Institute of Cartography and Geoinformatics | Leibniz Universität Hannover

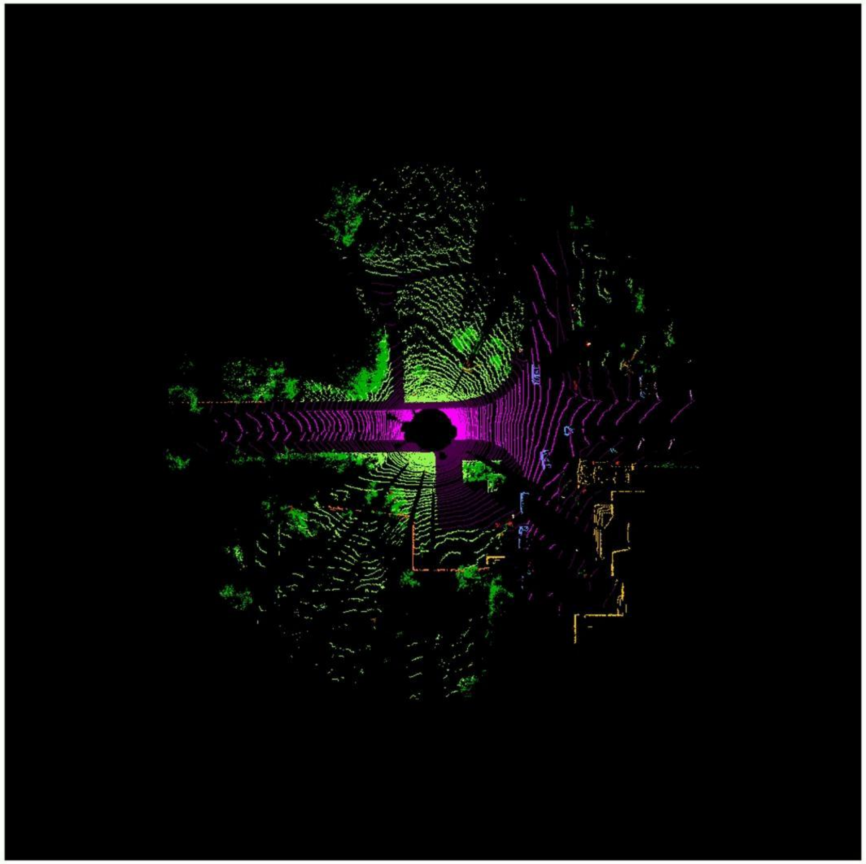
Uncertainty Estimation in LiDAR Semantic Segmentation Using Gaussian Mixture Models

Hanieh Shojaei
Hanieh.shojaei@ikg.uni-hannover.de

Leibniz
Universität
Hannover

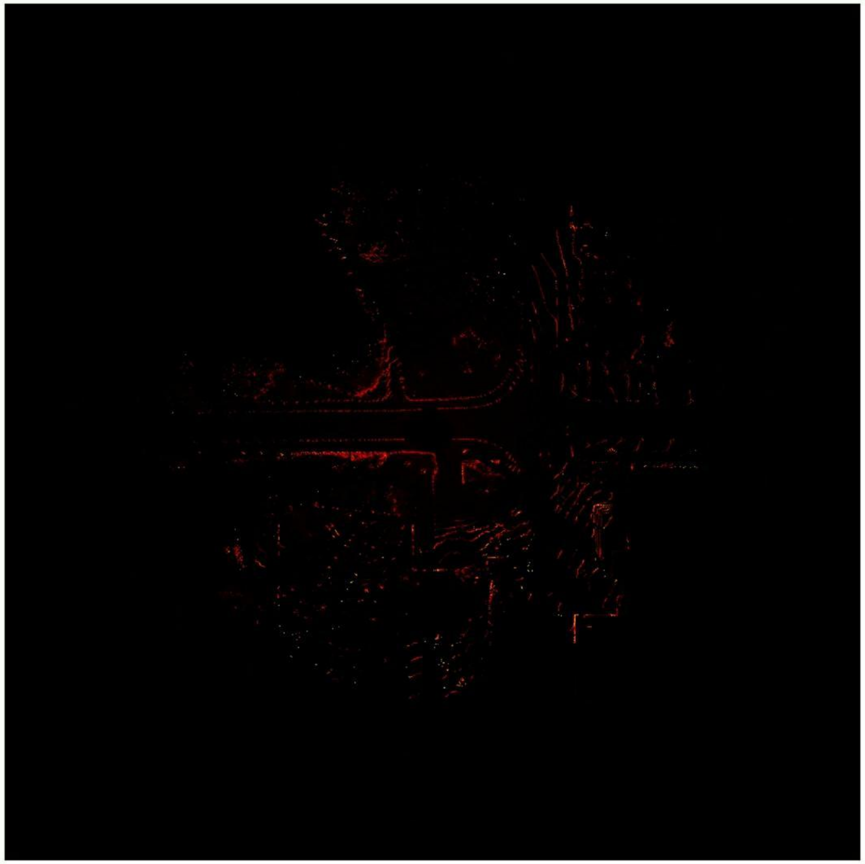
Uncertainty-aware Scene Understanding

Scene Prediction



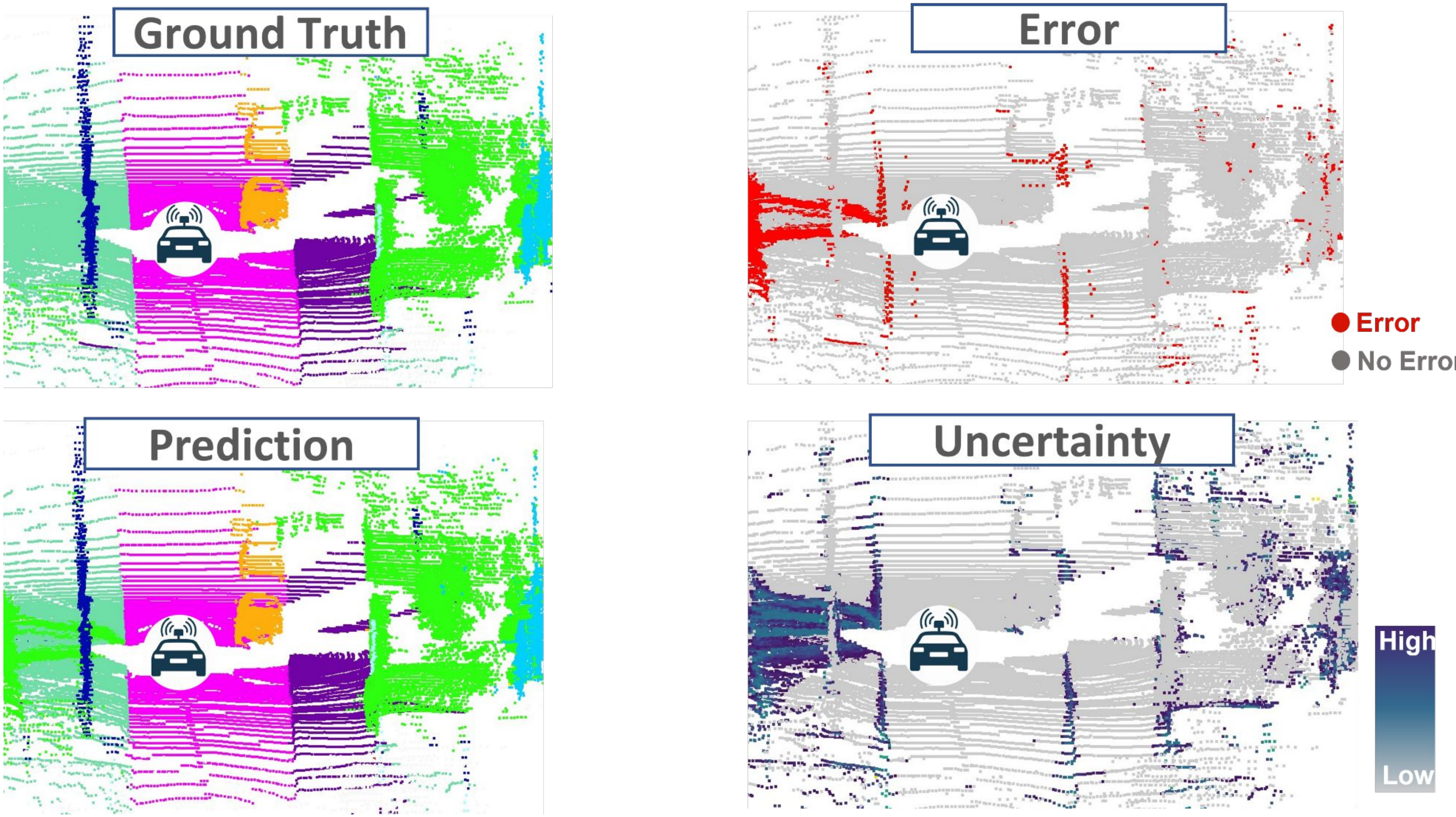
Street
Building
Vegetation
Sidewalk
Car

Uncertainty



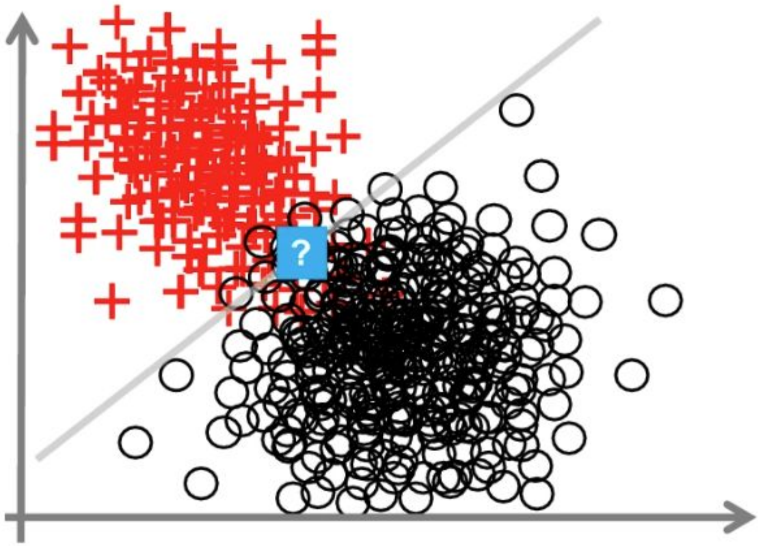
High Uncertainty

Uncertainty-aware Scene Understanding

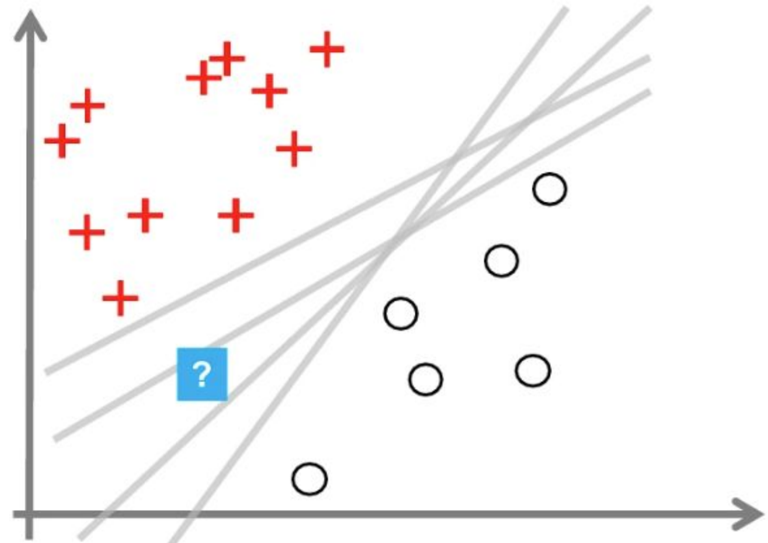


Uncertainty types in Deep Learning

- Aleatoric and Epistemic Uncertainty



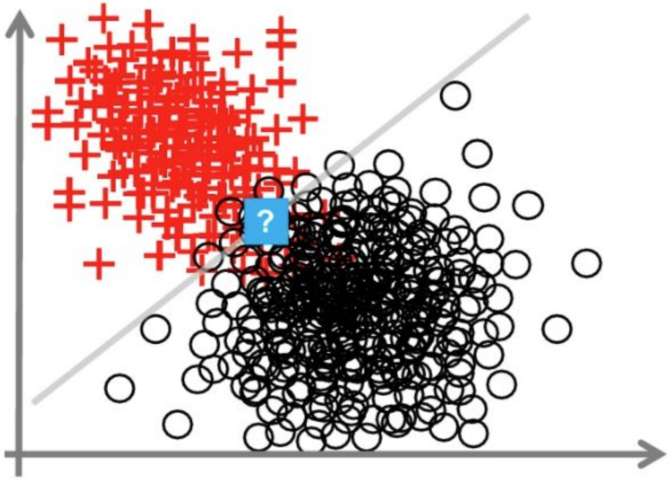
Aleatoric Uncertainty
Overlapping classes



Epistemic Uncertainty
Insufficient data

[1] Hüllermeier, Eyke, and Willem Waegeman. "Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods." *Machine learning* 110, no. 3 (2021): 457-506.

Aleatoric Uncertainty



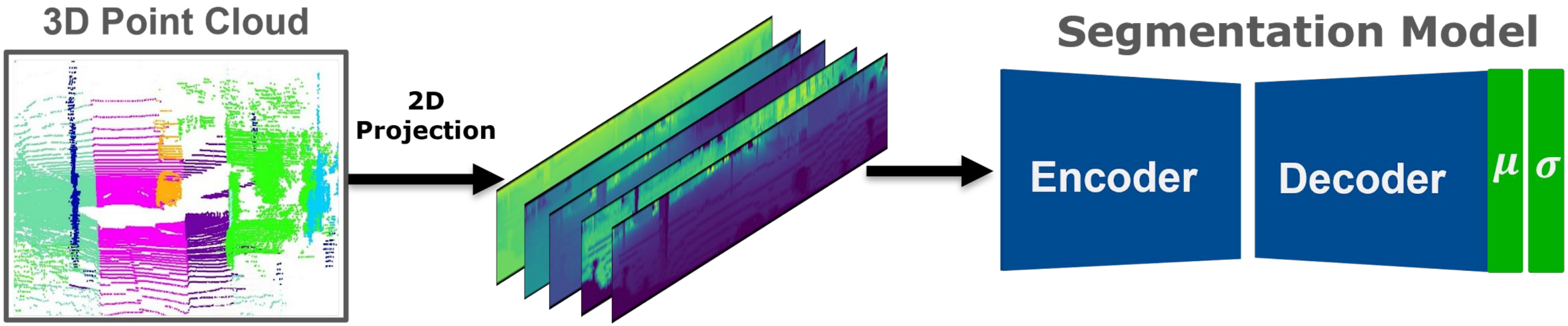
CORL

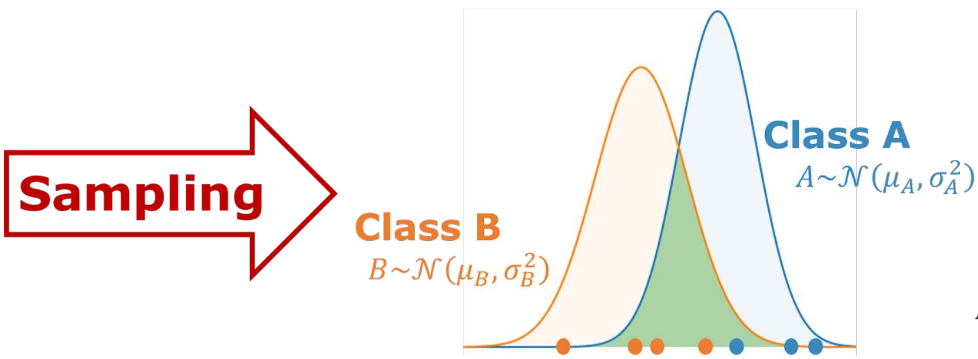
2025 Conference on Robot Learning

Hanieh Shojaei, Q. Zou, and C. Brenner. "Calibrated and Efficient Sampling-Free Confidence Estimation for LiDAR Scene Semantic Segmentation." *CORL 2025*

When the input data is noisy or ambiguous, the model predicts not just a single output, but also a variance!

Aleatoric Uncertainty

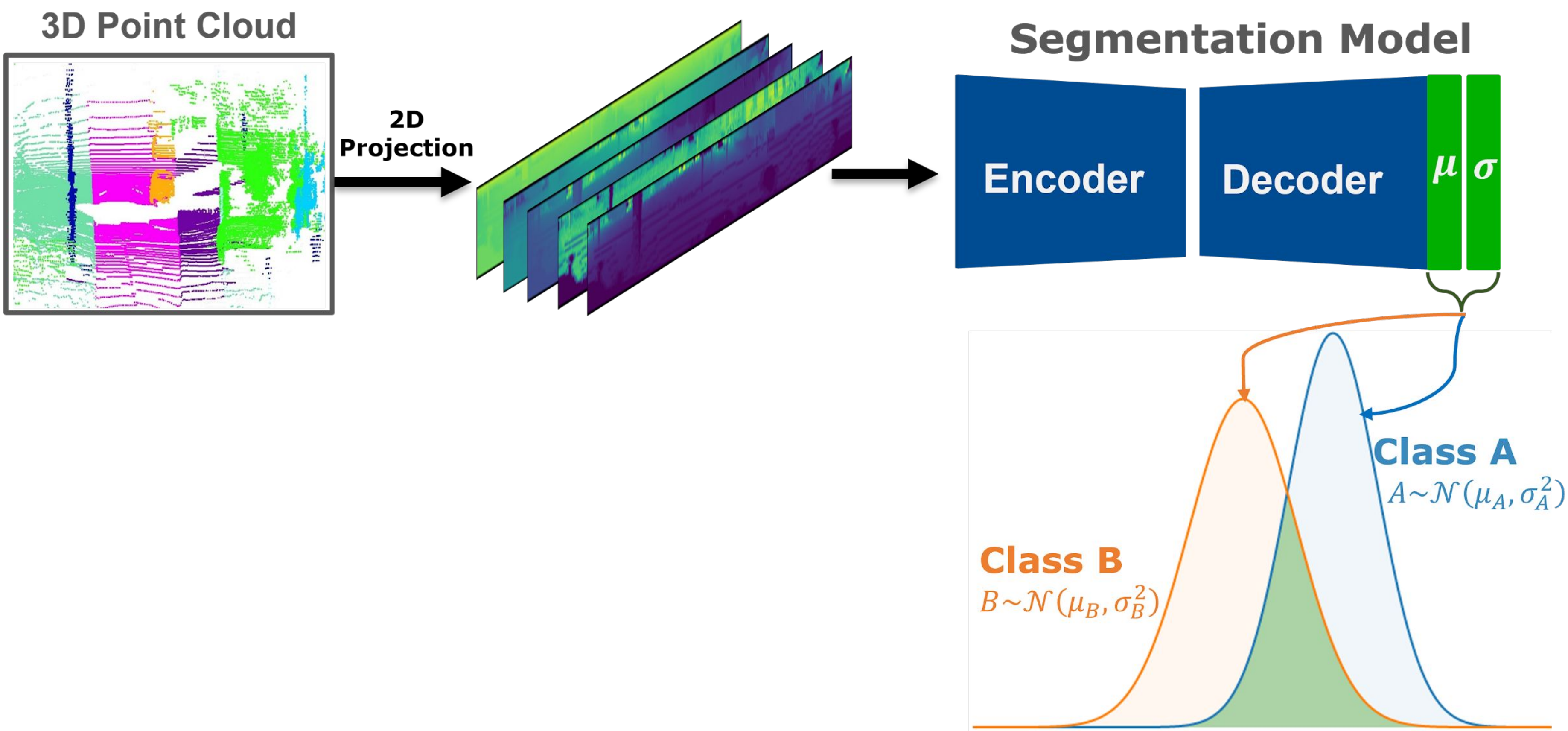




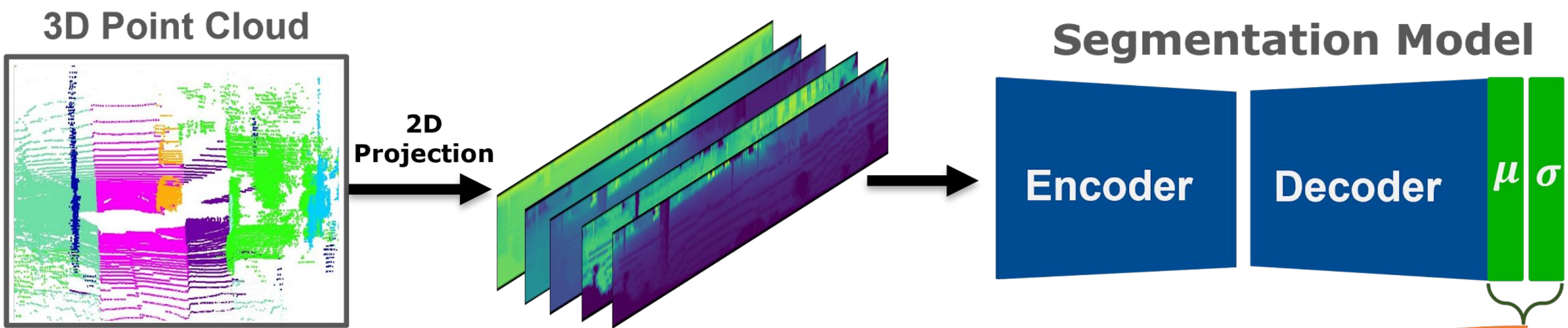
$$logit_{t,A} = \mu_A + \sigma_A \, \epsilon_t$$
$$logit_{t,B} = \mu_B + \sigma_B \, \epsilon_t$$
$$\epsilon_t \sim \mathcal{N}(0, I)$$

$$Loss_x = -\sum_i \log \frac{1}{T} \sum_t softmax(logit_{t,c})$$

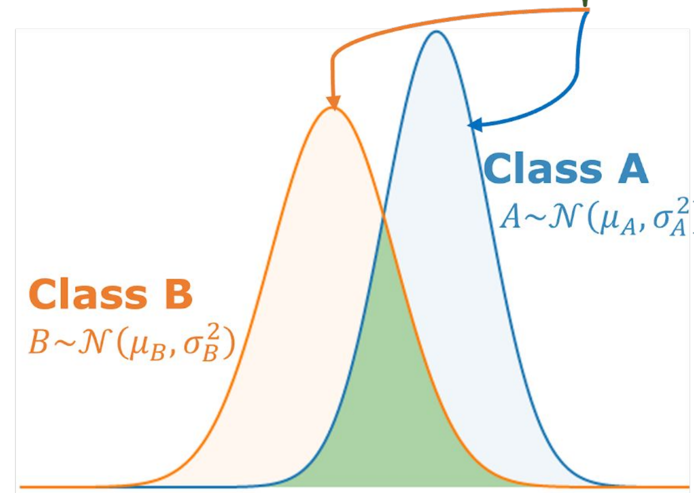
Aleatoric Uncertainty



Aleatoric Uncertainty

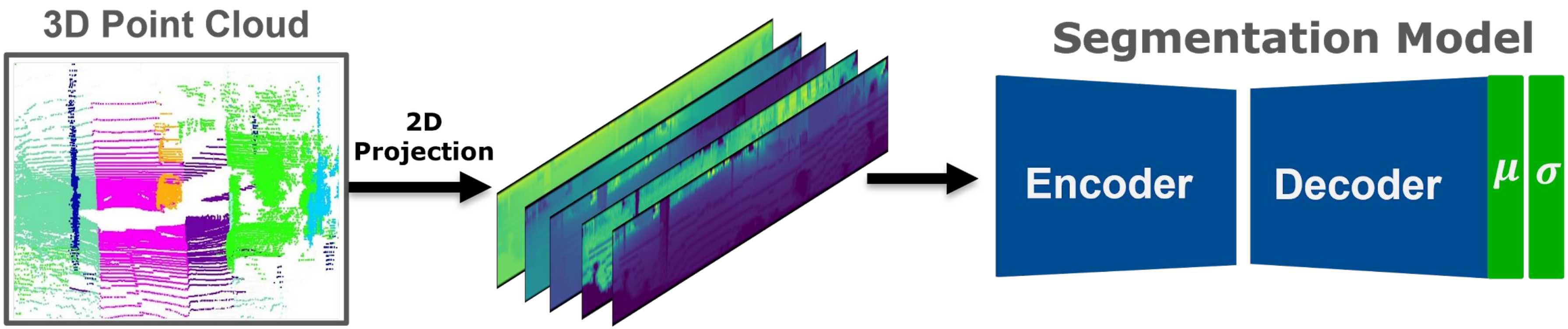


Which class do you **predict**? **Class A**
How **confidence** are you? $\boldsymbol{P(A > B)}$**=** $\boldsymbol{\Phi_{AB}}$



This overlap makes the classification uncertain!

Aleatoric Uncertainty

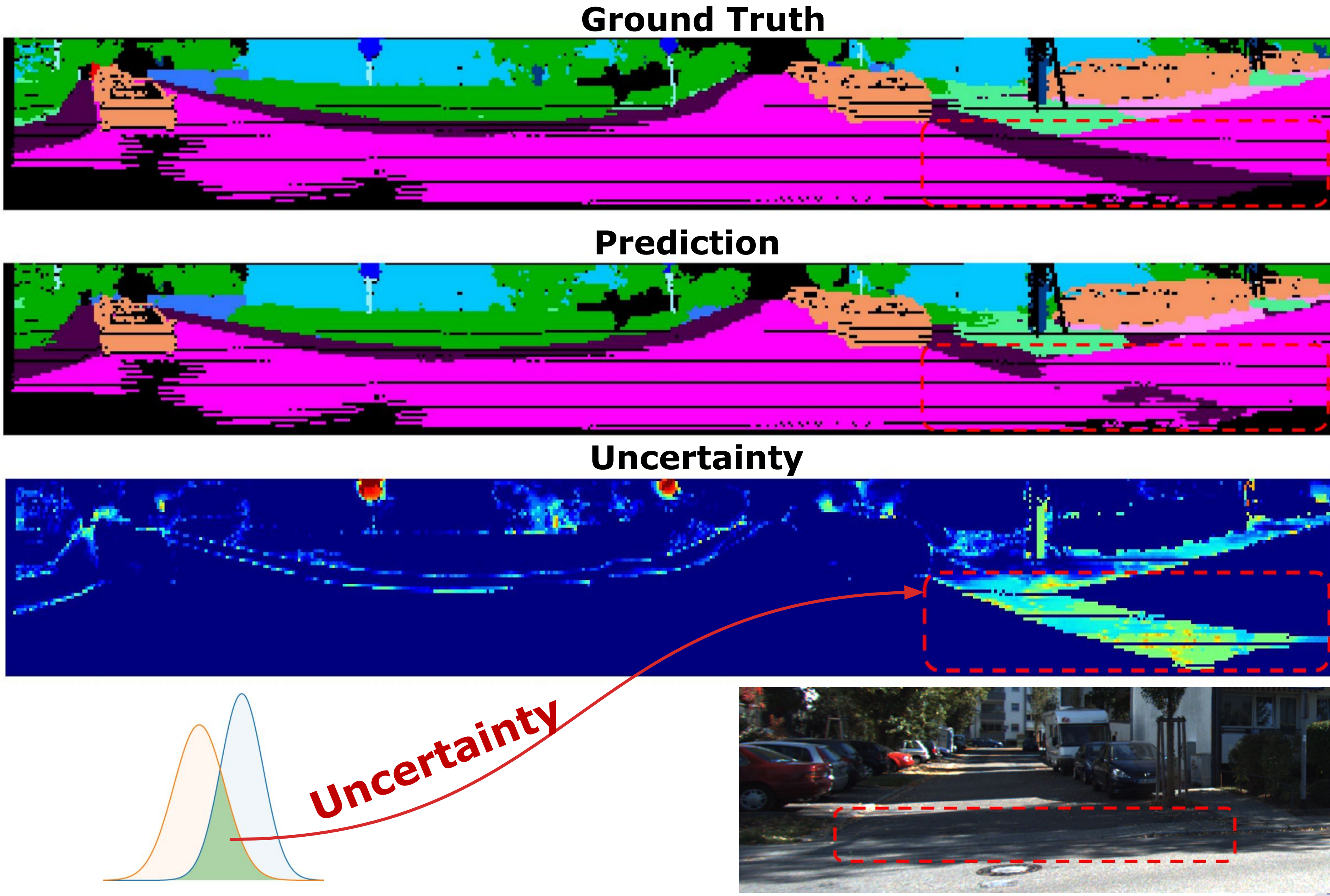


How confident are you that Class A is the correct prediction?

$$\boldsymbol{P}(\boldsymbol{A} > \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}, \boldsymbol{E}) = ?$$

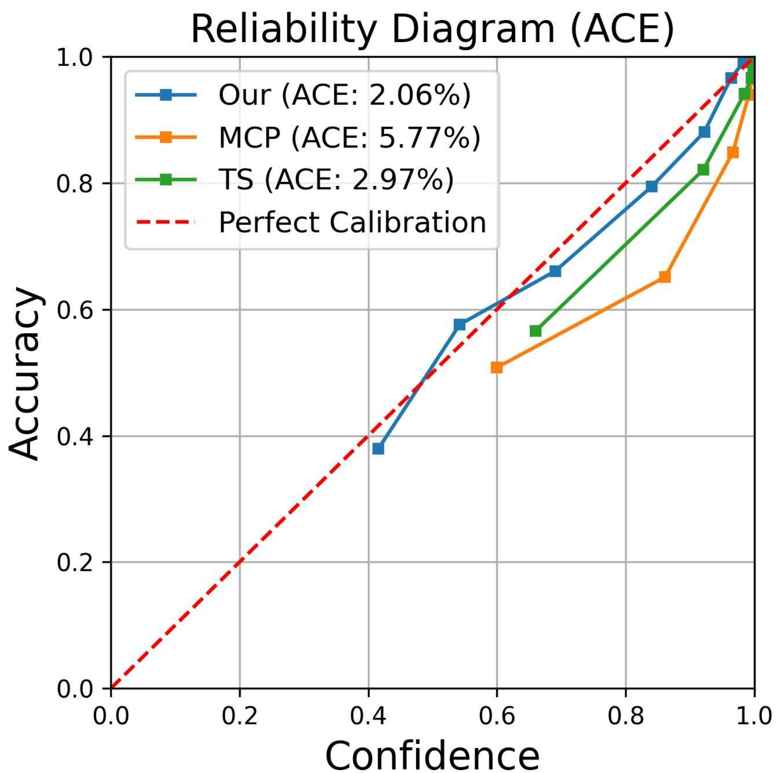
$$P\left(X_1 > \max_{i \geqq 2} X_i\right) = \prod_{i=2}^{C} \Phi_{1i}$$

Aleatoric Uncertainty

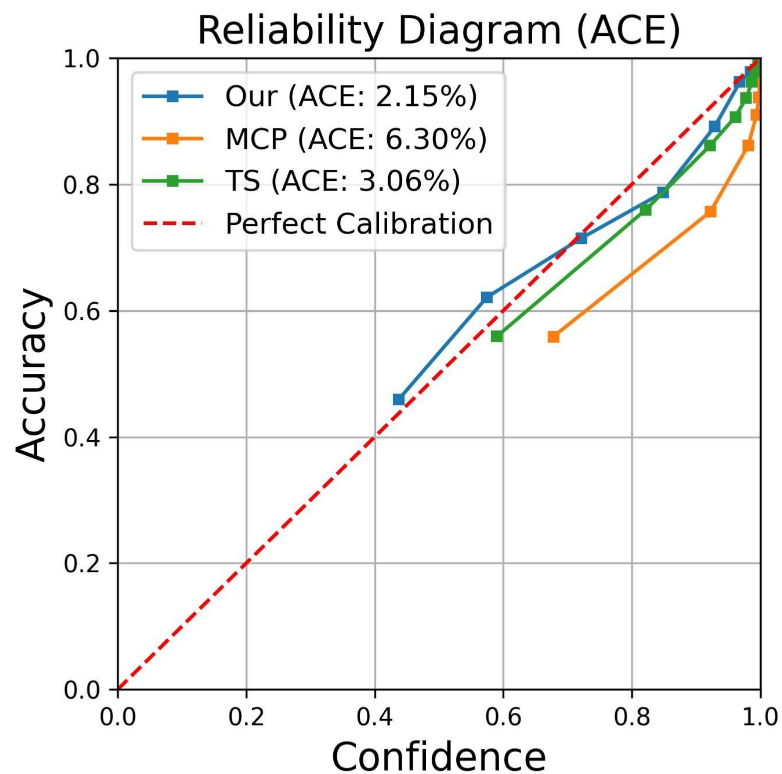


Aleatoric Uncertainty

- **Our work is better calibrated compared to the Baselines!**



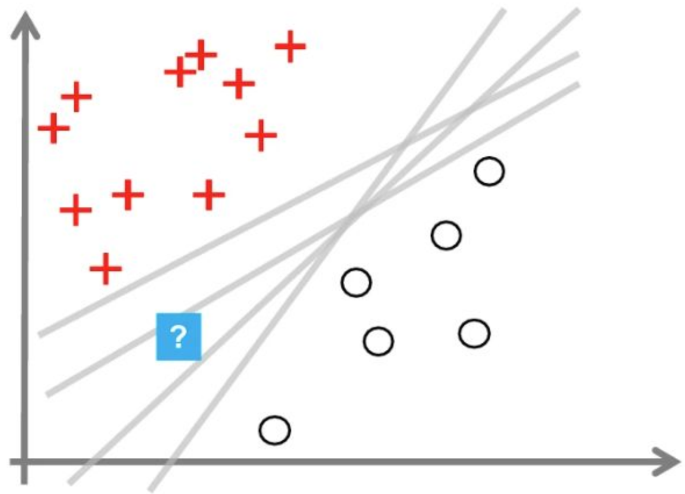
Backbone: RangeViT



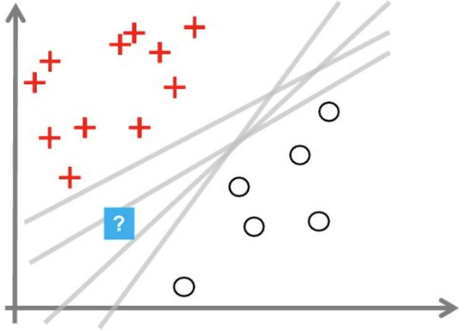
Backbone: SalsaNext

Hanieh Shojaei, Q. Zou, and C. Brenner. "Calibrated and Efficient Sampling-Free Confidence Estimation for LiDAR Scene Semantic Segmentation." *CORL 2025*

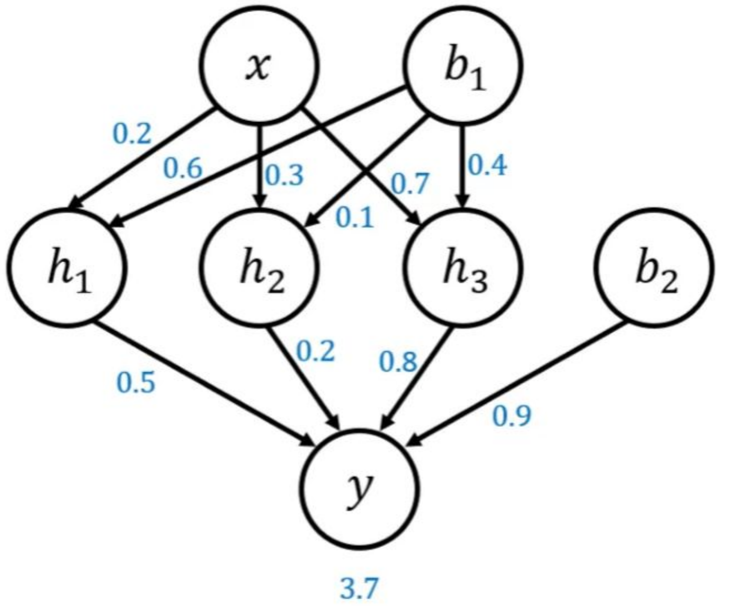
Epistemic Uncertainty



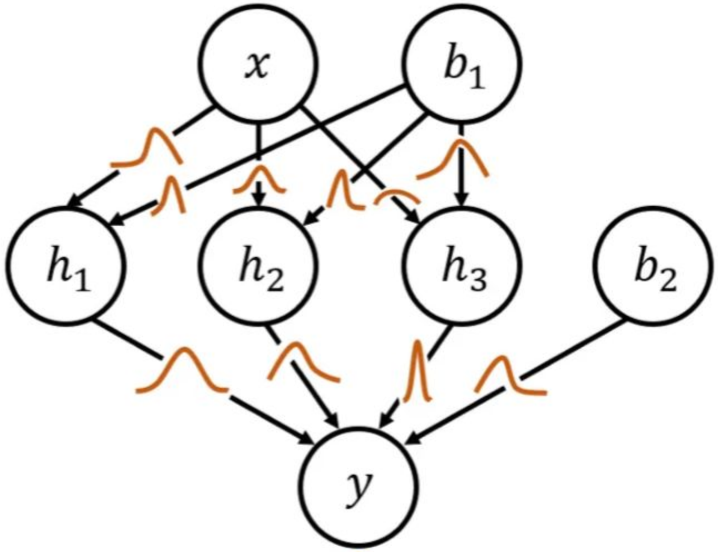
Epistemic Uncertainty



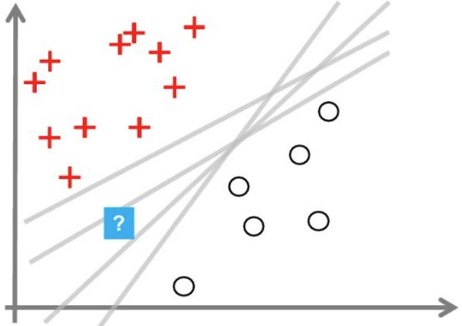
Standard Neural Network



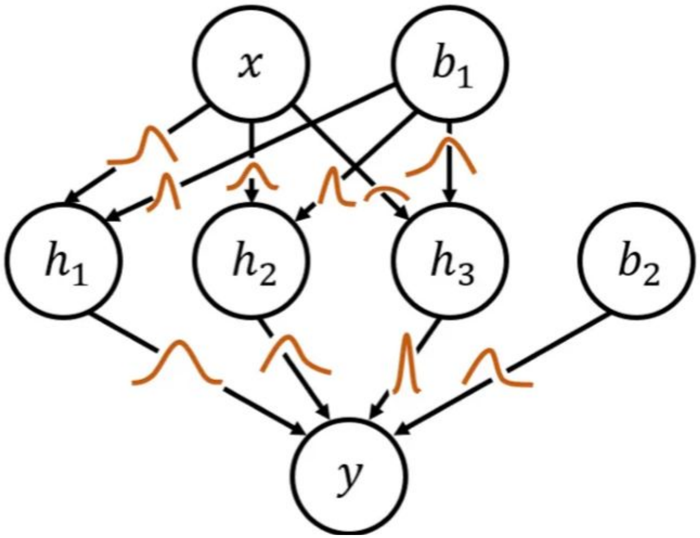
Bayesian Neural Network (BNN)



Epistemic Uncertainty



Bayesian Neural Network (BNN)



$$p(weights\,|data) = ?$$

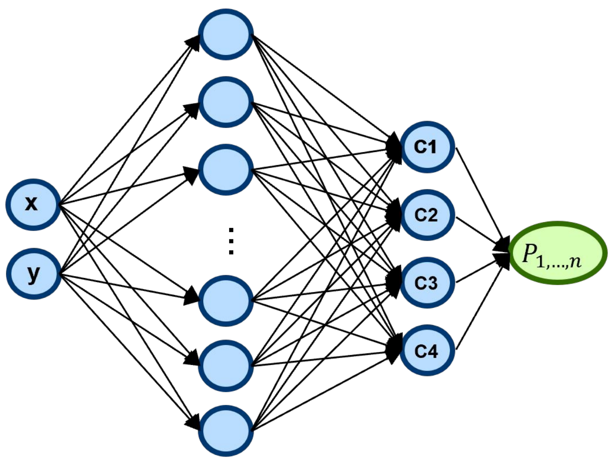
$$p(weights|data) = ?$$

$$\underbrace{p(weights\,|data)}_{\textbf{posterior}} \propto \underbrace{p(data\,|weights)}_{\textbf{likelihood}}\,.\;\underbrace{p(weights)}_{\textbf{prior}}$$

Epistemic Uncertainty

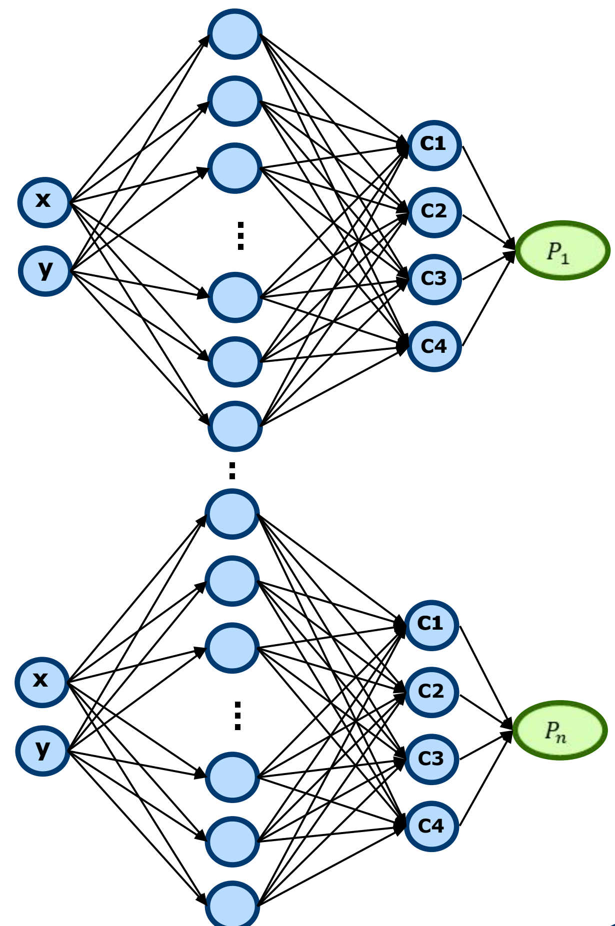
MC Dropout

n-times repetitions with Dropout



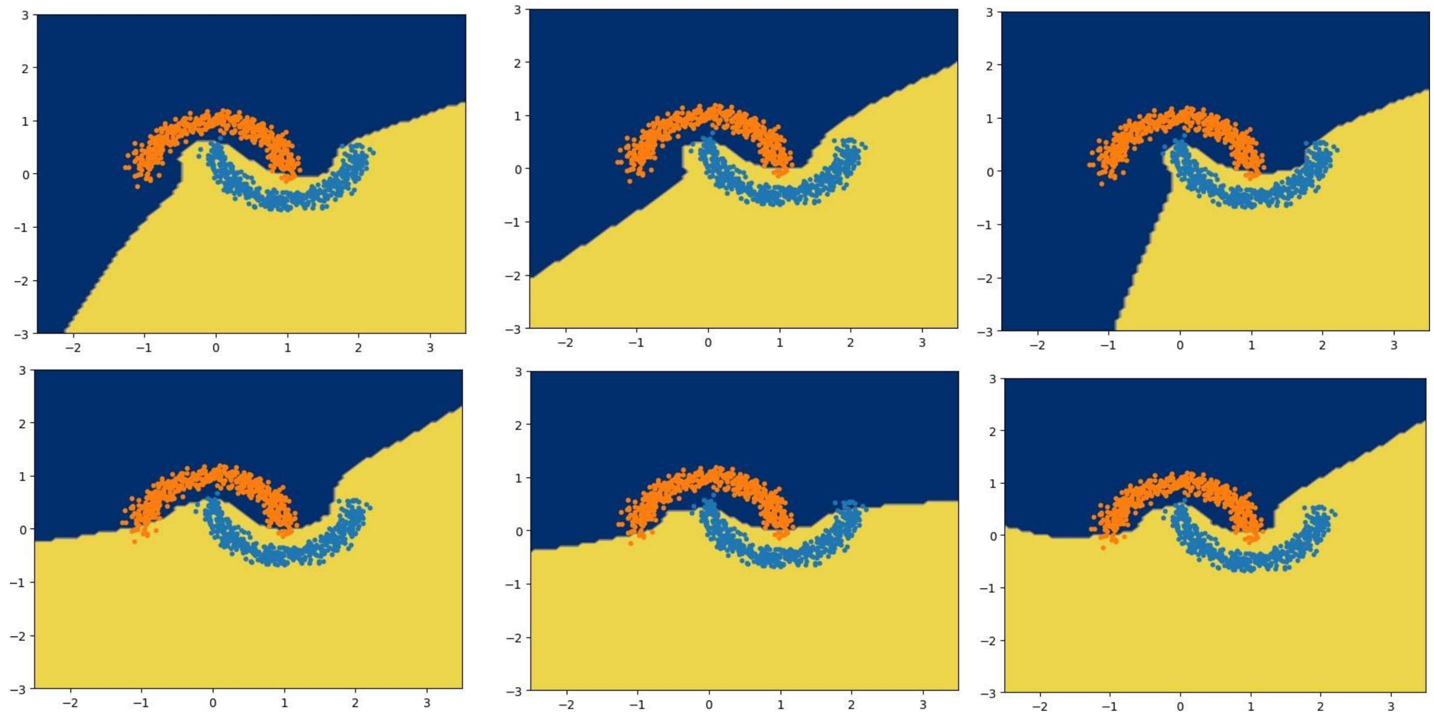
Deep Ensembles

n-times repetitions with n networks

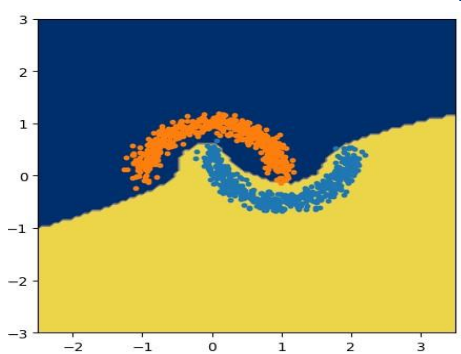


Epistemic Uncertainty

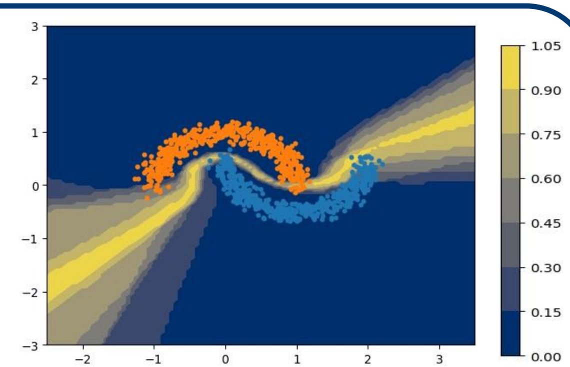
Multiple Predictions



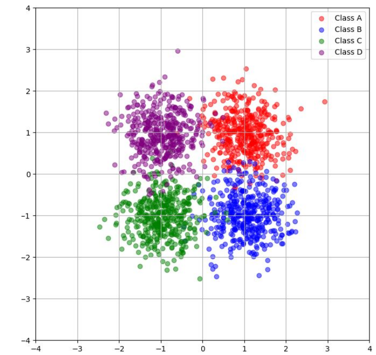
Final Prediction
Averaged Prediction



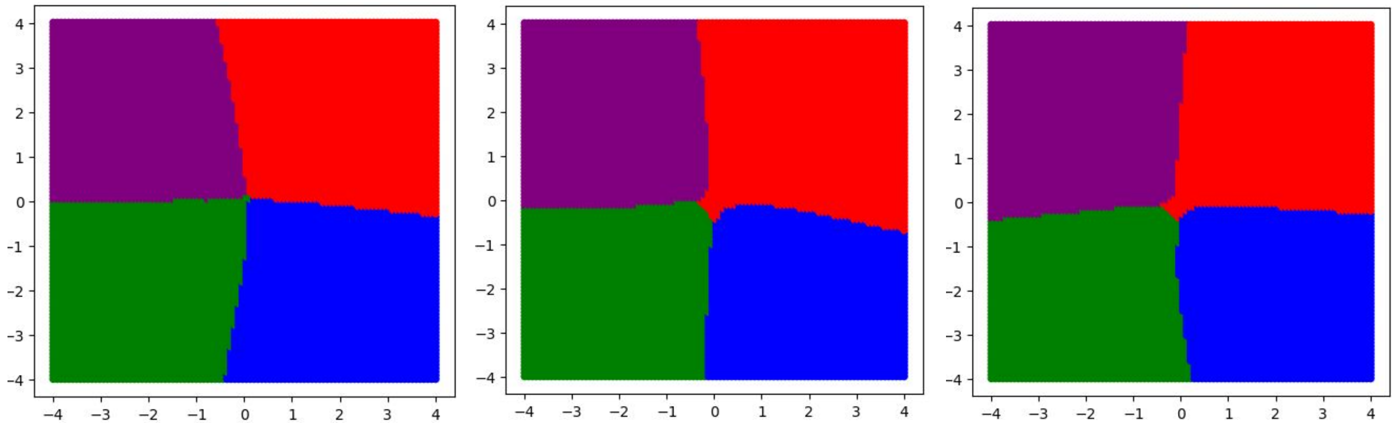
Epistemic Uncertainty
Entropy of Final prediction



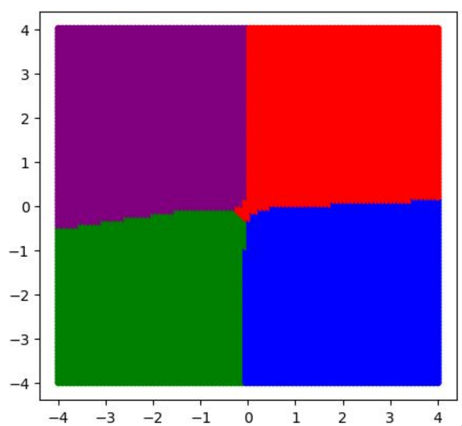
Epistemic Uncertainty



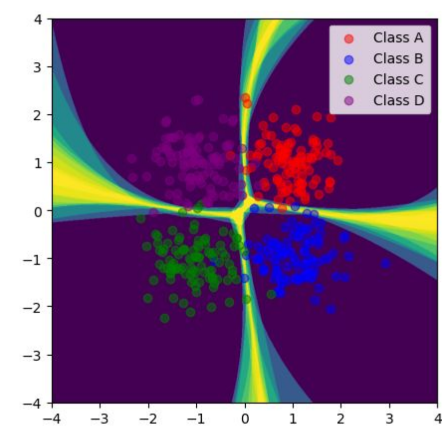
Multiple Predictions



Final Prediction
Averaged Prediction

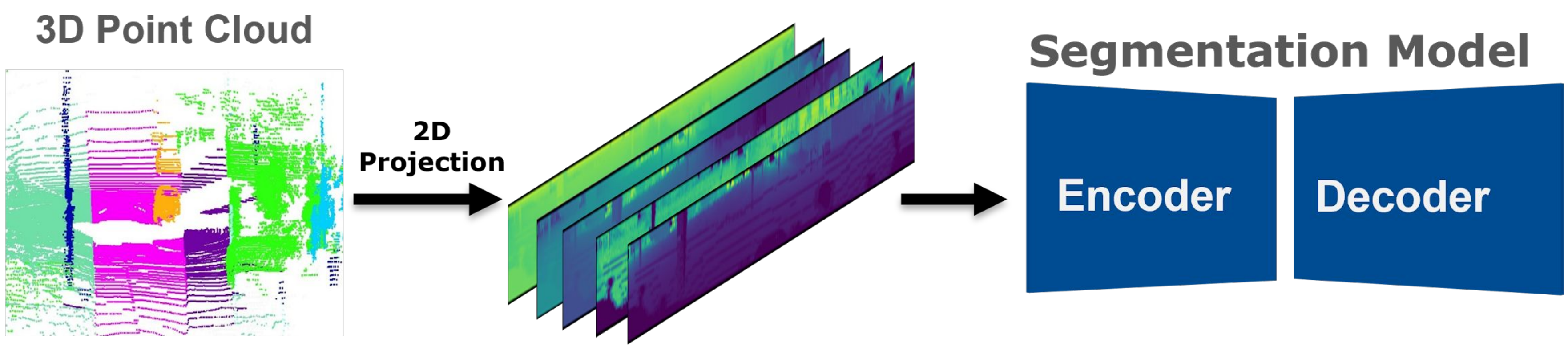


Epistemic Uncertainty
Entropy of Final prediction



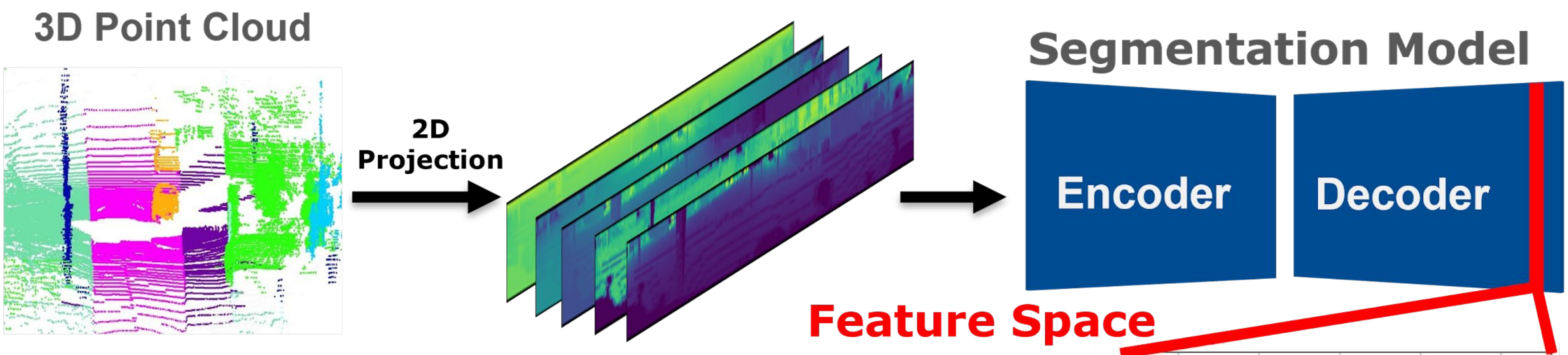
GMM:Generative Classifier

- GMM-GIG



GMM-GIG: Bayesian Generative Classifier

- GMM-GIG class-conditional densities

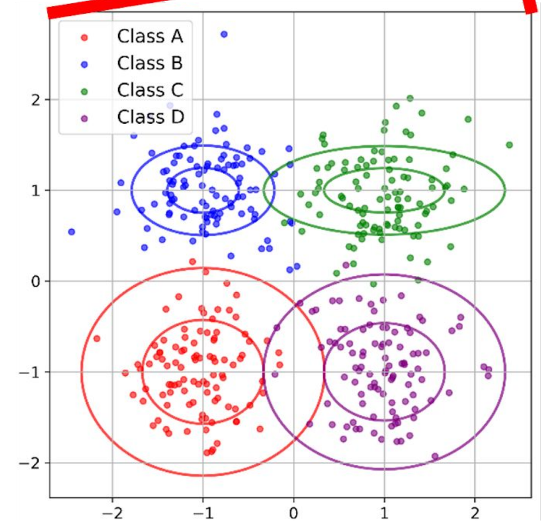


discriminative classifiers: p(class | features)

Generative Classifiers: p(features, class)
class-conditional densities

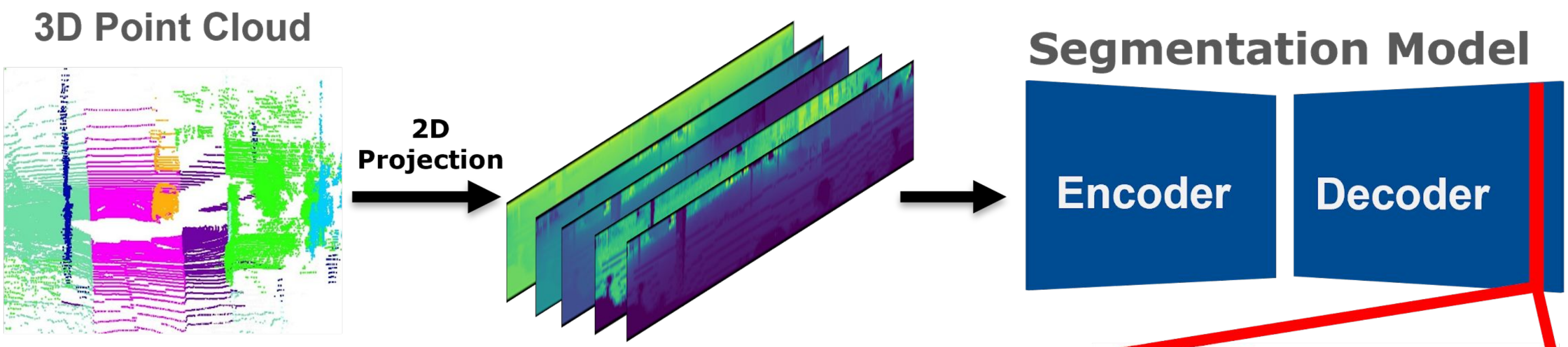
$$p(c \mid \mathbf{z}) = \frac{p(\mathbf{z} \mid c)*}{\sum_{c'} p(\mathbf{z} \mid c')}$$

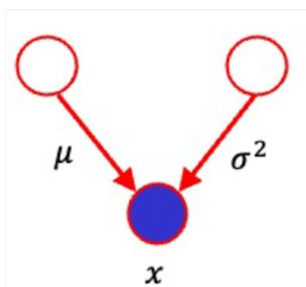
$$* \; p(\mathbf{z} \mid c) = \sum_{k=1}^{K} \pi_k^{(c)} \mathcal{N}(\mathbf{z} \mid \boldsymbol{\mu}_k^{(c)}, \boldsymbol{\Sigma}_k^{(c)})$$

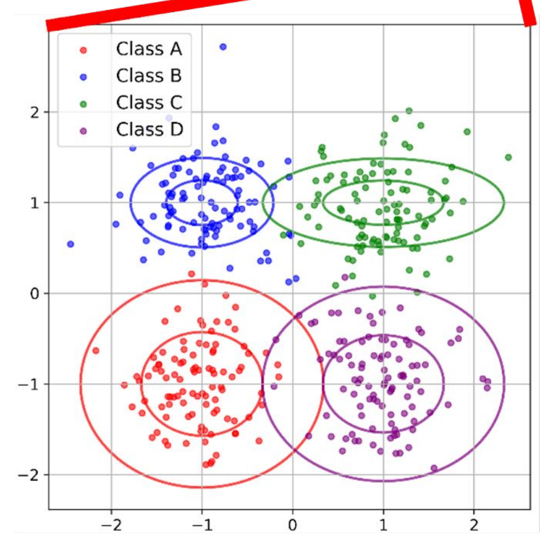


GMM-GIG in LiDAR Semantic Segmentation

- GMM-GIG

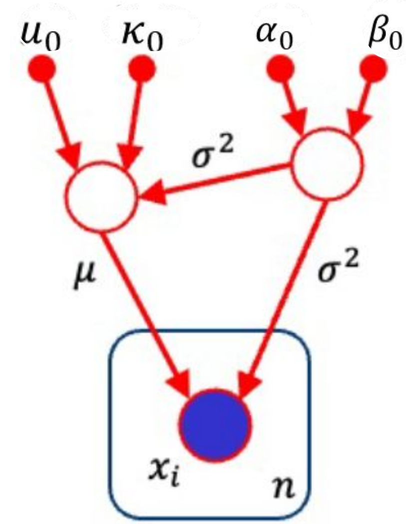






$$\underbrace{p(\mu,\sigma^2|data)}_{\text{posterior}} \propto \underbrace{p(data\,|\mu,\sigma^2)}_{\text{likelihood}} .\ \underbrace{p(\mu,\sigma^2)}_{\text{prior}}$$

GMM-GIG Epistemic Uncertainty

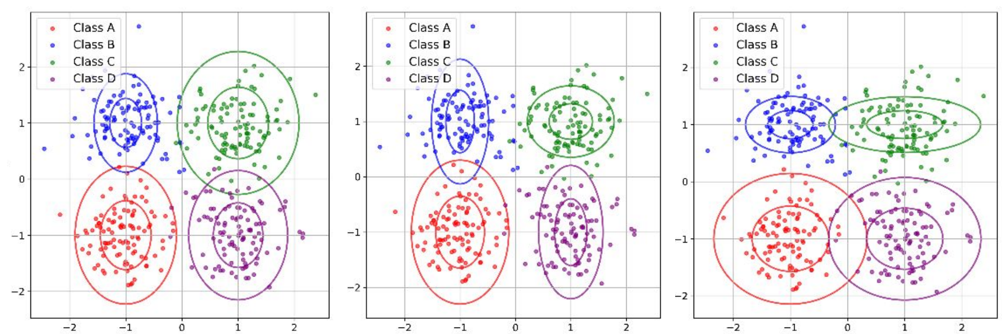


- **Prior**: Normal-Inverse Gamma: $\mu_0, \kappa_0, \alpha_0, \beta_0$

$$p(\mu, \sigma^2) = \mathcal{N}\left(\mu \mid \mu_0, \frac{\sigma^2}{\kappa_0}\right). Inv - Gamma(\sigma^2 \mid \alpha_0, \beta_0)$$

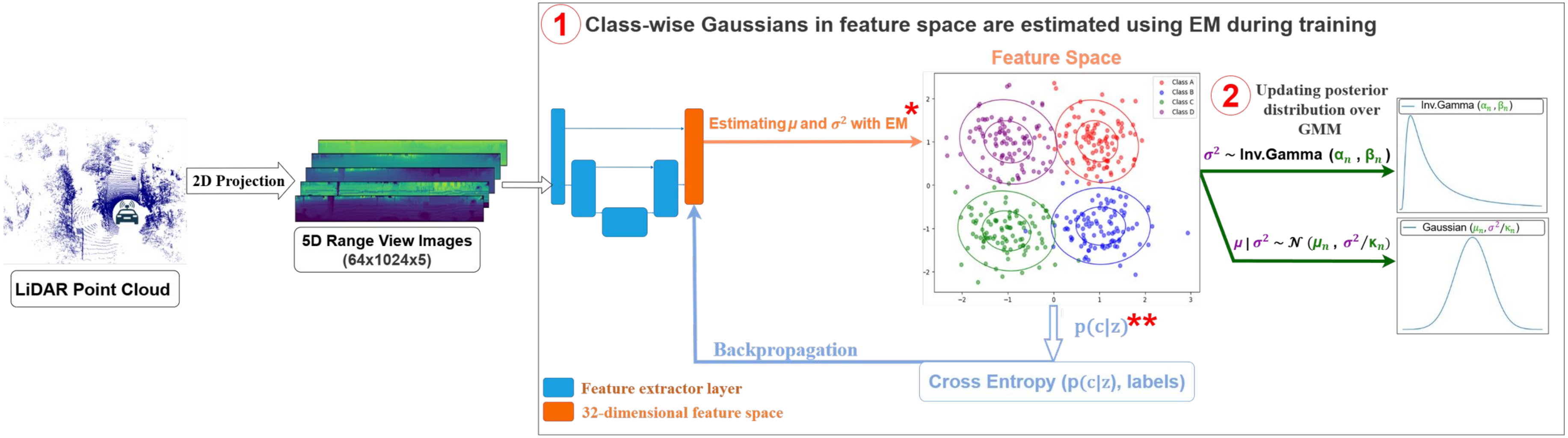
- Observe n samples x_i :
  - **Posterior** is updated after the model is trained: μ and σ^2 are defined
- But μ and σ^2 are not just one value (point estimation):
  - We have the distribution over the μ and σ^2
  - $\sigma^2 \sim Inv - Gamma(\sigma^2 \mid \alpha_0, \beta_0)$
  - $\mu \sim \mathcal{N}\left(\mu \mid \mu_0, \frac{\sigma^2}{\kappa_0}\right)$
- Sampling μ and σ^2:

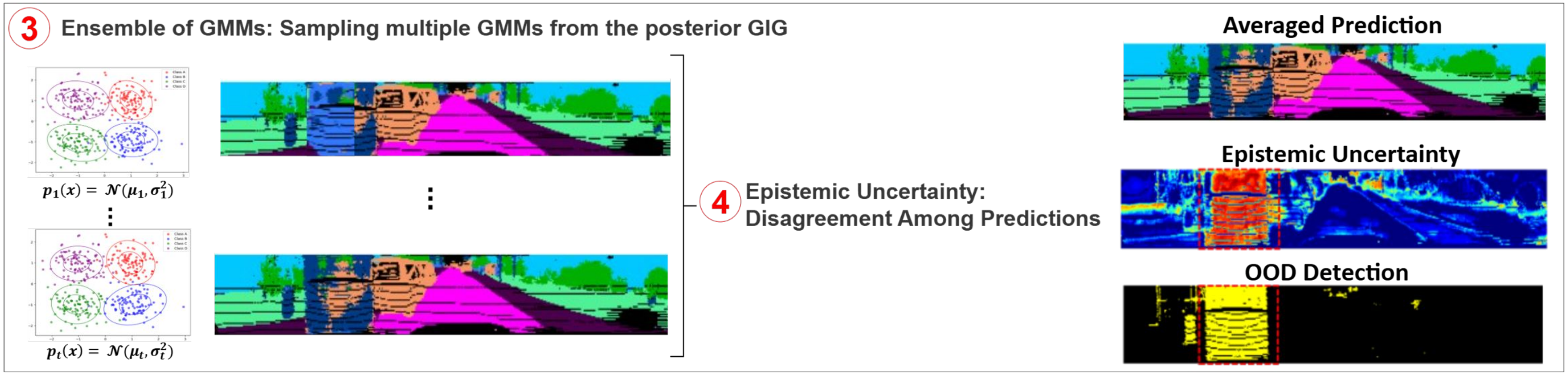
Ensemble of GMMs

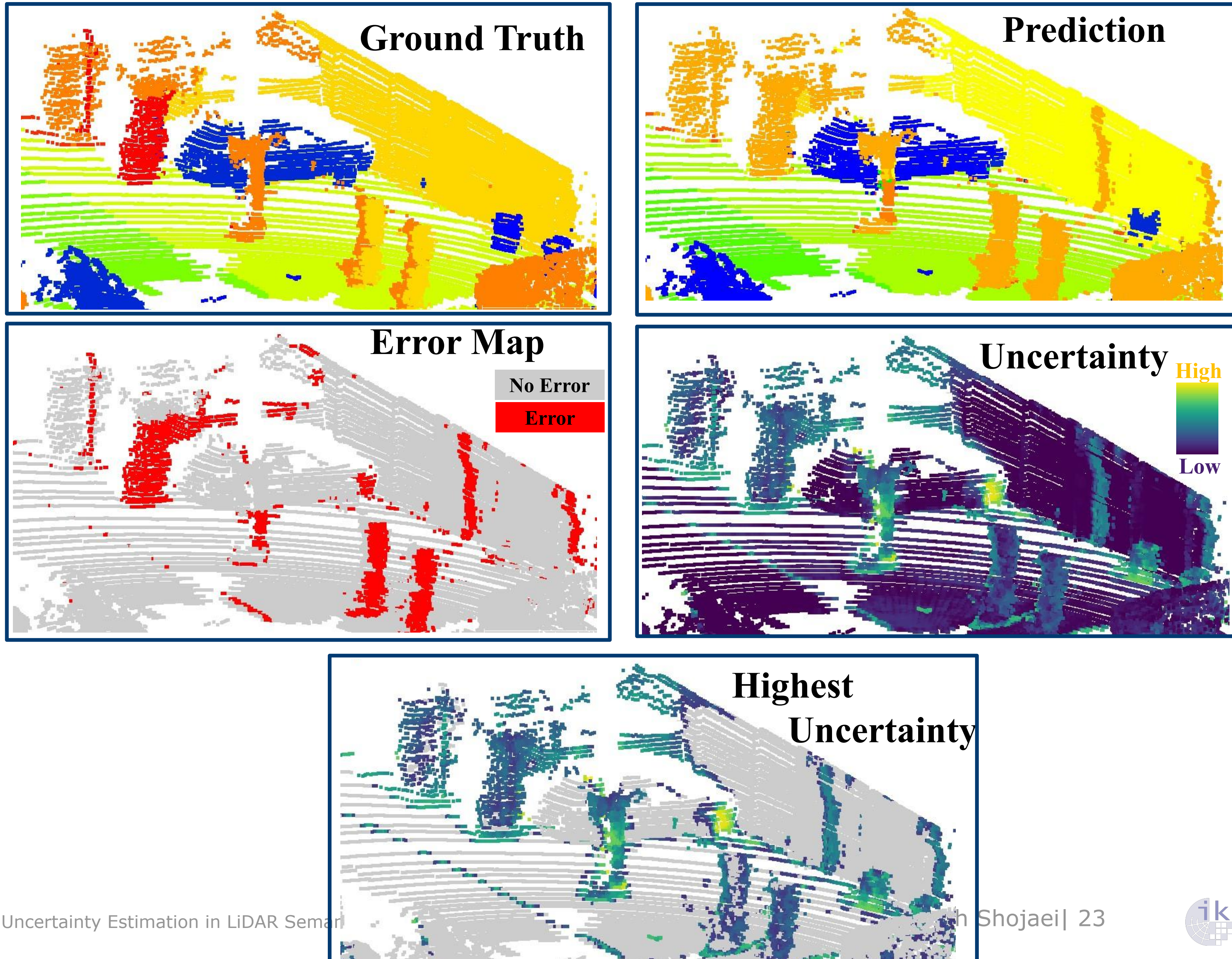


GMM-GIG in LiDAR Semantic Segmentation

- GMM training via Expectation Maximization to find GMM parameters
- Modeling the distribution over GMM parameters by GIG

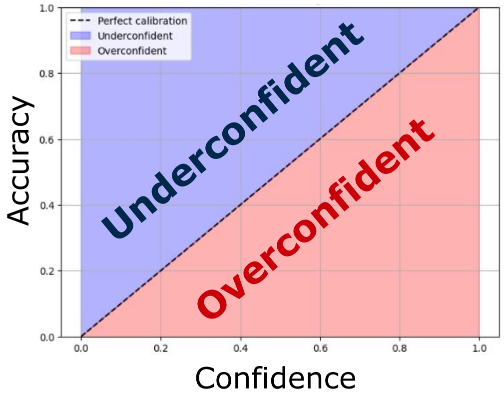


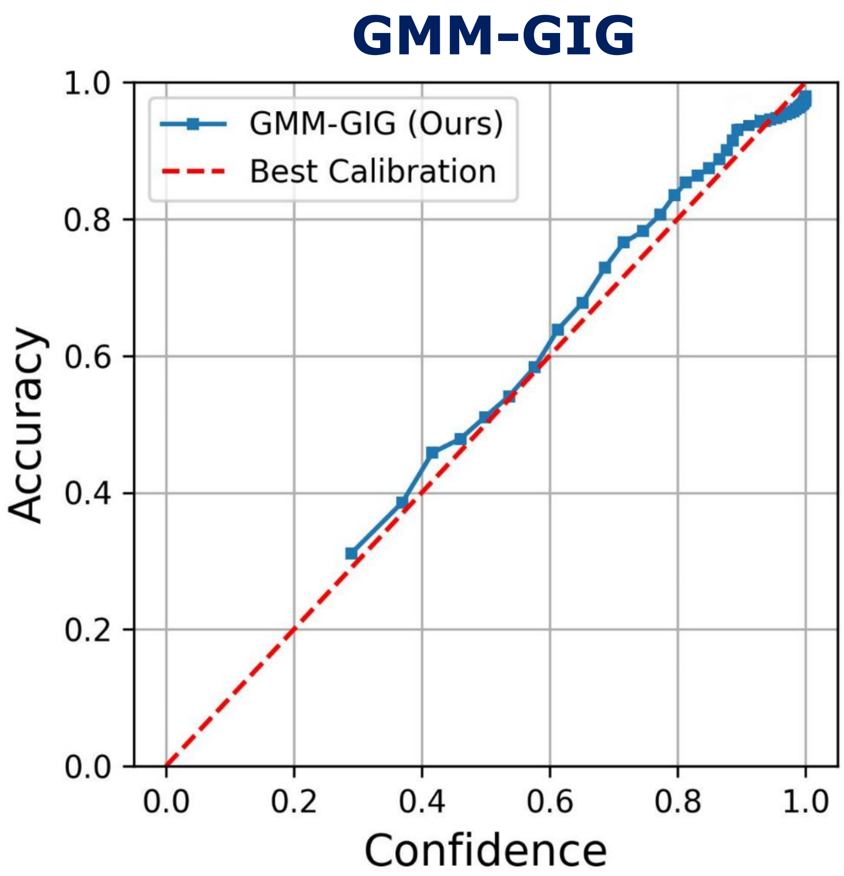


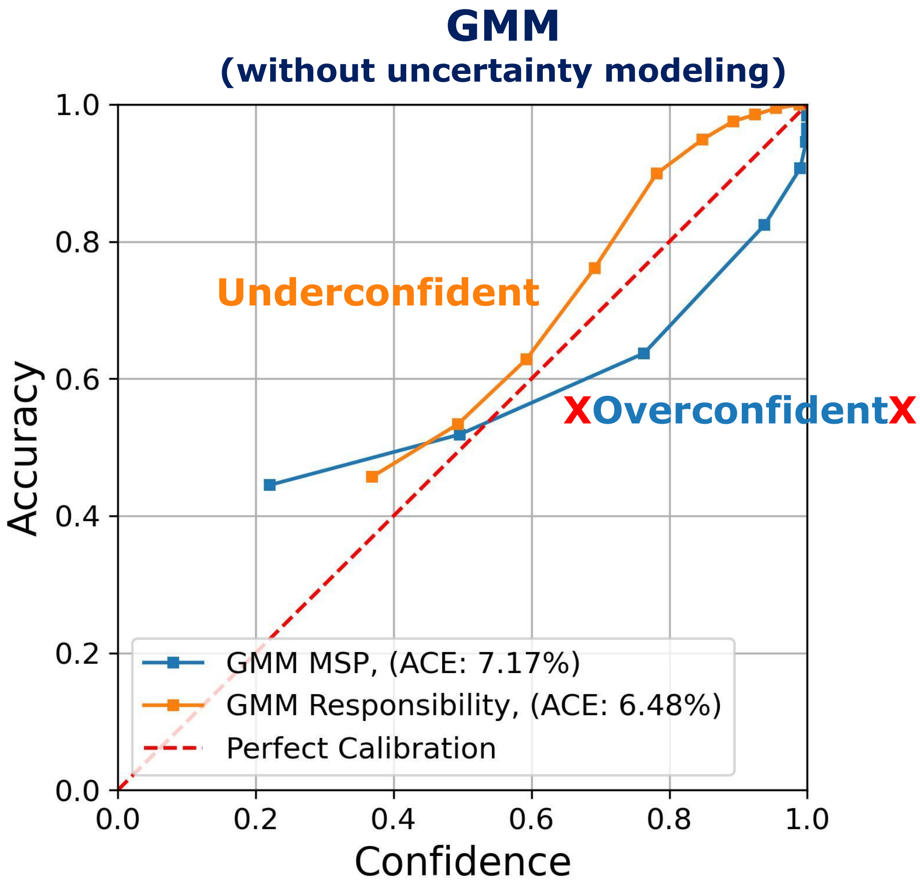

Ground Truth
Prediction
Error Map
No Error
Error
Uncertainty
High
Low
Highest
Uncertainty

Experimental Results for GMM-GIG

- Model Calibration (ACE)
  - Confidence Histogram (Reliability Diagram)







Experimental Results for GMM-GIG

- Improved Classification Accuracy (mIoU ↑)
- Improved Calibration Error (ACE ↓)
- Improved Failure Prediction (AUROC ↑, AUPR ↑)

Backbone	Method	Segmentation	Calibration	Failure Prediction	
		mIoU (%) ↑	ACE (%) ↓	AUROC (%) ↑	AUPR (%) ↑
RangeViT	MSP	57.58	5.42	81.07	35.01
	+ DE	61.88	4.30	87.90	41.03
	+ MC Dropout	56.58	3.88	87.21	40.31
	+ EDL	57.33	5.83	80.02	36.03
	+ GMM	61.01	6.02	81.34	41.67
	+ GMM-GIG (Ours)	**62.88**	**2.68**	**88.56**	**45.36**
SalsaNext	MSP	55.21	6.03	81.60	38.70
	+ DE	57.83	3.33	81.77	41.78
	+ MC Dropout	55.10	4.43	84.75	40.90
	+ EDL	55.63	5.32	79.91	35.03
	+ GMM	57.01	6.67	86.11	53.00
	+ GMM-GIG (Ours)	**57.96**	**2.55**	**86.32**	**53.68**

Experiments on SemanticKITTI dataset

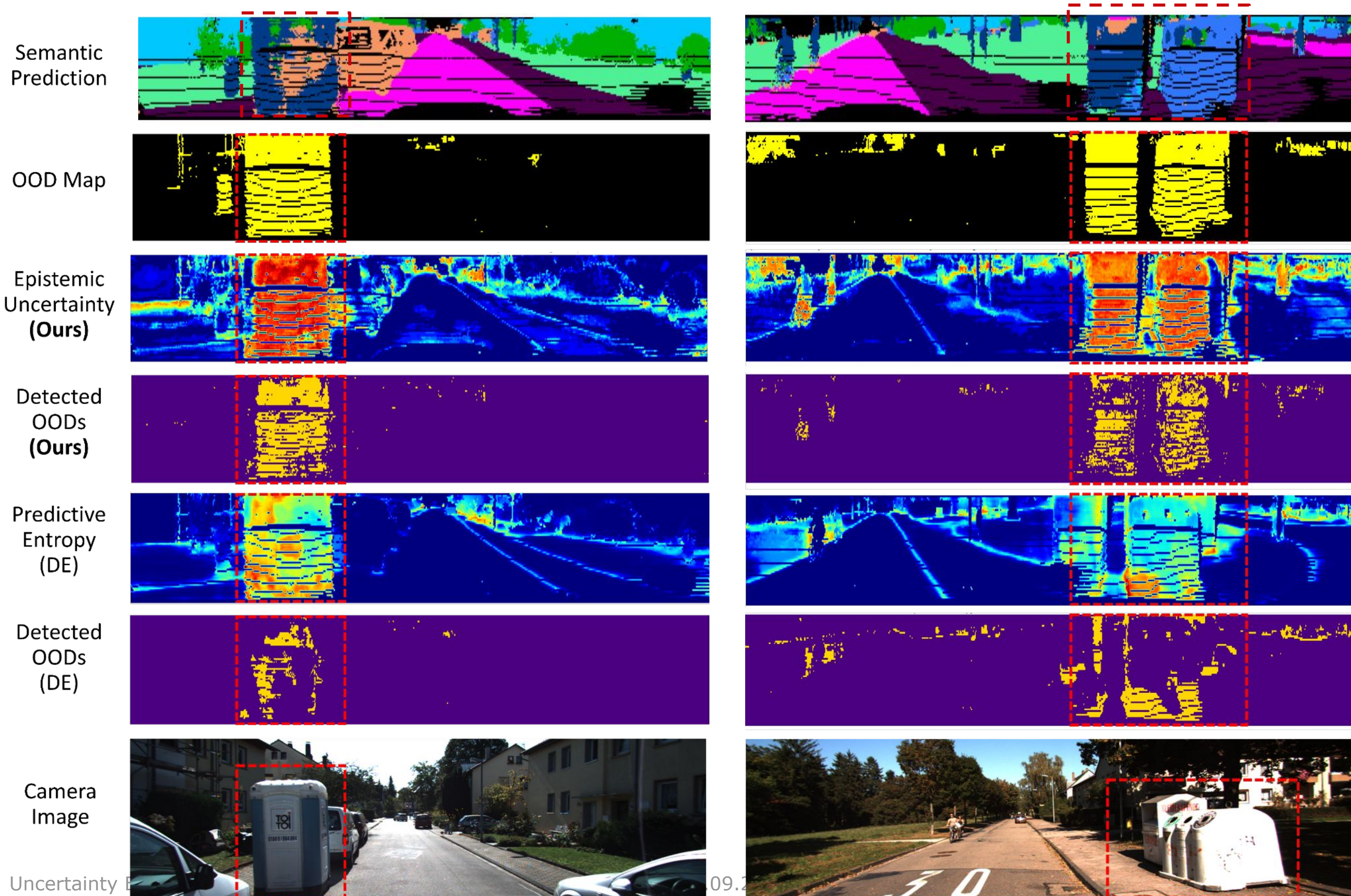
Experimental Results for GMM-GIG

- Improved Out-of-Distribution Detection*

Method	Uncertainty Type	↑AUROC	↑AUPRC(AP)	↓FPR95	↑mIoU
MSP	Deterministic Entropy	70.41	10.90	76.00	56.37
ODIN	Deterministic Entropy	73.74	12.45	75.54	56.37
MCDropout	Predictive Entropy	73.64	13.65	75.92	57.15
DE	Predictive Entropy	73.03	16.14	76.48	57.17
GMMSeg	Deterministic Entropy	87.62	26.14	48.84	57.60
Ours	Epistemic Uncertainty	**91.06**	**37.67**	**40.14**	**57.96**

* **Hanieh Shojaei, Claus Brenner. "Out-of-Distribution Detection in LiDAR Semantic Segmentation Using Epistemic Uncertainty from Hierarchical GMMs". GCPR 2025**

Experimental Results for GMM-GIG

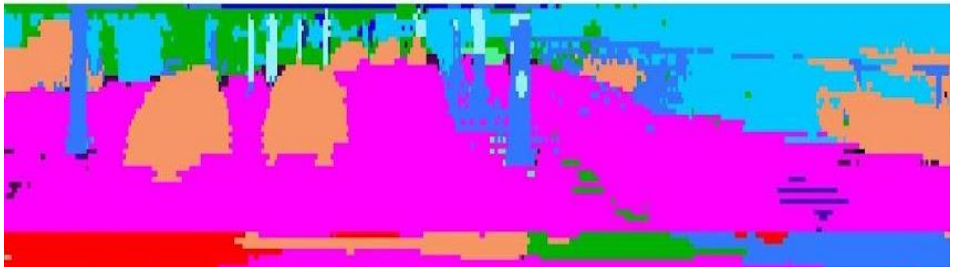


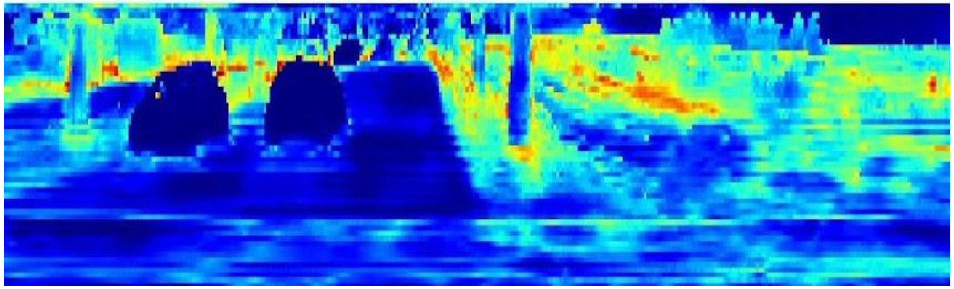
Conclusion

- GMM-GIG : Ensemble of GMMs, instead of one deterministic GMM
  - Epistemic uncertainty estimation
  - Better calibrated predictions
  - More accurate classification
- Applications of epistemic uncertainty
  - Failure detection
  - OOD detection

Future plans

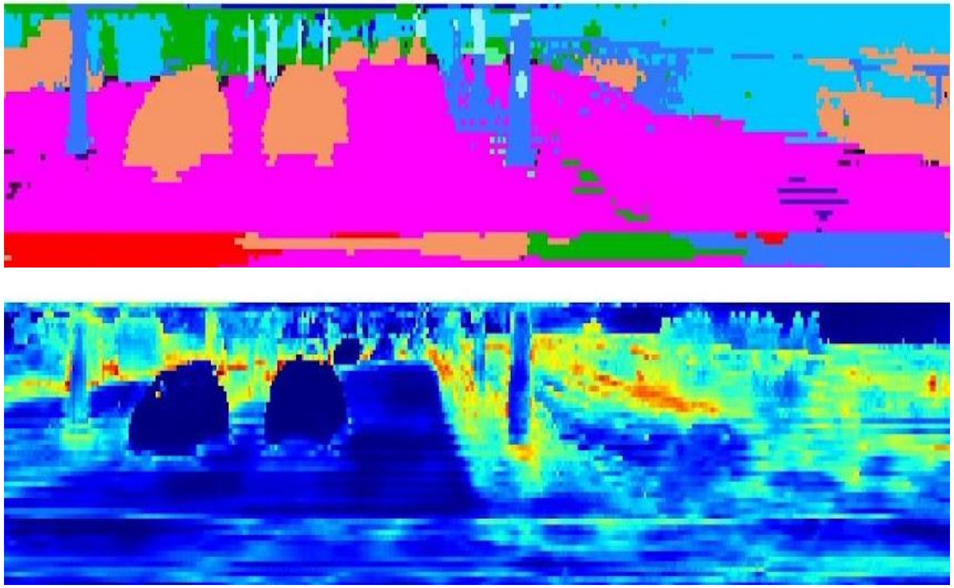
- Label prediction for our LiDARs: Hesai64 and Ouster128
  - without annotated data
    - Performance is poor

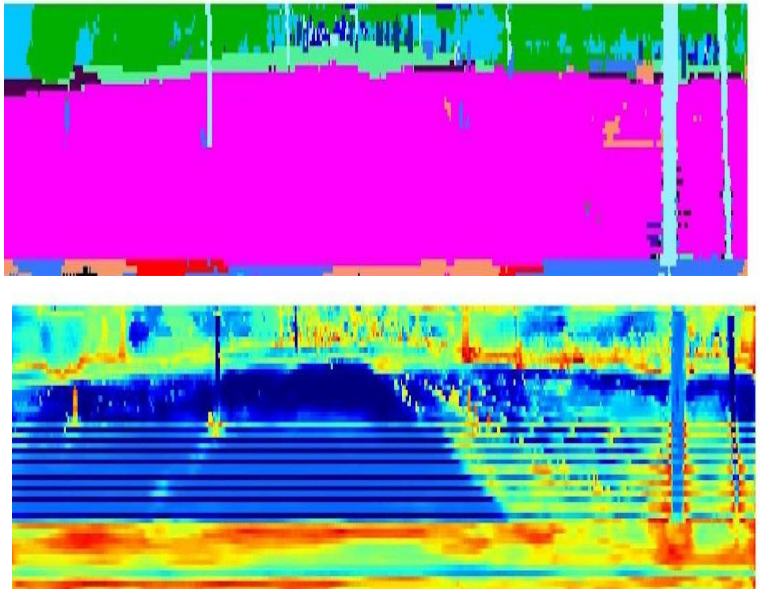




Future plans

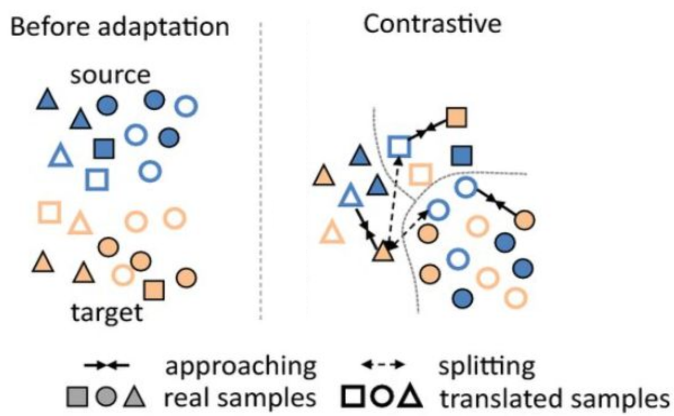
- Label prediction for our LiDARs: Hesai64 and Ouster128
  - without annotated data
    - Performance is poor





Future plans

- Label prediction for our LiDARs: Hesai64 and Ouster128
  - without annotated data
    - Performance is poor:
      - **Test-time entropy minimization**
      - **Unsupervised contrastive adaptation**



  - With only a few annotated scans
    - Fine-tune GMM parameters using few-shot updates

Thank you for your attention!

Hanieh Shojaei

hanieh.shojaei@ikg.uni-hannover.de