JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Data-Free Sketch-Based Image Retrieval
(WED-AM-368)


Abhra Chaudhuri


Ayan Kumar Bhunia


Yi-Zhe Song


Anjan Dutta

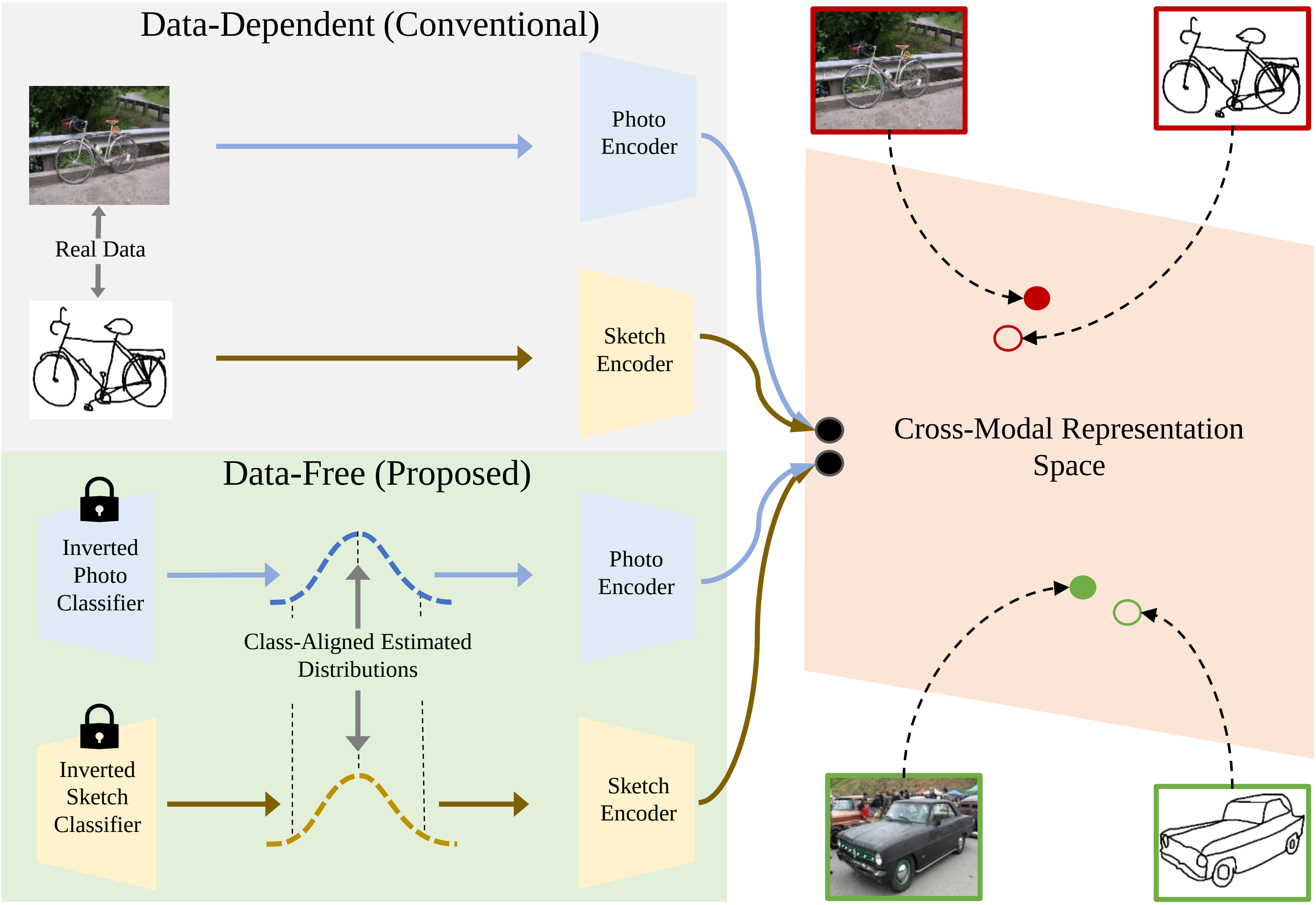
University of Exeter

UNIVERSITY OF SURREY

People-Centred AI
UNIVERSITY OF SURREY

Preview

- **Motivation:** Data-dependent SBIR requires *expensive photo-sketch pairs* for training.
- **Observation:** Pre-trained photo/sketch classifiers *implicitly encode* their train set distributions.
- **Action:** Controlled reconstruction of such distributions for *training cross-modal photo-sketch encoders*.
- **Results:** *SBIR* can be performed in a *Data-Free* manner with considerable accuracy.



Outline

- Preliminaries
- Problem Definition
- Methodology
- Experiments
- Conclusion

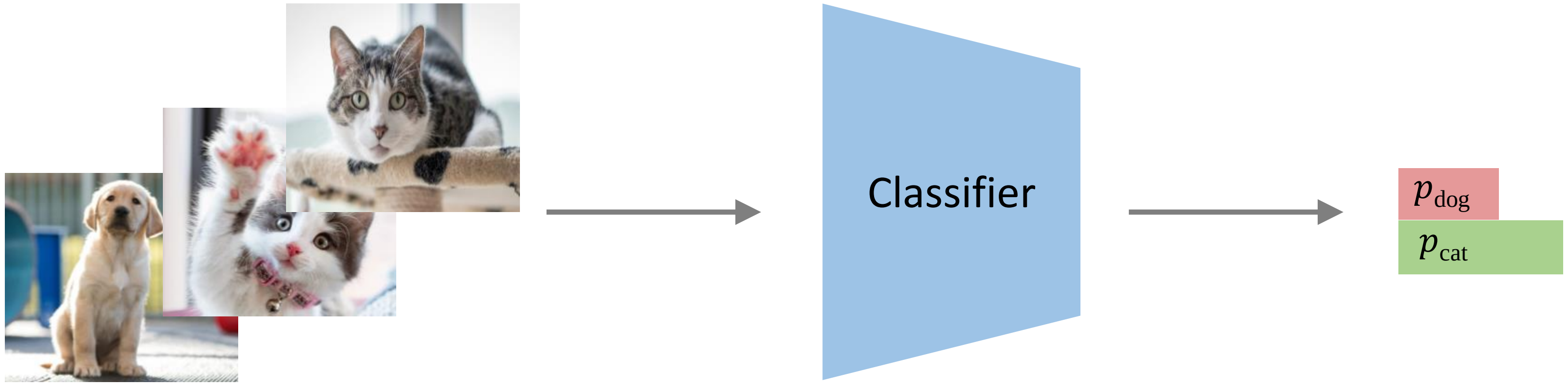
Outline

- Preliminaries
- Problem Definition
- Methodology
- Experiments
- Conclusion

Preliminaries

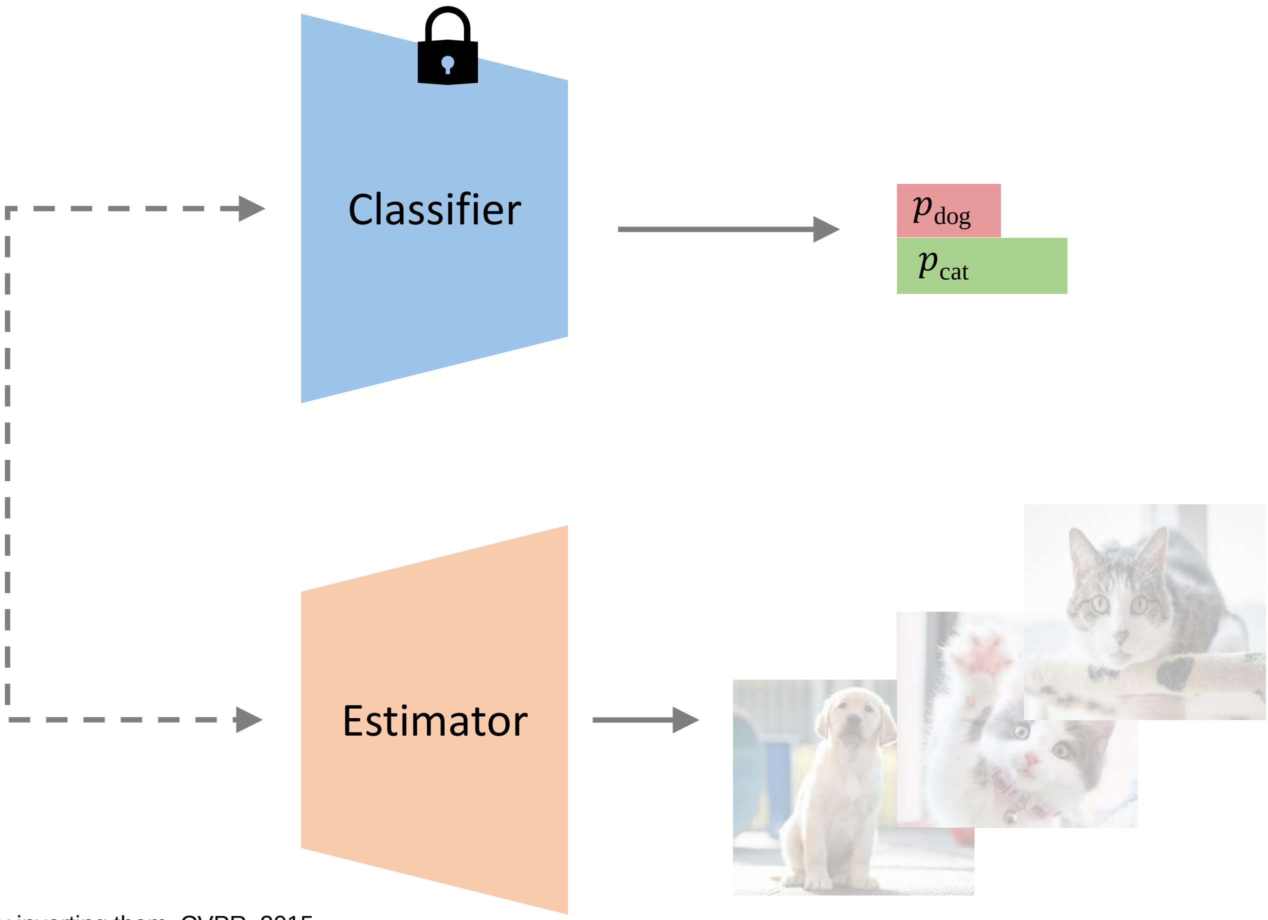
- Model Inversion
- Data-Free Learning
- Adversarial Data-Free Distillation

Model Inversion



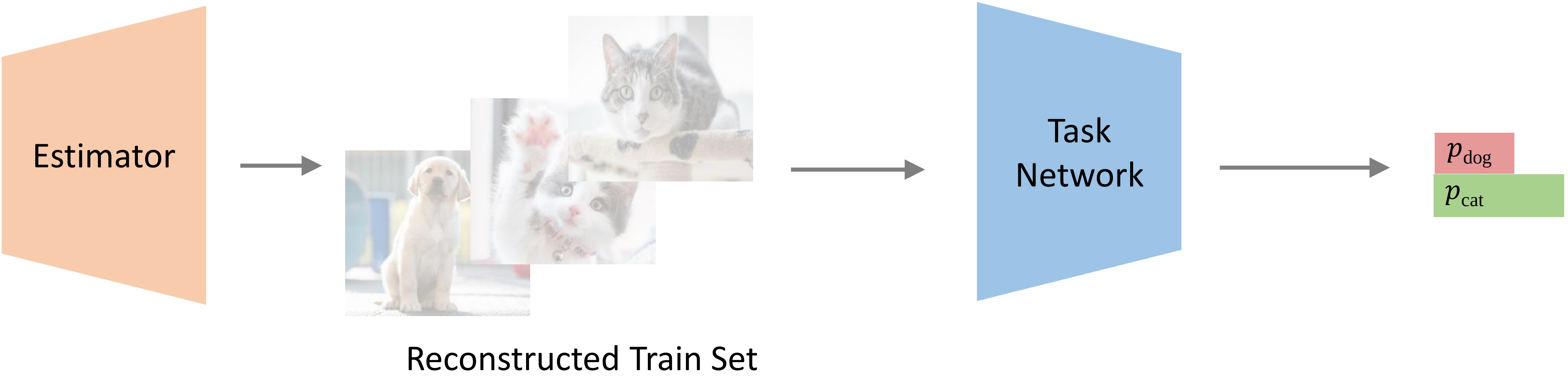
Mahendran & Vedaldi. Understanding deep image representations by inverting them, CVPR, 2015.

Model Inversion



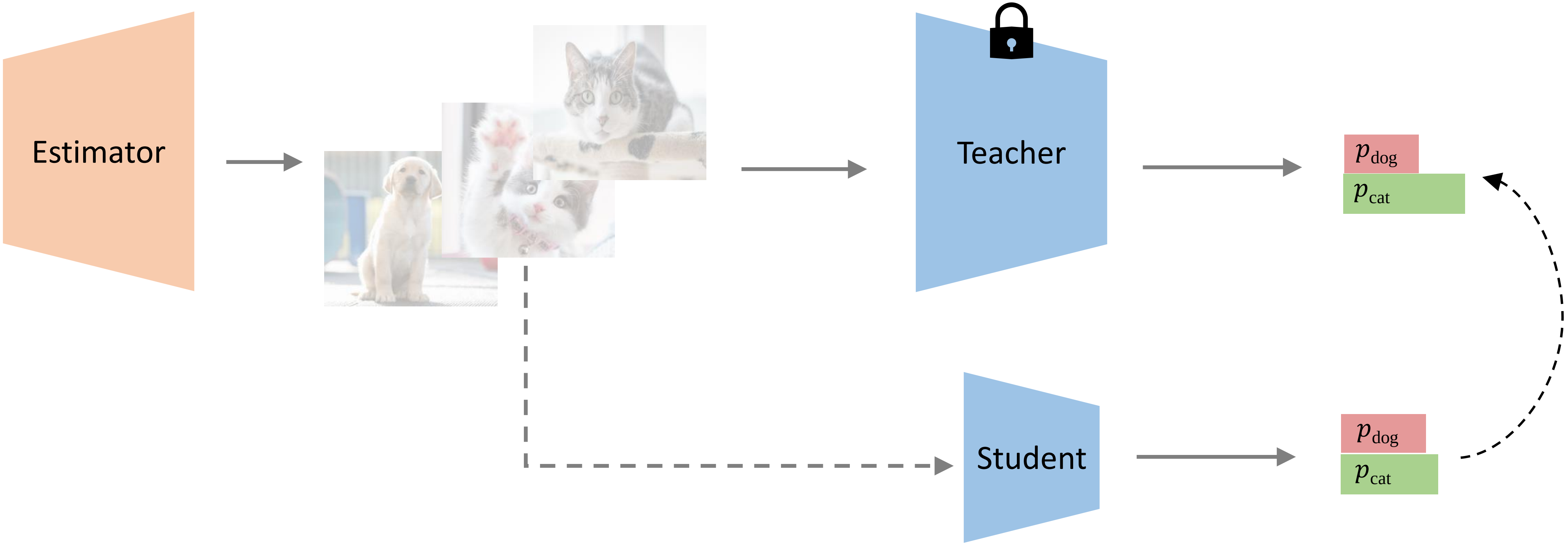
Mahendran & Vedaldi. Understanding deep image representations by inverting them, CVPR, 2015.

Data-Free Learning (DFL)



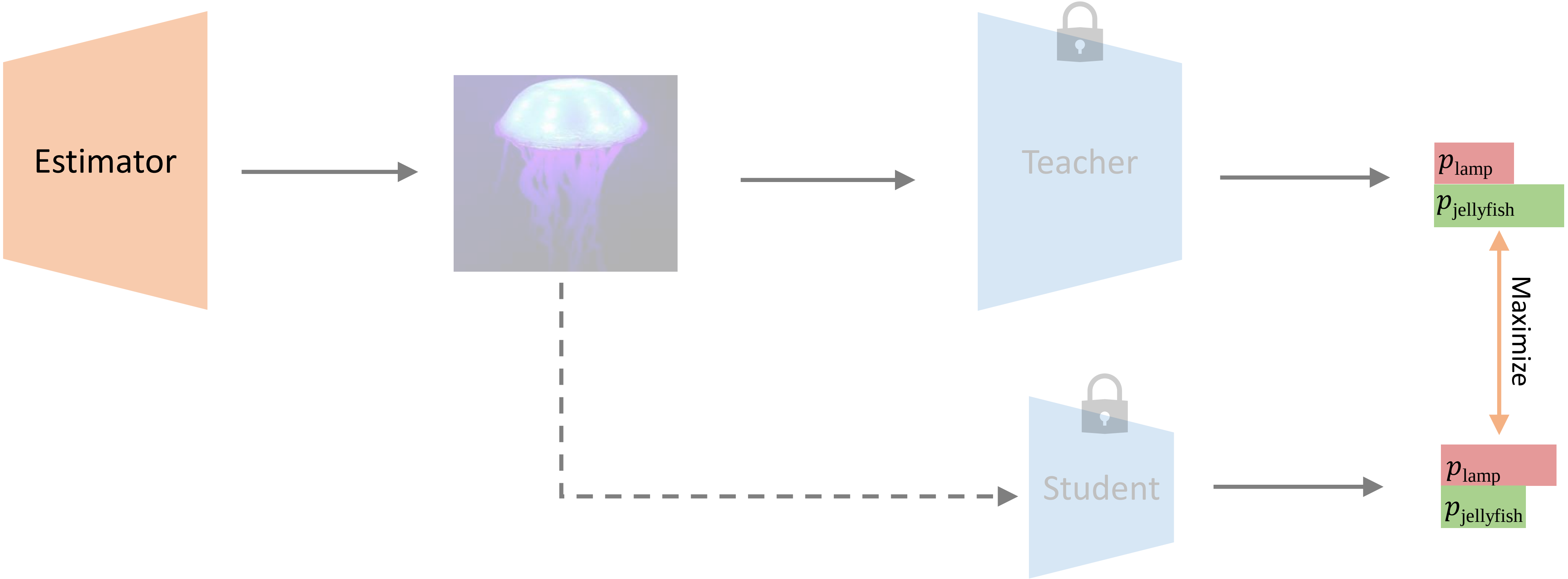
Chen *et al.* DAFL: Data-free learning of student networks, ICCV, 2019.

Data-Free Knowledge Distillation



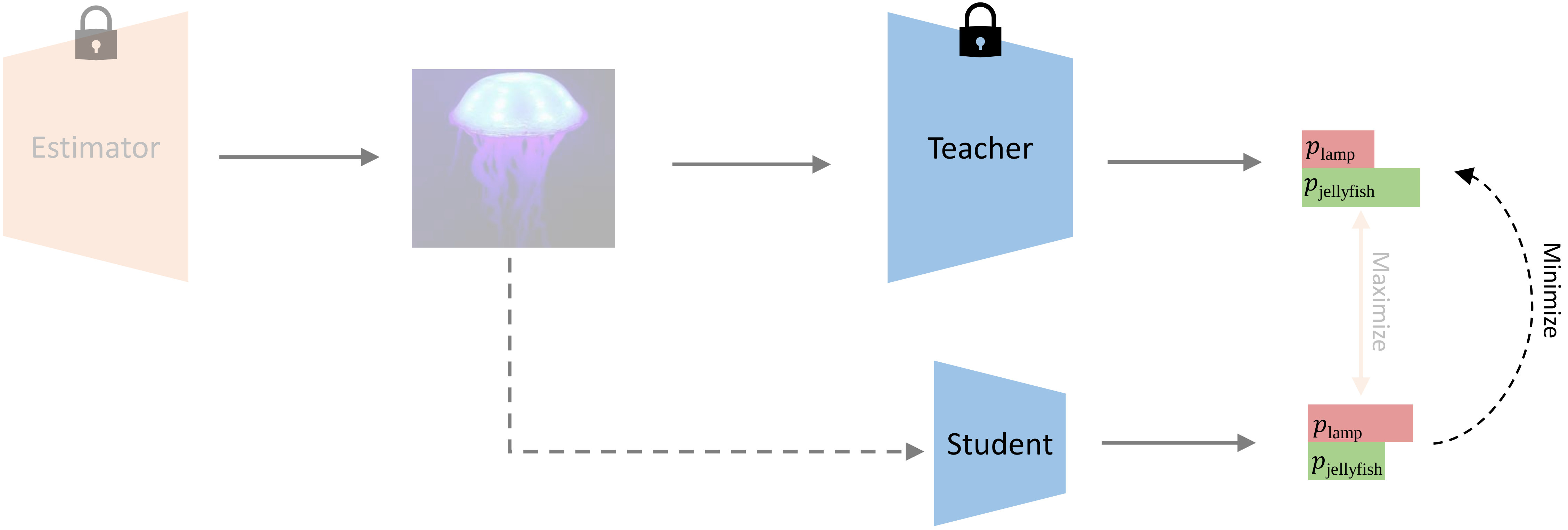
Chen *et al.* DAFL: Data-free learning of student networks, ICCV, 2019.

Adversarial Data-Free Distillation



Choi *et al*. Data-free network quantization with adversarial knowledge distillation, CVPRW, 2020.

Adversarial Data-Free Distillation

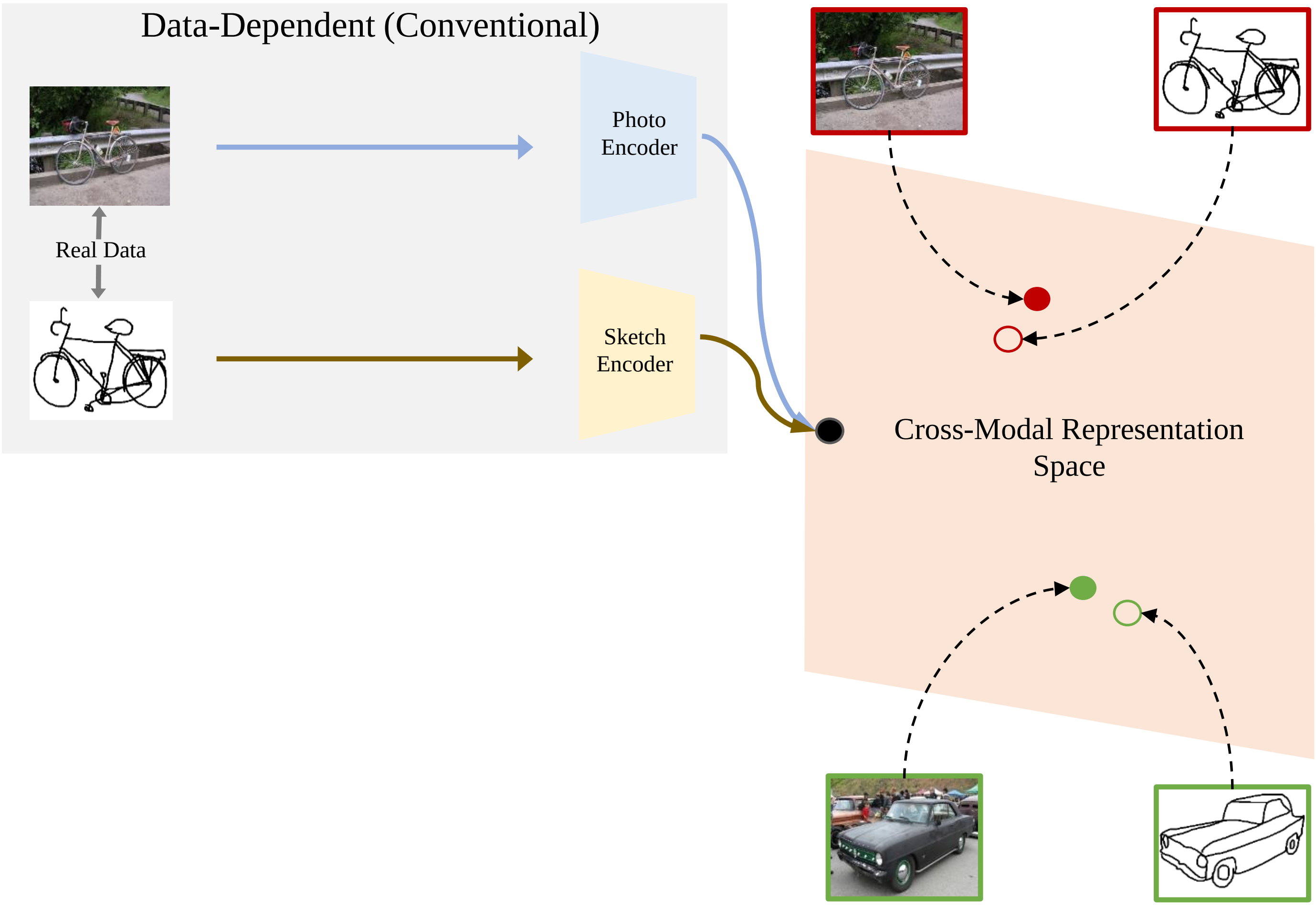


Choi *et al*. Data-free network quantization with adversarial knowledge distillation, CVPRW, 2020.

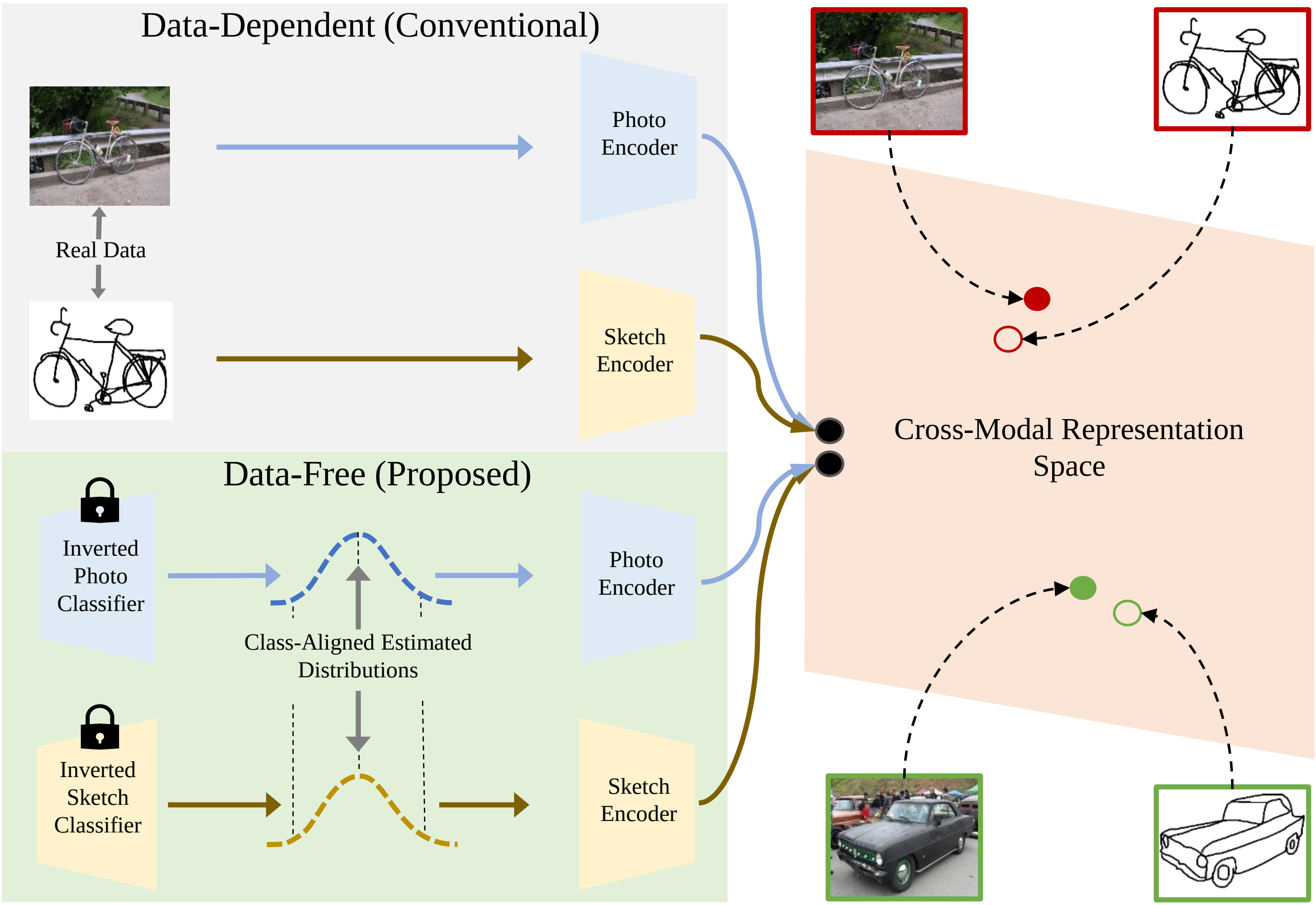
Outline

- Preliminaries
- Problem Definition
- Methodology
- Experiments
- Conclusion

Problem Definition – Data-Free SBIR



Problem Definition – Data-Free SBIR

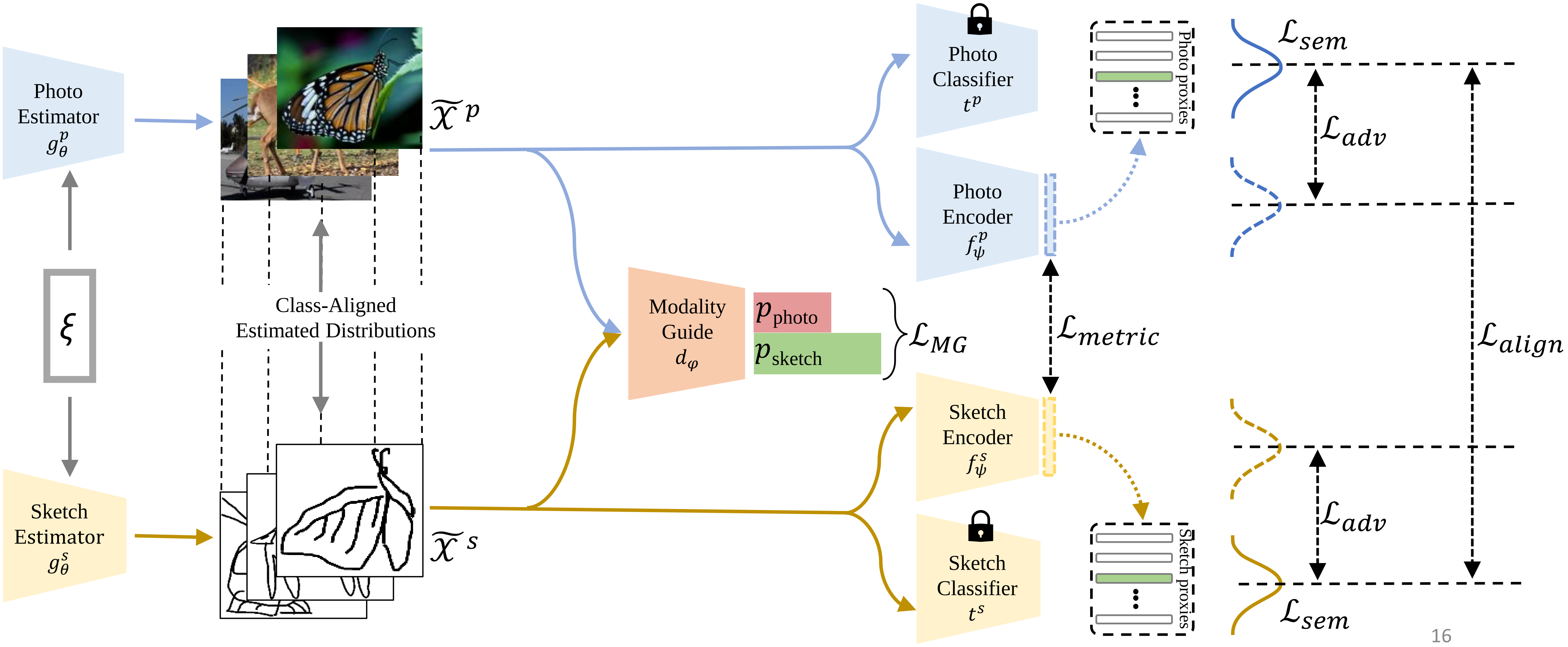


Outline

- Preliminaries
- Problem Definition
- Methodology
- Experiments
- Conclusion

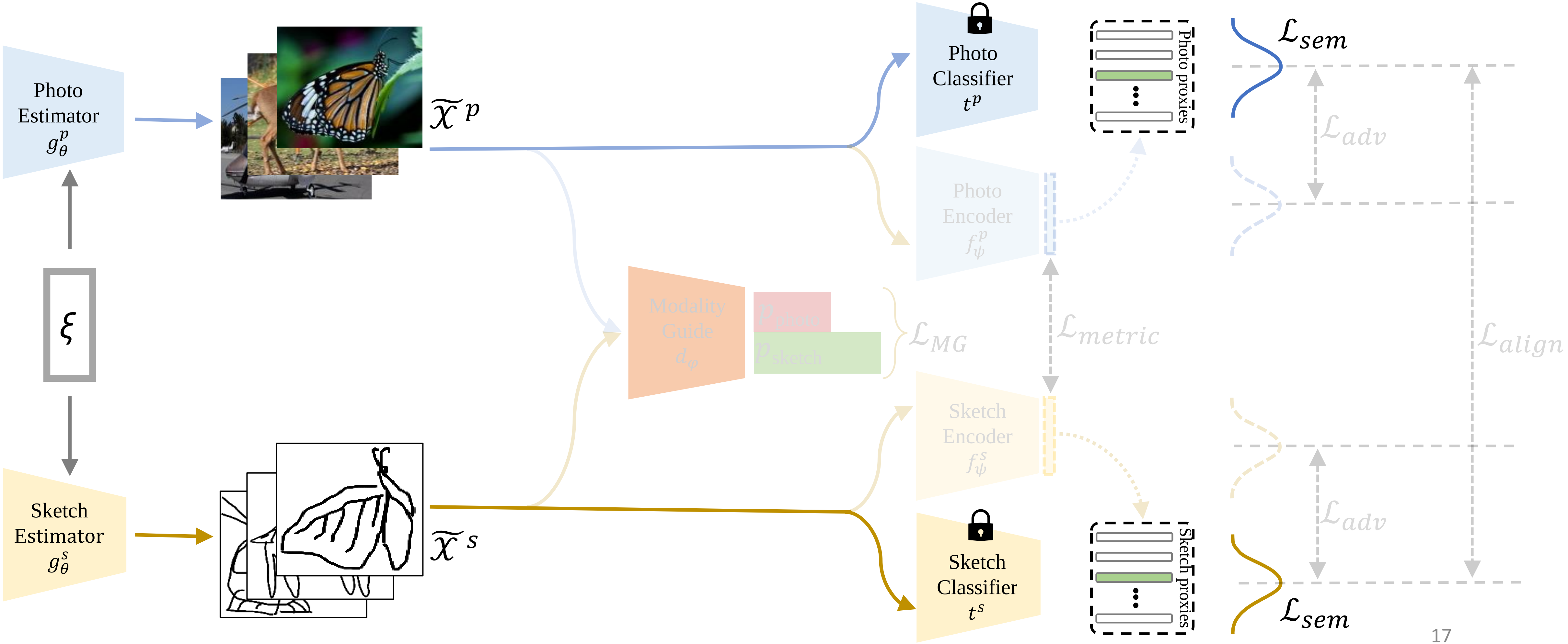
DFL Across Modalities and Metric Spaces (CrossX-DFL)

End-to-end architecture of CrossX-DFL for Data-Free SBIR.



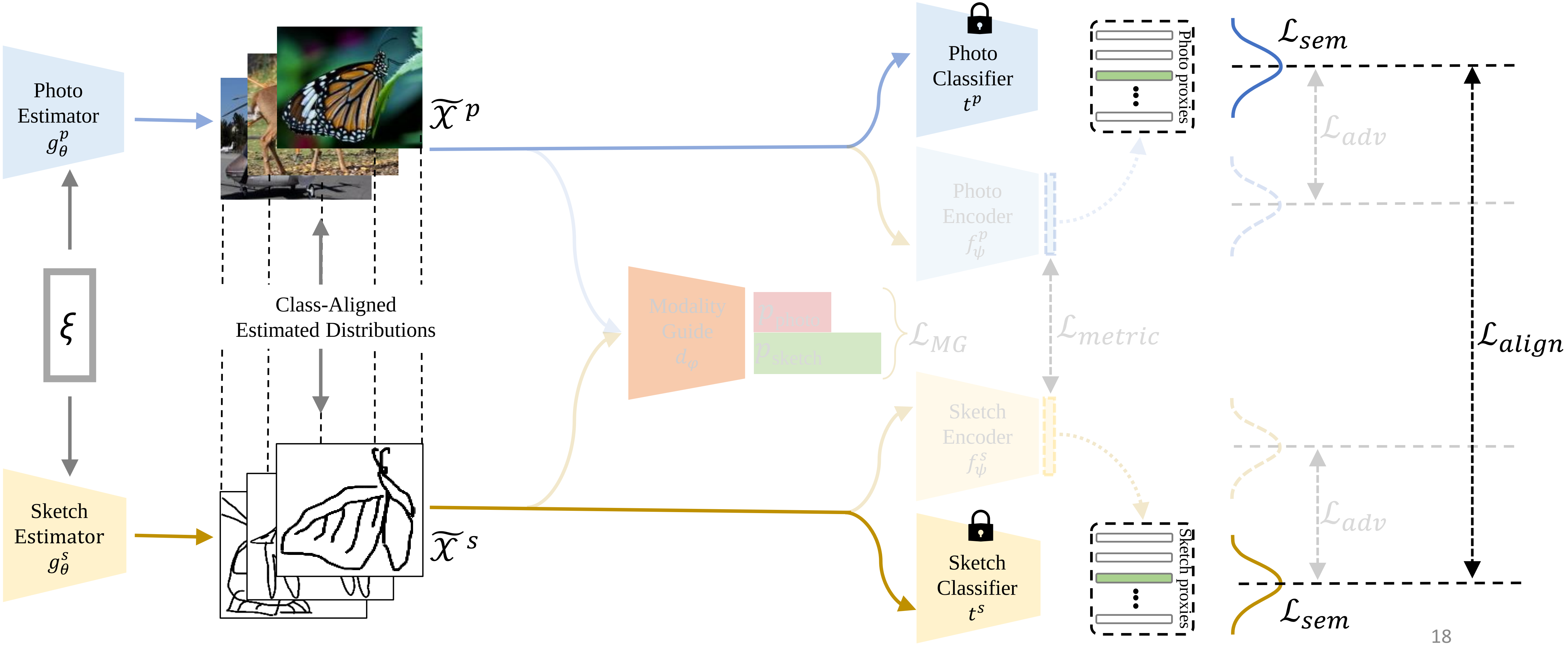
DFL Across Modalities and Metric Spaces (CrossX-DFL)

Semantic Consistency Loss: Ensures that the reconstructions belong to concrete, unambiguous classes.



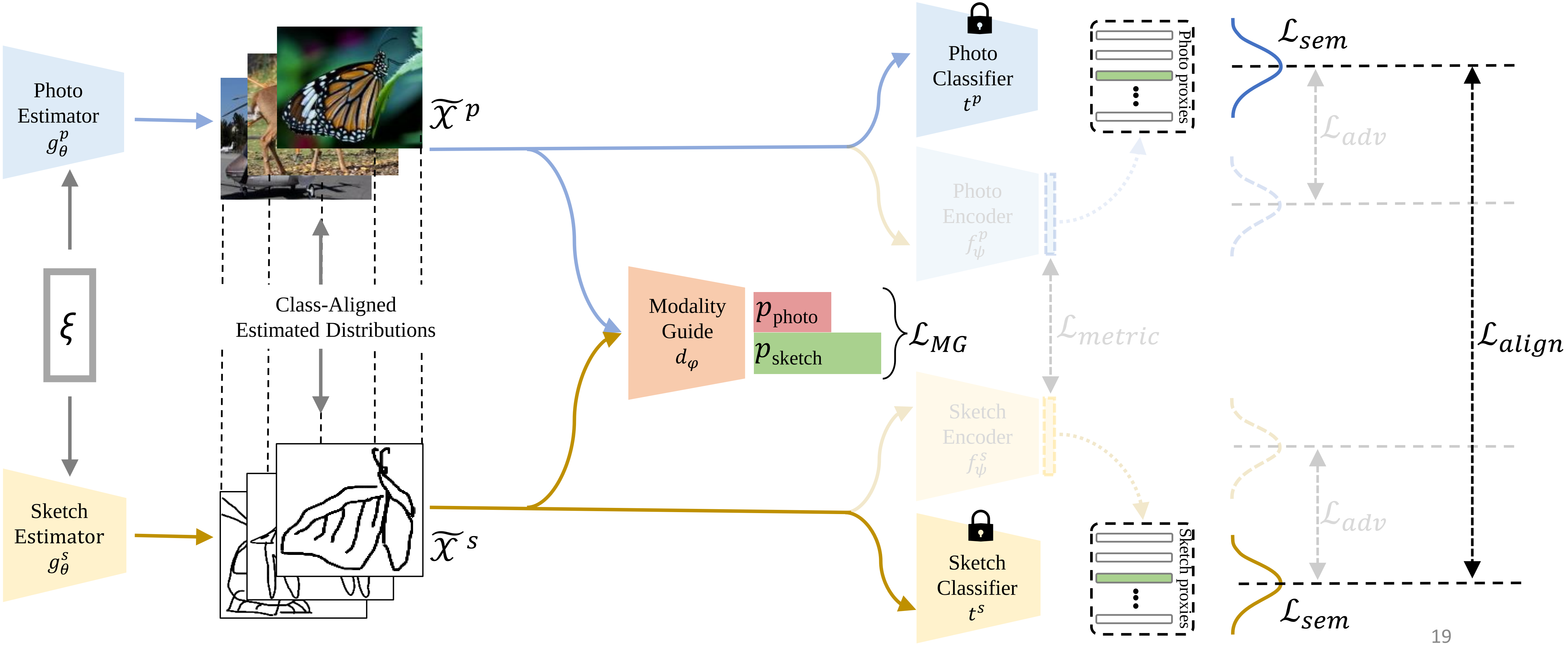
DFL Across Modalities and Metric Spaces (CrossX-DFL)

Class Alignment: Incentivises a common noise vector to induce similar label distributions across the two classifiers.



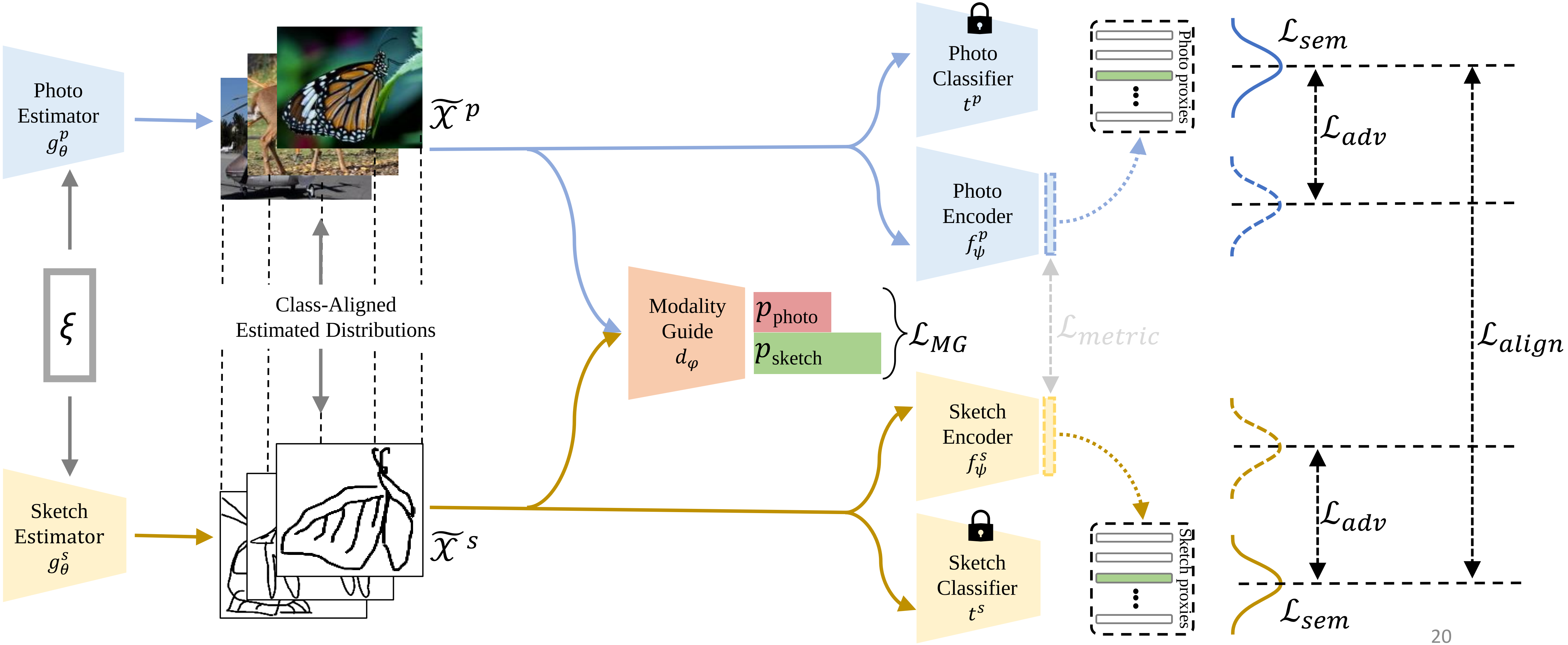
DFL Across Modalities and Metric Spaces (CrossX-DFL)

Modality Guidance: Restricts the estimators to produce modality-specific reconstructions.



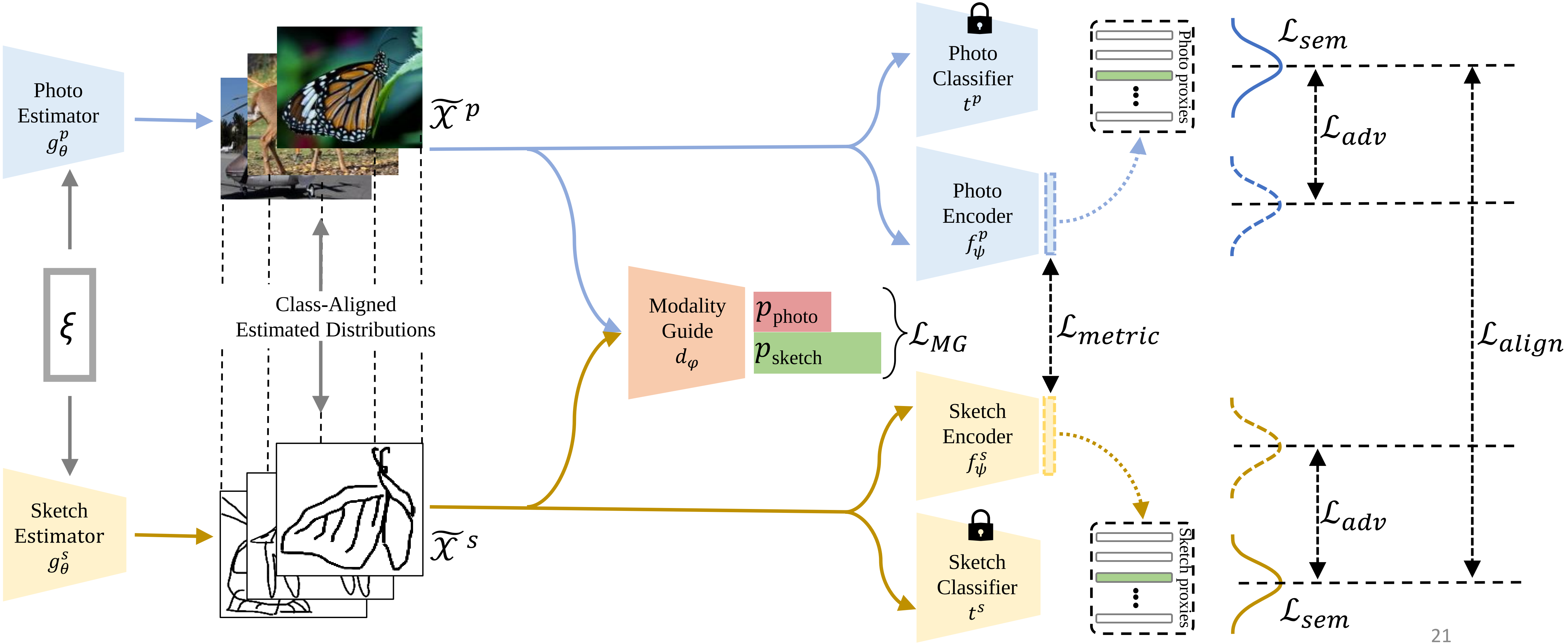
DFL Across Modalities and Metric Spaces (CrossX-DFL)

Metric-Agnostic Adversarial Estimation: Adversarial reconstruction across probabilistic and Euclidean metric spaces.



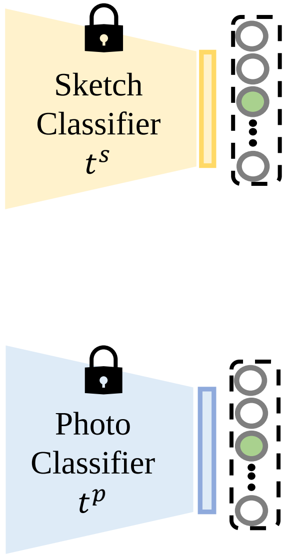
DFL Across Modalities and Metric Spaces (CrossX-DFL)

Cross-Modal Contrastive Learning: Queue-based Info-NCE minimization on the reconstructed photo-sketch pairs.



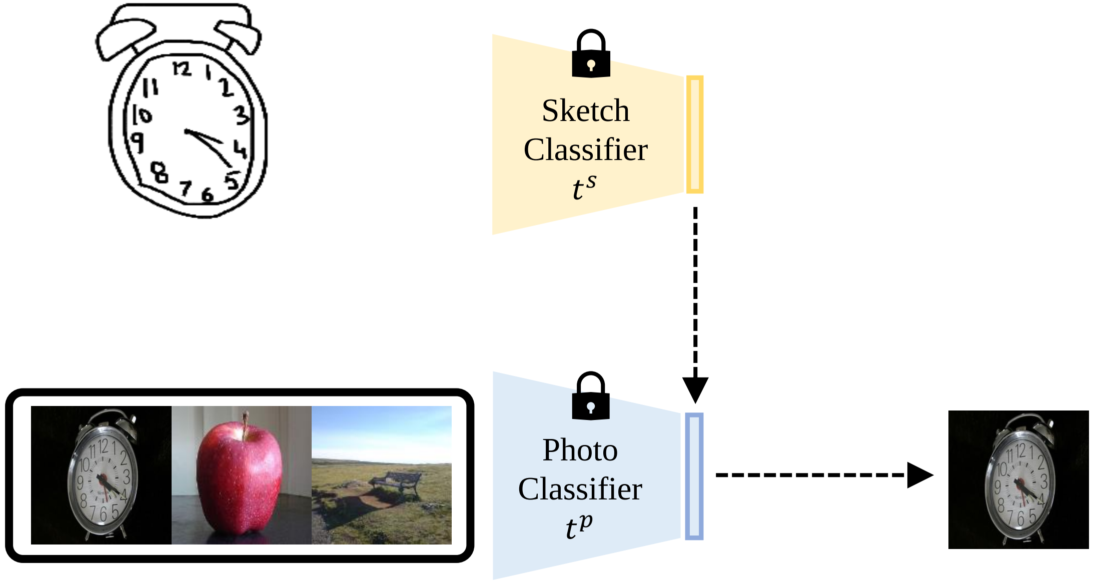
Outline

Comparison with Baselines



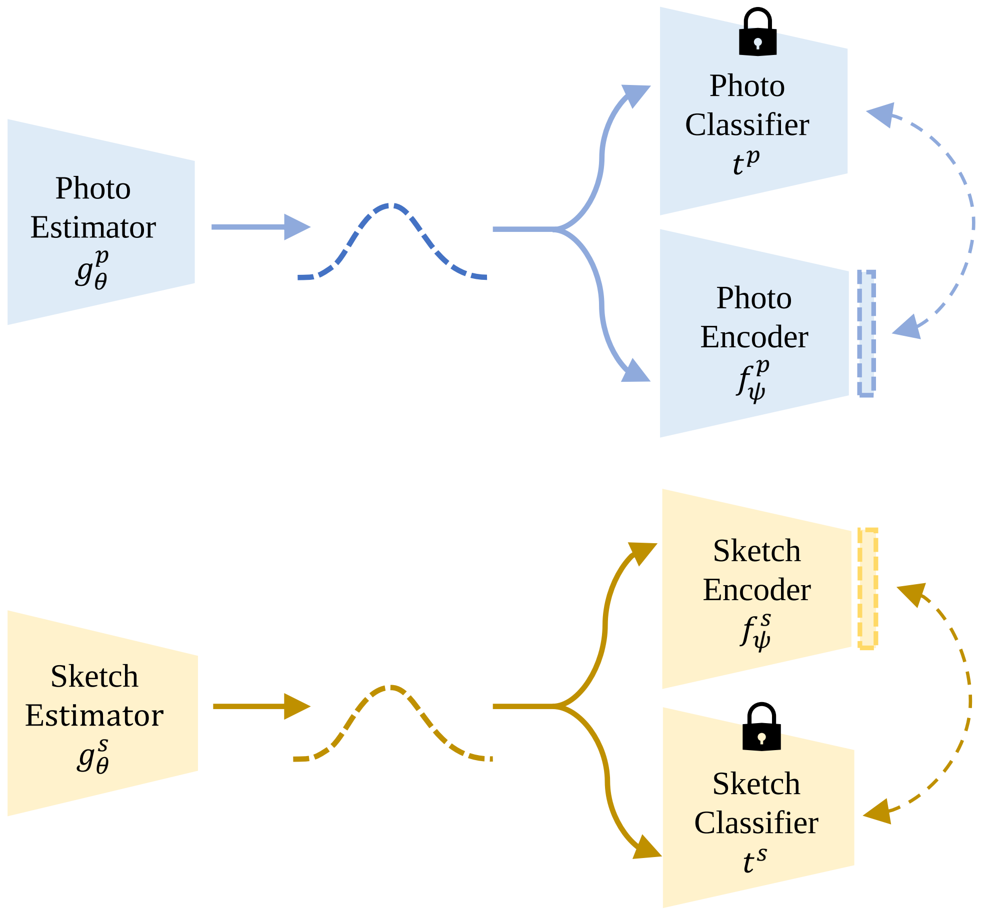
Method	Sketchy		TU-Berlin		QuickDraw	
	mAP@all	Prec@200	mAP@all	Prec@200	mAP@all	Prec@200
Classifier Only	0.530	0.542	0.330	0.338	0.160	0.180
Uni-Modal Distillation	0.529	0.537	0.291	0.295	0.130	0.140
Gaussian Prior	0.365	0.391	0.110	0.126	0.080	0.110
Averaging Weights	0.625	0.630	0.450	0.473	0.300	0.320
Meta-Data	0.573	0.576	0.380	0.395	0.200	0.221
Alternative Data	0.656	0.680	0.510	0.530	0.290	0330
Ours (CrossX-DFL)	**0.827**	**0.831**	**0.680**	**0.693**	**0.400**	**0.410**

Comparison with Baselines



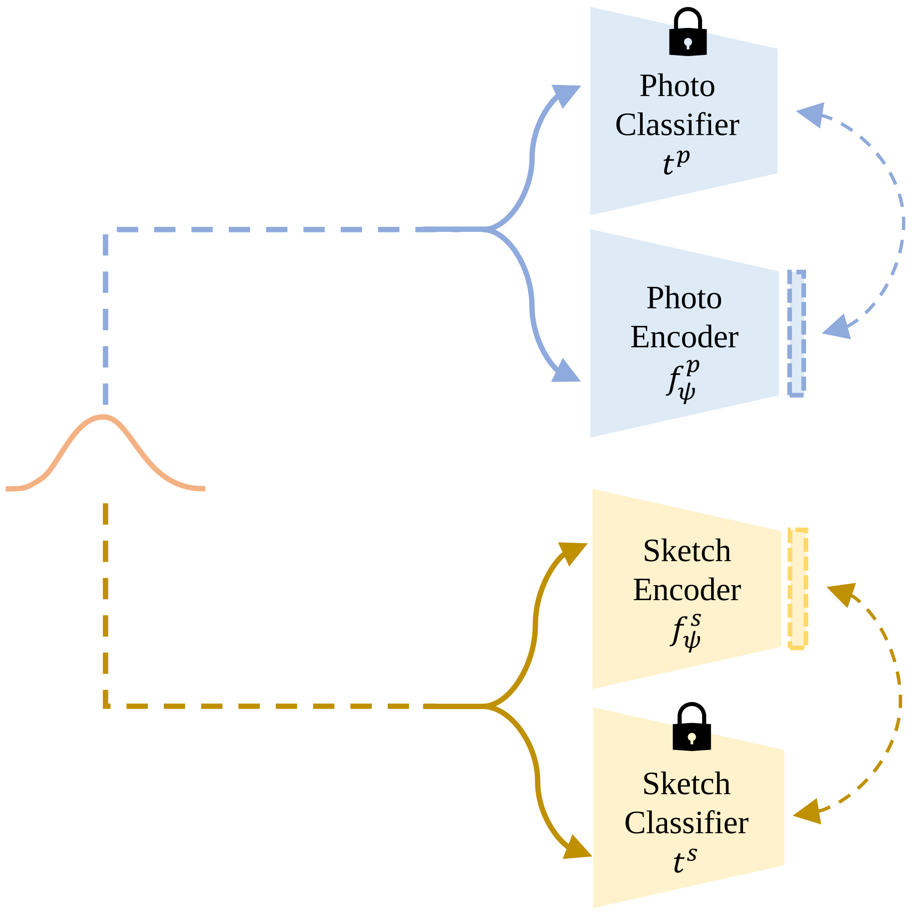
Method	Sketchy		TU-Berlin		QuickDraw	
	mAP@all	Prec@200	mAP@all	Prec@200	mAP@all	Prec@200
Classifier Only	0.530	0.542	0.330	0.338	0.160	0.180
Uni-Modal Distillation	0.529	0.537	0.291	0.295	0.130	0.140
Gaussian Prior	0.365	0.391	0.110	0.126	0.080	0.110
Averaging Weights	0.625	0.630	0.450	0.473	0.300	0.320
Meta-Data	0.573	0.576	0.380	0.395	0.200	0.221
Alternative Data	0.656	0.680	0.510	0.530	0.290	0330
Ours (CrossX-DFL)	**0.827**	**0.831**	**0.680**	**0.693**	**0.400**	**0.410**

Comparison with Baselines



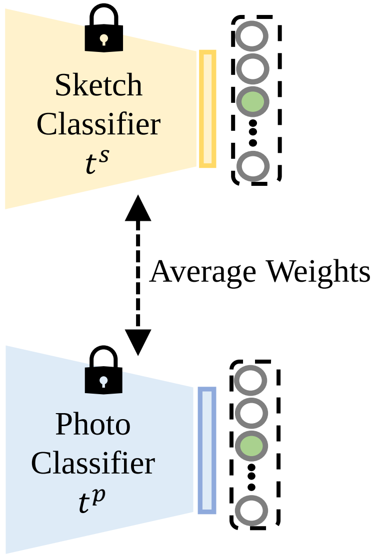
Method	Sketchy		TU-Berlin		QuickDraw	
	mAP@all	Prec@200	mAP@all	Prec@200	mAP@all	Prec@200
Classifier Only	0.530	0.542	0.330	0.338	0.160	0.180
Uni-Modal Distillation	0.529	0.537	0.291	0.295	0.130	0.140
Gaussian Prior	0.365	0.391	0.110	0.126	0.080	0.110
Averaging Weights	0.625	0.630	0.450	0.473	0.300	0.320
Meta-Data	0.573	0.576	0.380	0.395	0.200	0.221
Alternative Data	0.656	0.680	0.510	0.530	0.290	0330
Ours (CrossX-DFL)	**0.827**	**0.831**	**0.680**	**0.693**	**0.400**	**0.410**

Comparison with Baselines



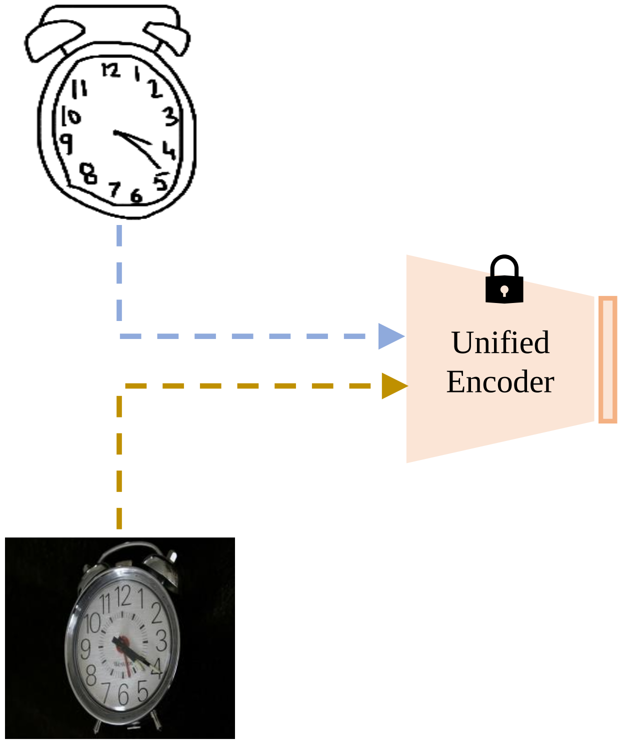
Method	Sketchy		TU-Berlin		QuickDraw	
	mAP@all	Prec@200	mAP@all	Prec@200	mAP@all	Prec@200
Classifier Only	0.530	0.542	0.330	0.338	0.160	0.180
Uni-Modal Distillation	0.529	0.537	0.291	0.295	0.130	0.140
Gaussian Prior	0.365	0.391	0.110	0.126	0.080	0.110
Averaging Weights	0.625	0.630	0.450	0.473	0.300	0.320
Meta-Data	0.573	0.576	0.380	0.395	0.200	0.221
Alternative Data	0.656	0.680	0.510	0.530	0.290	0330
Ours (CrossX-DFL)	**0.827**	**0.831**	**0.680**	**0.693**	**0.400**	**0.410**

Comparison with Baselines



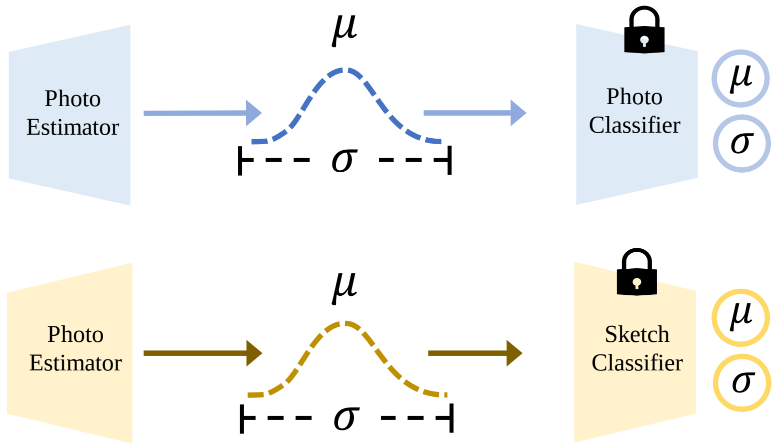
Method	Sketchy		TU-Berlin		QuickDraw	
	mAP@all	Prec@200	mAP@all	Prec@200	mAP@all	Prec@200
Classifier Only	0.530	0.542	0.330	0.338	0.160	0.180
Uni-Modal Distillation	0.529	0.537	0.291	0.295	0.130	0.140
Gaussian Prior	0.365	0.391	0.110	0.126	0.080	0.110
Averaging Weights	0.625	0.630	0.450	0.473	0.300	0.320
Meta-Data	0.573	0.576	0.380	0.395	0.200	0.221
Alternative Data	0.656	0.680	0.510	0.530	0.290	0330
Ours (CrossX-DFL)	**0.827**	**0.831**	**0.680**	**0.693**	**0.400**	**0.410**

Comparison with Baselines



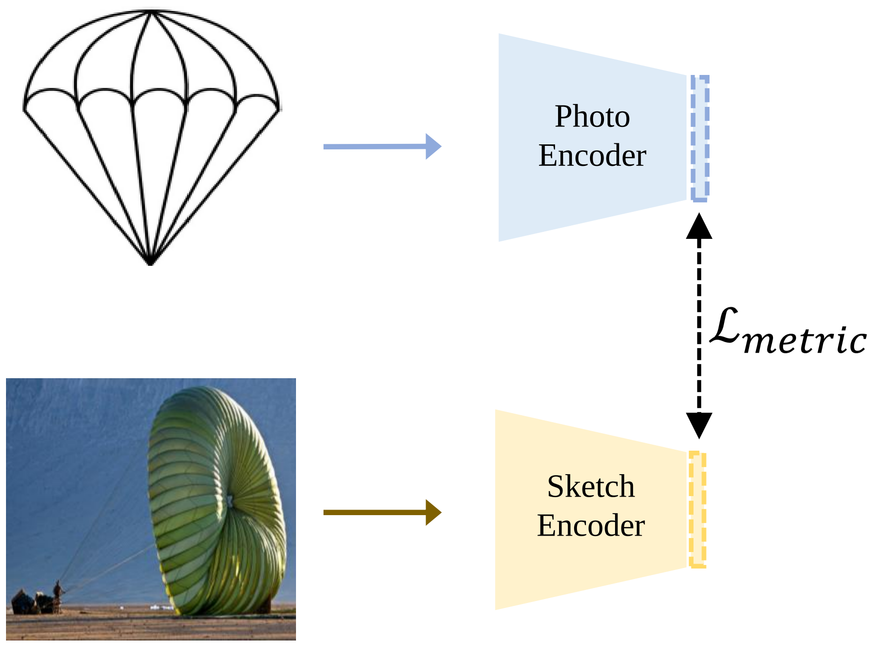
Method	Sketchy		TU-Berlin		QuickDraw	
	mAP@all	Prec@200	mAP@all	Prec@200	mAP@all	Prec@200
Classifier Only	0.530	0.542	0.330	0.338	0.160	0.180
Uni-Modal Distillation	0.529	0.537	0.291	0.295	0.130	0.140
Gaussian Prior	0.365	0.391	0.110	0.126	0.080	0.110
Averaging Weights	0.625	0.630	0.450	0.473	0.300	0.320
Meta-Data	0.573	0.576	0.380	0.395	0.200	0.221
Alternative Data	0.656	0.680	0.510	0.530	0.290	0330
Ours (CrossX-DFL)	**0.827**	**0.831**	**0.680**	**0.693**	**0.400**	**0.410**

Comparison with Baselines



Method	Sketchy		TU-Berlin		QuickDraw	
	mAP@all	Prec@200	mAP@all	Prec@200	mAP@all	Prec@200
Classifier Only	0.530	0.542	0.330	0.338	0.160	0.180
Uni-Modal Distillation	0.529	0.537	0.291	0.295	0.130	0.140
Gaussian Prior	0.365	0.391	0.110	0.126	0.080	0.110
Averaging Weights	0.625	0.630	0.450	0.473	0.300	0.320
Meta-Data	0.573	0.576	0.380	0.395	0.200	0.221
Alternative Data	0.656	0.680	0.510	0.530	0.290	0330
Ours (CrossX-DFL)	0.827	0.831	0.680	0.693	0.400	0.410

Comparison with Baselines



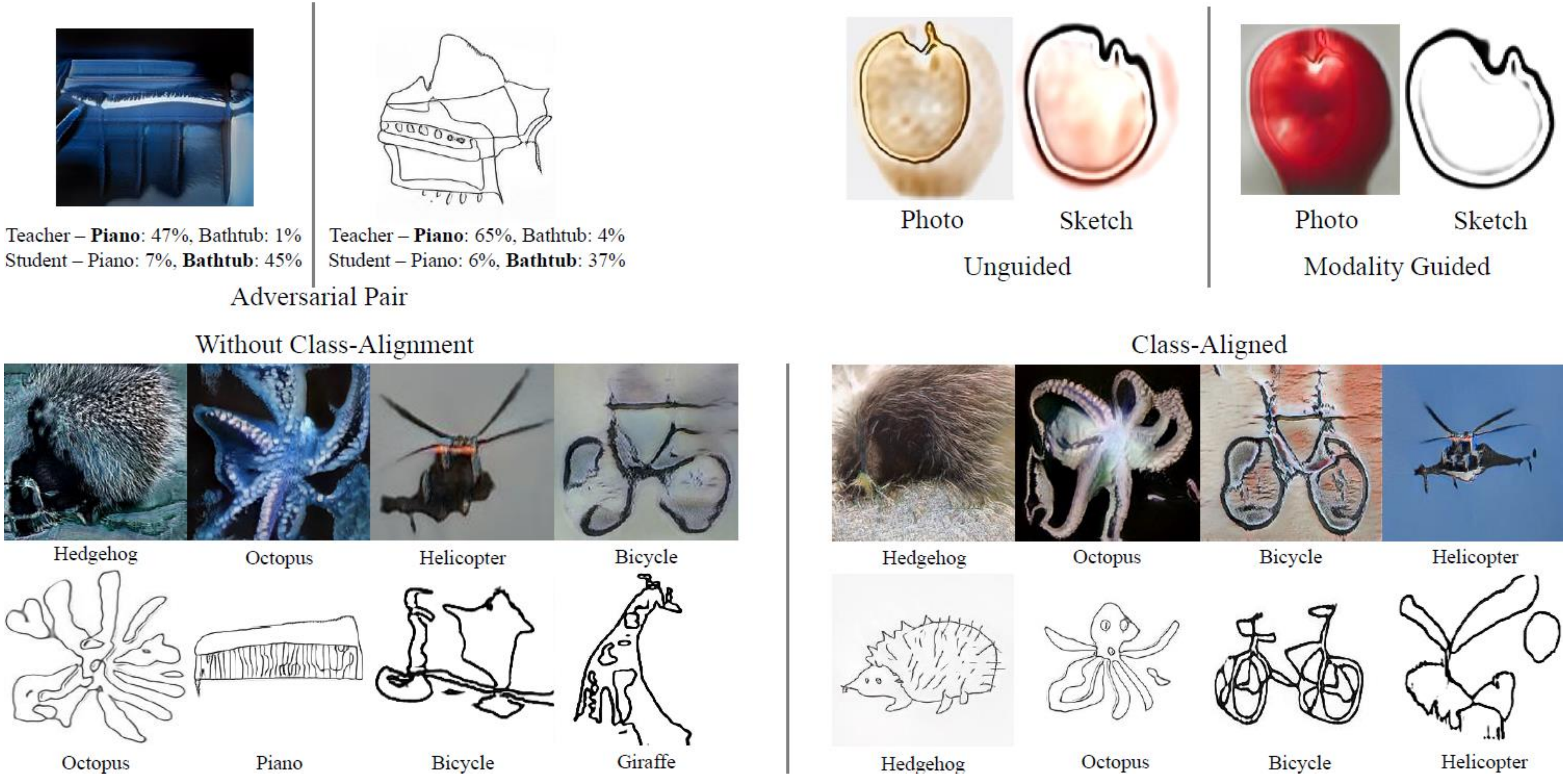
Method	Sketchy		TU-Berlin		QuickDraw	
	mAP@all	Prec@200	mAP@all	Prec@200	mAP@all	Prec@200
Classifier Only	0.530	0.542	0.330	0.338	0.160	0.180
Uni-Modal Distillation	0.529	0.537	0.291	0.295	0.130	0.140
Gaussian Prior	0.365	0.391	0.110	0.126	0.080	0.110
Averaging Weights	0.625	0.630	0.450	0.473	0.300	0.320
Meta-Data	0.573	0.576	0.380	0.395	0.200	0.221
Alternative Data	0.656	0.680	0.510	0.530	0.290	0330
Ours (CrossX-DFL)	0.827	0.831	0.680	0.693	0.400	0.410

Qualitative Reconstruction Results

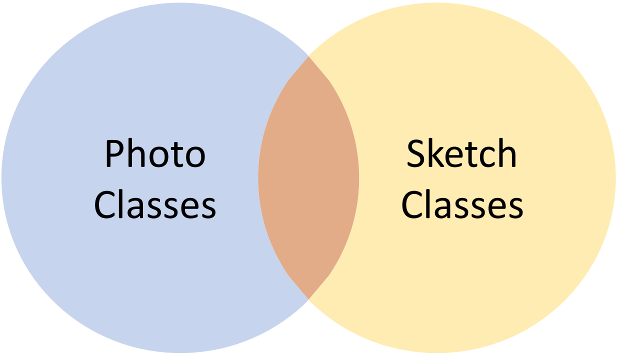
Comparison with Data-Dependent Settings

Objective	**Data-Dependent**	**Data-Free**	Δ
Siamese	0.715	0.679	0.036
Triplet	0.772	0.750	0.022
MIB	0.871	0.815	0.056
Ours (CrossX-DFL)	0.862	0.827	0.035

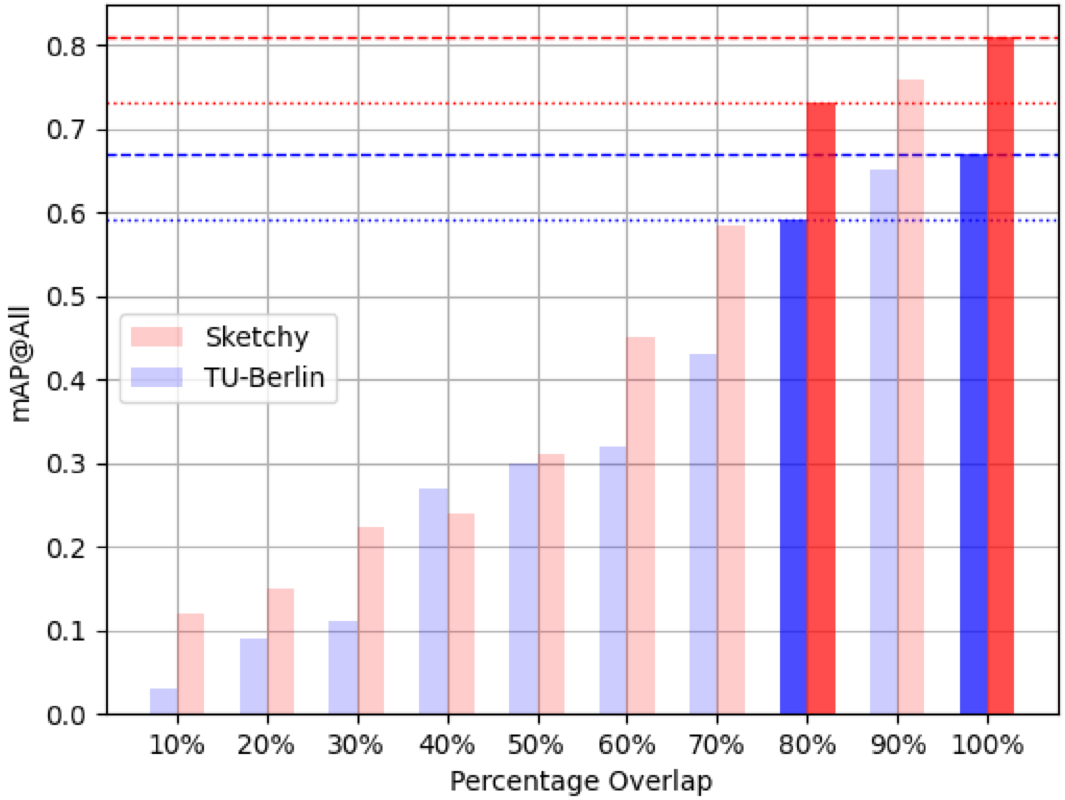
Qualitative Ablations



Teachers (Classifiers) with Partial Class Overlap



- Some classes are shared between the photo and sketch classifiers, while others are not.
- Retrieval is performed on the set of classes obtained via the union of photo and sketch classifier domains.



Outline

- Preliminaries
- Problem Definition
- Methodology
- Experiments
- Conclusion

Summary

- We presented CrossXDFL, an approach to train photo and sketch encoders for performing Sketch-Based Image Retrieval in a *Data Free* manner.
- We achieved the above by controlled reconstruction of the train set distributions of pretrained photo and sketch classifiers.
- Specifically, the following were the key components of our model – ***Class-Alignment*** (for paired photo-sketch generation); ***Modality Guidance*** (for modality specific reconstruction); and ***Metric-Agnostic Adversarial Estimation*** (for generating hard sample that help in training robust encoders).
- CrossXDFL performs significantly better than existing DFL baselines, and competitively with respect to data-driven approaches.

Summary

- We presented CrossXDFL, an approach to train photo and sketch encoders for performing Sketch-Based Image Retrieval in a Data Free manner.
- We achieved the above by controlled reconstruction of the train set distributions of pretrained photo and sketch classifiers.
- Specifically, the following were the key components of our model – ***Class-Alignment*** (for paired photo-sketch generation); ***Modality Guidance*** (for modality specific reconstruction); and ***Metric-Agnostic Adversarial Estimation*** (for generating hard sample that help in training robust encoders).
- CrossXDFL performs significantly better than existing DFL baselines, and competitively with respect to data-driven approaches.

Summary

- We presented CrossXDFL, an approach to train photo and sketch encoders for performing Sketch-Based Image Retrieval in a Data Free manner.
- We achieved the above by controlled reconstruction of the train set distributions of pretrained photo and sketch classifiers.
- Specifically, the following were the key components of our model – ***Class-Alignment*** (for paired photo-sketch generation); ***Modality Guidance*** (for modality specific reconstruction); and ***Metric-Agnostic Adversarial Estimation*** (for generating hard sample that help in training robust encoders).
- CrossXDFL performs significantly better than existing DFL baselines, and competitively with respect to data-driven approaches.

Summary

- We presented CrossXDFL, an approach to train photo and sketch encoders for performing Sketch-Based Image Retrieval in a Data Free manner.
- We achieved the above by controlled reconstruction of the train set distributions of pretrained photo and sketch classifiers.
- Specifically, the following were the key components of our model – ***Class-Alignment*** (for paired photo-sketch generation); ***Modality Guidance*** (for modality specific reconstruction); and ***Metric-Agnostic Adversarial Estimation*** (for generating hard sample that help in training robust encoders).
- CrossXDFL performs significantly better than existing DFL baselines, and competitively with respect to data-driven approaches.

Data-Free Sketch-Based Image Retrieval

Get in touch:
Abhra Chaudhuri
ac1151@exeter.ac.uk

https://arxiv.org/abs/2303.07775

https://github.com/abhrac/data-free-sbir