ICCV23
PARIS

Team **YSVnL**

TTA-DAME: **T**est-**T**ime **A**daptation with **D**omain **A**ugmentation and **M**odel **E**nsemble for Dynamic Driving Conditions


Dongjae Jeon


Taeheon Kim


Seongwon Cho


Jonghyun Choi

Yonsei University

Different Setups of Domain Adaptation

ICCV23
PARIS

	Domain Adaptation	Test-time Adaptation	Continual Test-time Adaptation
Accessibility to Source Domain Data	O	X	X
Adaptation Time	Training	Inference	Inference
Continual shifting target domains	X (e.g Night)	X (e.g Night)	O (e.g Day => Night)

Continual Test-time Adaptation

Goal

- Adapt the pretrained model continuously to shifting domains.

SHIFT Dataset [1]

- Discrete, Continuous annotated images with various domains.

Train dataset example

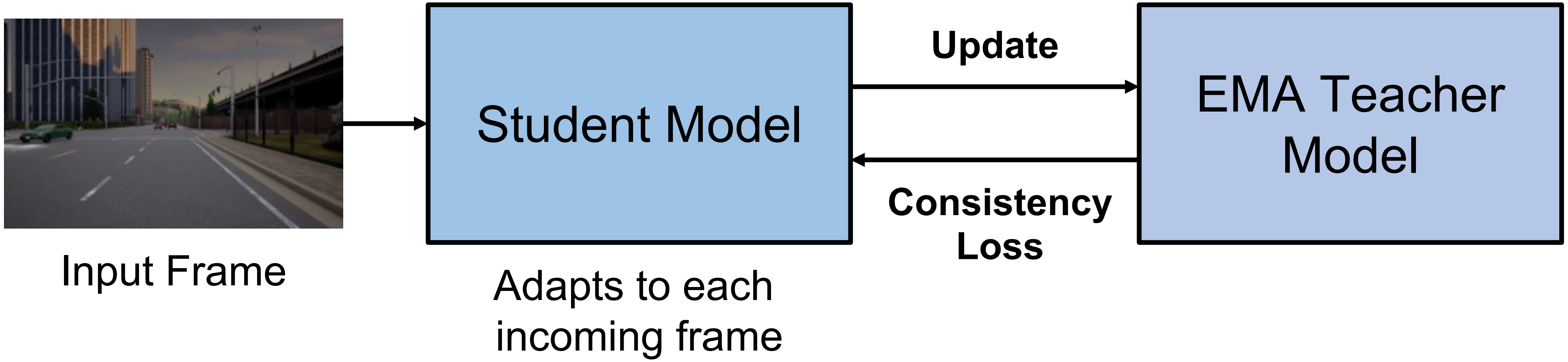
[1] Sun, Tao, et al. "SHIFT: a synthetic driving dataset for continuous multi-task domain adaptation", *CVPR*, 2022.

How To Use Unlabeled Data for Training?

Generating pseudo labels could change into supervised task.

Mean-teacher [1]

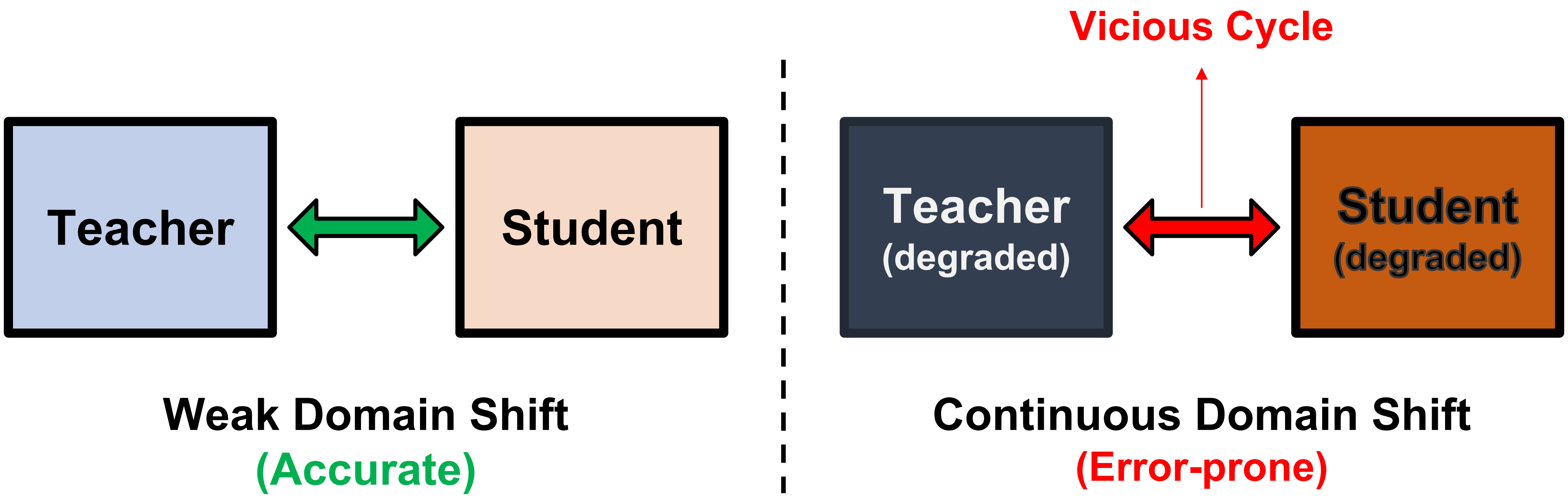
: EMA teacher (ensemble of prev. models) generate pseudo labels.



[1] Tarvainen, Antti, and Harri Valpola. "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results." Advances in neural information processing systems 30 (2017).

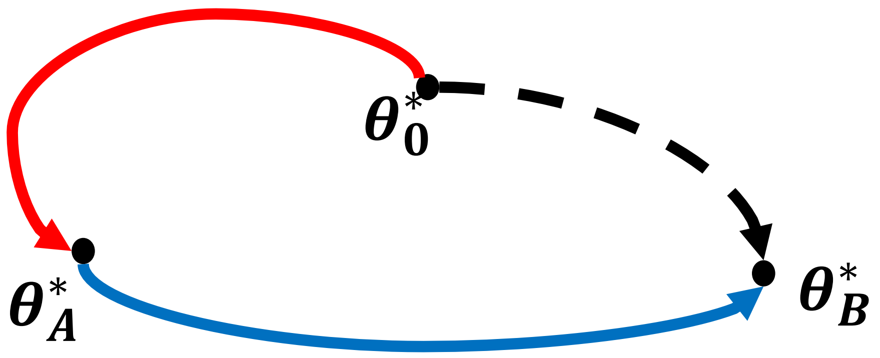
Domain Shift Error Accumulation

: Mean-Teacher generates poor labels on continuously shifting domains.
Vicious cycle break down entire training.



Over Adaptation Hinders Adaptability

: Strong adaptation to a specific domain restricts adaptability[1]
Prefer generalizable weights to domain specific weights.



Adaptation from initial parameter to domain A.

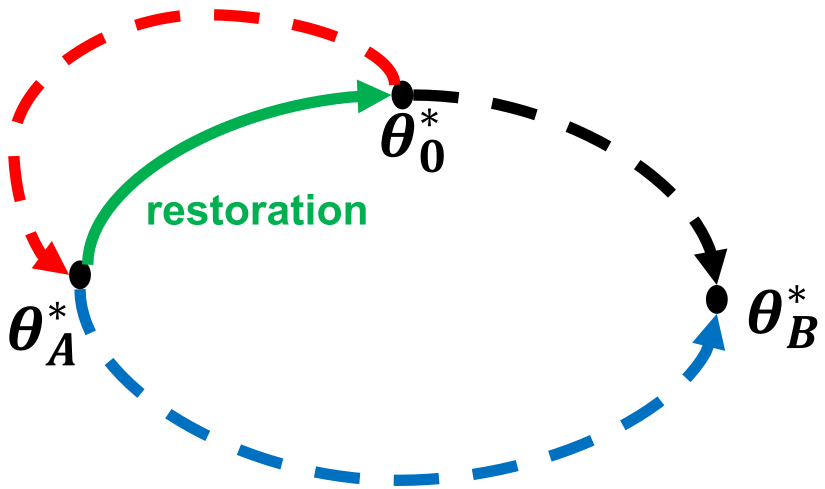
Adaptation from domain A to domain B.

Direct adaptation from initial parameter to domain B. **(Preferred)**

[1] Wang, Qin, et al. "Continual test-time domain adaptation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022.

Stochastic Restoration[1] To Mitigate Issues

: Randomly resets the weights from student model to initial(pretrained) model

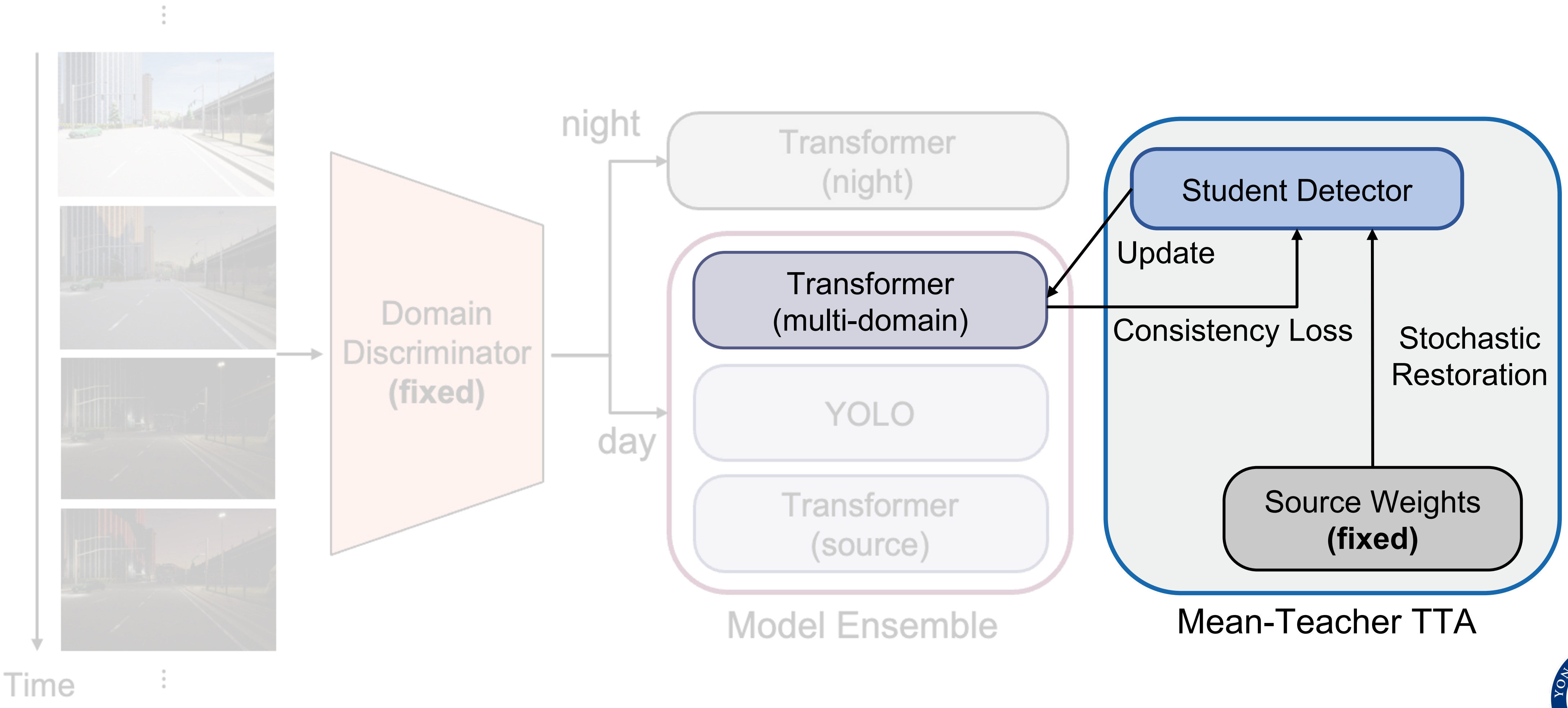


- Adaptation from initial parameter to domain A.
- Adaptation from domain A to domain B.
- Direct adaptation from initial parameter to domain B.
- **Stochastic restoration to initial parameter θ_0^***

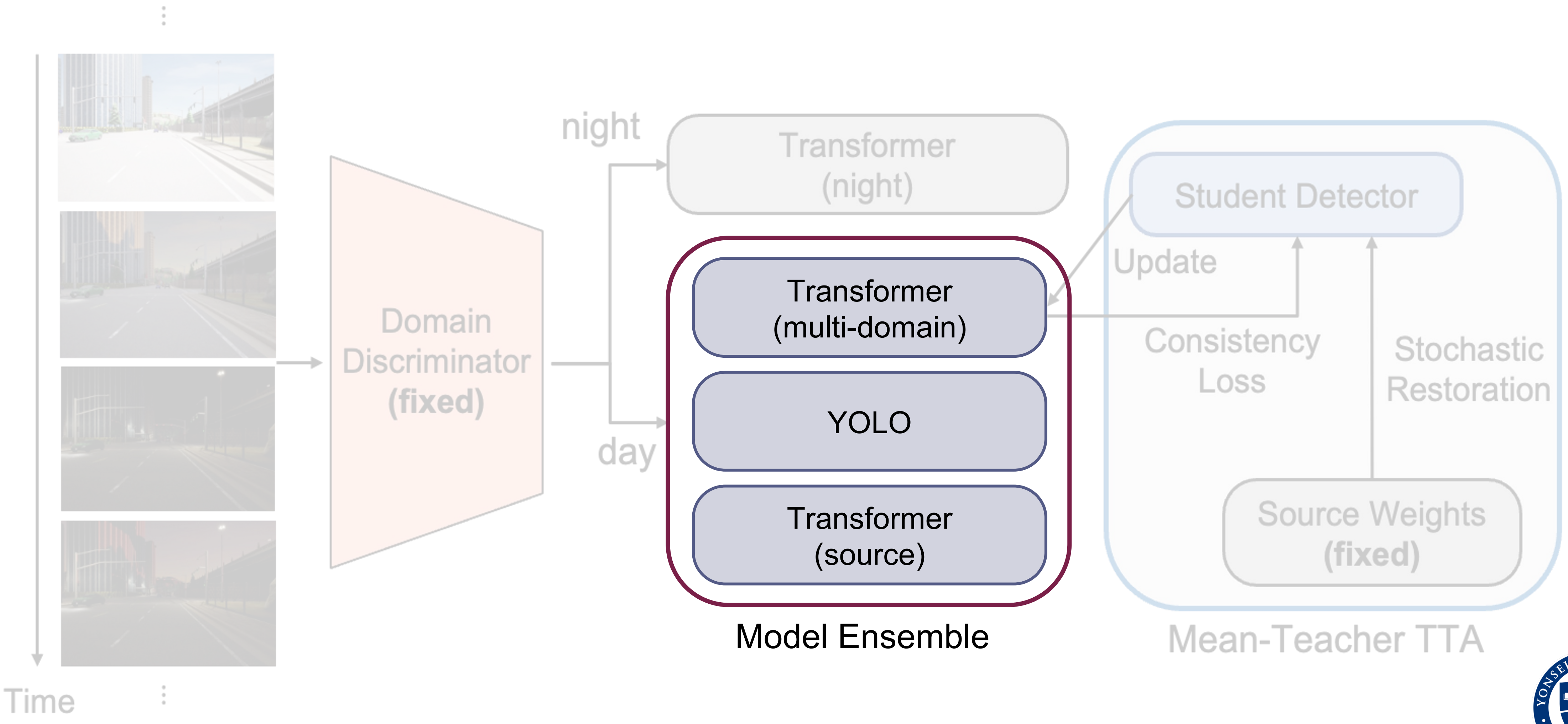
Method	AP(%)	AR(%)
Baseline (DINO[12] with MT)	47.1	57.7
+ Stochastic Restoration	47.8	58.4

[1] Wang, Qin, et al. "Continual test-time domain adaptation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022.

Mean-Teacher With Stochastic Restoration Is Adapted As A Main Model

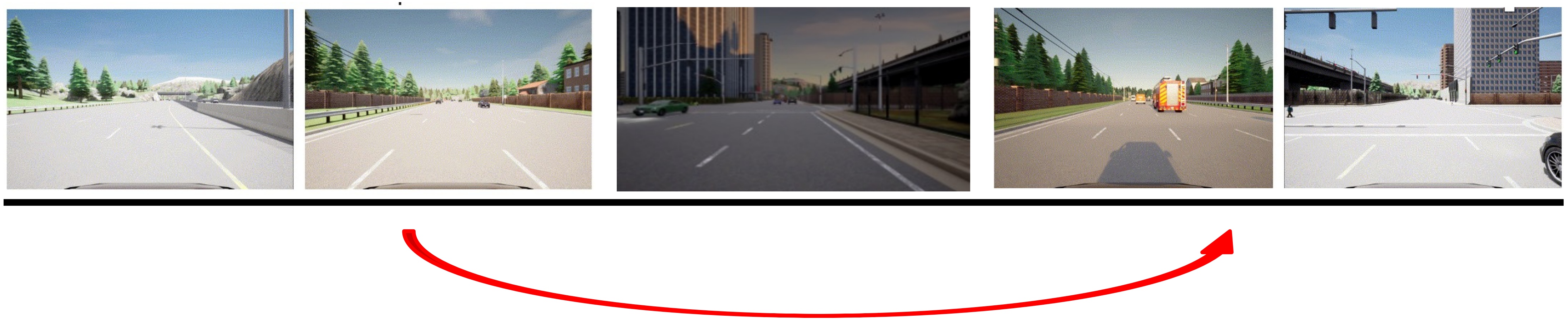


Two Distinguished Models for Additional Support



Retaining Source Knowledge Is Important

Validation data example



Time

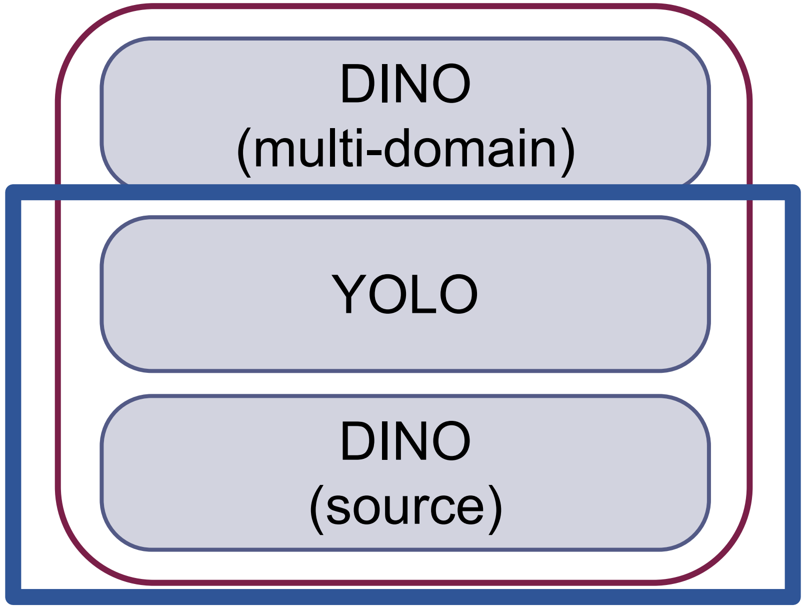
Returns to Source Domain

- Need to keep source knowledge along the sequences.
- Stochastic Restoration does not alleviate catastrophic forgetting

Source knowledge is regained via Ensemble

Two fixed models support the main model

- DINO [1] (multi-domain) (Adapt)
- **YOLOv8 (source) (Fixed)**
- **DINO [1] (source) (Fixed)**



Gradual Contribution Equalization (GCE)

- Gradually weights to DINO(source) to prevent catastrophic forgetting of source knowledge.
- Equalize contribution of all models at last.

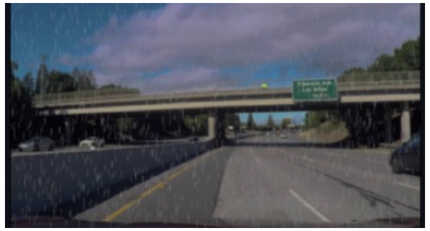
[1] Zhang, Hao, et al. "Dino: Detr with improved denoising anchor boxes for end-to-end object detection." arXiv preprint arXiv:2203.03605 (2022).

Model Suffers The Most On Night Domain

- We have evaluated possible target domains using data augmentation.
- Among them, model suffers in “Night” most.

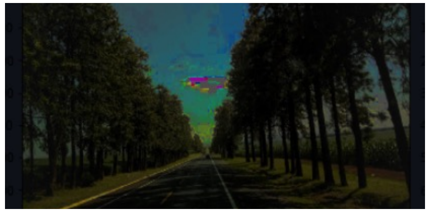
Transformer
(multi-domain)

…



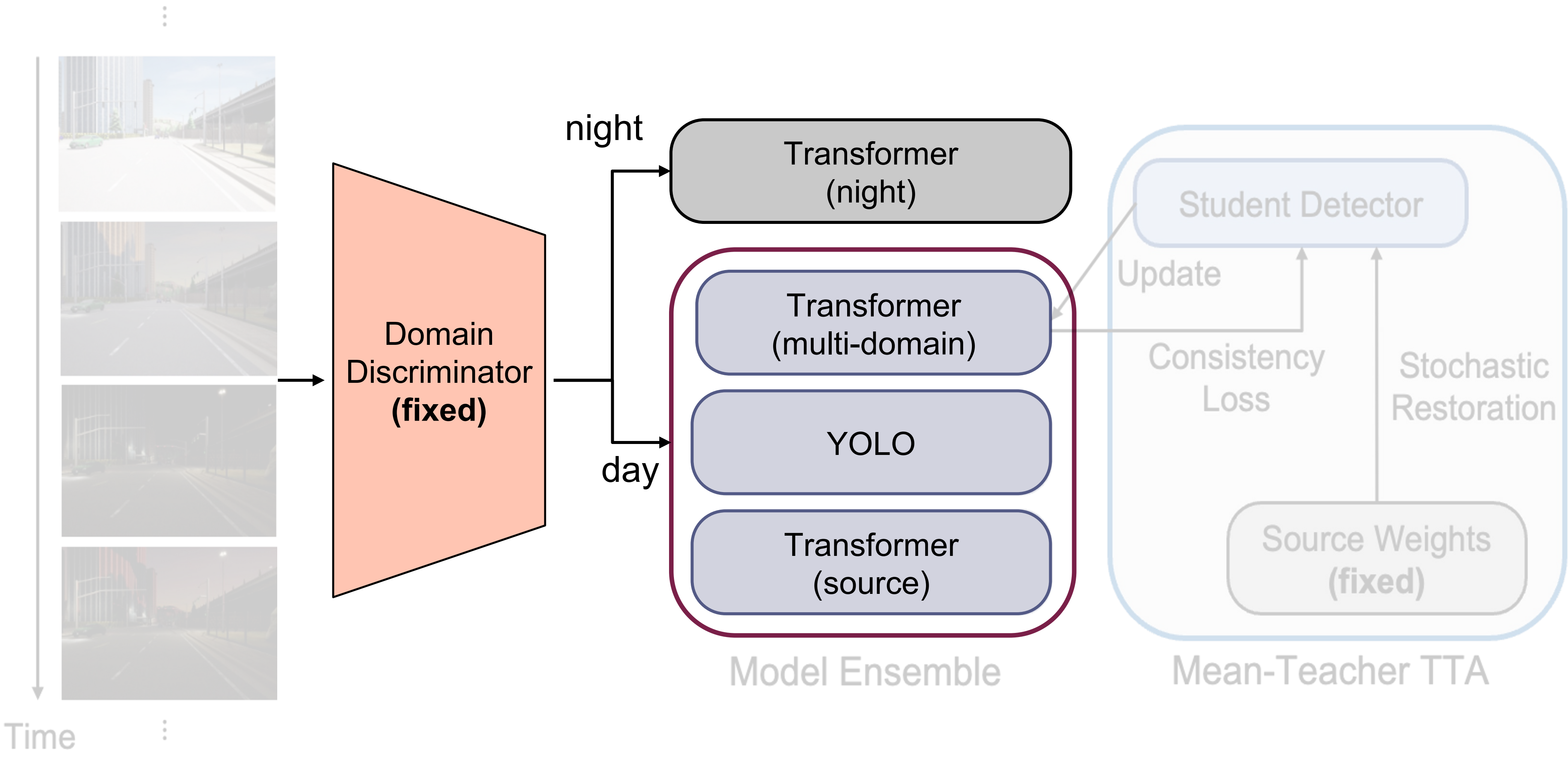
Rainy (AP: 45.3)

Snowy (AP: 42.1)



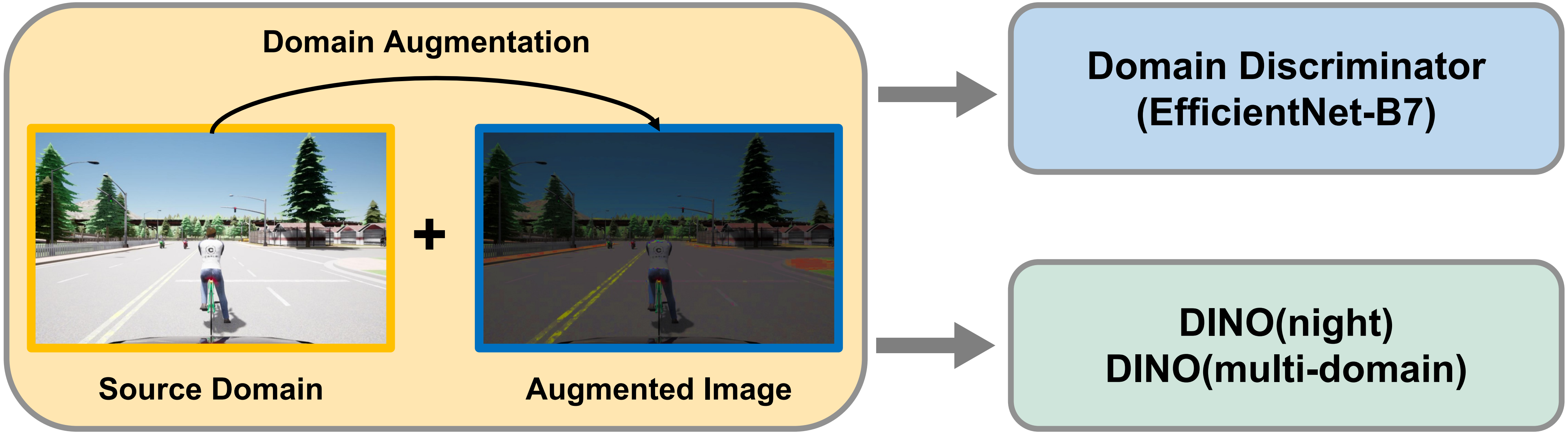
Night **(AP: 36.4)**

Night Images Are Separated By Domain Discriminator



Domain Augmentation for Specialization

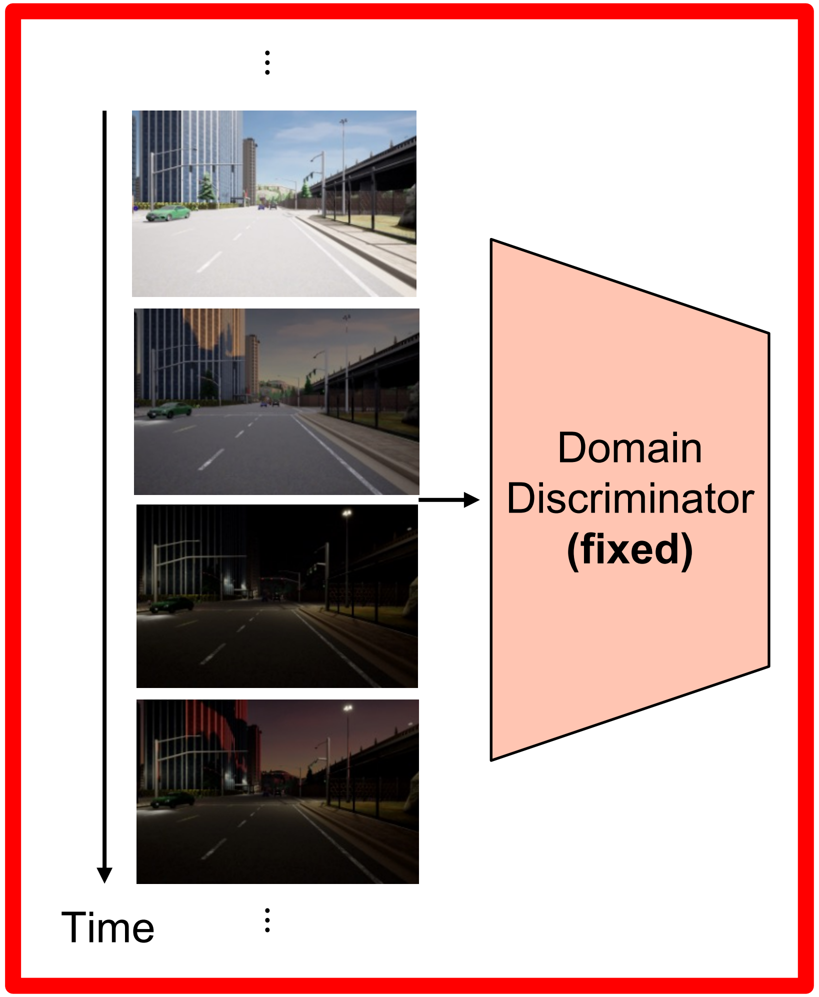
- Utilized ‘automold’ library to transform images into diverse time & weather conditions.
- Used in **DINO(multi-domain), DINO(night), Domain Discriminator** training.



[1] https://github.com/UjjwalSaxena/Automold--Road-Augmentation-Library

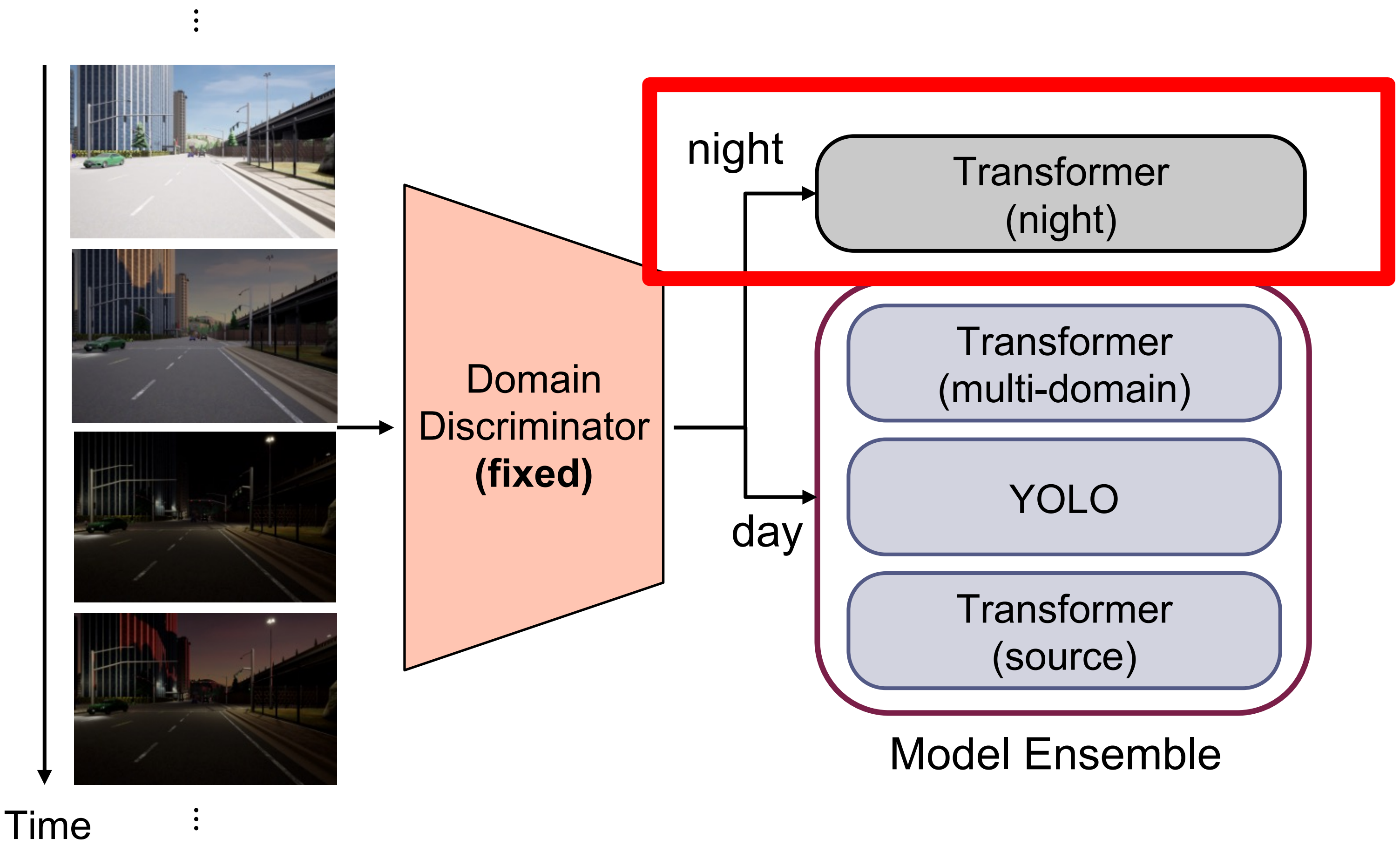
Summary

Classify the input image as day / night using the trained domain discriminator.



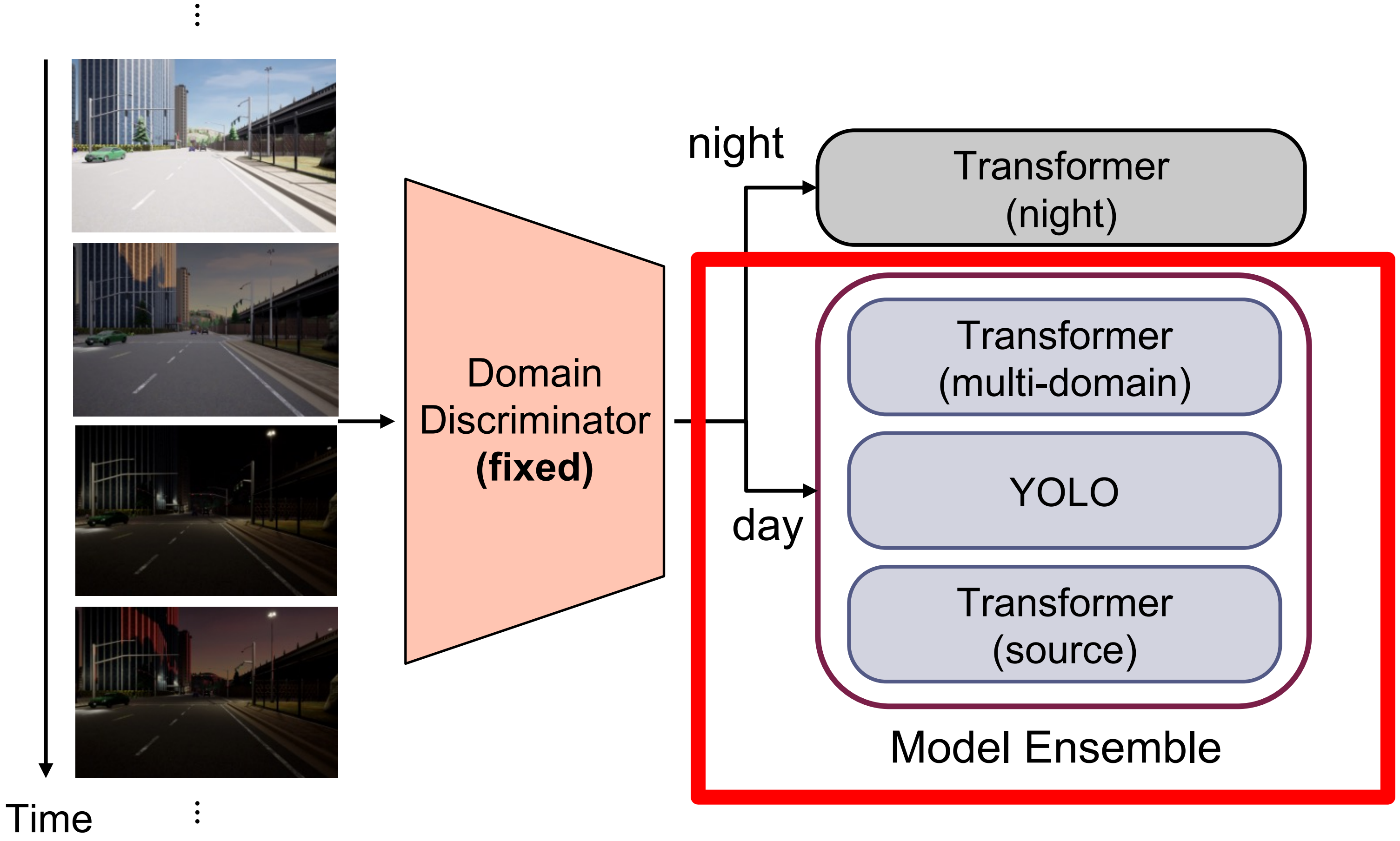
Summary

If the domain is '**night**,' the trained nighttime detector is employed for predictions.



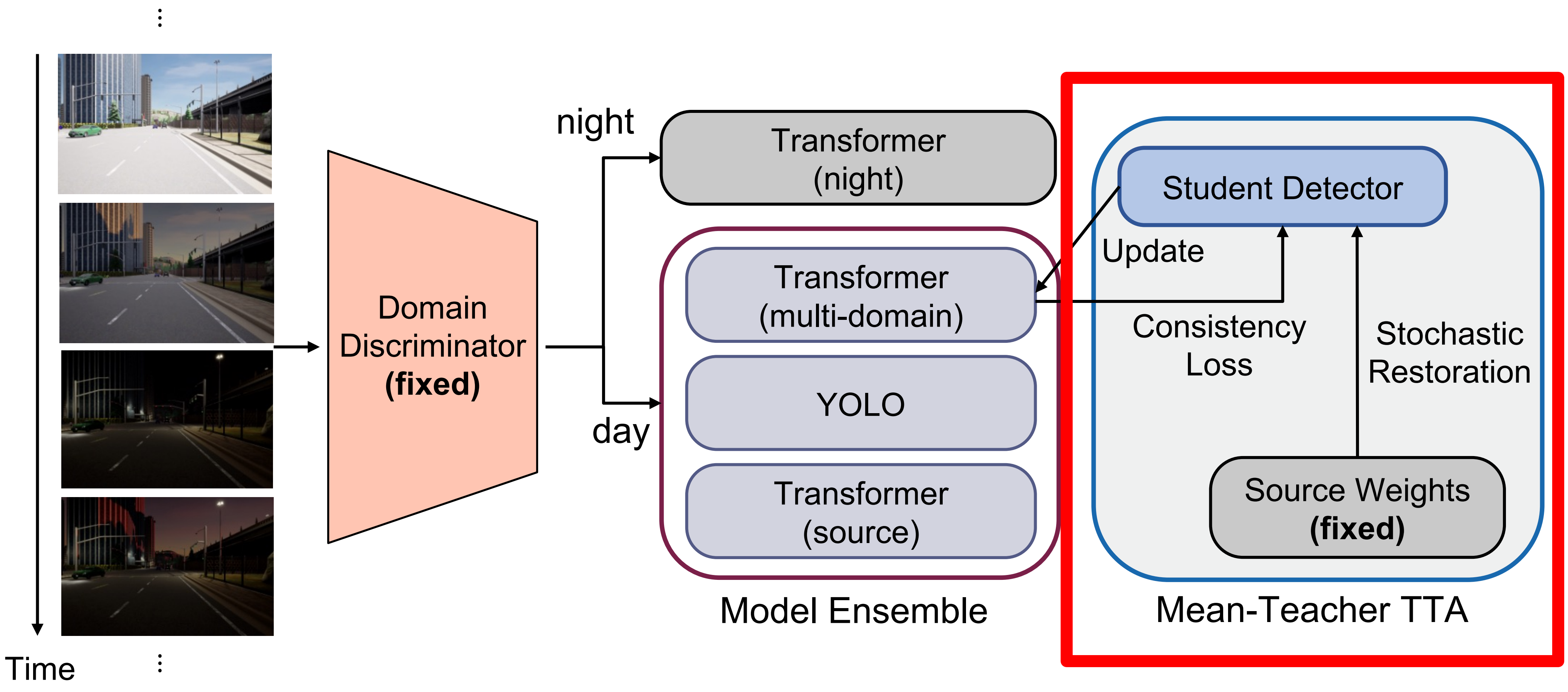
Summary

If not '**night**,' the Mean-teacher model combines with multiple detectors for predictions.



Summary

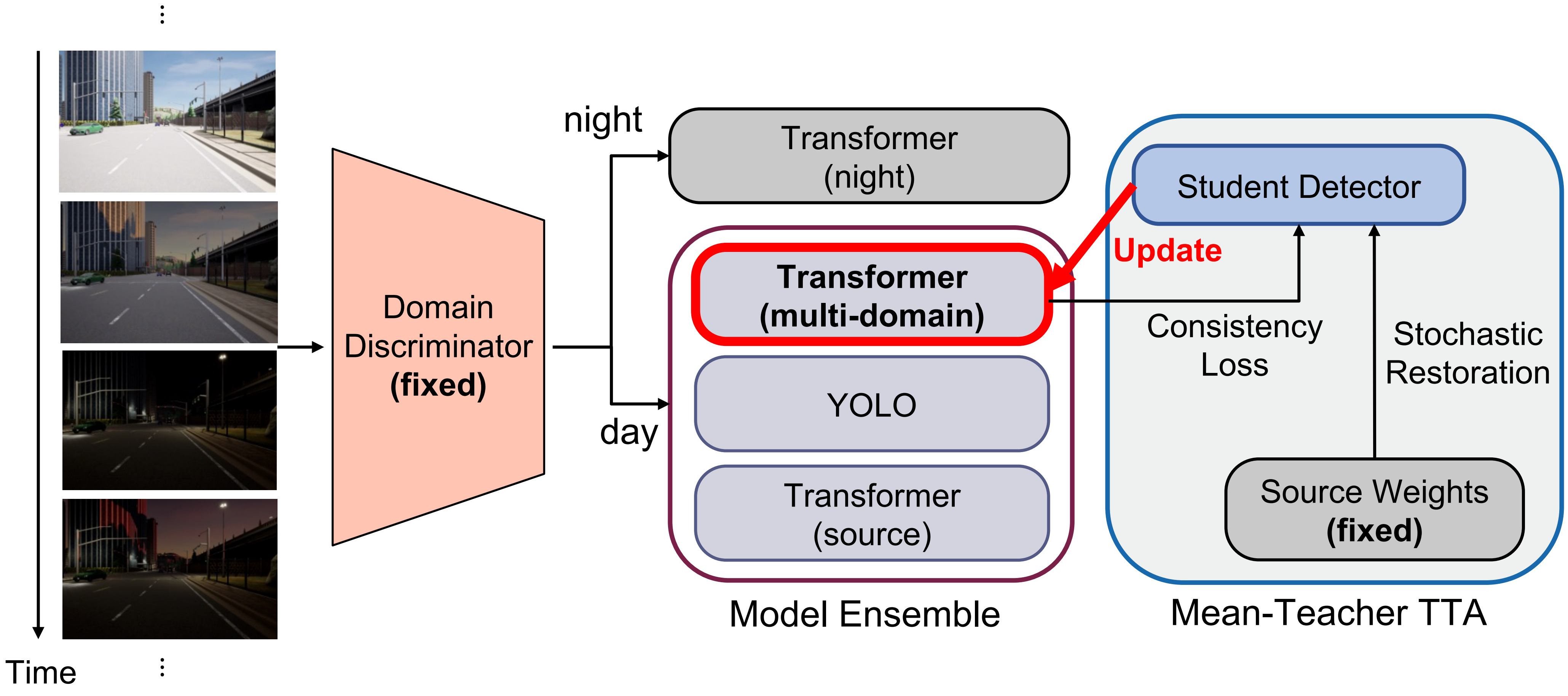
Adapts the student model with pseudo labels generated from Mean-Teacher



Summary

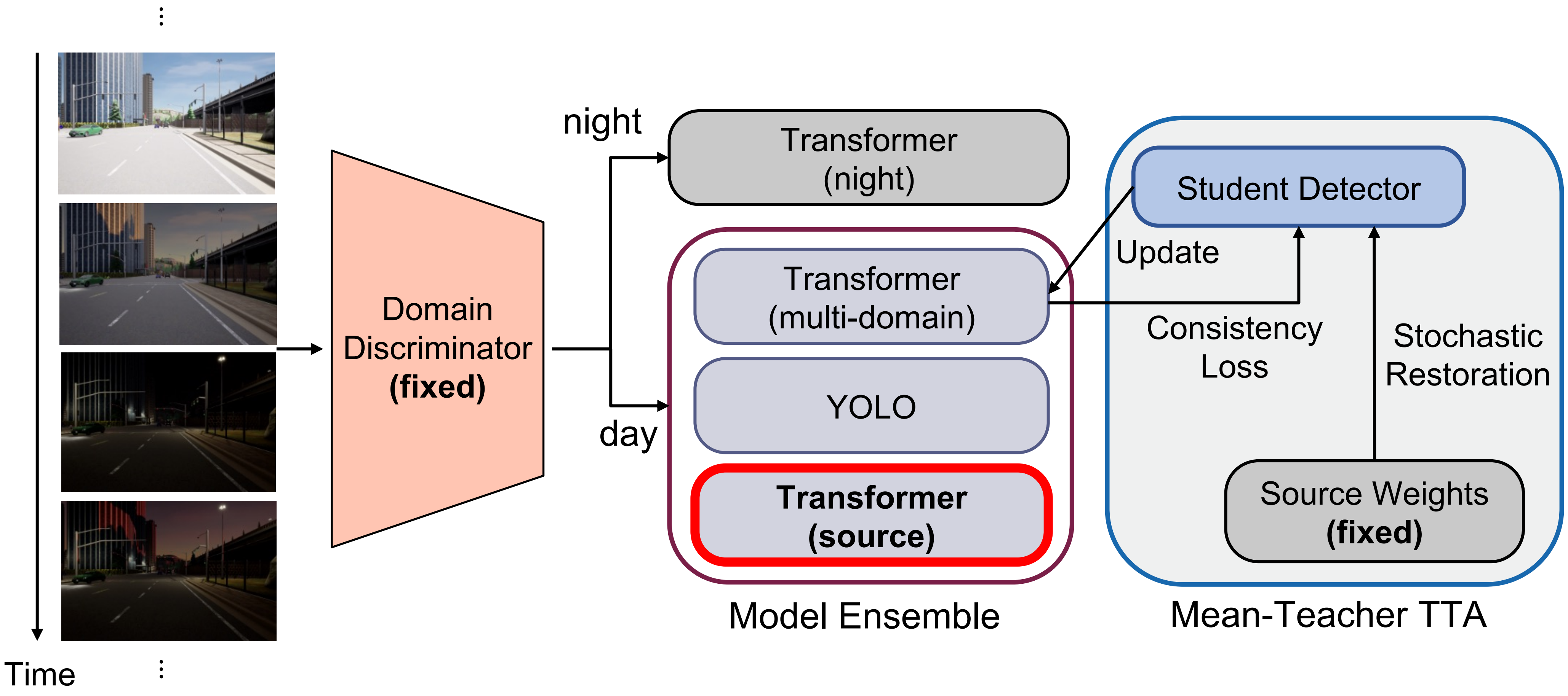
ICCV23 PARIS

The Mean-Teacher model adapts through EMA



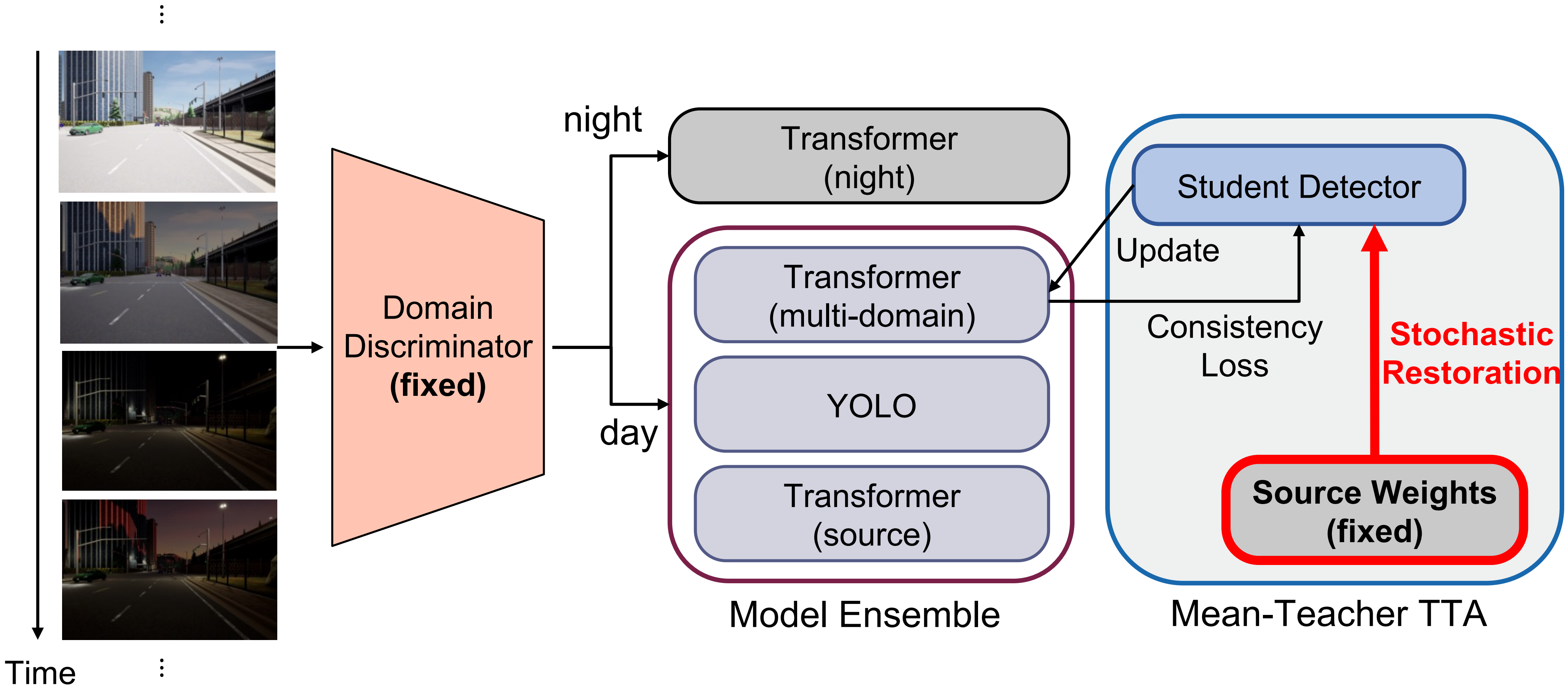
Summary

Source model plays a role in preventing catastrophic forgetting



Summary

Periodically reset the student's weights to prevent error accumulation and over-adapting



Quantitative Results

Method	AP(%)	AR(%)
Baseline (DINO[12] with MT)	47.1	57.7
+ Stochastic Restoration	47.8	58.4
+ Domain Augmentation	48.1	58.9
+ Domain Discriminator	48.5	59.1
+ Model Ensemble (TTA-DAME)	**49.4**	**62.2**

2.3%p increase in Average Precision
4.5%p increase in Average Recall

Qualitative Results

ICCV23 PARIS

Baseline (DINO MT)

TTA-DAME (Ours)

pedestrian car truck bus motorcycle bicycle

ICCV23
PARIS

Q&A

YONSEI UNIVERSITY
1885
연세대학교

Summary (backup)

Integrating Mean Teacher Method & Stochastic Restoration