ActMAD: Activation Matching to Align Distributions for Test-Time Training

CVPR 2023

THU-PM-339

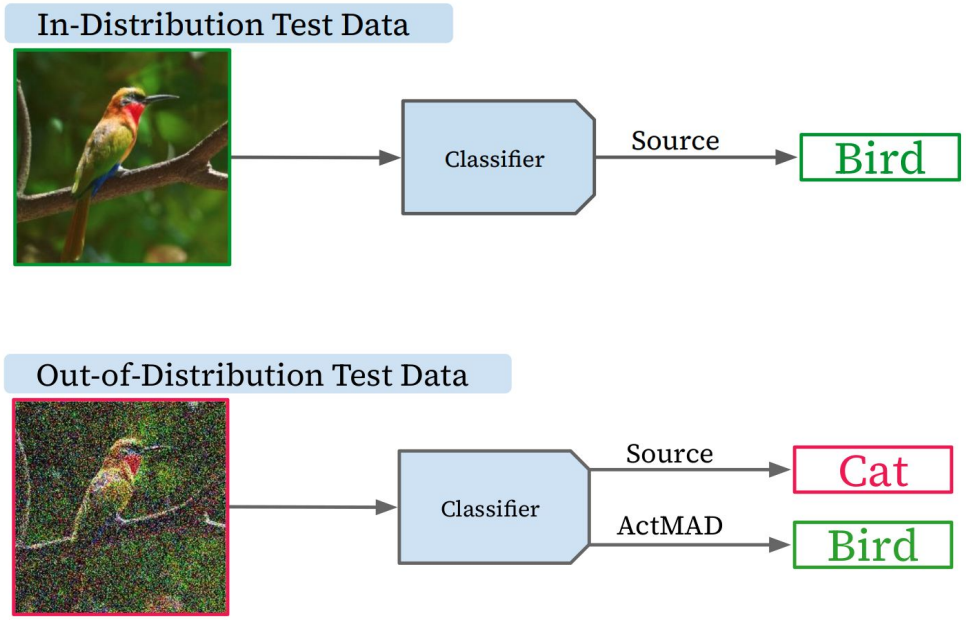
Muhammad Jehanzeb Mirza[1,2], Pol Jane Soneria[3], Wei Lin[1], Mateusz Kozinski[1], Horst Possegger[1], Horst Bischof[1,2]

[1] Institute for Computer Graphics and Vision (ICG), Graz University of Technology.

[2] Christian Doppler Laboratory for Embedded Machine Learning.

[3] Karlsruhe Institute for Technology (KIT), Germany.

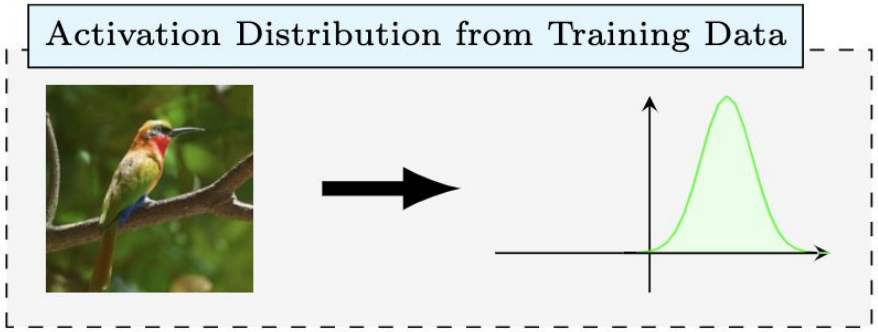
Motivation

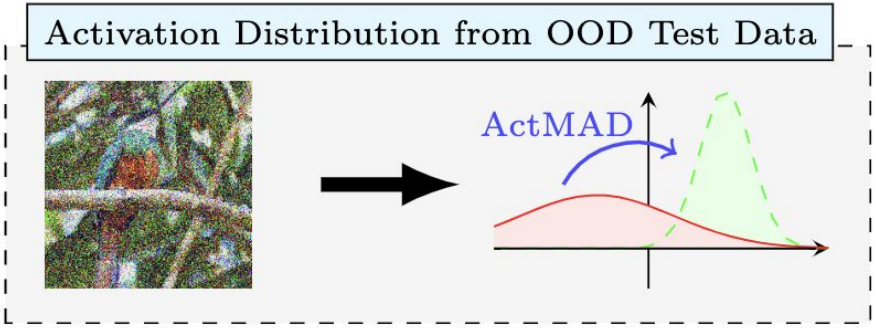


Test-Time-Training [1]

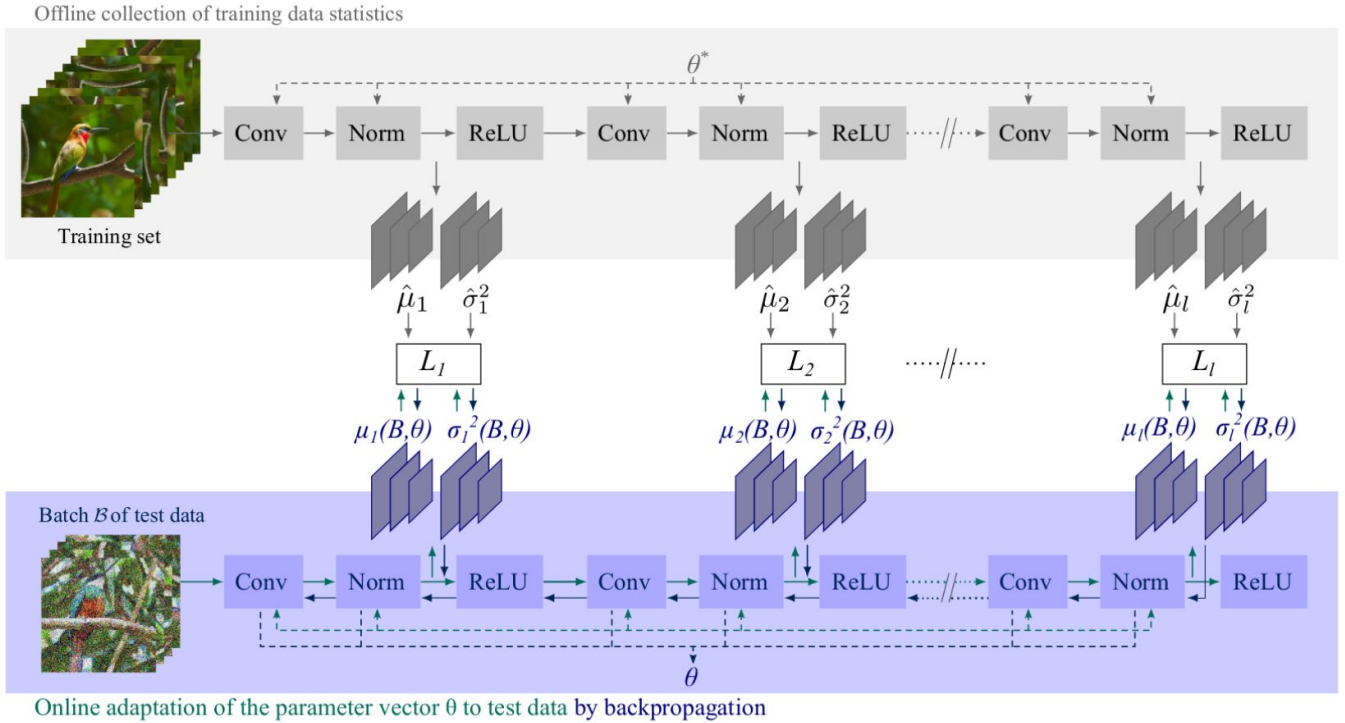
[1] Sun et al. "Test-time training with self-supervision for generalization under distribution shifts." ICML, 2020.

Overview

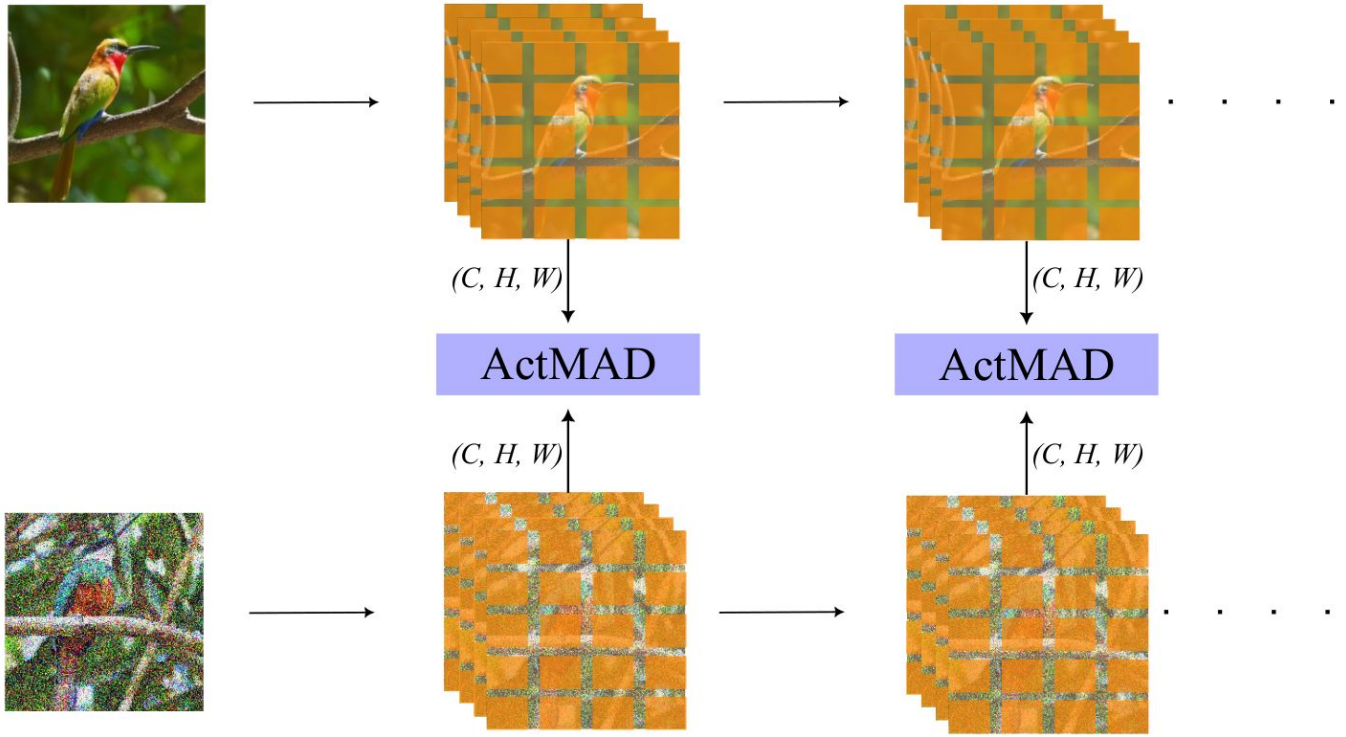

Activation Distribution from Training Data


Activation Distribution from OOD Test Data
ActMAD

Method


Offline collection of training data statistics
Training set
θ^*
Conv
Norm
ReLU
Conv
Norm
ReLU
Conv
Norm
ReLU
$\hat{\mu}_1$
$\hat{\sigma}_1^2$
$\hat{\mu}_2$
$\hat{\sigma}_2^2$
$\hat{\mu}_l$
$\hat{\sigma}_l^2$
L_1
L_2
L_l
$\mu_1(B,\theta)$ $\sigma_1^2(B,\theta)$
$\mu_2(B,\theta)$ $\sigma_2^2(B,\theta)$
$\mu_l(B,\theta)$ $\sigma_l^2(B,\theta)$
Batch $\mathcal{B}$ of test data
Conv
Norm
ReLU
Conv
Norm
ReLU
Conv
Norm
ReLU
θ
Online adaptation of the parameter vector θ to test data by backpropagation

Location Aware Feature Alignment



Results - Image Classification

Corruptions:	Gauss	Shot	Impul	Defcs	Gls	Mtn	Zm	Snw	Frst	Fg	Brt	Cnt	Els	Px	Jpg	Mean
Source	98.4	97.7	98.4	90.6	92.5	89.8	81.8	89.5	85.0	86.3	51.1	97.2	85.3	76.9	71.7	86.2
SHOT† (Offline)	73.8	70.5	72.2	79.2	80.6	58.5	54.0	53.6	63.0	47.3	39.2	97.7	48.7	46.1	53.0	62.5
TTT†	96.9	95.5	96.5	89.9	93.2	86.5	81.5	82.9	82.1	80.0	53.0	85.6	79.1	77.2	74.7	83.6
DUA*	89.4	87.6	88.1	88.0	88.6	84.7	74.3	77.8	78.4	68.6	45.6	95.9	72.2	66.5	67.4	78.2
NORM*	87.1	90.6	89.5	87.6	93.4	80.0	71.9	70.6	81.5	65.9	46.8	89.8	73.5	63.2	67.5	77.3
T3A*	85.5	84.0	85.0	86.6	85.9	76.1	65.4	70.3	71.0	58.7	41.3	86.8	60.5	54.4	61.0	71.5
SHOT† (Online)	83.9	82.3	83.7	83.9	83.8	72.6	61.9	65.7	68.6	54.8	39.4	85.9	58.1	53.1	62.3	69.3
P-L*	82.0	79.7	81.5	84.2	83.0	71.0	60.7	65.4	68.6	52.9	41.7	82.6	55.5	51.1	55.7	67.7
TENT*	80.8	78.6	80.4	82.5	82.5	72.1	60.5	63.7	66.7	52.1	**39.2**	84.2	55.5	50.8	58.2	67.2
CFA†	78.2	**76.4**	78.2	81.9	80.4	**69.6**	**60.1**	**63.4**	67.6	**52.0**	41.5	79.5	**54.3**	**50.2**	**55.1**	**65.9**
ActMAD†	**76.3**	77.4	**77.4**	**76.1**	**75.4**	72.0	62.8	66.6	**65.8**	55.8	40.9	**78.8**	55.7	51.4	57.6	66.0

Top-1 Classification Error (%) for all corruptions in ImageNet-C (level 5). Lower is better. All results are for a ResNet-18 network pre-trained on the clean train set. Source denotes its performance on the corrupted test data without any adaptation.

(* - Fully Test-Time Approaches, † - Approaches requiring some supervision from the training data)

Results - Object Detection

(a) KITTI-Clear → KITTI-Fog

	car	van	truck	ped	sit	cyc	tram	misc	Mean
Source	31.3	15.0	6.0	34.8	33.6	20.2	6.7	9.1	19.6
TTT†	42.6	19.5	10.5	49.7	51.4	31.0	10.5	20.2	29.4
DUA*	51.4	13.5	9.1	48.1	57.3	36.3	14.5	18.0	31.0
NORM*	50.1	27.6	12.6	47.6	50.0	30.9	17.7	21.7	32.3
ActMAD†	**67.0**	**41.2**	**25.5**	**62.2**	**68.7**	**50.9**	**30.5**	**35.7**	**47.7**

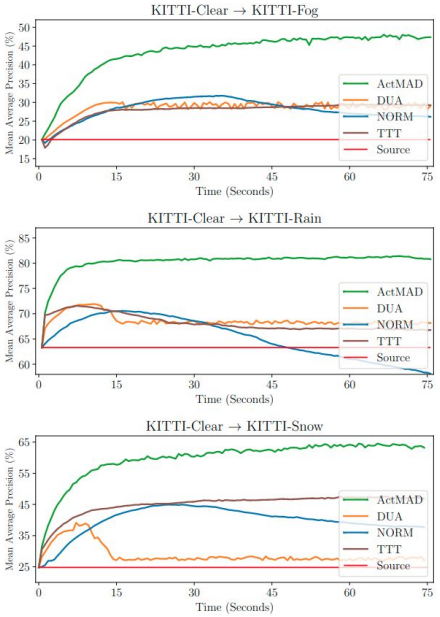
(b) KITTI-Clear → KITTI-Rain

Source	86.4	69.6	58.6	68.6	63.7	60.2	64.5	60.4	66.5
TTT†	86.4	76.1	68.0	68.7	66.6	66.3	75.0	65.2	71.5
DUA*	88.3	70.4	70.4	70.8	**67.7**	66.8	73.5	67.5	71.9
NORM*	88.3	77.0	65.7	69.1	61.5	66.7	69.1	67.1	70.6
ActMAD†	**94.2**	**89.2**	**87.3**	**74.1**	65.6	**77.9**	**82.5**	**80.1**	**81.4**

(c) KITTI-Clear → KITTI-Snow

Source	54.8	27.8	31.7	35.7	1.3	15.7	18.2	13.3	24.8
TTT†	77.2	53.2	60.6	48.4	29.7	37.1	43.2	31.1	47.5
DUA*	64.6	38.9	49.3	44.0	20.8	22.8	27.8	25.4	36.7
NORM*	75.5	51.0	51.7	46.8	21.7	34.9	43.4	34.3	44.9
ActMAD†	**89.5**	**78.0**	**82.6**	**57.8**	**38.0**	**53.4**	**58.8**	**58.1**	**64.5**

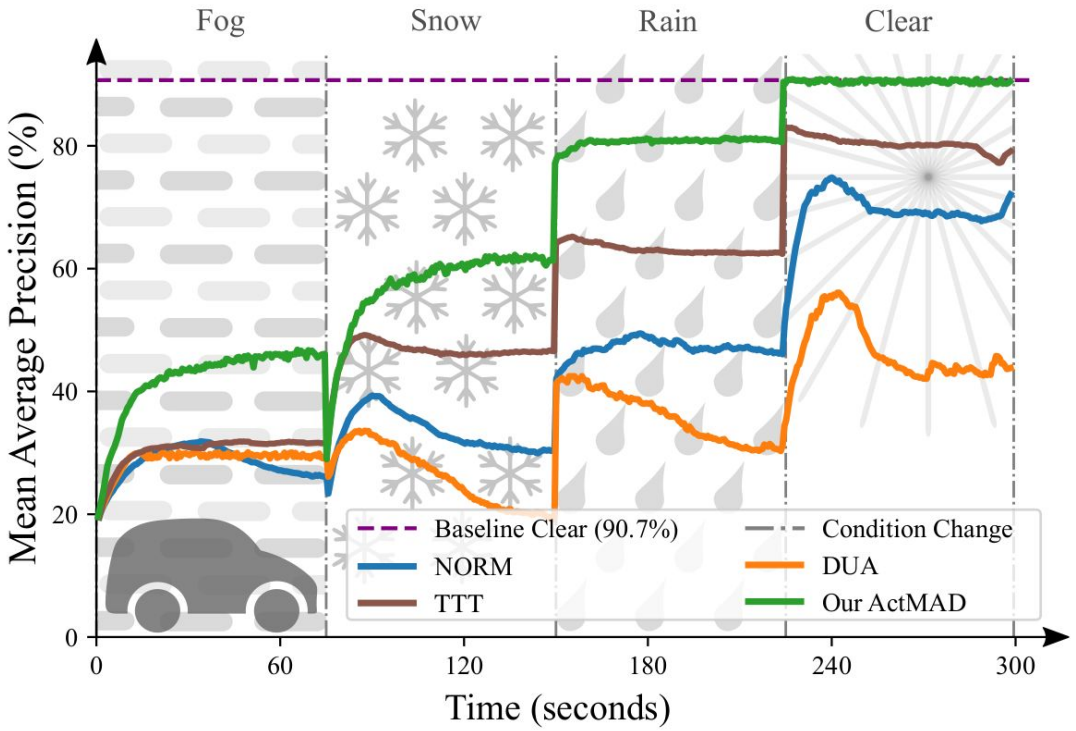
Mean Average Precision (mAP@50) for a KITTI pre-trained YOLOv3 tested on rain, fog and snow datasets. Higher is better. a) Results for the most severe fog level, i.e. only 30m visibility. b) Results for the most severe rain level, i.e. 200 mm/hr rain intensity. c) Results for snow. Best mAP is shown in bold, while the second best is underlined.



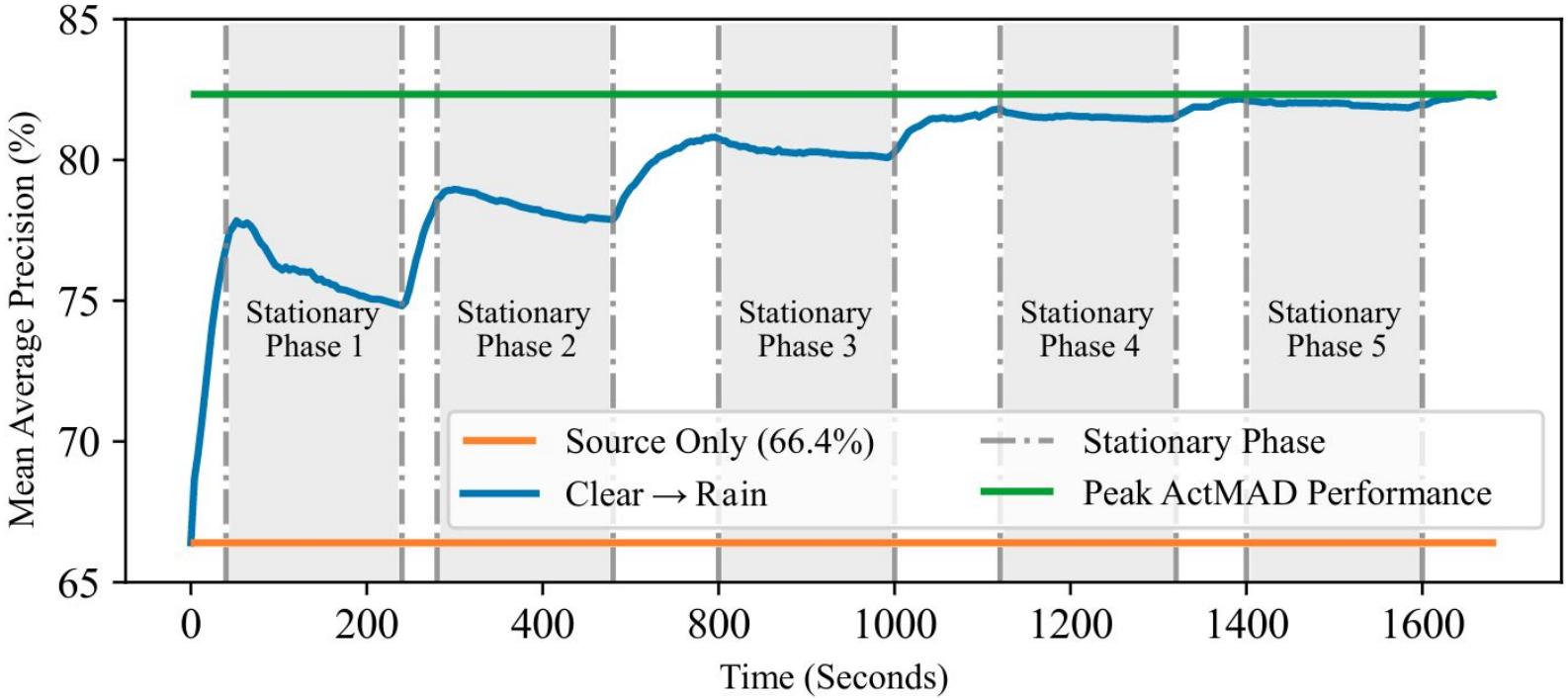
Online adaptation for each individual weather condition and comparison with baselines. We again report the Mean Average Precision (mAP@50) averaged over all the 8 classes in the KITTI dataset.

(* - Fully Test-Time Approaches, † - Approaches requiring some supervision from the training data)

Goal: Driving Continuously in Bad Weather



Adaptation in Stationary Scenarios

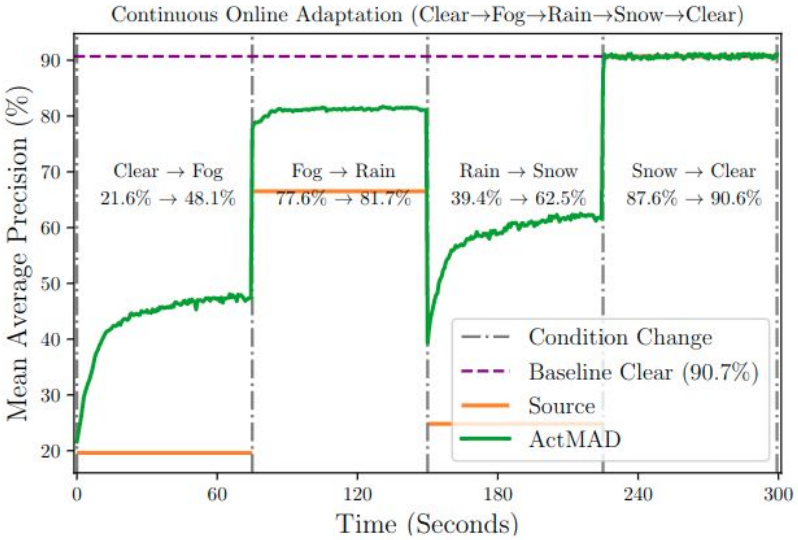


Ablation on Design Choices

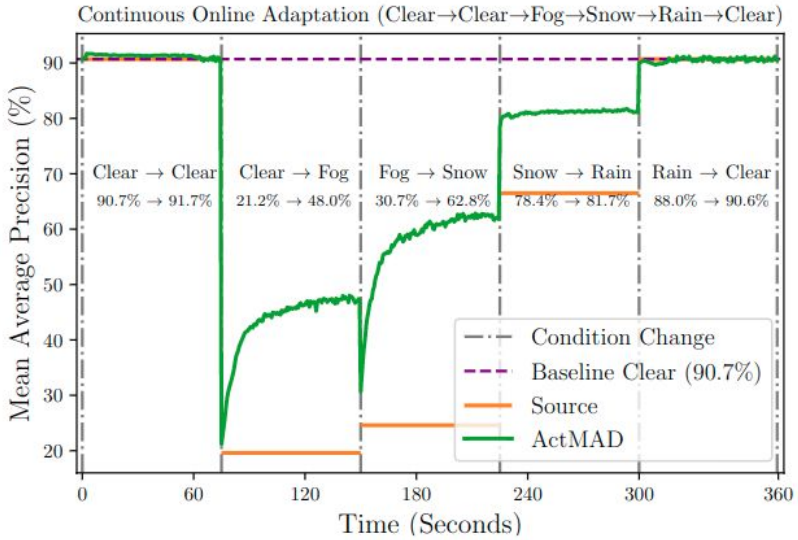
				CIFAR-10C		KITTI-FOG	
				% Error ↓	Change	mAP@50 ↑	Change
Source (no adaptation)				18.3		19.6	
Full ActMAD				10.4	0	47.7	0
Replace	*multi-layer alignment*	by	*last layer alignment*	12.3	+1.9	36.0	-11.7
Replace	*pixel statistics*	by	*channel averaged statistics*	11.5	+1.1	38.6	- 9.1
Replace	*mean and variance alignment*	by	*central moment difference*	10.5	+0.1	41.2	- 6.5
Replace	*full parameter update*	by	*only affine parameter update*	10.6	+0.2	45.2	- 2.5

Ablation study on design choices of ActMAD. We use a Wide-ResNet-40-2 for CIFAR-10C, and report the mean performance over 10 experiments for each of the 15 corruptions. We use YOLOv3 for the KITTI-FOG experiment.

Order Does Not Matter



(a) Clear→Fog→Snow→Rain→Clear



(b) Clear→Clear→Fog→Snow→Rain→Clear

Object Detection - STF - YOLOv3

(a) STF-Clear → STF-Fog

	car	ped	cyc	Mean
Source	70.2	67.5	49.5	62.4
TTT	72.5	69.8	53.4	65.2
DUA	70.4	67.7	50.6	62.9
NORM	70.5	67.4	50.1	62.7
ActMAD	**79.0**	**76.5**	**60.1**	**71.9**

(b) STF-Clear → STF-Snow

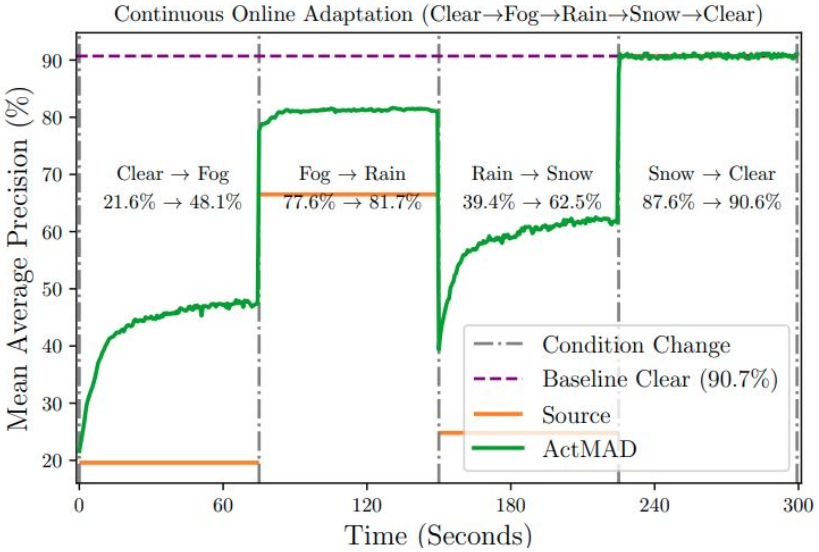
	car	ped	cyc	Mean
Source	70.3	76.0	44.1	63.5
TTT	71.1	76.3	46.6	64.7
DUA	70.2	75.9	43.9	63.3
NORM	70.8	76.0	44.1	63.6
ActMAD	**72.8**	**79.4**	**52.0**	**68.1**

Object Detection - CityScapes - Faster-RCNN

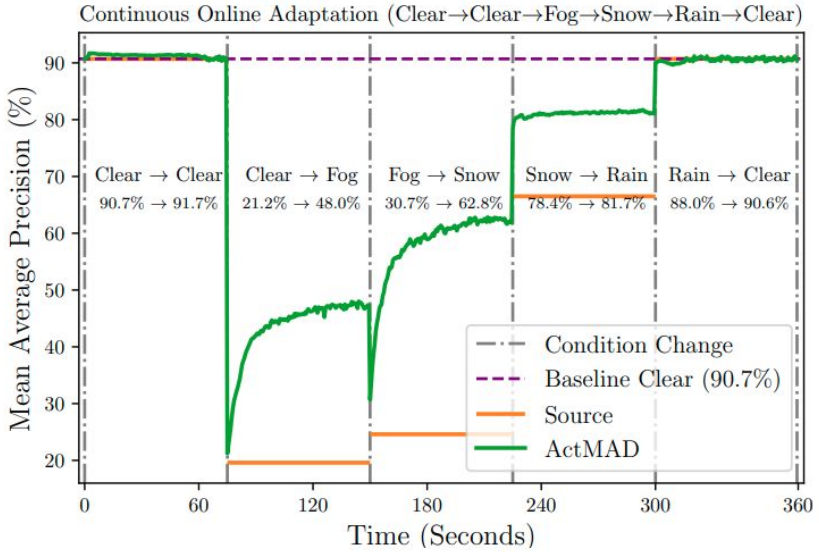
CityScapes → FoggyCityScapes

	persn	ridr	car	trck	bus	train	mcyc	bicyc	Mean
Source	29.3	34.1	35.8	15.4	26.0	9.1	22.4	29.7	25.2
TTT	30.1	36.3	36.7	19.5	29.0	13.5	27.9	37.0	28.8
DUA	31.2	37.0	36.9	20.4	27.3	15.5	**29.3**	35.5	29.1
MemCLR	32.1	41.4	43.5	21.4	33.1	11.5	25.5	32.9	30.2
ActMAD	**36.5**	**42.7**	**47.1**	**28.2**	**35.4**	**23.3**	28.2	**38.3**	**35.0**

Order of Weather Does Not Matter



(a) Clear→Fog→Snow→Rain→Clear



(b) Clear→Clear→Fog→Snow→Rain→Clear

Object Classification Results - CIFAR100

corruption:	gauss	shot	impul	defcs	gls	mtn	zm	snw	frst	fg	brt	cnt	els	px	jpg	Mean
Source	65.7	60.1	59.1	32.0	51.0	33.6	32.4	41.4	45.2	51.4	31.6	55.5	40.3	59.7	42.4	46.7
SHOT (Offline)	37.2	36.2	36.7	27.5	38.2	28.5	27.8	31.8	32.0	33.4	25.8	29.6	34.5	29.8	37.2	32.4
TTT++ (Offline)	40.7	36.4	41.5	27.5	47.8	31.1	25.1	36.5	34.7	33.7	23.3	24.7	40.2	30.5	33.3	33.8
DUA	42.2	40.9	41.0	30.5	44.8	32.2	29.9	38.9	37.2	43.6	29.5	39.2	39.0	35.3	41.2	37.7
NORM	42.5	41.8	42.6	29.7	43.9	30.6	29.7	35.7	34.6	42.2	26.9	32.8	38.1	35.5	40.9	36.5
T3A	42.4	41.8	42.5	29.7	44.3	30.5	29.5	35.9	34.5	42.1	26.8	32.8	38.0	35.9	40.7	36.5
P-L	41.3	40.5	42.5	29.6	43.1	30.3	29.4	35.8	34.3	41.7	26.7	32.4	37.8	33.5	40.8	36.0
TTT++	43.9	40.0	56.3	32.5	54.2	35.9	29.9	42.2	39.4	39.7	27.5	29.6	44.2	37.0	**37.4**	39.3
CFA	40.4	39.3	42.1	29.4	42.3	30.2	29.2	35.1	34.1	39.8	26.7	32.1	37.6	32.8	40.6	35.5
SHOT (Online)	39.7	38.9	42.1	29.0	41.9	30.2	29.3	34.8	34.2	39.7	26.7	32.2	37.2	32.5	40.4	35.3
TENT	39.9	39.1	42.2	29.0	42.0	30.2	29.3	34.9	34.2	39.7	26.7	32.3	37.4	32.4	40.4	35.3
EATA	39.1	38.5	41.2	28.9	41.8	29.2	29.1	34.0	33.8	39.1	26.6	31.9	36.6	32.1	40.1	34.8
ActMAD	39.6	38.4	39.5	29.1	41.5	30.0	29.1	34.0	33.2	40.2	26.4	31.5	36.4	31.4	38.9	34.6 ± 0.1
ActMAD+TENT	38.0	37.1	37.9	28.1	39.8	29.2	28.0	32.5	**32.3**	34.8	26.4	29.6	35.5	30.0	37.6	33.1 ± 0.3
ActMAD+EATA	**37.5**	**36.8**	**37.1**	**27.3**	**39.1**	**29.1**	**27.9**	**31.9**	32.7	**33.9**	**25.9**	**29.3**	**35.4**	**29.7**	37.5	**32.7** ± 0.5

Object Classification Results - ImageNet

corruption:	gauss	shot	impul	defcs	gls	mtn	zm	snw	frst	fg	brt	cnt	els	px	jpg	Mean
Source	98.4	97.7	98.4	90.6	92.5	89.8	81.8	89.5	85.0	86.3	51.1	97.2	85.3	76.9	71.7	86.2
SHOT (Offline)	73.8	70.5	72.2	79.2	80.6	58.5	54.0	53.6	63.0	47.3	39.2	97.7	48.7	46.1	53.0	62.5
TTT	96.9	95.5	96.5	89.9	93.2	86.5	81.5	82.9	82.1	80.0	53.0	85.6	79.1	77.2	74.7	83.6
DUA	89.4	87.6	88.1	88.0	88.6	84.7	74.3	77.8	78.4	68.6	45.6	95.9	72.2	66.5	67.4	78.2
NORM	87.1	90.6	89.5	87.6	93.4	80.0	71.9	70.6	81.5	65.9	46.8	89.8	73.5	63.2	67.5	77.3
T3A	85.5	84.0	85.0	86.6	85.9	76.1	65.4	70.3	71.0	58.7	41.3	86.8	60.5	54.4	61.0	71.5
SHOT (Online)	82.0	79.3	81.8	85.3	83.7	72.9	63.3	66.8	70.9	55.9	46.1	81.6	57.4	54.3	58.6	69.3
P-L	82.0	79.7	81.5	84.2	83.0	71.0	60.7	65.4	68.6	52.9	41.7	82.6	55.5	51.1	55.7	67.7
CFA	78.2	76.4	78.2	81.9	80.4	69.6	60.1	63.4	67.6	52.0	41.5	79.5	54.3	50.2	55.1	65.9
TENT	80.8	78.6	80.4	82.5	82.5	72.1	60.5	63.7	66.7	52.1	39.2	84.2	55.5	50.8	58.2	67.2
EATA	75.0	73.3	75.2	77.6	77.0	65.8	57.8	61.0	65.1	50.7	41.5	**70.8**	52.7	49.4	54.1	63.1
ActMAD	76.3	77.4	77.4	76.1	75.4	72.0	62.8	66.6	65.8	55.8	40.9	78.8	55.7	51.4	57.6	66.0 ± 0.1
ActMAD+TENT	74.1	71.6	74.2	77.9	75.9	63.9	55.0	58.3	64.4	47.7	39.1	81.3	49.2	46.0	50.9	62.0 ± 0.5
ActMAD+EATA	**70.7**	**69.2**	**72.3**	**75.9**	**74.4**	**60.8**	**53.4**	**55.5**	**60.2**	**46.4**	**38.5**	78.4	**48.0**	**45.2**	**49.5**	**59.9** ± 0.8