ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

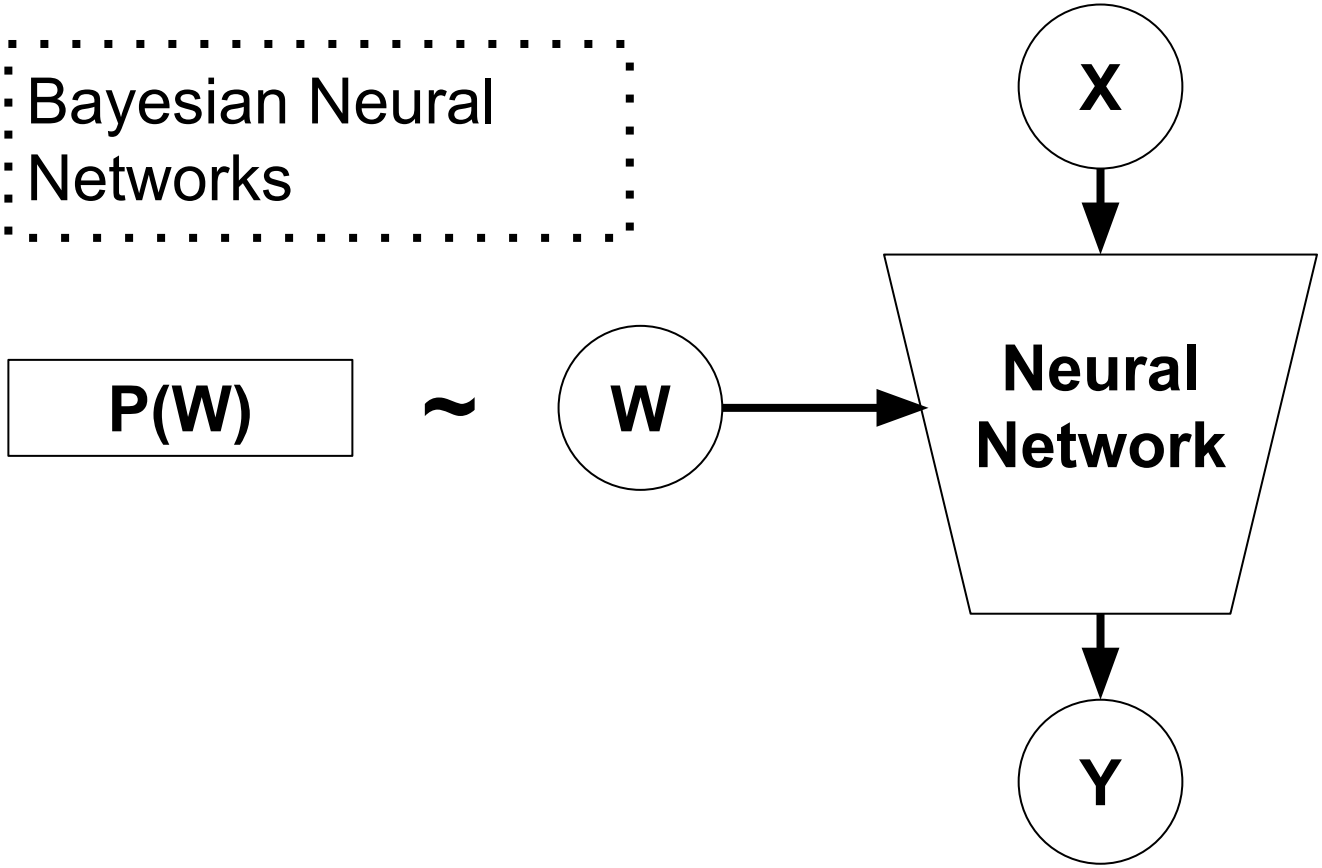
CVL Computer Vision Lab

On the Practicality of Deterministic Epistemic Uncertainty

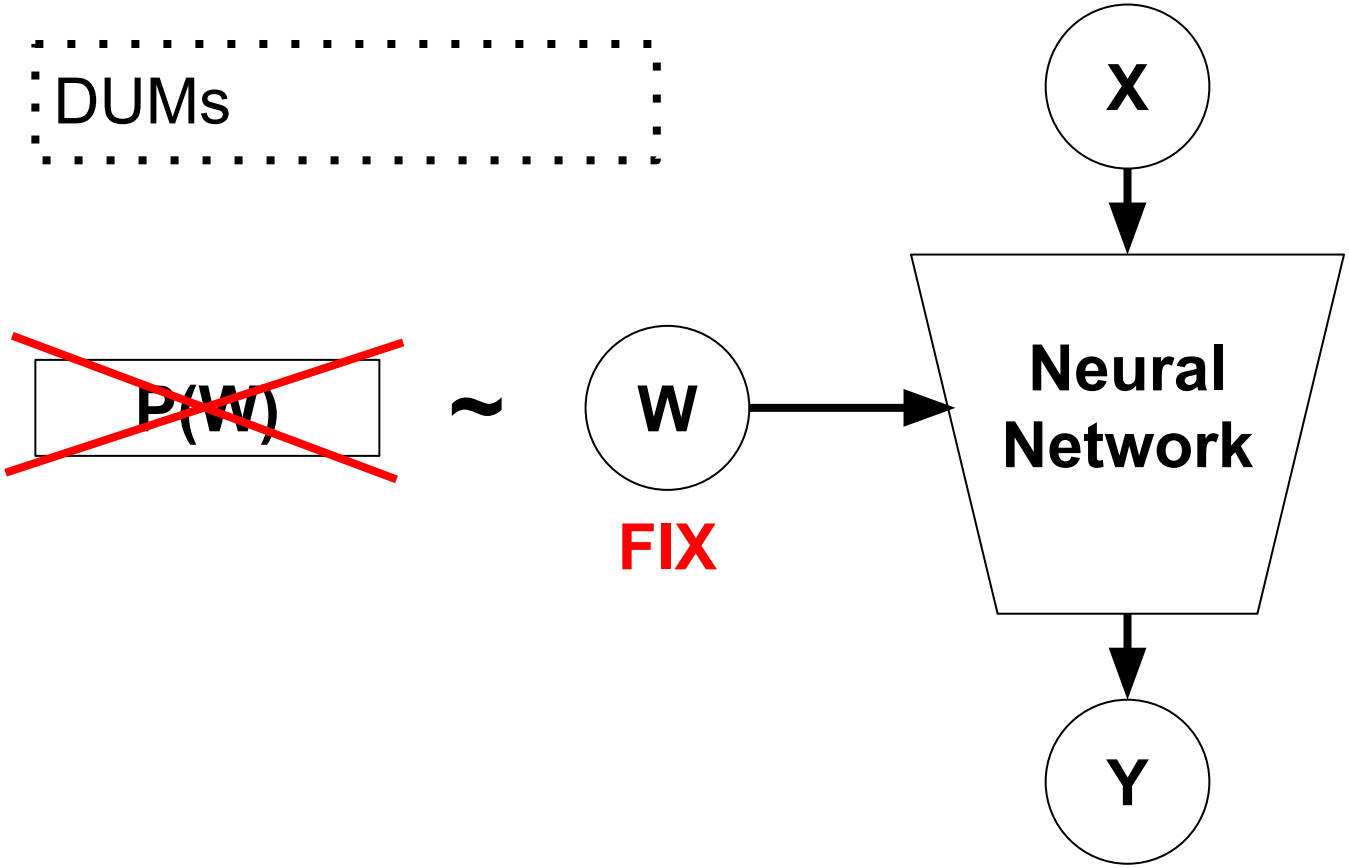
Janis Postels* [1], Mattia Segu* [1,2], Tao Sun [1], Luca Daniel Sieber [1], Luc Van Gool [1], Fisher Yu [1], Federico Tombari [3,4]

[1] ETH Zurich [2] MPI Informatics [3] Technical University Munich [4] Google

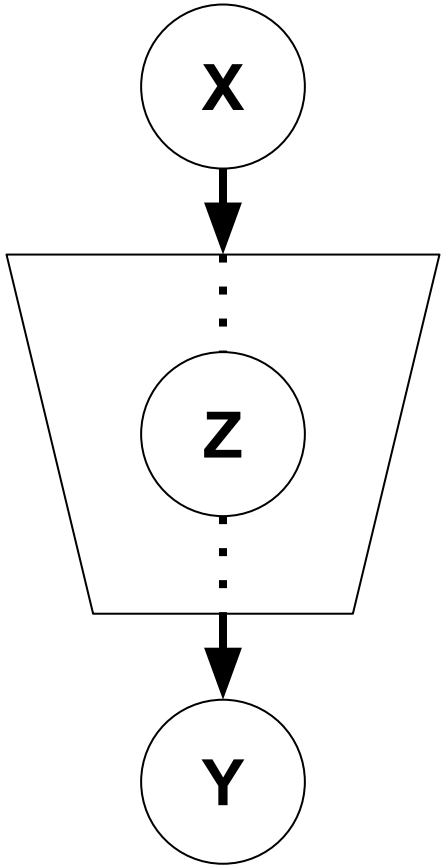
Deterministic Uncertainty Methods (DUMs)



Deterministic Uncertainty Methods (DUMs)

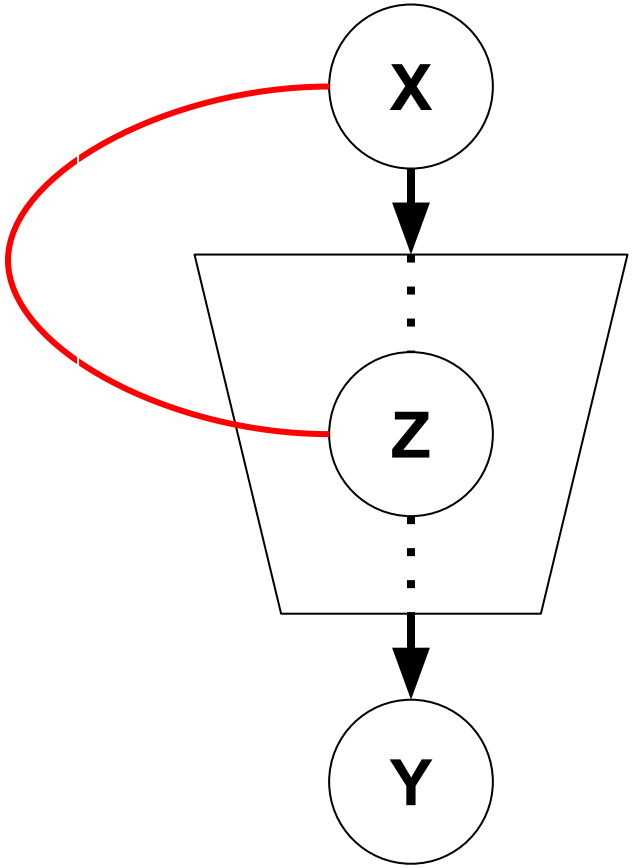


How to Regularize DUMs?



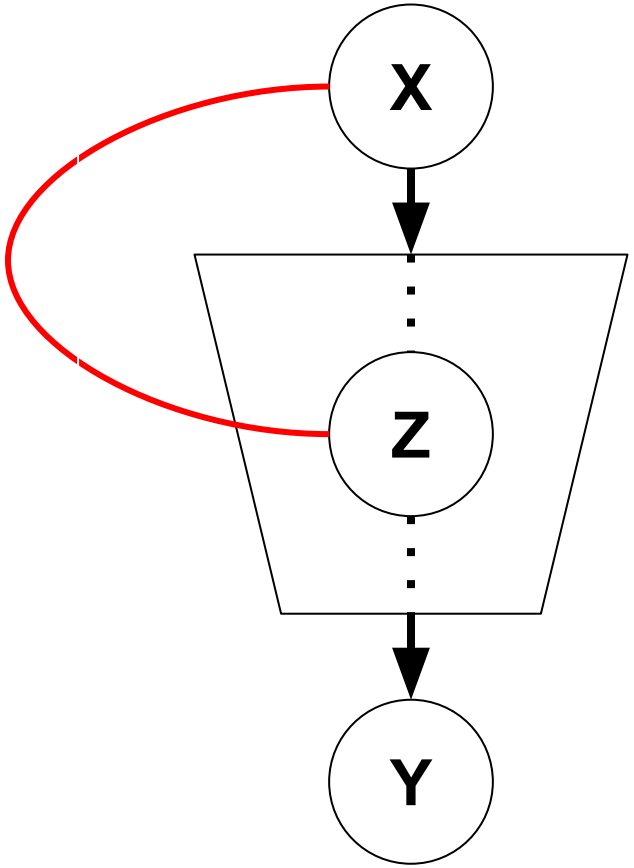
How to Regularize DUMs?

Distance-aware Representations
|Z1 - Z2| ~ |X1 - X2|

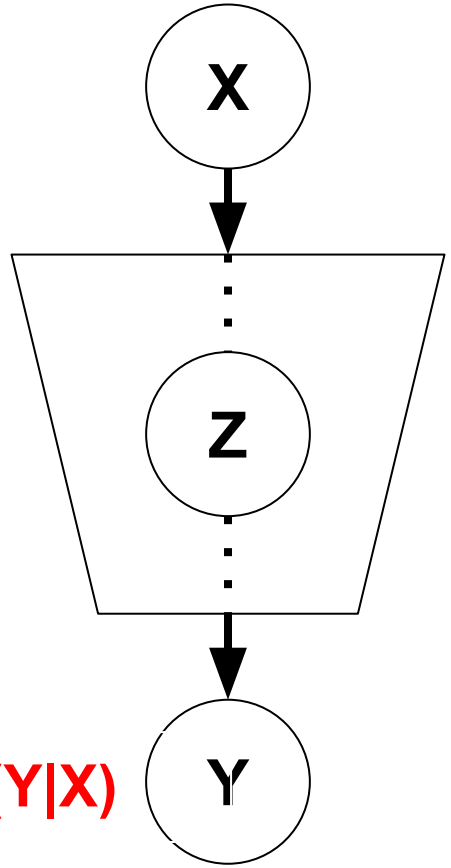


How to Regularize DUMs?

Informative Representations
max MI(X, Z)



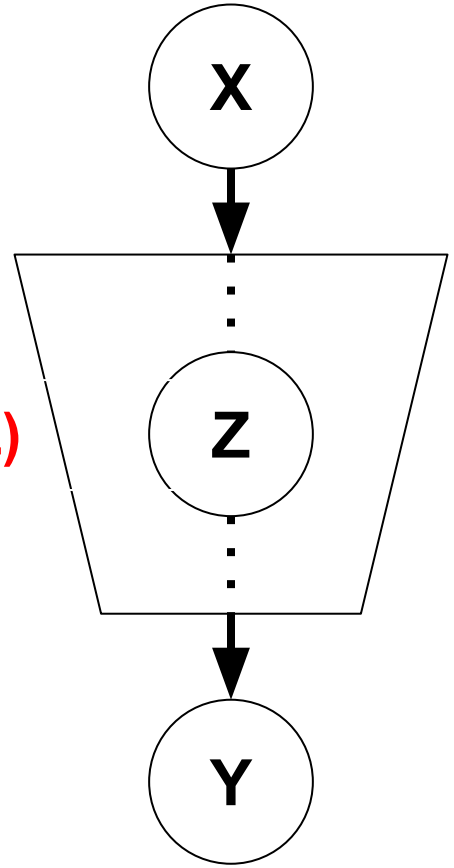
How Do DUMs Estimate Uncertainty?



Discriminative P(Y|X)

How Do DUMs Estimate Uncertainty?

Generative P(Z)



Desired Properties of an Epistemic Uncertainty Estimate

Desired Properties of an Epistemic Uncertainty Estimate

Out-of-Distribution Detection

- Residing far away from the training data distribution
- Random model performance

Desired Properties of an Epistemic Uncertainty Estimate

Out-of-Distribution Detection

- Residing far away from the training data distribution
- Random model performance

Calibration

- Model can generalize close to training data distribution
- Epistemic uncertainty should contain information about expected model performance

Research Questions regarding DUMs

Research Questions regarding DUMs

Do DUMs scale to realistic vision tasks?

Research Questions regarding DUMs

Do DUMs scale to realistic vision tasks?

Is the epistemic uncertainty from DUMs well calibrated?

Measuring the Calibration of DUMs

ECE / Brier Score not applicable

Measuring the Calibration of DUMs

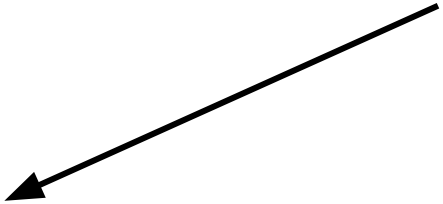
ECE / Brier Score not applicable

Property of well-calibrated uncertainty:
Can order samples according to predictive quality

Measuring the Calibration of DUMs

ECE / Brier Score not applicable

Property of well-calibrated uncertainty:
Can order samples according to predictive quality

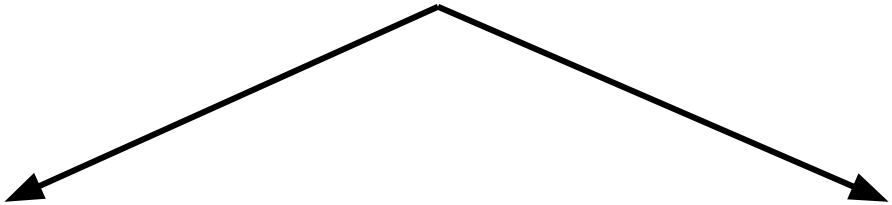


AUROC

Measuring the Calibration of DUMs

ECE / Brier Score not applicable

Property of well-calibrated uncertainty:
Can order samples according to predictive quality



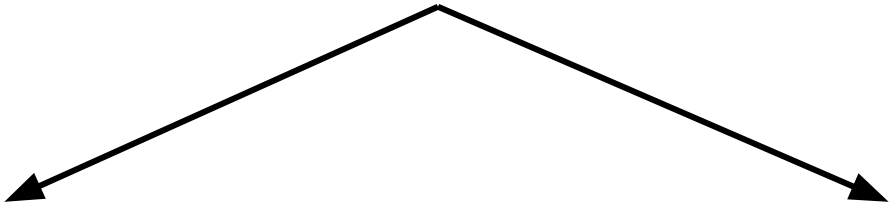
AUROC

rAULC
(Relative Area Under the Lift Curve)

Measuring the Calibration of DUMs

ECE / Brier Score not applicable

Property of well-calibrated uncertainty:
Can order samples according to predictive quality



AUROC

rAULC
(Relative Area Under the Lift Curve)

➢ Measures improvement over randomly ordered samples

Datasets & Distributional Shifts

Datasets & Distributional Shifts

Image classification:

- CIFAR10/100-C

Datasets & Distributional Shifts

Image classification:

- CIFAR10/100-C

Semantic Segmentation:

- CITYSCAPES-C
- CARLA-C
  - Uses SHIFT [1] toolkit for data collection in CARLA

[1] Sun & Segu et al. SHIFT: A Synthetic Driving Dataset for Continuous Multi-Task Domain Adaptation. CVPR 2022

Summary

DUMs scale to realistic vision tasks

Summary

DUMs scale to realistic vision tasks

Epistemic uncertainty from DUMs is not well calibrated

Summary

DUMs scale to realistic vision tasks

Epistemic uncertainty from DUMs is not well calibrated

Distance sensitivity does not correlate well with calibration or OOD detection

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

CVL Computer Vision Lab

Thanks for Your Attention!

Tue Jul 19 06:30 PM -- 08:30 PM
@ Hall E #506

Paper: