2	30.1	28.1	27.7	41.4	49.4	47.2	41.5	57.7	67.4	29.2	54.8	58.5	52.4	43.1
• EATA	34.9	37.1	35.8	33.4	33.0	47.1	52.7	51.6	45.7	60.0	**68.1**	44.4	57.9	60.6	55.1	47.8
• SAR	30.6	30.6	31.3	28.5	28.5	41.9	49.4	47.1	42.2	57.5	67.3	37.8	54.6	58.4	52.1	43.9
• DeYO (ours)	**35.6**$_{\pm0.2}$	**37.9**$_{\pm0.1}$	**37.1**$_{\pm0.1}$	**33.8**$_{\pm0.2}$	**34.1**$_{\pm0.2}$	**48.5**$_{\pm0.1}$	**52.8**$_{\pm0.1}$	**52.7**$_{\pm0.0}$	**46.4**$_{\pm0.1}$	**60.6**$_{\pm0.0}$	68.0$_{\pm0.1}$	**46.1**$_{\pm0.1}$	**58.4**$_{\pm0.1}$	**61.5**$_{\pm0.1}$	**55.7**$_{\pm0.1}$	**48.6**$_{\pm0.0}$

- TTA on a spurious correlations shift (biased) scenario (ResNet-18/50-BN)

Table 2: Comparisons with baselines on ColoredMNIST regarding accuracy (%).

Biased	Avg Acc	Worst-Group Acc
ResNet-18-BN	63.40	20.05
• Tent	57.06	9.80
• MEMO	63.77	6.23
• SENTRY	63.23	15.78
• EATA	60.81	17.98
• SAR	58.37	12.36
• DeYO (ours)	**78.24**	**67.39**

Table 3: Comparisons with baselines on WaterBirds regarding accuracy (%).

Biased	Avg Acc	Worst-Group Acc
ResNet-50-BN	83.16	64.90
• Tent	82.95	54.14
• MEMO	82.34	50.47
• SENTRY	85.77	60.90
• EATA	82.38	52.38
• SAR	82.60	53.41
• DeYO (ours)	**87.42**	**73.92**

Experiments (cont'd)

- TTA on wild scenarios (ResNet-50-GN, ViT-16/B)
  - Temporally-correlated label shifts
  - Batch size 1

Table 5: Comparisons with baselines on ImageNet-C at severity level 5 under online imbalanced label shifts (imbalance ratio = ∞) or under batch size 1 regarding accuracy (%).

	Noise			Blur				Weather				Digital				
Label Shifts	Gauss.	Shot	Impul.	Defoc.	Glass	Motion	Zoom	Snow	Frost	Fog	Brit.	Contr.	Elastic	Pixel	JPEG	Avg.
ResNet-50-GN	17.9	19.9	17.9	19.7	11.3	21.3	24.9	40.4	47.4	33.6	69.3	36.3	18.7	28.4	52.2	30.6
• MEMO	18.4	20.6	18.4	17.1	12.7	21.8	26.9	40.7	46.9	34.8	69.6	36.4	19.2	32.2	53.4	31.3
• Tent	3.6	4.2	4.4	16.5	5.9	26.9	28.4	17.9	26.2	2.3	72.2	46.1	7.3	52.3	56.2	24.7
• EATA	25.7	28.6	24.8	18.5	19.6	24.1	28.4	35.3	33.0	**41.2**	65.2	33.3	28.0	42.4	43.1	32.7
• SAR	33.7	36.9	35.3	19.3	**20.3**	33.8	**29.8**	21.9	44.7	34.9	71.9	46.7	6.6	52.3	56.2	36.3
• DeYO (ours)	**42.5** $_{\pm0.5}$	**44.9** $_{\pm0.2}$	**43.8** $_{\pm0.3}$	**22.2** $_{\pm0.0}$	16.3 $_{\pm10.2}$	**41.0** $_{\pm0.2}$	13.2 $_{\pm9.8}$	**52.2** $_{\pm0.4}$	**51.5** $_{\pm0.5}$	39.7 $_{\pm27.4}$	**73.4** $_{\pm0.1}$	**52.6** $_{\pm0.2}$	**46.9** $_{\pm1.2}$	**59.3** $_{\pm0.1}$	**59.3** $_{\pm0.0}$	**43.9** $_{\pm2.0}$
VitBase-LN	9.4	6.7	8.3	29.1	23.4	34.0	27.1	15.8	26.4	47.4	54.7	44.0	30.5	44.5	47.6	29.9
• MEMO	21.6	17.4	20.6	37.1	29.6	40.6	34.4	25.0	34.8	55.2	65.0	54.9	37.4	55.5	57.7	39.1
• Tent	33.9	1.8	27.2	54.8	52.9	58.6	**54.3**	12.4	11.7	69.7	76.3	66.3	59.6	69.7	66.6	47.7
• EATA	36.2	34.7	35.5	43.4	44.3	49.3	48.5	53.2	53.5	62.3	72.7	18.8	58.0	64.7	62.8	49.2
• SAR	42.3	34.9	44.1	50.0	50.5	55.6	53.1	59.7	47.2	66.2	75.2	50.3	60.1	67.3	65.0	54.8
• DeYO (ours)	**53.5** $_{\pm0.5}$	**36.0** $_{\pm25.2}$	**54.6** $_{\pm0.8}$	**57.6** $_{\pm0.2}$	**58.7** $_{\pm0.2}$	**63.7** $_{\pm0.1}$	46.2 $_{\pm18.7}$	**67.6** $_{\pm0.1}$	**66.0** $_{\pm0.1}$	**73.2** $_{\pm0.2}$	**77.9** $_{\pm0.1}$	**66.7** $_{\pm0.1}$	**69.0** $_{\pm0.1}$	**73.5** $_{\pm0.1}$	**70.3** $_{\pm0.2}$	**62.3** $_{\pm1.7}$

	Noise			Blur				Weather				Digital				
Batch Size 1	Gauss.	Shot	Impul.	Defoc.	Glass	Motion	Zoom	Snow	Frost	Fog	Brit.	Contr.	Elastic	Pixel	JPEG	Avg.
ResNet-50-GN	18.0	19.8	17.9	19.8	11.4	21.4	24.9	40.4	47.3	33.6	69.3	36.3	18.6	28.4	52.3	30.6
• MEMO	18.5	20.5	18.4	17.1	12.6	21.8	26.9	40.4	47.0	34.4	69.5	36.5	19.2	32.1	53.3	31.2
• Tent	3.1	4.2	4.0	16.5	5.3	27.4	**30.3**	17.7	24.9	2.0	72.1	46.2	7.8	52.6	56.3	24.7
• EATA	24.8	27.9	25.8	17.9	17.3	28.7	29.3	44.7	44.4	**40.2**	71.0	44.5	27.0	46.8	55.6	36.4
• SAR	23.3	26.6	23.9	18.5	15.2	28.6	**30.3**	44.0	44.7	29.0	72.3	44.6	13.1	46.8	56.1	34.5
• DeYO (ours)	**41.8** $_{\pm0.7}$	**44.7** $_{\pm0.4}$	**43.0** $_{\pm0.7}$	**22.5** $_{\pm0.1}$	**24.7** $_{\pm0.3}$	**41.8** $_{\pm0.1}$	24.4 $_{\pm9.8}$	**54.5** $_{\pm0.2}$	**52.2** $_{\pm0.2}$	20.7 $_{\pm26.8}$	**73.5** $_{\pm0.0}$	**53.5** $_{\pm0.2}$	**48.5** $_{\pm0.3}$	**60.2** $_{\pm0.0}$	**59.8** $_{\pm0.1}$	**44.4** $_{\pm1.2}$
VitBase-LN	9.5	6.8	8.2	29.0	23.5	33.9	27.1	15.9	26.5	47.2	54.7	44.1	30.5	44.5	47.8	29.9
• MEMO	21.6	17.3	20.6	37.1	29.6	40.4	34.4	24.9	34.7	55.1	64.8	54.9	37.4	55.4	57.6	39.1
• Tent	43.0	1.6	43.9	52.8	48.8	55.9	51.3	22.9	21.1	66.9	75.1	65.0	54.0	67.0	64.3	48.9
• EATA	32.2	26.7	30.3	43.8	40.1	47.7	42.6	35.7	43.4	60.8	65.6	61.1	46.5	60.5	58.2	46.3
• SAR	40.6	36.9	41.9	53.7	50.5	57.4	52.8	58.9	52.7	68.9	76.0	65.8	57.9	68.9	65.8	56.6
• DeYO (ours)	**54.0** $_{\pm0.7}$	**52.1** $_{\pm3.6}$	**55.1** $_{\pm0.8}$	**58.8** $_{\pm0.1}$	**59.5** $_{\pm0.1}$	**64.2** $_{\pm0.1}$	**53.5** $_{\pm5.5}$	**68.2** $_{\pm0.1}$	**66.4** $_{\pm0.0}$	**73.7** $_{\pm0.1}$	**78.3** $_{\pm0.0}$	**68.2** $_{\pm0.1}$	**68.9** $_{\pm0.1}$	**73.8** $_{\pm0.1}$	**70.8** $_{\pm0.3}$	**64.4** $_{\pm0.7}$

Thank you!

- **TL;DR**
  - Address the limitations of relying solely on entropy as a confidence metric for TTA.
- **Summary**
  - Theoretically prove why entropy is not enough for TTA.
    - Entropy cannot discern the CPR and TRAP factors.
  - Introduce an effective TTA method based on the proposed novel confidence metric.
  - Achieve state-of-the-art performances in various TTA scenarios.
- More details can be found:
  - Paper: https://openreview.net/forum?id=9w3iw8wDuE
  - Project page: https://whitesnowdrop.github.io/DeYO/
  - Code: https://github.com/Jhyun17/DeYO
  - Poster Session: Tue 7 May 10:45 am - 12:45 pm at Halle B