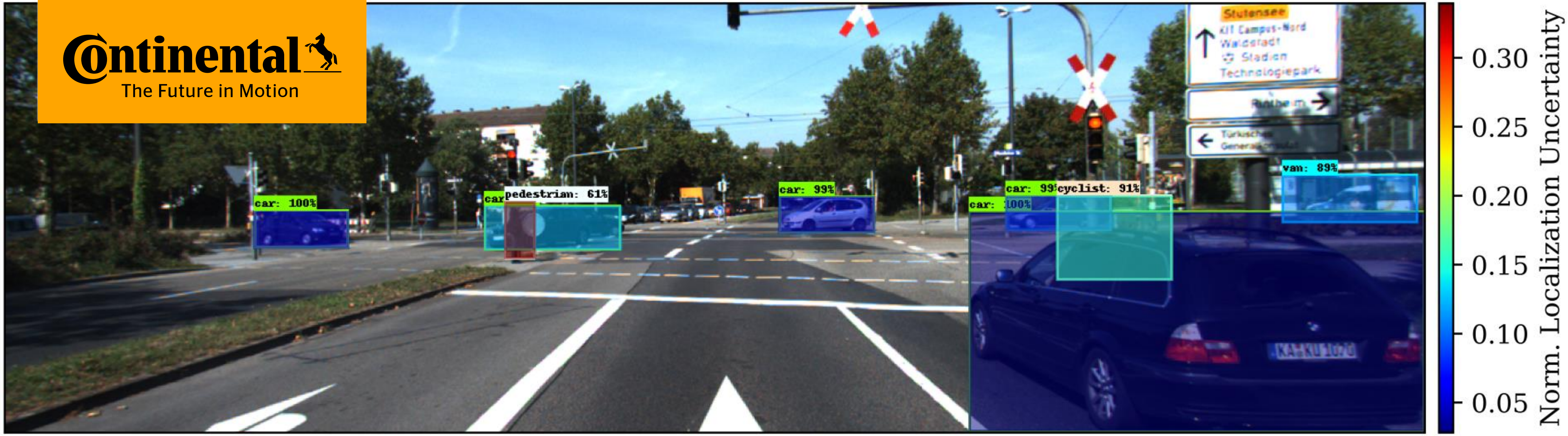
Continental
The Future in Motion



Overcoming the Limitations of Localization Uncertainty

Efficient & Exact Non-Linear Post-Processing and Calibration

Moussa Kassem Sbeyti, Michelle Karg, Christian Wirth, Azarm Nowzad and Sahin Albayrak

TECHNISCHE UNIVERSITÄT BERLIN

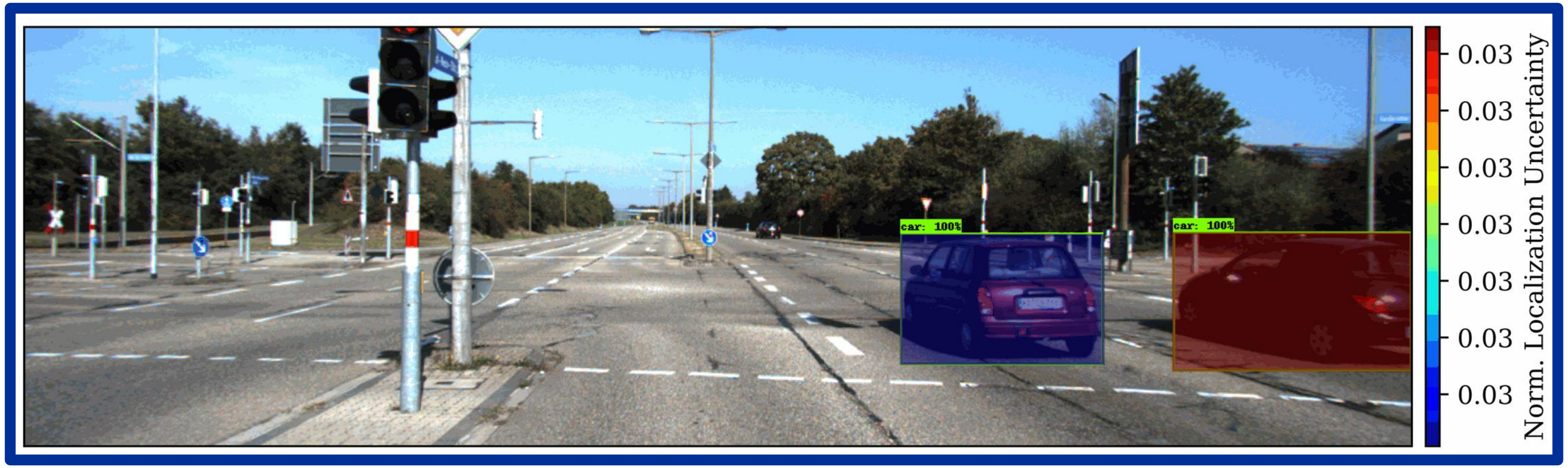
DAI-Labor
Distributed Artificial Intelligence Laboratory

ECML PKDD 2023

Results
Visualization – KITTI

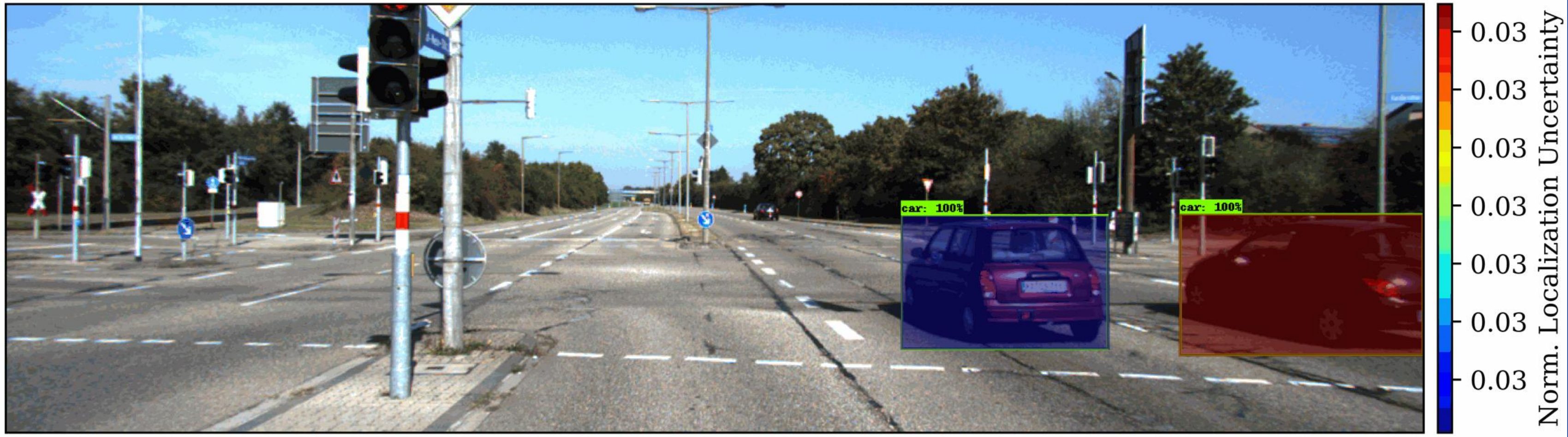
Top 10 frames with lowest and highest uncertainty out of 100 frames

Low uncertainty for close and clear objects

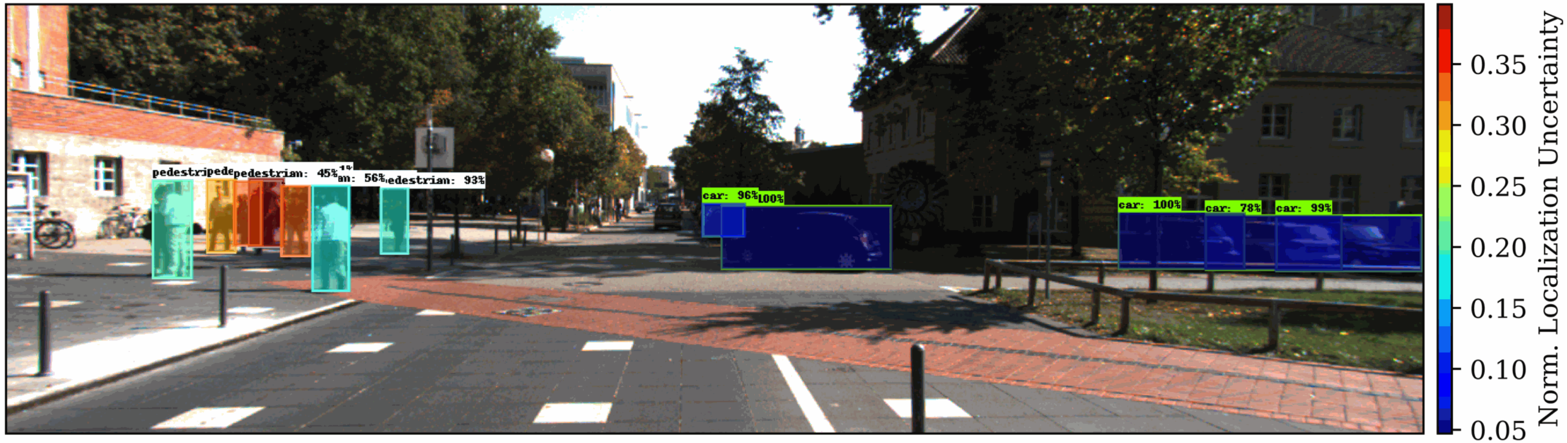


Results
Visualization – KITTI

Top 10 frames with lowest and highest uncertainty out of 100 frames



Low uncertainty for close and clear objects



High uncertainty for far, occluded and poorly detected objects

Results
Visualization – BDD100K

Top 10 frames with lowest and highest uncertainty out of 100 frames

Low uncertainty for close and clear objects

Results

Visualization – BDD100K

Top 10 frames with lowest and highest uncertainty out of 100 frames

Low uncertainty for close and clear objects

High uncertainty for far, occluded and poorly detected objects

Loss Attenuation

Aleatoric Uncertainty

Observation noise,
weather conditions,
misleading situations

Input

Loss Attenuation

Aleatoric Uncertainty

Observation noise, weather conditions, misleading situations

Input → Network

Rewrite box loss with loss attenuation

Negative log-likelihood $\mathcal{L}_{\mathrm{NN}} = \frac{1}{2N} \sum_{i=1}^{N} \frac{\|\mathrm{y}_i^* - \mathrm{f}(\mathrm{x}_i)\|^2}{\sigma(\mathrm{x}_i)^2} + \log \sigma(\mathrm{x}_i)^2$

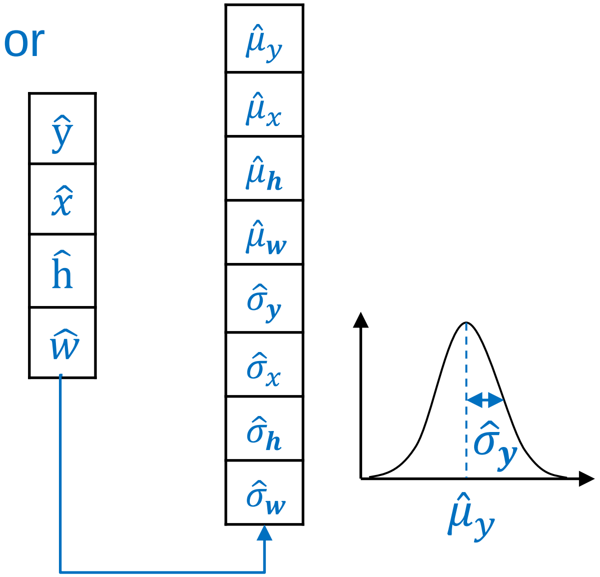
Loss Attenuation

Aleatoric Uncertainty

Observation noise, weather conditions, misleading situations

Extend output to 8 values per anchor

Input → Network → Output



Rewrite box loss with loss attenuation

Negative log-likelihood $\mathcal{L}_{\mathrm{NN}} = \frac{1}{2N}\sum_{i=1}^{N} \frac{\|\mathrm{y}_i^* - \mathrm{f}(\mathrm{x}_i)\|^2}{\sigma(\mathrm{x}_i)^2} + \log \sigma(\mathrm{x}_i)^2$

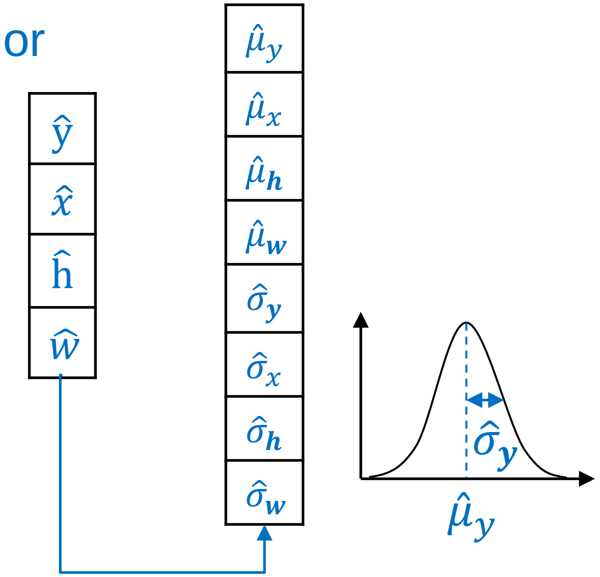
Loss Attenuation

Aleatoric Uncertainty

Observation noise, weather conditions, misleading situations

Input → Network → Output

Extend output to 8 values per anchor



Rewrite box loss with loss attenuation

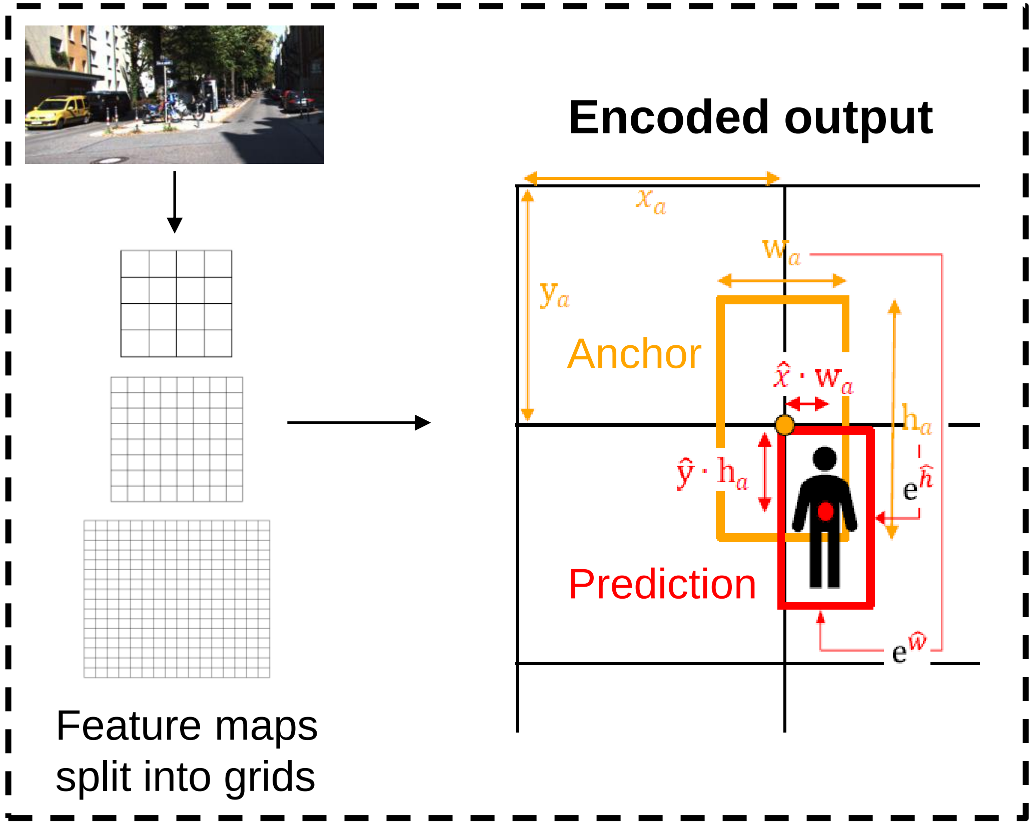
Negative log-likelihood $\mathcal{L}_{\mathrm{NN}} = \frac{1}{2N}\sum_{i=1}^{N}\frac{\|\mathrm{y}_i^* - \mathrm{f}(\mathrm{x}_i)\|^2}{\sigma(\mathrm{x}_i)^2} + \log\sigma(\mathrm{x}_i)^2$

The uncertainty is learned as a function of the data

Use Case Definition

How is the output distribution $\mathcal{N}(\mu, \sigma^2)$ propagated through non-linear functions?

Network Output

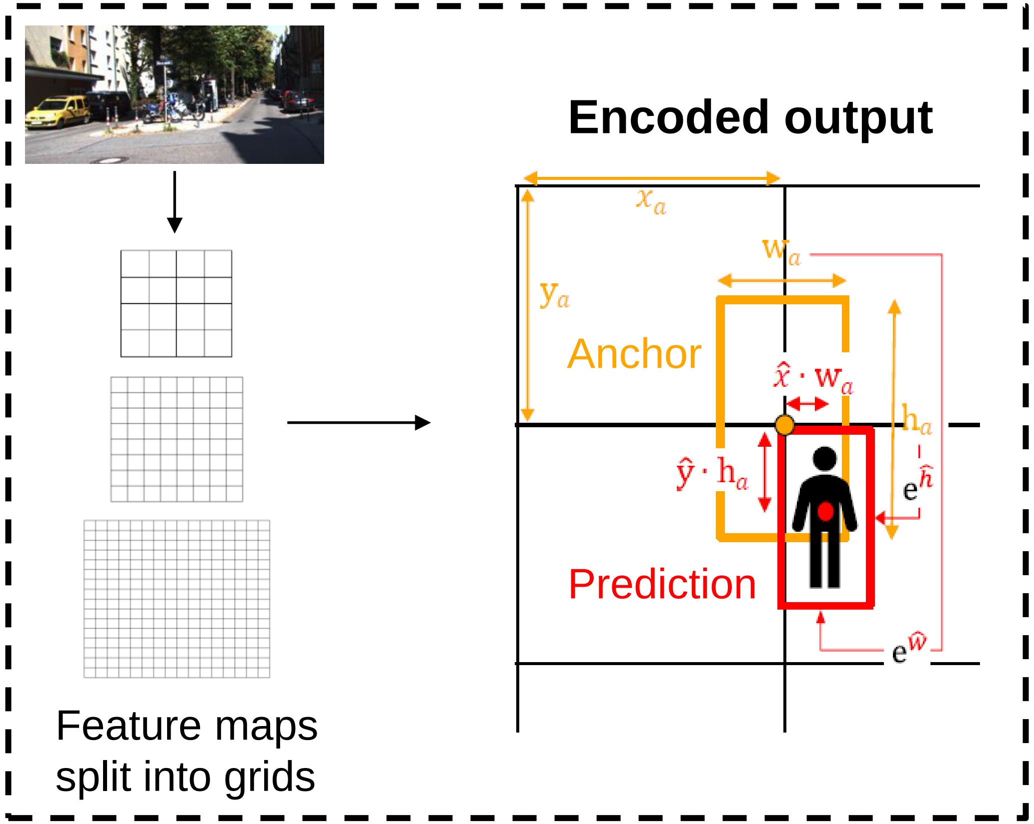


Anchor-relative center coordinates $(\hat{x}, \hat{y})$, width $\hat{w}$ and height $\hat{h}$.
Anchor center coordinates (x_a, y_a), width w_a and height h_a.

Use Case Definition

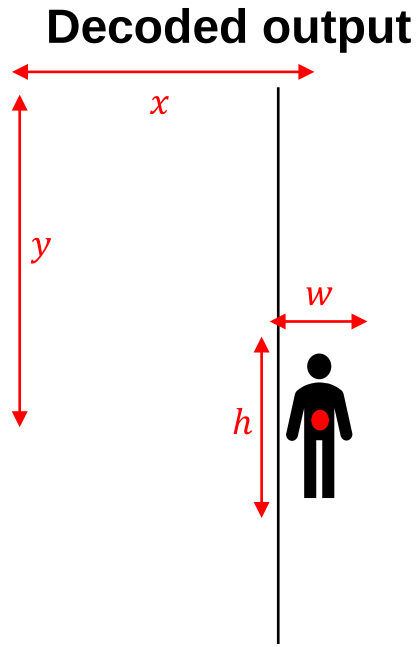
How is the output distribution $\mathcal{N}(\mu, \sigma^2)$ propagated through non-linear functions?

Network Output



Post-processing

$$x = \hat{x} \cdot w_a + x_a$$
$$y = \hat{y} \cdot h_a + y_a$$
$$w = e^{\hat{w}} \cdot w_a$$
$$h = e^{\hat{h}} \cdot h_a$$

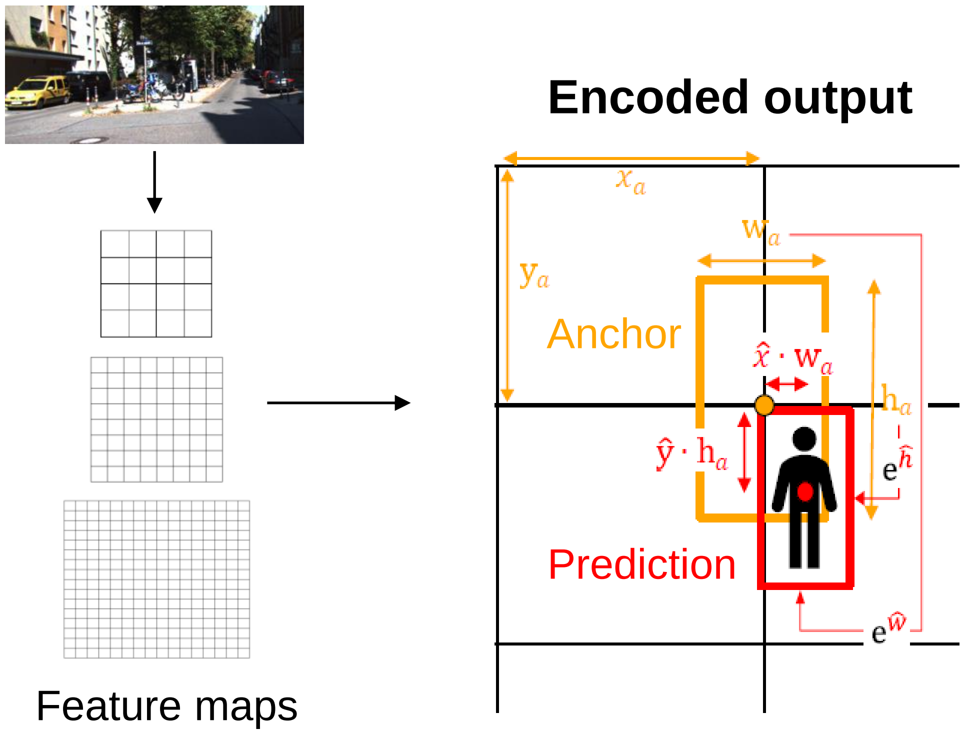


Anchor-relative center coordinates $(\hat{x}, \hat{y})$, width $\hat{w}$ and height $\hat{h}$.
Anchor center coordinates (x_a, y_a), width w_a and height h_a.

Use Case Definition

How is the output distribution $\mathcal{N}(\mu, \sigma^2)$ propagated through non-linear functions?

Network Output



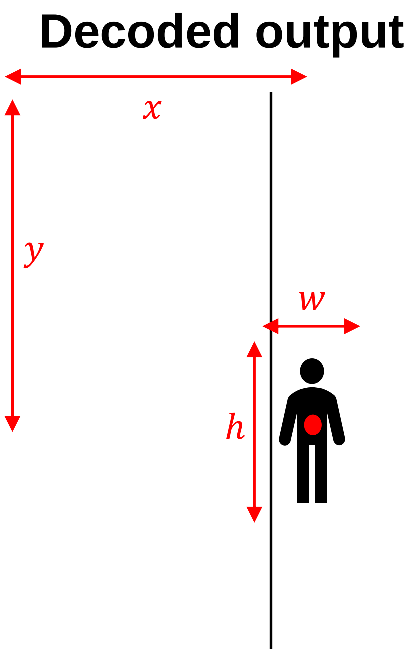
Feature maps
split into grids

Post-processing

Non-linear transformations

$$x = \hat{x} \cdot w_a + x_a$$
$$y = \hat{y} \cdot h_a + y_a$$
$$w = \boxed{e^{\hat{w}}} \cdot w_a$$
$$h = \boxed{e^{\hat{h}}} \cdot h_a$$

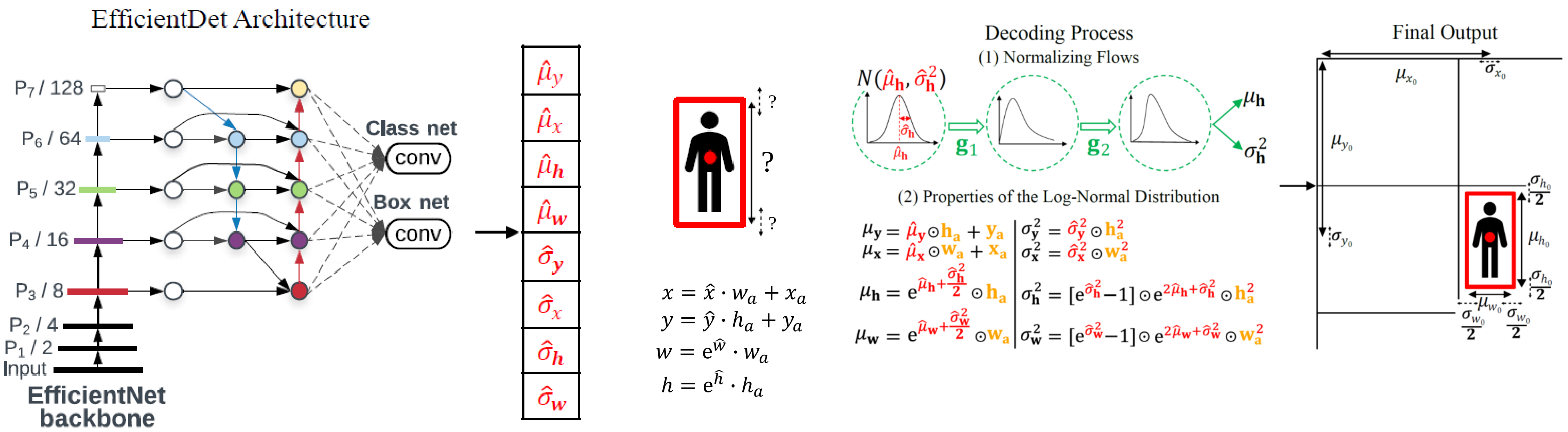
?



Anchor-relative center coordinates $(\hat{x}, \hat{y})$, width $\hat{w}$ and height $\hat{h}$.
Anchor center coordinates (x_a, y_a), width w_a and height h_a.

Propagation Methods
Illustration

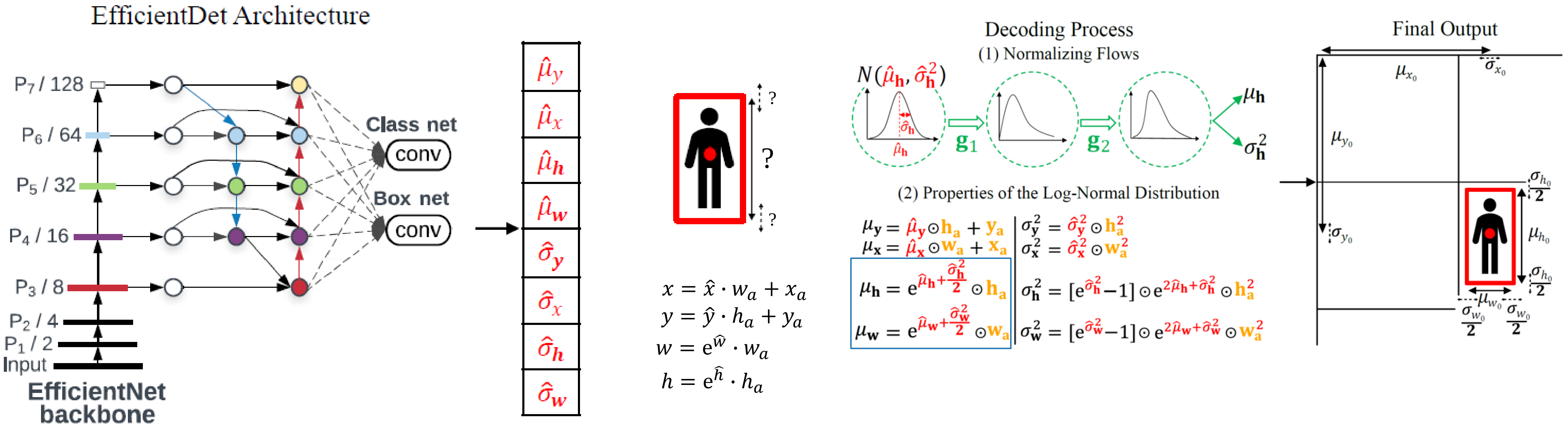
Loss attenuation in EfficientDet $\mathcal{L}_{\mathrm{NN}} = \frac{1}{2 \cdot 4N_{pos}} \sum_{i=1}^{N} \sum_{j=1}^{4} \left(\frac{\| \mathbf{y}_{ij}^{*} - \hat{\mu}_{\mathbf{j}}(\mathbf{x}_i) \|^2}{\hat{\sigma}_{\mathbf{j}}(\mathbf{x}_i)^2} + \log \hat{\sigma}_{\mathbf{j}}(\mathbf{x}_i)^2 \right) \cdot \mathbf{m}(\mathbf{y}_{ij}^{*})$



Propagation Methods

Illustration

Loss attenuation in EfficientDet $\mathcal{L}_{\mathrm{NN}} = \frac{1}{2 \cdot 4N_{pos}} \sum_{i=1}^{N} \sum_{j=1}^{4} \left(\frac{\| \mathbf{y}_{ij}^{*} - \hat{\mu}_{\mathbf{j}}(\mathbf{x}_i) \|^2}{\hat{\sigma}_{\mathbf{j}}(\mathbf{x}_i)^2} + \log \hat{\sigma}_{\mathbf{j}}(\mathbf{x}_i)^2 \right) \cdot \mathbf{m}(\mathbf{y}_{ij}^{*})$ with $\| \mathbf{y}_{ij}^{*} - [\hat{\mu}_{\mathbf{j}}(\mathbf{x}_i) + \frac{\hat{\sigma}_{\mathbf{j}}(\mathbf{x}_i)^2}{2}] \|^2$ for $j = 3,4$



Results

Propagation Methods

- **EfficientDet-D0** pre-trained on **COCO**
- Input resolution: **1024x512**
- Soft-NMS output is **reordered** based on **lowest MSE**
- **BDD100K:** 10 classes
- **KITTI:** 7 classes

$$\mu_w = e^{\hat{\mu}_w + \cancel{\frac{\hat{\sigma}_w^2}{2}}} \odot w_a$$

KITTI

Method	AP↑	RMSE↓	mIoU↑	NLL↓	ET↓ (s)	IT↓ (ms)
Baseline	72.8 ± 0.1	**5.07 ± 0.1**	90.1 ± 0.1	-	**115.6 ± 3**	34.8 ± 4
FalseDec	73.1 ± 0.5	5.27 ± 0.1	90.3 ± 0.1	4.27 ± 0.1	116.0 ± 3	31.1 ± 3
L-norm	**73.3** ± 0.5	5.17 ± 0.2	90.3 ± 0.0	3.22 ± 0.0	**115.6 ± 2**	**31.0 ± 3**
N-flow	**73.3** ± 0.5	5.17 ± 0.2	90.3 ± 0.0	3.22 ± 0.0	116.6 ± 1	31.6 ± 3
Samp30	68.6 ± 0.4	5.43 ± 0.1	88.7 ± 0.1	**3.19 ± 0.0**	118.8 ± 2	34.5 ± 3
Samp100	71.8 ± 0.5	5.23 ± 0.1	90.1 ± 0.0	3.20 ± 0.0	117.4 ± 4	47.0 ± 3
Samp1000	73.1 ± 0.5	5.18 ± 0.2	**90.4 ± 0.0**	3.21 ± 0.0	117.9 ± 4	187.4 ± 4

$$\mu_w = e^{\hat{\mu}_w + \cancel{\frac{\hat{\sigma}_w^2}{2}}} \odot w_a$$

BDD

Method	AP↑	RMSE↓	mIoU↑	NLL↓	ET↓ (s)	IT↓ (ms)
Baseline	**24.7 ± 0.1**	8.96 ± 0.2	66.6 ± 1.6	-	115.7 ± 3	33.0 ± 4
FalseDec	23.9 ± 0.2	8.81 ± 0.2	67.3 ± 0.0	4.40 ± 0.1	115.9 ± 2	**30.4 ± 4**
L-norm	24.4 ± 0.1	**8.53 ± 0.2**	**67.7 ± 0.0**	**3.69 ± 0.0**	**115.3 ± 1**	30.6 ± 4
N-flow	24.4 ± 0.1	**8.53 ± 0.2**	**67.7 ± 0.0**	**3.69 ± 0.0**	116.4 ± 1	31.0 ± 3
Samp30	21.0 ± 0.1	9.02 ± 0.2	64.7 ± 0.0	3.70 ± 0.0	118.0 ± 3	33.6 ± 4
Samp100	23.2 ± 0.1	8.68 ± 0.2	66.7 ± 0.0	**3.69 ± 0.0**	117.0 ± 3	45.4 ± 4
Samp1000	24.2 ± 0.1	8.55 ± 0.2	67.6 ± 0.1	**3.69 ± 0.0**	118.4 ± 3	187.3 ± 4

AP: Average precision
RMSE: Root-mean-square error
mIoU: Average intersection over union
NLL: Negative log-likelihood
ET: Model exporting time in seconds
IT: Inference time in milliseconds

Results

Propagation Methods

- **EfficientDet-D0** pre-trained on **COCO**
- Input resolution: **1024x512**
- Soft-NMS output is **reordered** based on **lowest MSE**
- **BDD100K:** 10 classes
- **KITTI:** 7 classes

$$\mu_w = e^{\hat{\mu}_w + \cancel{\frac{\hat{\sigma}_w^2}{2}}} \odot w_a$$

Our methods

Method	AP↑	RMSE↓	mIoU↑	NLL↓	ET↓ (s)	IT↓ (ms)	
Baseline	72.8 ± 0.1	**5.07 ± 0.1**	90.1 ± 0.1	-	**115.6 ± 3**	34.8 ± 4	
FalseDec	73.1 ± 0.5	5.27 ± 0.1	90.3 ± 0.1	4.27 ± 0.1	116.0 ± 3	31.1 ± 3	
L-norm	**73.3** ± 0.5	5.17 ± 0.2	90.3 ± 0.0	3.22 ± 0.0	**115.6 ± 2**	**31.0 ± 3**	KITTI
N-flow	**73.3** ± 0.5	5.17 ± 0.2	90.3 ± 0.0	3.22 ± 0.0	116.6 ± 1	31.6 ± 3	
Samp30	68.6 ± 0.4	5.43 ± 0.1	88.7 ± 0.1	**3.19 ± 0.0**	118.8 ± 2	34.5 ± 3	
Samp100	71.8 ± 0.5	5.23 ± 0.1	90.1 ± 0.0	3.20 ± 0.0	117.4 ± 4	47.0 ± 3	
Samp1000	73.1 ± 0.5	5.18 ± 0.2	**90.4 ± 0.0**	3.21 ± 0.0	117.9 ± 4	187.4 ± 4	
Baseline	**24.7 ± 0.1**	8.96 ± 0.2	66.6 ± 1.6	-	115.7 ± 3	33.0 ± 4	
FalseDec	23.9 ± 0.2	8.81 ± 0.2	67.3 ± 0.0	4.40 ± 0.1	115.9 ± 2	**30.4 ± 4**	
L-norm	24.4 ± 0.1	**8.53 ± 0.2**	**67.7 ± 0.0**	**3.69 ± 0.0**	**115.3 ± 1**	30.6 ± 4	BDD
N-flow	24.4 ± 0.1	**8.53 ± 0.2**	**67.7 ± 0.0**	**3.69 ± 0.0**	116.4 ± 1	31.0 ± 3	
Samp30	21.0 ± 0.1	9.02 ± 0.2	64.7 ± 0.0	3.70 ± 0.0	118.0 ± 3	33.6 ± 4	
Samp100	23.2 ± 0.1	8.68 ± 0.2	66.7 ± 0.0	**3.69 ± 0.0**	117.0 ± 3	45.4 ± 4	
Samp1000	24.2 ± 0.1	8.55 ± 0.2	67.6 ± 0.1	**3.69 ± 0.0**	118.4 ± 3	187.3 ± 4	

$$\mu_w = e^{\hat{\mu}_w + \cancel{\frac{\hat{\sigma}_w^2}{2}}} \odot w_a$$

Our methods

AP: Average precision
RMSE: Root-mean-square error
mIoU: Average intersection over union
NLL: Negative log-likelihood
ET: Model exporting time in seconds
IT: Inference time in milliseconds

Results

Propagation Methods

- **EfficientDet-D0** pre-trained on **COCO**
- Input resolution: **1024x512**
- Soft-NMS output is **reordered** based on **lowest MSE**
- **BDD100K:** 10 classes
- **KITTI:** 7 classes

$$\mu_w = e^{\hat{\mu}_w + \cancel{\frac{\hat{\sigma}_w^2}{2}}} \odot w_a$$

Our methods

Method	AP↑	RMSE↓	mIoU↑	NLL↓	ET↓ (s)	IT↓ (ms)
Baseline	72.8 ± 0.1	**5.07 ± 0.1**	90.1 ± 0.1	-	**115.6 ± 3**	34.8 ± 4
FalseDec	73.1 ± 0.5	5.27 ± 0.1	90.3 ± 0.1	4.27 ± 0.1	116.0 ± 3	31.1 ± 3
L-norm	**73.3** ± 0.5	5.17 ± 0.2	90.3 ± 0.0	3.22 ± 0.0	**115.6 ± 2**	**31.0 ± 3**
N-flow	**73.3** ± 0.5	5.17 ± 0.2	90.3 ± 0.0	3.22 ± 0.0	116.6 ± 1	31.6 ± 3
Samp30	68.6 ± 0.4	5.43 ± 0.1	88.7 ± 0.1	**3.19 ± 0.0**	118.8 ± 2	34.5 ± 3
Samp100	71.8 ± 0.5	5.23 ± 0.1	90.1 ± 0.0	3.20 ± 0.0	117.4 ± 4	47.0 ± 3
Samp1000	73.1 ± 0.5	5.18 ± 0.2	**90.4 ± 0.0**	3.21 ± 0.0	117.9 ± 4	187.4 ± 4

KITTI

$$\mu_w = e^{\hat{\mu}_w + \cancel{\frac{\hat{\sigma}_w^2}{2}}} \odot w_a$$

Our methods

Method	AP↑	RMSE↓	mIoU↑	NLL↓	ET↓ (s)	IT↓ (ms)
Baseline	**24.7 ± 0.1**	8.96 ± 0.2	66.6 ± 1.6	-	115.7 ± 3	33.0 ± 4
FalseDec	23.9 ± 0.2	8.81 ± 0.2	67.3 ± 0.0	4.40 ± 0.1	115.9 ± 2	**30.4 ± 4**
L-norm	24.4 ± 0.1	**8.53 ± 0.2**	**67.7 ± 0.0**	**3.69 ± 0.0**	**115.3 ± 1**	30.6 ± 4
N-flow	24.4 ± 0.1	**8.53 ± 0.2**	**67.7 ± 0.0**	**3.69 ± 0.0**	116.4 ± 1	31.0 ± 3
Samp30	21.0 ± 0.1	9.02 ± 0.2	64.7 ± 0.0	3.70 ± 0.0	118.0 ± 3	33.6 ± 4
Samp100	23.2 ± 0.1	8.68 ± 0.2	66.7 ± 0.0	**3.69 ± 0.0**	117.0 ± 3	45.4 ± 4
Samp1000	24.2 ± 0.1	8.55 ± 0.2	67.6 ± 0.1	**3.69 ± 0.0**	118.4 ± 3	187.3 ± 4

BDD

AP: Average precision
RMSE: Root-mean-square error
mIoU: Average intersection over union
NLL: Negative log-likelihood
ET: Model exporting time in seconds
IT: Inference time in milliseconds

Results

Propagation Methods

- **EfficientDet-D0** pre-trained on **COCO**
- Input resolution: **1024x512**
- Soft-NMS output is **reordered** based on **lowest MSE**
- **BDD100K:** 10 classes
- **KITTI:** 7 classes

$$\boxed{\mu_w = e^{\hat{\mu}_w + \cancel{\frac{\hat{\sigma}_w^2}{2}}} \odot w_a}$$

Dataset	Group	Method	AP↑	RMSE↓	mIoU↑	NLL↓	ET↓ (s)	IT↓ (ms)
KITTI		Baseline	72.8 ± 0.1	**5.07 ± 0.1**	90.1 ± 0.1	-	**115.6 ± 3**	34.8 ± 4
	Our methods	FalseDec	73.1 ± 0.5	5.27 ± 0.1	90.3 ± 0.1	4.27 ± 0.1	116.0 ± 3	31.1 ± 3
	Our methods	L-norm	**73.3** ± 0.5	5.17 ± 0.2	90.3 ± 0.0	3.22 ± 0.0	**115.6 ± 2**	**31.0 ± 3**
	Our methods	N-flow	**73.3** ± 0.5	5.17 ± 0.2	90.3 ± 0.0	3.22 ± 0.0	116.6 ± 1	31.6 ± 3
	SOTA: Sampling	Samp30	68.6 ± 0.4	5.43 ± 0.1	88.7 ± 0.1	**3.19 ± 0.0**	118.8 ± 2	34.5 ± 3
	SOTA: Sampling	Samp100	71.8 ± 0.5	5.23 ± 0.1	90.1 ± 0.0	3.20 ± 0.0	117.4 ± 4	47.0 ± 3
	SOTA: Sampling	Samp1000	73.1 ± 0.5	5.18 ± 0.2	**90.4 ± 0.0**	3.21 ± 0.0	117.9 ± 4	187.4 ± 4
BDD		Baseline	**24.7 ± 0.1**	8.96 ± 0.2	66.6 ± 1.6	-	115.7 ± 3	33.0 ± 4
	Our methods	FalseDec	23.9 ± 0.2	8.81 ± 0.2	67.3 ± 0.0	4.40 ± 0.1	115.9 ± 2	**30.4 ± 4**
	Our methods	L-norm	24.4 ± 0.1	**8.53 ± 0.2**	**67.7 ± 0.0**	**3.69 ± 0.0**	**115.3 ± 1**	30.6 ± 4
	Our methods	N-flow	24.4 ± 0.1	**8.53 ± 0.2**	**67.7 ± 0.0**	**3.69 ± 0.0**	116.4 ± 1	31.0 ± 3
	SOTA : Sampling	Samp30	21.0 ± 0.1	9.02 ± 0.2	64.7 ± 0.0	3.70 ± 0.0	118.0 ± 3	33.6 ± 4
	SOTA : Sampling	Samp100	23.2 ± 0.1	8.68 ± 0.2	66.7 ± 0.0	**3.69 ± 0.0**	117.0 ± 3	45.4 ± 4
	SOTA : Sampling	Samp1000	24.2 ± 0.1	8.55 ± 0.2	67.6 ± 0.1	**3.69 ± 0.0**	118.4 ± 3	187.3 ± 4

AP: Average precision
RMSE: Root-mean-square error
mIoU: Average intersection over union
NLL: Negative log-likelihood
ET: Model exporting time in seconds
IT: Inference time in milliseconds

Results

Propagation Methods

- **EfficientDet-D0** pre-trained on **COCO**
- Input resolution: **1024x512**
- Soft-NMS output is **reordered** based on **lowest MSE**
- **BDD100K:** 10 classes
- **KITTI:** 7 classes

$$\mu_w = e^{\hat{\mu}_w + \cancel{\frac{\hat{\sigma}_w^2}{2}}} \odot w_a$$

	Method	AP↑	RMSE↓	mIoU↑	NLL↓	ET↓ (s)	IT↓ (ms)	
	Baseline	72.8 ± 0.1	**5.07 ± 0.1**	90.1 ± 0.1	-	**115.6 ± 3**	34.8 ± 4	KITTI
	FalseDec	73.1 ± 0.5	5.27 ± 0.1	90.3 ± 0.1	4.27 ± 0.1	116.0 ± 3	31.1 ± 3	
Our methods	L-norm	**73.3** ± 0.5	5.17 ± 0.2	90.3 ± 0.0	3.22 ± 0.0	**115.6 ± 2**	**31.0 ± 3**	
	N-flow	**73.3** ± 0.5	5.17 ± 0.2	90.3 ± 0.0	3.22 ± 0.0	116.6 ± 1	31.6 ± 3	
SOTA: Sampling	Samp30	68.6 ± 0.4	5.43 ± 0.1	88.7 ± 0.1	**3.19 ± 0.0**	118.8 ± 2	34.5 ± 3	
	Samp100	71.8 ± 0.5	5.23 ± 0.1	90.1 ± 0.0	3.20 ± 0.0	117.4 ± 4	47.0 ± 3	
	Samp1000	73.1 ± 0.5	5.18 ± 0.2	**90.4 ± 0.0**	3.21 ± 0.0	117.9 ± 4	187.4 ± 4	
	Baseline	**24.7 ± 0.1**	8.96 ± 0.2	66.6 ± 1.6	-	115.7 ± 3	33.0 ± 4	BDD
	FalseDec	23.9 ± 0.2	8.81 ± 0.2	67.3 ± 0.0	4.40 ± 0.1	115.9 ± 2	**30.4 ± 4**	
Our methods	L-norm	24.4 ± 0.1	**8.53 ± 0.2**	**67.7 ± 0.0**	**3.69 ± 0.0**	**115.3 ± 1**	30.6 ± 4	
	N-flow	24.4 ± 0.1	**8.53 ± 0.2**	**67.7 ± 0.0**	**3.69 ± 0.0**	116.4 ± 1	31.0 ± 3	
SOTA : Sampling	Samp30	21.0 ± 0.1	9.02 ± 0.2	64.7 ± 0.0	3.70 ± 0.0	118.0 ± 3	33.6 ± 4	
	Samp100	23.2 ± 0.1	8.68 ± 0.2	66.7 ± 0.0	**3.69 ± 0.0**	117.0 ± 3	45.4 ± 4	
	Samp1000	24.2 ± 0.1	8.55 ± 0.2	67.6 ± 0.1	**3.69 ± 0.0**	118.4 ± 3	187.3 ± 4	

Increases localization performance

AP: Average precision
RMSE: Root-mean-square error
mIoU: Average intersection over union
NLL: Negative log-likelihood
ET: Model exporting time in seconds
IT: Inference time in milliseconds

Results

Propagation Methods

- **EfficientDet-D0** pre-trained on **COCO**
- Input resolution: **1024x512**
- Soft-NMS output is **reordered** based on **lowest MSE**
- **BDD100K:** 10 classes
- **KITTI:** 7 classes

$$\mu_w = e^{\hat{\mu}_w + \cancel{\frac{\hat{\sigma}_w^2}{2}}} \odot w_a$$

	Method	AP↑	RMSE↓	mIoU↑	NLL↓	ET↓ (s)	IT↓ (ms)	
	Baseline	72.8 ± 0.1	**5.07** ± **0.1**	90.1 ± 0.1	-	**115.6** ± **3**	34.8 ± 4	KITTI
	FalseDec	73.1 ± 0.5	5.27 ± 0.1	90.3 ± 0.1	4.27 ± 0.1	116.0 ± 3	31.1 ± 3	KITTI
Our methods	L-norm	**73.3** ± 0.5	5.17 ± 0.2	90.3 ± 0.0	3.22 ± 0.0	**115.6** ± **2**	**31.0** ± **3**	KITTI
Our methods	N-flow	**73.3** ± 0.5	5.17 ± 0.2	90.3 ± 0.0	3.22 ± 0.0	116.6 ± 1	31.6 ± 3	KITTI
SOTA: Sampling	Samp30	68.6 ± 0.4	5.43 ± 0.1	88.7 ± 0.1	**3.19** ± **0.0**	118.8 ± 2	34.5 ± 3	KITTI
SOTA: Sampling	Samp100	71.8 ± 0.5	5.23 ± 0.1	90.1 ± 0.0	3.20 ± 0.0	117.4 ± 4	47.0 ± 3	KITTI
SOTA: Sampling	Samp1000	73.1 ± 0.5	5.18 ± 0.2	**90.4** ± **0.0**	3.21 ± 0.0	117.9 ± 4	187.4 ± 4	KITTI
	Baseline	**24.7** ± **0.1**	8.96 ± 0.2	66.6 ± 1.6	-	115.7 ± 3	33.0 ± 4	BDD
	FalseDec	23.9 ± 0.2	8.81 ± 0.2	67.3 ± 0.0	4.40 ± 0.1	115.9 ± 2	**30.4** ± **4**	BDD
Our methods	L-norm	24.4 ± 0.1	**8.53** ± **0.2**	**67.7** ± **0.0**	**3.69** ± **0.0**	**115.3** ± **1**	30.6 ± 4	BDD
Our methods	N-flow	24.4 ± 0.1	**8.53** ± **0.2**	**67.7** ± **0.0**	**3.69** ± **0.0**	116.4 ± 1	31.0 ± 3	BDD
SOTA : Sampling	Samp30	21.0 ± 0.1	9.02 ± 0.2	64.7 ± 0.0	3.70 ± 0.0	118.0 ± 3	33.6 ± 4	BDD
SOTA : Sampling	Samp100	23.2 ± 0.1	8.68 ± 0.2	66.7 ± 0.0	**3.69** ± **0.0**	117.0 ± 3	45.4 ± 4	BDD
SOTA : Sampling	Samp1000	24.2 ± 0.1	8.55 ± 0.2	67.6 ± 0.1	**3.69** ± **0.0**	118.4 ± 3	187.3 ± 4	BDD

$$\mu_w = e^{\hat{\mu}_w + \cancel{\frac{\hat{\sigma}_w^2}{2}}} \odot w_a$$

Increases localization performance
Faster than baseline and sampling

AP: Average precision
RMSE: Root-mean-square error
mIoU: Average intersection over union
NLL: Negative log-likelihood
ET: Model exporting time in seconds
IT: Inference time in milliseconds

Results
Propagation Methods

- **EfficientDet-D0** pre-trained on **COCO**
- Input resolution: **1024x512**
- Soft-NMS output is **reordered** based on **lowest MSE**
- **BDD100K:** 10 classes
- **KITTI:** 7 classes

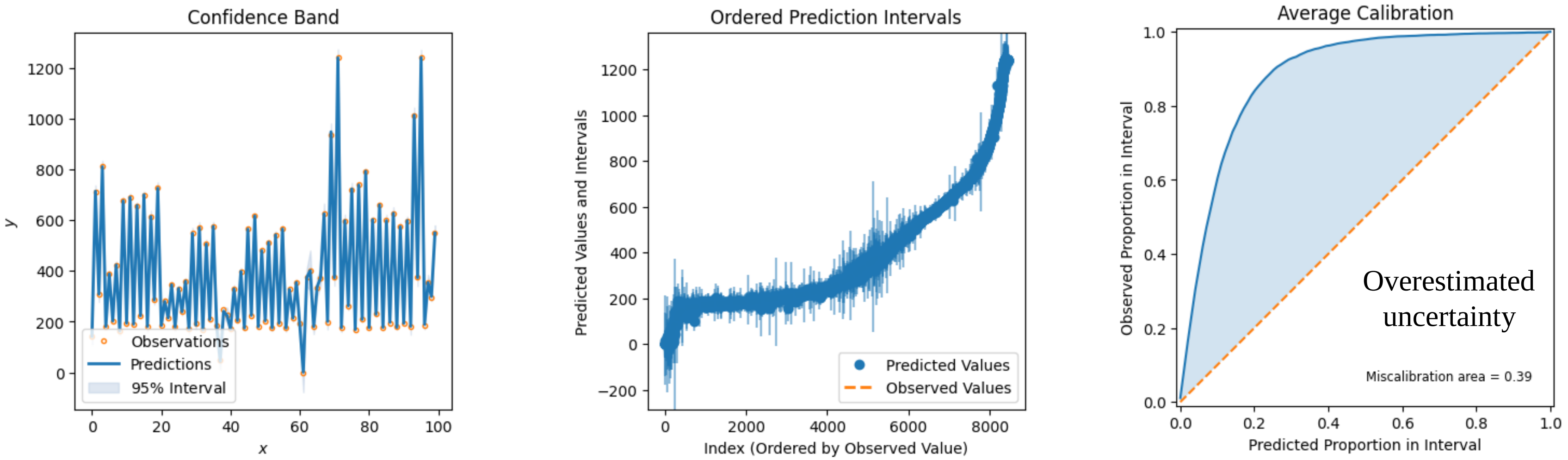
	Method	AP↑	RMSE↓	mIoU↑	NLL↓	ET↓ (s)	IT↓ (ms)	
$\mu_w = e^{\hat{\mu}_w + \cancel{\frac{\hat{\sigma}_w^2}{2}}} \odot w_a$	Baseline	72.8 ± 0.1	**5.07** ± **0.1**	90.1 ± 0.1	-	**115.6** ± **3**	34.8 ± 4	KITTI
	FalseDec	73.1 ± 0.5	5.27 ± 0.1	90.3 ± 0.1	4.27 ± 0.1	116.0 ± 3	31.1 ± 3	KITTI
Our methods	L-norm	**73.3** ± 0.5	5.17 ± 0.2	90.3 ± 0.0	3.22 ± 0.0	**115.6** ± **2**	**31.0** ± **3**	KITTI
Our methods	N-flow	**73.3** ± 0.5	5.17 ± 0.2	90.3 ± 0.0	3.22 ± 0.0	116.6 ± 1	31.6 ± 3	KITTI
SOTA: Sampling	Samp30	68.6 ± 0.4	5.43 ± 0.1	88.7 ± 0.1	**3.19** ± **0.0**	118.8 ± 2	34.5 ± 3	KITTI
SOTA: Sampling	Samp100	71.8 ± 0.5	5.23 ± 0.1	90.1 ± 0.0	3.20 ± 0.0	117.4 ± 4	47.0 ± 3	KITTI
SOTA: Sampling	Samp1000	73.1 ± 0.5	5.18 ± 0.2	**90.4** ± **0.0**	3.21 ± 0.0	117.9 ± 4	187.4 ± 4	KITTI
$\mu_w = e^{\hat{\mu}_w + \cancel{\frac{\hat{\sigma}_w^2}{2}}} \odot w_a$	Baseline	**24.7** ± **0.1**	8.96 ± 0.2	66.6 ± 1.6	-	115.7 ± 3	33.0 ± 4	BDD
	FalseDec	23.9 ± 0.2	8.81 ± 0.2	67.3 ± 0.0	4.40 ± 0.1	115.9 ± 2	**30.4** ± **4**	BDD
Our methods	L-norm	24.4 ± 0.1	**8.53** ± **0.2**	**67.7** ± **0.0**	**3.69** ± **0.0**	**115.3** ± **1**	30.6 ± 4	BDD
Our methods	N-flow	24.4 ± 0.1	**8.53** ± **0.2**	**67.7** ± **0.0**	**3.69** ± **0.0**	116.4 ± 1	31.0 ± 3	BDD
SOTA : Sampling	Samp30	21.0 ± 0.1	9.02 ± 0.2	64.7 ± 0.0	3.70 ± 0.0	118.0 ± 3	33.6 ± 4	BDD
SOTA : Sampling	Samp100	23.2 ± 0.1	8.68 ± 0.2	66.7 ± 0.0	**3.69** ± **0.0**	117.0 ± 3	45.4 ± 4	BDD
SOTA : Sampling	Samp1000	24.2 ± 0.1	8.55 ± 0.2	67.6 ± 0.1	**3.69** ± **0.0**	118.4 ± 3	187.3 ± 4	BDD

AP: Average precision
RMSE: Root-mean-square error
mIoU: Average intersection over union
NLL: Negative log-likelihood
ET: Model exporting time in seconds
IT: Inference time in milliseconds

Increases localization performance
Faster than baseline and sampling
Addresses the drawbacks of sampling

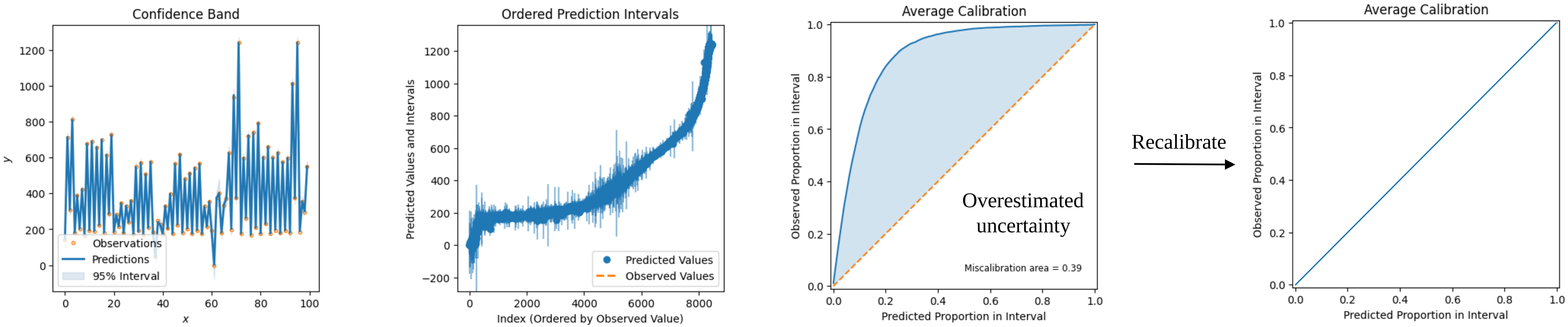
Calibration

Is the predicted uncertainty well-calibrated?



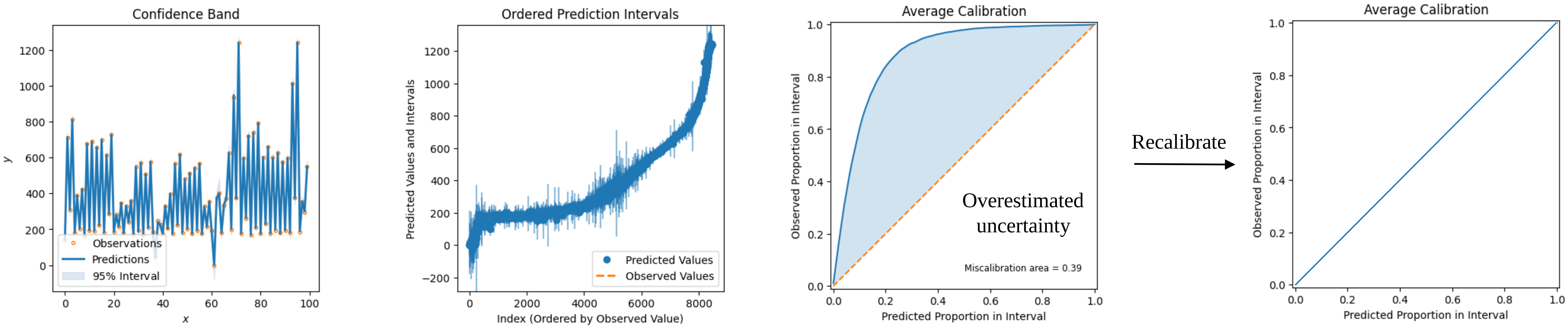
Calibration

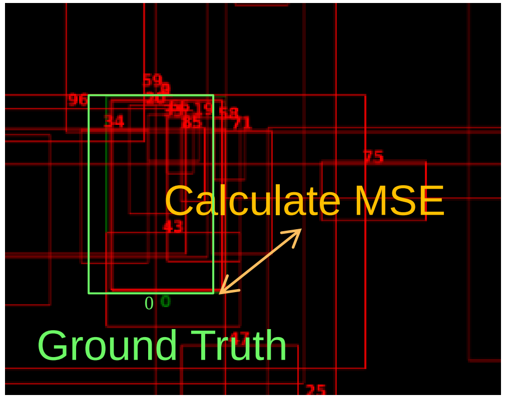
Is the predicted uncertainty well-calibrated?



Calibration

Is the predicted uncertainty well-calibrated?



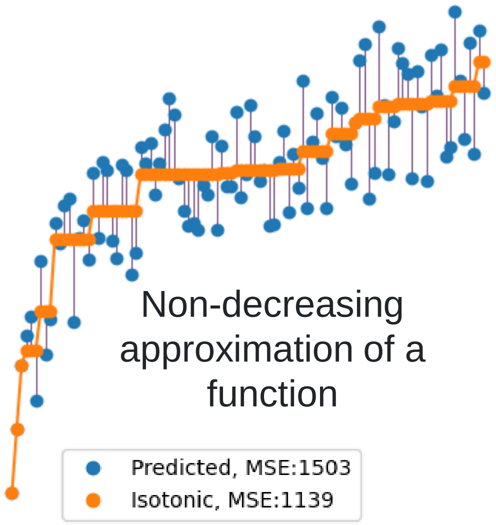


Prediction allocation via **nearest-neighbor** rather than thresholding:

- For every **label**, a **corresponding box** is present.
- Usually the one with the **highest classification score** is designated.
- Does not necessarily correspond to the **ground truth** based on **localization**.

Calibration
Calibration Methods

Isotonic Regression



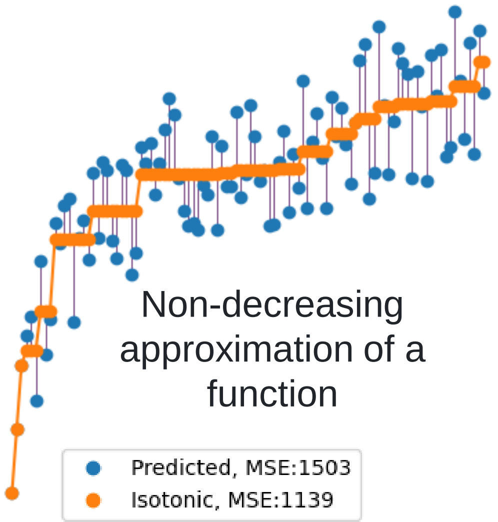
Non-decreasing approximation of a function

From one model to a model:
- Per-coordinate
- Per-class
- Per-coordinate + per-class

Calibration

Calibration Methods

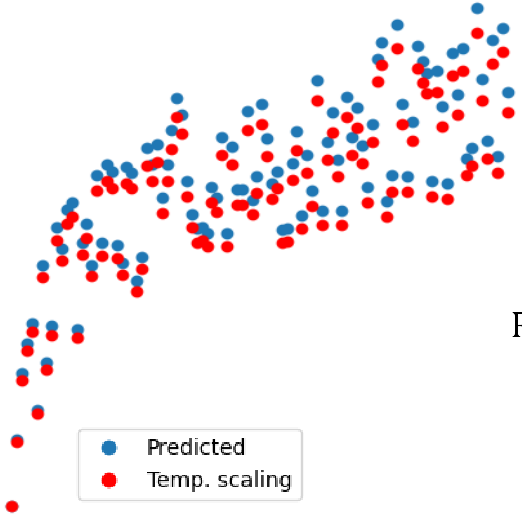
Isotonic Regression



Non-decreasing approximation of a function

From one model to a model:
- Per-coordinate
- Per-class
- Per-coordinate + per-class

Factor Scaling



$$\mathcal{N}(\mu, s \cdot \sigma)$$

$$\Delta_i = |y_i^* - \mu_i|$$ Residuals

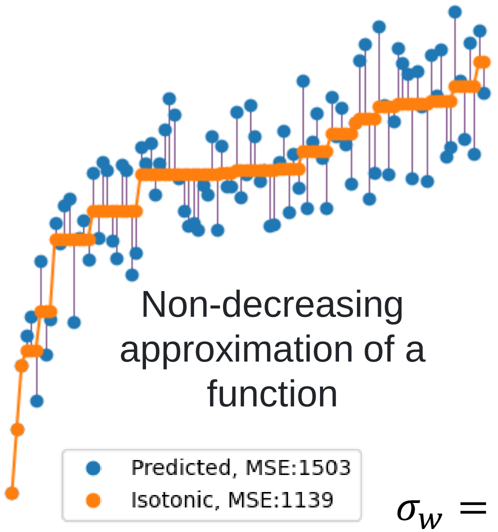
$$\mathrm{RMSE}_{unc} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\Delta_i - \sigma_i)^2}$$

$$\mathrm{MAE}_{unc} = \frac{1}{N}\sum_{i=1}^{N}|\Delta_i - \sigma_i|$$

Calibration

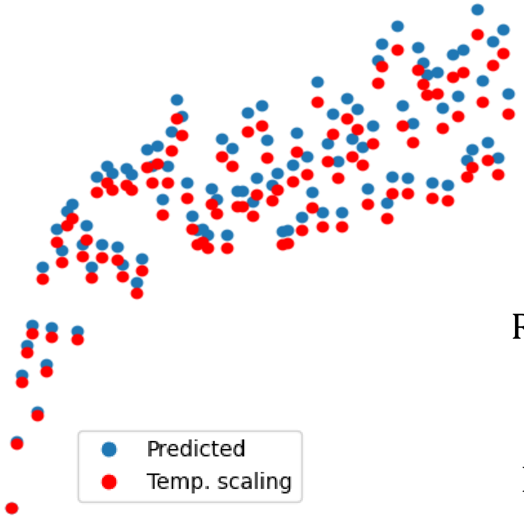
Calibration Methods

Isotonic Regression



From one model to a model:
- Per-coordinate
- Per-class
- Per-coordinate + per-class

Factor Scaling

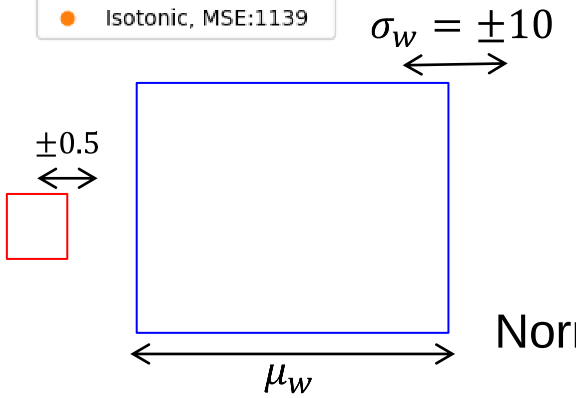


$\mathcal{N}(\mu, s \cdot \sigma)$

$\Delta_i = |y_i^* - \mu_i|$ Residuals

$$\text{RMSE}_{unc} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\Delta_i - \sigma_i)^2}$$

$$\text{MAE}_{unc} = \frac{1}{N}\sum_{i=1}^{N}|\Delta_i - \sigma_i|$$

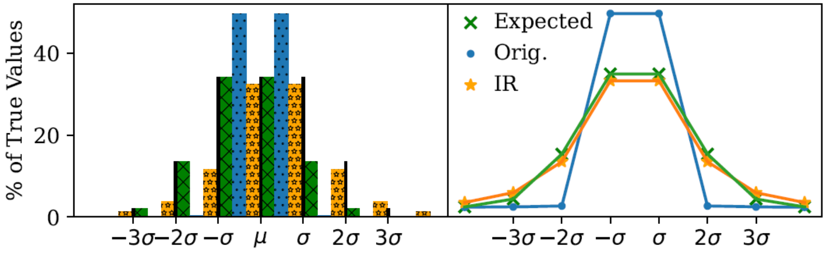


Relative Calibration

Normalize Δ_i, σ_i with the corresponding width and height before calibration

Results

Calibration Methods

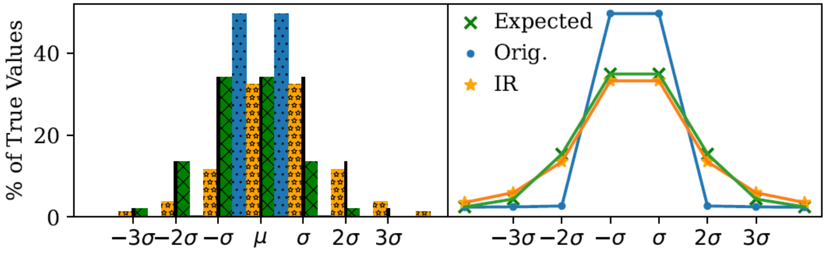


Distribution of the localization error; expected is a Gaussian

Uncertainty is mis-calibrated

Results

Calibration Methods



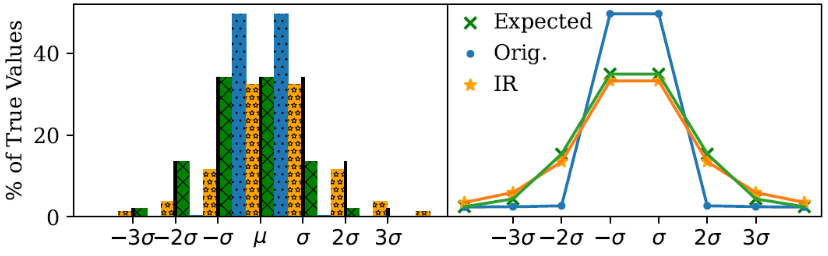
Distribution of the localization error; expected is a Gaussian

	KITTI				BDD			
Method	**RMSUE↓**	**ECE↓**	**NLL↓**	**Sharp↓**	**RMSUE↓**	**ECE↓**	**NLL↓**	**Sharp↓**
Uncalibrated	13.0 ± 0.0	0.384 ± 0.000	3.22 ± 0.0	14.9 ± 0.0	15.1 ± 0.1	0.323 ± 0.000	3.69 ± 0.0	17.22 ± 0.0
FS NLL	5.0 ± 0.3	0.194 ± 0.021	**2.51 ± 0.1**	4.7 ± 0.5	9.4 ± 0.4	0.217 ± 0.008	**3.46 ± 0.0**	9.72 ± 0.4
FS MAUE	4.6 ± 0.2	0.047 ± 0.001	3.14 ± 0.4	**2.5 ± 0.0**	7.5 ± 0.3	0.026 ± 0.001	4.72 ± 0.2	4.28 ± 0.0
FS RMSUE	4.6 ± 0.2	0.088 ± 0.003	2.79 ± 0.2	3.0 ± 0.0	7.6 ± 0.3	0.074 ± 0.000	6.43 ± 0.3	**3.21 ± 0.0**
Rel. FS RMSUE	7.2 ± 0.1	0.306 ± 0.002	2.74 ± 0.0	8.3 ± 0.1	8.5 ± 0.3	0.175 ± 0.003	3.50 ± 0.1	8.06 ± 0.1
Abs. IR	4.5 ± 0.2	0.032 ± 0.001	3.15 ± 0.3	**2.5 ± 0.0**	7.5 ± 0.3	0.027 ± 0.001	4.60 ± 0.1	4.09 ± 0.0
Abs. IR CL	4.4 ± 0.2	0.029 ± 0.001	2.86 ± 0.2	2.7 ± 0.0	7.4 ± 0.3	0.026 ± 0.001	4.39 ± 0.1	4.23 ± 0.0
Abs. IR PCo	4.5 ± 0.2	0.032 ± 0.001	3.03 ± 0.2	2.6 ± 0.0	7.5 ± 0.3	0.027 ± 0.001	4.57 ± 0.2	4.11 ± 0.0
Abs. IR PCo CL	**4.3 ± 0.2**	0.028 ± 0.000	2.70 ± 0.1	2.9 ± 0.0	7.4 ± 0.3	0.025 ± 0.001	4.36 ± 0.1	4.33 ± 0.0
Rel. IR	4.5 ± 0.2	0.027 ± 0.001	3.06 ± 0.3	**2.5 ± 0.0**	7.4 ± 0.3	0.018 ± 0.001	4.52 ± 0.1	4.07 ± 0.0
Rel. IR CL	4.4 ± 0.2	0.026 ± 0.001	2.78 ± 0.2	3.1 ± 0.4	**7.3 ± 0.3**	**0.017 ± 0.000**	4.29 ± 0.1	4.24 ± 0.0
Rel. IR PCo	4.5 ± 0.2	0.027 ± 0.001	3.03 ± 0.2	**2.5 ± 0.1**	7.4 ± 0.3	0.018 ± 0.000	4.49 ± 0.1	4.08 ± 0.0
Rel. IR PCo CL	4.4 ± 0.3	**0.025 ± 0.000**	2.69 ± 0.2	3.2 ± 0.5	**7.3 ± 0.3**	**0.017 ± 0.000**	4.23 ± 0.1	4.27 ± 0.0

Uncertainty is mis-calibrated

Results

Calibration Methods



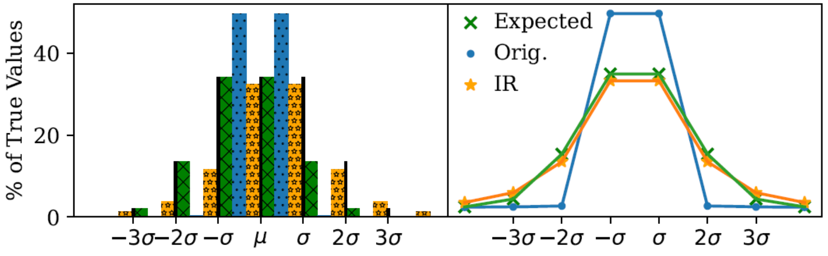
Distribution of the localization error; expected is a Gaussian

	KITTI				BDD			
Method	**RMSUE↓**	**ECE↓**	**NLL↓**	**Sharp↓**	**RMSUE↓**	**ECE↓**	**NLL↓**	**Sharp↓**
Uncalibrated	13.0 ± 0.0	0.384 ± 0.000	3.22 ± 0.0	14.9 ± 0.0	15.1 ± 0.1	0.323 ± 0.000	3.69 ± 0.0	17.22 ± 0.0
FS NLL	5.0 ± 0.3	0.194 ± 0.021	**2.51 ± 0.1**	4.7 ± 0.5	9.4 ± 0.4	0.217 ± 0.008	**3.46 ± 0.0**	9.72 ± 0.4
FS MAUE	4.6 ± 0.2	0.047 ± 0.001	3.14 ± 0.4	**2.5 ± 0.0**	7.5 ± 0.3	0.026 ± 0.001	4.72 ± 0.2	4.28 ± 0.0
FS RMSUE	4.6 ± 0.2	0.088 ± 0.003	2.79 ± 0.2	3.0 ± 0.0	7.6 ± 0.3	0.074 ± 0.000	6.43 ± 0.3	**3.21 ± 0.0**
Rel. FS RMSUE	7.2 ± 0.1	0.306 ± 0.002	2.74 ± 0.0	8.3 ± 0.1	8.5 ± 0.3	0.175 ± 0.003	3.50 ± 0.1	8.06 ± 0.1
Abs. IR	4.5 ± 0.2	0.032 ± 0.001	3.15 ± 0.3	**2.5 ± 0.0**	7.5 ± 0.3	0.027 ± 0.001	4.60 ± 0.1	4.09 ± 0.0
Abs. IR CL	4.4 ± 0.2	0.029 ± 0.001	2.86 ± 0.2	2.7 ± 0.0	7.4 ± 0.3	0.026 ± 0.001	4.39 ± 0.1	4.23 ± 0.0
Abs. IR PCo	4.5 ± 0.2	0.032 ± 0.001	3.03 ± 0.2	2.6 ± 0.0	7.5 ± 0.3	0.027 ± 0.001	4.57 ± 0.2	4.11 ± 0.0
Abs. IR PCo CL	**4.3 ± 0.2**	0.028 ± 0.000	2.70 ± 0.1	2.9 ± 0.0	7.4 ± 0.3	0.025 ± 0.001	4.36 ± 0.1	4.33 ± 0.0
Rel. IR	4.5 ± 0.2	0.027 ± 0.001	3.06 ± 0.3	**2.5 ± 0.0**	7.4 ± 0.3	0.018 ± 0.001	4.52 ± 0.1	4.07 ± 0.0
Rel. IR CL	4.4 ± 0.2	0.026 ± 0.001	2.78 ± 0.2	3.1 ± 0.4	**7.3 ± 0.3**	**0.017 ± 0.000**	4.29 ± 0.1	4.24 ± 0.0
Rel. IR PCo	4.5 ± 0.2	0.027 ± 0.001	3.03 ± 0.2	**2.5 ± 0.1**	7.4 ± 0.3	0.018 ± 0.000	4.49 ± 0.1	4.08 ± 0.0
Rel. IR PCo CL	4.4 ± 0.3	**0.025 ± 0.000**	2.69 ± 0.2	3.2 ± 0.5	**7.3 ± 0.3**	**0.017 ± 0.000**	4.23 ± 0.1	4.27 ± 0.0

Uncertainty is mis-calibrated
Our FS losses outperform SOTA

Results

Calibration Methods



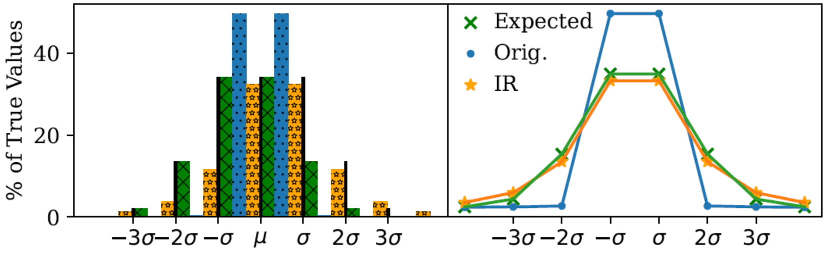
Distribution of the localization error; expected is a Gaussian

	KITTI				BDD			
Method	**RMSUE↓**	**ECE↓**	**NLL↓**	**Sharp↓**	**RMSUE↓**	**ECE↓**	**NLL↓**	**Sharp↓**
Uncalibrated	13.0 ± 0.0	0.384 ± 0.000	3.22 ± 0.0	14.9 ± 0.0	15.1 ± 0.1	0.323 ± 0.000	3.69 ± 0.0	17.22 ± 0.0
FS NLL	5.0 ± 0.3	0.194 ± 0.021	**2.51 ± 0.1**	4.7 ± 0.5	9.4 ± 0.4	0.217 ± 0.008	**3.46 ± 0.0**	9.72 ± 0.4
FS MAUE	4.6 ± 0.2	0.047 ± 0.001	3.14 ± 0.4	**2.5 ± 0.0**	7.5 ± 0.3	0.026 ± 0.001	4.72 ± 0.2	4.28 ± 0.0
FS RMSUE	4.6 ± 0.2	0.088 ± 0.003	2.79 ± 0.2	3.0 ± 0.0	7.6 ± 0.3	0.074 ± 0.000	6.43 ± 0.3	**3.21 ± 0.0**
Rel. FS RMSUE	7.2 ± 0.1	0.306 ± 0.002	2.74 ± 0.0	8.3 ± 0.1	8.5 ± 0.3	0.175 ± 0.003	3.50 ± 0.1	8.06 ± 0.1
Abs. IR	4.5 ± 0.2	0.032 ± 0.001	3.15 ± 0.3	**2.5 ± 0.0**	7.5 ± 0.3	0.027 ± 0.001	4.60 ± 0.1	4.09 ± 0.0
Abs. IR CL	4.4 ± 0.2	0.029 ± 0.001	2.86 ± 0.2	2.7 ± 0.0	7.4 ± 0.3	0.026 ± 0.001	4.39 ± 0.1	4.23 ± 0.0
Abs. IR PCo	4.5 ± 0.2	0.032 ± 0.001	3.03 ± 0.2	2.6 ± 0.0	7.5 ± 0.3	0.027 ± 0.001	4.57 ± 0.2	4.11 ± 0.0
Abs. IR PCo CL	**4.3 ± 0.2**	0.028 ± 0.000	2.70 ± 0.1	2.9 ± 0.0	7.4 ± 0.3	0.025 ± 0.001	4.36 ± 0.1	4.33 ± 0.0
Rel. IR	4.5 ± 0.2	0.027 ± 0.001	3.06 ± 0.3	**2.5 ± 0.0**	7.4 ± 0.3	0.018 ± 0.001	4.52 ± 0.1	4.07 ± 0.0
Rel. IR CL	4.4 ± 0.2	0.026 ± 0.001	2.78 ± 0.2	3.1 ± 0.4	**7.3 ± 0.3**	**0.017 ± 0.000**	4.29 ± 0.1	4.24 ± 0.0
Rel. IR PCo	4.5 ± 0.2	0.027 ± 0.001	3.03 ± 0.2	**2.5 ± 0.1**	7.4 ± 0.3	0.018 ± 0.000	4.49 ± 0.1	4.08 ± 0.0
Rel. IR PCo CL	4.4 ± 0.3	**0.025 ± 0.000**	2.69 ± 0.2	3.2 ± 0.5	**7.3 ± 0.3**	**0.017 ± 0.000**	4.23 ± 0.1	4.27 ± 0.0

Uncertainty is mis-calibrated
Our FS losses outperform SOTA
Per-coordinate and per-class calibration IR outperforms other methods

Results

Calibration Methods



Distribution of the localization error; expected is a Gaussian

	KITTI				BDD			
Method	**RMSUE↓**	**ECE↓**	**NLL↓**	**Sharp↓**	**RMSUE↓**	**ECE↓**	**NLL↓**	**Sharp↓**
Uncalibrated	13.0 ± 0.0	0.384 ± 0.000	3.22 ± 0.0	14.9 ± 0.0	15.1 ± 0.1	0.323 ± 0.000	3.69 ± 0.0	17.22 ± 0.0
FS NLL	5.0 ± 0.3	0.194 ± 0.021	**2.51 ± 0.1**	4.7 ± 0.5	9.4 ± 0.4	0.217 ± 0.008	**3.46 ± 0.0**	9.72 ± 0.4
FS MAUE	4.6 ± 0.2	0.047 ± 0.001	3.14 ± 0.4	**2.5 ± 0.0**	7.5 ± 0.3	0.026 ± 0.001	4.72 ± 0.2	4.28 ± 0.0
FS RMSUE	4.6 ± 0.2	0.088 ± 0.003	2.79 ± 0.2	3.0 ± 0.0	7.6 ± 0.3	0.074 ± 0.000	6.43 ± 0.3	**3.21 ± 0.0**
Rel. FS RMSUE	7.2 ± 0.1	0.306 ± 0.002	2.74 ± 0.0	8.3 ± 0.1	8.5 ± 0.3	0.175 ± 0.003	3.50 ± 0.1	8.06 ± 0.1
Abs. IR	4.5 ± 0.2	0.032 ± 0.001	3.15 ± 0.3	**2.5 ± 0.0**	7.5 ± 0.3	0.027 ± 0.001	4.60 ± 0.1	4.09 ± 0.0
Abs. IR CL	4.4 ± 0.2	0.029 ± 0.001	2.86 ± 0.2	2.7 ± 0.0	7.4 ± 0.3	0.026 ± 0.001	4.39 ± 0.1	4.23 ± 0.0
Abs. IR PCo	4.5 ± 0.2	0.032 ± 0.001	3.03 ± 0.2	2.6 ± 0.0	7.5 ± 0.3	0.027 ± 0.001	4.57 ± 0.2	4.11 ± 0.0
Abs. IR PCo CL	**4.3 ± 0.2**	0.028 ± 0.000	2.70 ± 0.1	2.9 ± 0.0	7.4 ± 0.3	0.025 ± 0.001	4.36 ± 0.1	4.33 ± 0.0
Rel. IR	4.5 ± 0.2	0.027 ± 0.001	3.06 ± 0.3	**2.5 ± 0.0**	7.4 ± 0.3	0.018 ± 0.001	4.52 ± 0.1	4.07 ± 0.0
Rel. IR CL	4.4 ± 0.2	0.026 ± 0.001	2.78 ± 0.2	3.1 ± 0.4	**7.3 ± 0.3**	**0.017 ± 0.000**	4.29 ± 0.1	4.24 ± 0.0
Rel. IR PCo	4.5 ± 0.2	0.027 ± 0.001	3.03 ± 0.2	**2.5 ± 0.1**	7.4 ± 0.3	0.018 ± 0.000	4.49 ± 0.1	4.08 ± 0.0
Rel. IR PCo CL	4.4 ± 0.3	**0.025 ± 0.000**	2.69 ± 0.2	3.2 ± 0.5	**7.3 ± 0.3**	**0.017 ± 0.000**	4.23 ± 0.1	4.27 ± 0.0

Uncertainty is mis-calibrated
Our FS losses outperform SOTA
Per-coordinate and per-class calibration IR outperforms other methods
Relative calibration affects small objects

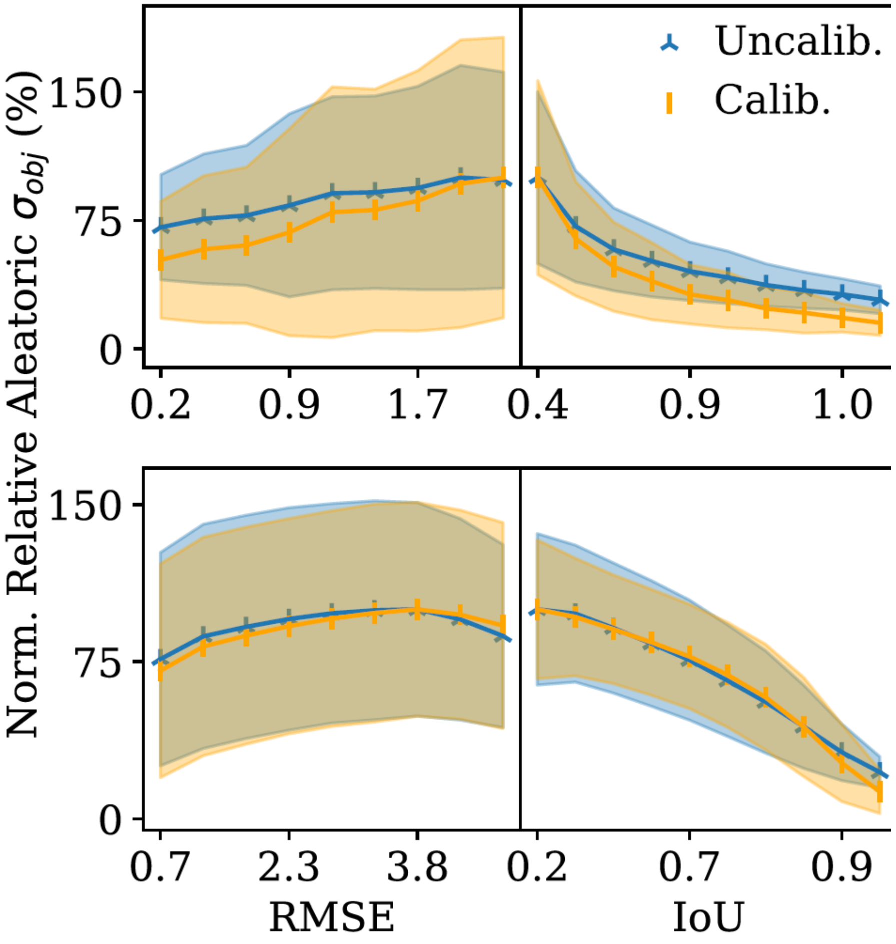
	Uncalibrated		**Absolute Calib.**		**Relative Calib.**	
	ECE	NLL	ECE	NLL	ECE	NLL
ALL	0.384 ± 0.000	3.22 ± 0.01	0.033 ± 0.000	3.15 ± 0.42	0.031 ± 0.000	2.93 ± 0.18
Small	0.383 ± 0.000	2.95 ± 0.06	0.049 ± 0.005	4.54 ± 2.59	0.029 ± 0.002	3.16 ± 1.36
Medium	0.382 ± 0.000	3.00 ± 0.00	0.039 ± 0.001	2.55 ± 0.12	0.033 ± 0.001	2.42 ± 0.09
Large	0.387 ± 0.000	3.79 ± 0.00	0.056 ± 0.002	3.45 ± 0.15	0.041 ± 0.001	3.73 ± 0.22

Results

What correlations exist between the data and the uncertainty?

KITTI

BDD



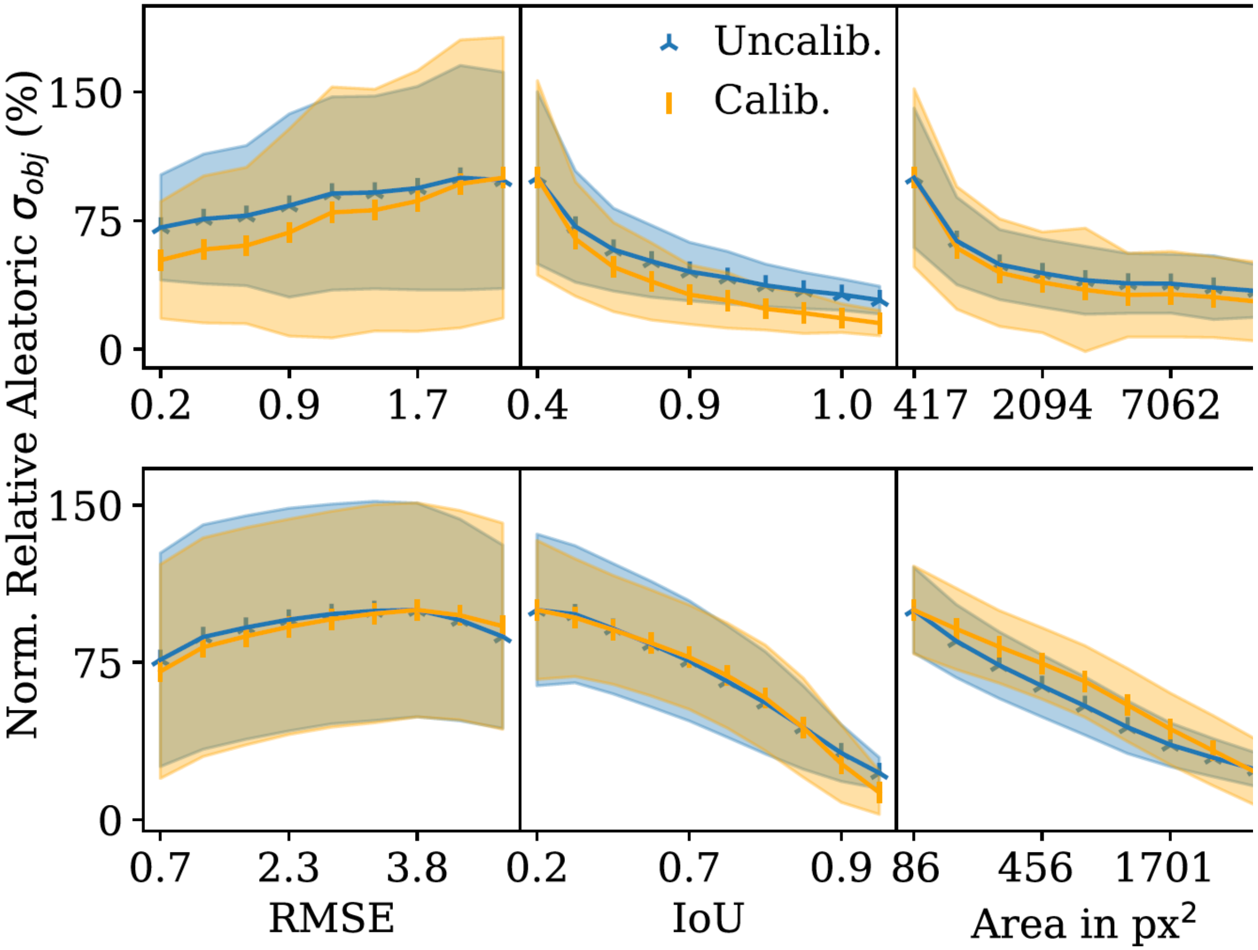
*Quantile-based binning, normalized by bin with highest uncertainty

Results

What correlations exist between the data and the uncertainty?

KITTI

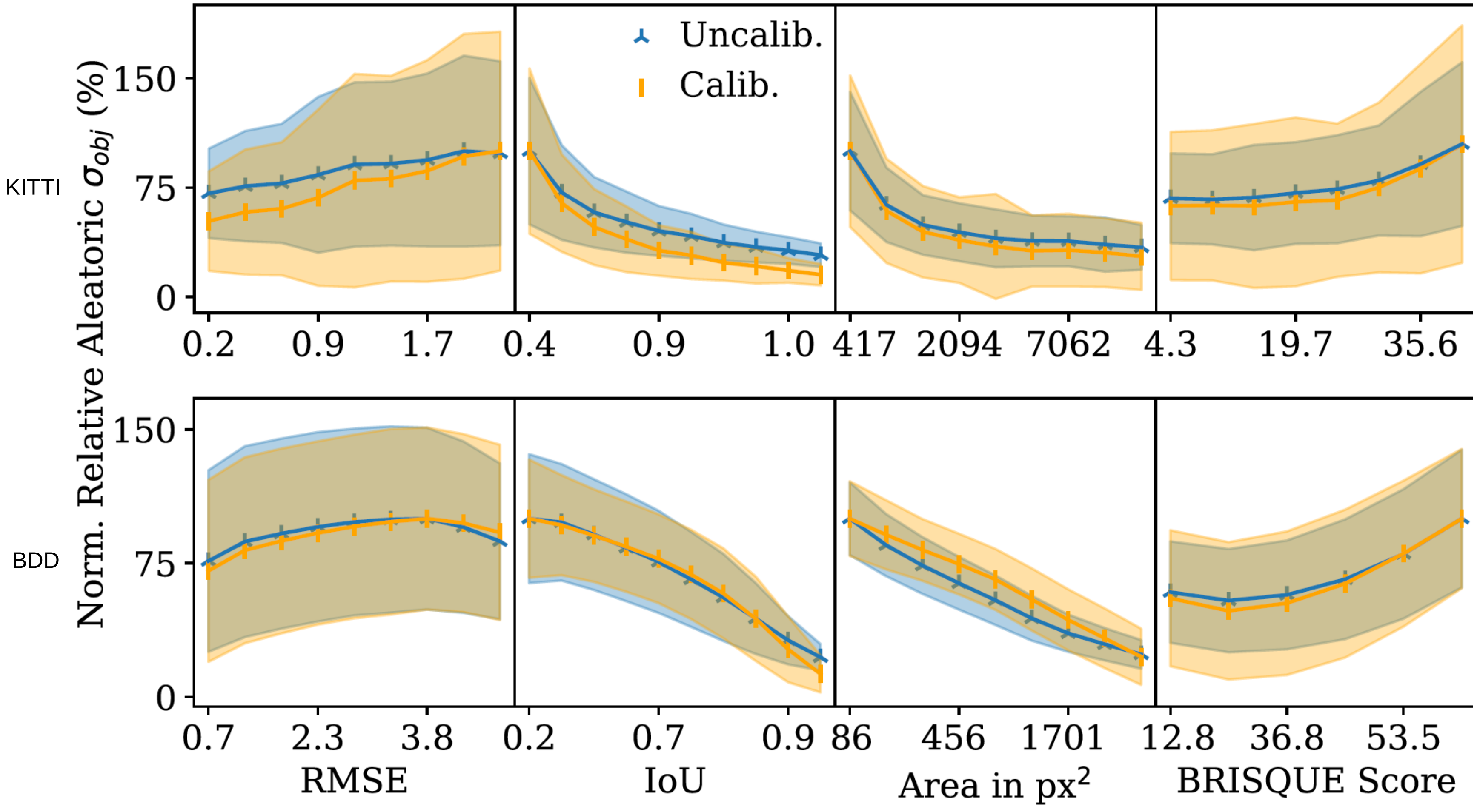
BDD



*Quantile-based binning, normalized by bin with highest uncertainty

Results

What correlations exist between the data and the uncertainty?



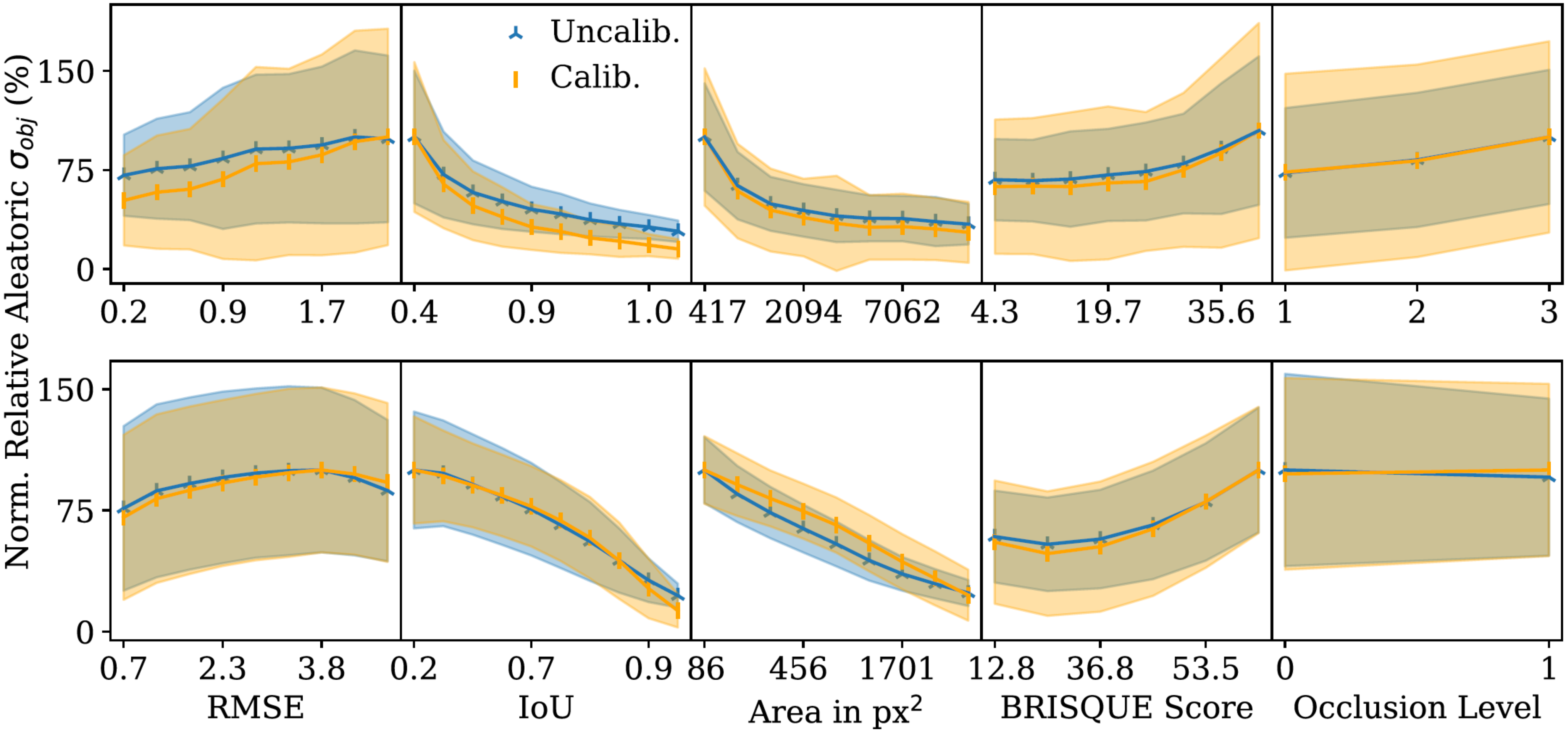
*Quantile-based binning, normalized by bin with highest uncertainty

Results

What correlations exist between the data and the uncertainty?

KITTI

BDD



1. Fully visible
2. Occluded less than 50%
3. Occluded more than 50%

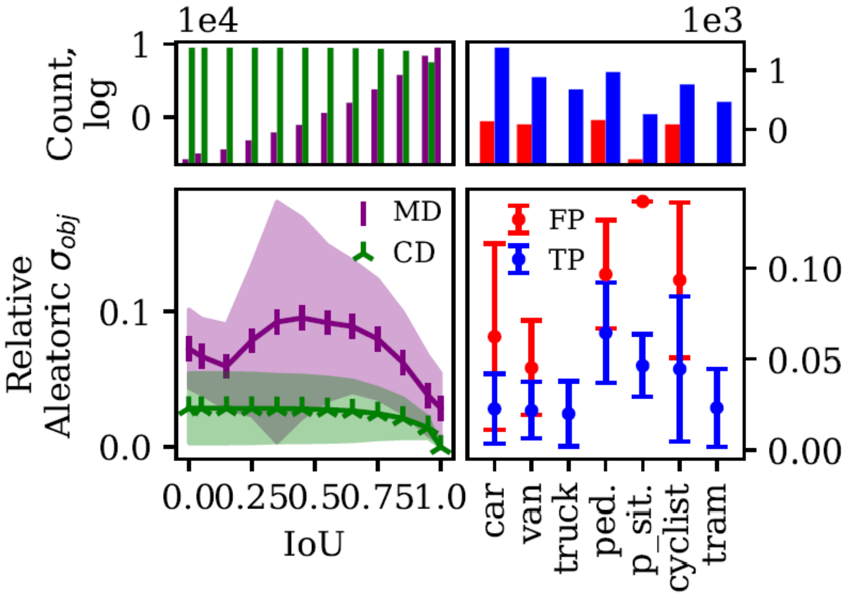
0. Visible
1. Occluded

*Quantile-based binning, normalized by bin with highest uncertainty

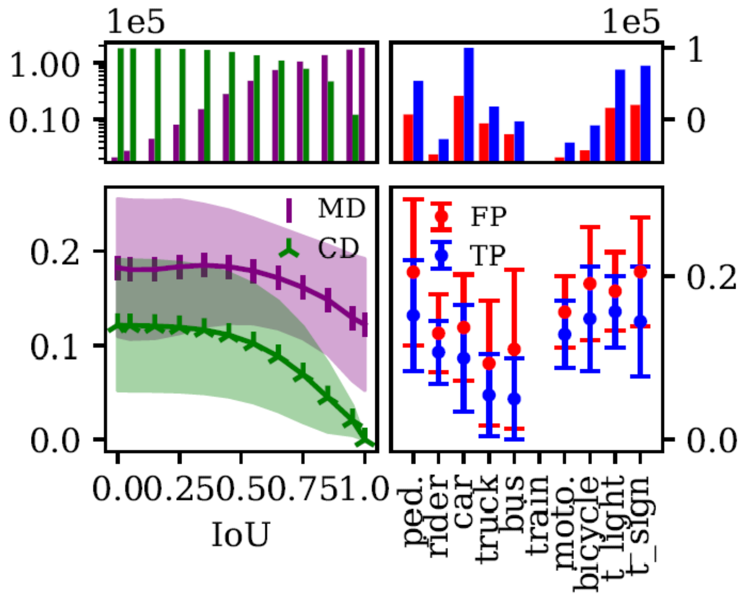
Results

Correlation Analysis – Misdetections

KITTI



BDD



Mis**D**etections: IoU ≤ IoU Threshold
Correct **D**etections: IoU > Threshold

Negative correlation with detection accuracy
Possible thresholding via aleatoric uncertainty

Summary

Loss attenuation for EfficientDet with **increased localization performance**.

Summary

Estimation

Propagation

Calibration

Explanation

Loss attenuation for EfficientDet with **increased localization performance**.

Two decoding methods:

- **Exact** and **fast** propagation.
- **Generalize** for any non-linear functions in regression networks, for **different equations** and **distributions**.

Summary

Estimation › Propagation › Calibration › Explanation

Estimation

Loss attenuation for EfficientDet with **increased localization performance**.

Propagation

Two decoding methods:

- **Exact** and **fast** propagation.
- **Generalize** for any non-linear functions in regression networks, for **different equations** and **distributions**.

Calibration

Extension of calibration methods: **Relative Per-coordinate**, **per-class calibration** with isotonic regression produces **well-calibrated uncertainties**.

Explanation

Summary

Estimation

Propagation

Calibration

Explanation

Loss attenuation for EfficientDet with **increased localization performance**.

Two decoding methods:
- **Exact** and **fast** propagation.
- **Generalize** for any non-linear functions in regression networks, for **different equations** and **distributions**.

Extension of calibration methods: **Relative Per-coordinate**, **per-class calibration** with isotonic regression produces **well-calibrated uncertainties**.

- Negative correlation with **detection performance** and **image quality**.
- Positive correlation with **occlusion** and **object distance** in the real-world.

Summary

Estimation → Propagation → Calibration → Explanation

Estimation

Loss attenuation for EfficientDet with **increased localization performance**.

Propagation

Two decoding methods:
- **Exact** and **fast** propagation.
- **Generalize** for any non-linear functions in regression networks, for **different equations** and **distributions**.

Calibration

Extension of calibration methods: **Relative Per-coordinate**, **per-class calibration** with isotonic regression produces **well-calibrated uncertainties**.

Explanation

- Negative correlation with **detection performance** and **image quality**.
- Positive correlation with **occlusion** and **object distance** in the real-world.

Future Work
- **Enhancing** object detection **safety** and **robustness**.

Summary

Estimation	Propagation	Calibration	Explanation
Loss attenuation for EfficientDet with **increased localization performance**.	Two decoding methods: › **Exact** and **fast** propagation. › **Generalize** for any non-linear functions in regression networks, for **different equations** and **distributions**.	Extension of calibration methods: **Relative Per-coordinate**, **per-class calibration** with isotonic regression produces **well-calibrated uncertainties**.	› Negative correlation with **detection performance** and **image quality**. › Positive correlation with **occlusion** and **object distance** in the real-world.

Future Work

- **Enhancing** object detection **safety** and **robustness**.
- **Deeper** model **understanding** with **more** uncertainty **types**.

Summary

Estimation | Propagation | Calibration | Explanation

Estimation

Loss attenuation for EfficientDet with **increased localization performance**.

Propagation

Two decoding methods:
- **Exact** and **fast** propagation.
- **Generalize** for any non-linear functions in regression networks, for **different equations** and **distributions**.

Calibration

Extension of calibration methods: **Relative Per-coordinate**, **per-class calibration** with isotonic regression produces **well-calibrated uncertainties**.

Explanation

- Negative correlation with **detection performance** and **image quality**.
- Positive correlation with **occlusion** and **object distance** in the real-world.

Future Work

- **Enhancing** object detection **safety** and **robustness**.
- **Deeper** model **understanding** with **more** uncertainty **types**.
- **Investigating** uncertainty **across domains**.

Summary

Estimation › **Propagation** › **Calibration** › **Explanation**

Estimation

Loss attenuation for EfficientDet with **increased localization performance**.

Propagation

Two decoding methods:

- **Exact** and **fast** propagation.
- **Generalize** for any non-linear functions in regression networks, for **different equations** and **distributions**.

Calibration

Extension of calibration methods: **Relative Per-coordinate**, **per-class calibration** with isotonic regression produces **well-calibrated uncertainties**.

Explanation

- Negative correlation with **detection performance** and **image quality**.
- Positive correlation with **occlusion** and **object distance** in the real-world.

Paper

Future Work

- **Enhancing** object detection **safety** and **robustness**.
- **Deeper** model **understanding** with **more** uncertainty **types**.
- **Investigating** uncertainty **across domains**.

Appendix
State of the Art

Open Questions

- How is the **output distribution N(μ, σ^2)** propagated through **non-linear** functions?
- Is the predicted **uncertainty well-calibrated**?
- What **correlations** exist between the **data** and **uncertainty**?

State of the Art

Appendix
State of the Art

Open Questions

- How is the **output distribution N(μ, σ^2)** propagated through **non-linear** functions?
- Is the predicted **uncertainty well-calibrated**?
- What **correlations** exist between the **data** and **uncertainty**?

State of the Art

- **Incorrect** propagation, **no mention**, **only** used during **training** [1,2,3,4]**.**
- **Sampling** [5]**.**

Missing propagation approach

Appendix
State of the Art

Open Questions

- How is the **output distribution N(μ, σ^2)** propagated through **non-linear** functions?
- Is the predicted **uncertainty well-calibrated**?
- What **correlations** exist between the **data** and **uncertainty**?

State of the Art

- **Incorrect** propagation, **no mention**, **only** used during **training** [1,2,3,4].
- **Sampling** [5].
- Uncertainty is **biased** [6,7,8].
- **Recalibration** via isotonic regression [4,8,9] and temperature scaling [7,10].

Missing propagation approach
Calibration not adapted to localization

Appendix
State of the Art

Open Questions

- How is the **output distribution N(μ, σ^2)** propagated through **non-linear** functions?
- Is the predicted **uncertainty well-calibrated**?
- What **correlations** exist between the **data** and **uncertainty**?

State of the Art

- **Incorrect** propagation, **no mention**, **only** used during **training** [1,2,3,4].
- **Sampling** [5].
- Uncertainty is **biased** [6,7,8].
- **Recalibration** via isotonic regression [4,8,9] and temperature scaling [7,10].
- **Occlusion**[4,11,12].
- **Distance** in LiDAR data[11,12].
- **Not** with detection accuracy [11,12].

Missing propagation approach
Calibration not adapted to localization
Uncertainty still unclear

Appendix
Propagation Methods

Normalizing Flows

A **normalizing flow** is a **transformation** of a **distribution** via a **sequence** of invertible and differentiable **mappings** [13]. *

$$h = \mathrm{e}^{\hat{h}} h_a$$

$$\mathrm{g}_1(\mathbf{y}) = \exp(\mathbf{y})$$
$$\mathrm{g}_2(\mathrm{y}) = c\mathrm{y} \text{ and } c \in \mathbb{R}$$
$$h = \mathrm{g}_2 \circ \mathrm{g}_1(\hat{h}) \text{ with } c = h_a$$

* $$\begin{aligned} p_{\mathbf{Y}}(\mathbf{y}) &= p_{\mathbf{Z}}(\mathbf{f}(\mathbf{y}))|\det \mathrm{Df}(\mathbf{y})| \\ &= p_{\mathbf{Z}}(\mathbf{f}(\mathbf{y}))|\det \mathrm{Dg}(\mathbf{f}(\mathbf{y}))|^{-1} \end{aligned}$$

Appendix
Propagation Methods

Normalizing Flows

A **normalizing flow** is a **transformation** of a **distribution** via a **sequence** of invertible and differentiable **mappings** [13]. *

$$h = \mathrm{e}^{\hat{h}} h_a$$

$$\mathrm{g}_1(\mathbf{y}) = \exp(\mathbf{y})$$
$$\mathrm{g}_2(\mathrm{y}) = c\mathrm{y} \text{ and } c \in \mathbb{R}$$
$$h = \mathrm{g}_2 \circ \mathrm{g}_1(\hat{h}) \text{ with } c = h_a$$

Log-Normal Approach

If $\hat{\boldsymbol{h}}$ follows a **normal** distribution, then $\mathbf{e}^{\hat{\boldsymbol{h}}}$ follow a **log-normal** distribution [14].

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma(x)} e^{\frac{-[\log(x)-\mu]^2}{2\sigma^2}}$$

$$E[\mathrm{e}^{\hat{h}}] = \mathrm{e}^{\mu_{\hat{h}} + \frac{\sigma_{\hat{h}}^2}{2}}$$
$$\mathrm{Var}[\mathrm{e}^{\hat{h}}] = \left[\mathrm{e}^{\sigma_{\hat{h}}^2} - 1\right] \mathrm{e}^{2\mu_{\hat{h}} + \sigma_{\hat{h}}^2}$$
$$\mu_h = E[\mathrm{e}^{\hat{h}}] \cdot h_a$$
$$\sigma_h^2 = \mathrm{Var}[\mathrm{e}^{\hat{h}}] \cdot h_a^2$$

$$^* \; p_{\mathbf{Y}}(\mathbf{y}) = p_{\mathbf{Z}}(\mathbf{f}(\mathbf{y}))|\det \mathrm{D}\mathbf{f}(\mathbf{y})| = p_{\mathbf{Z}}(\mathbf{f}(\mathbf{y}))|\det \mathrm{D}\mathbf{g}(\mathbf{f}(\mathbf{y}))|^{-1}$$

Appendix
Calibration Example Analysis

Appendix
Calibration Example Analysis

Calibration reduces uncertainty

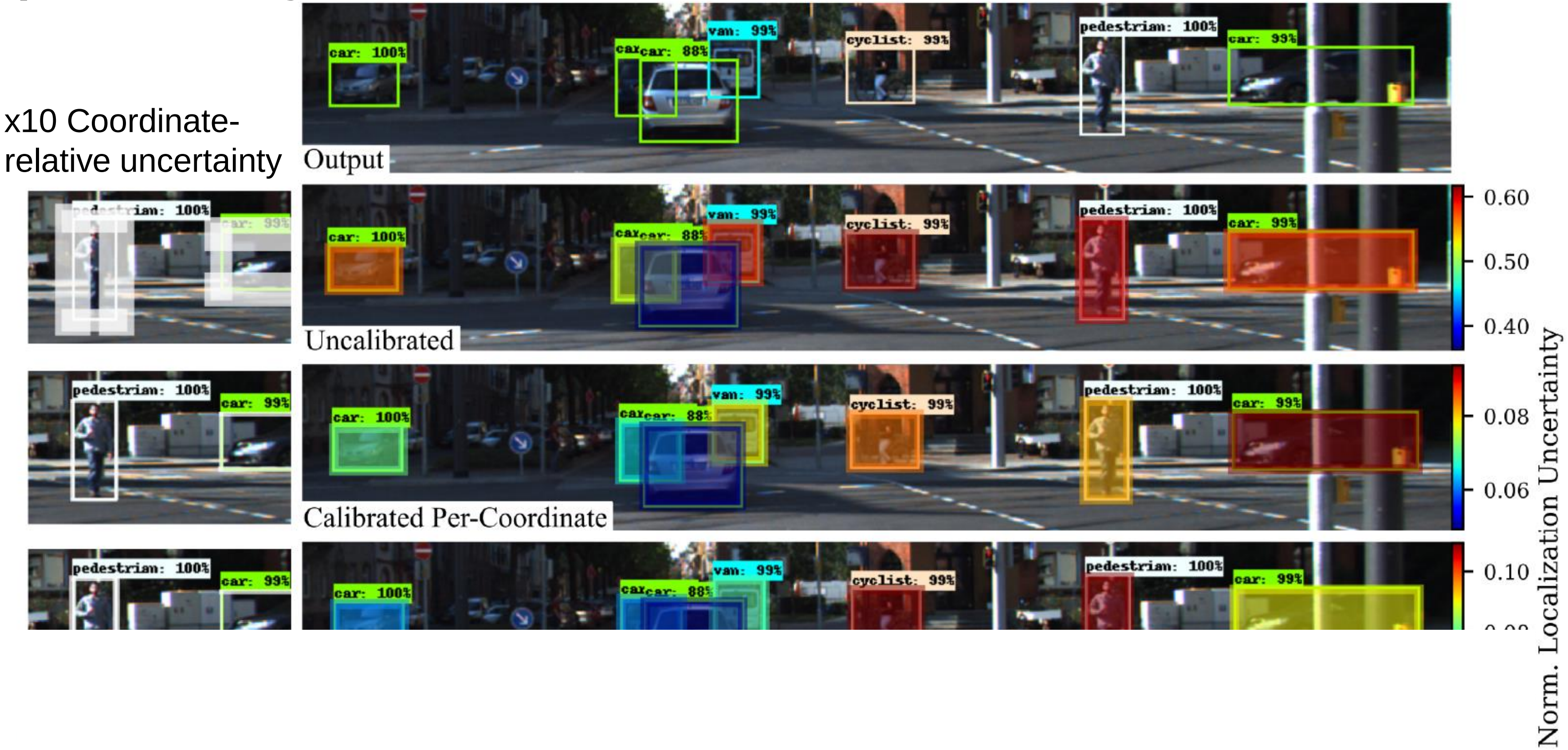
x10 Coordinate-relative uncertainty

Appendix
Calibration Example Analysis

Calibration reduces uncertainty

Per-class calibration shifts uncertainty towards classes with lower performance

x10 Coordinate-relative uncertainty



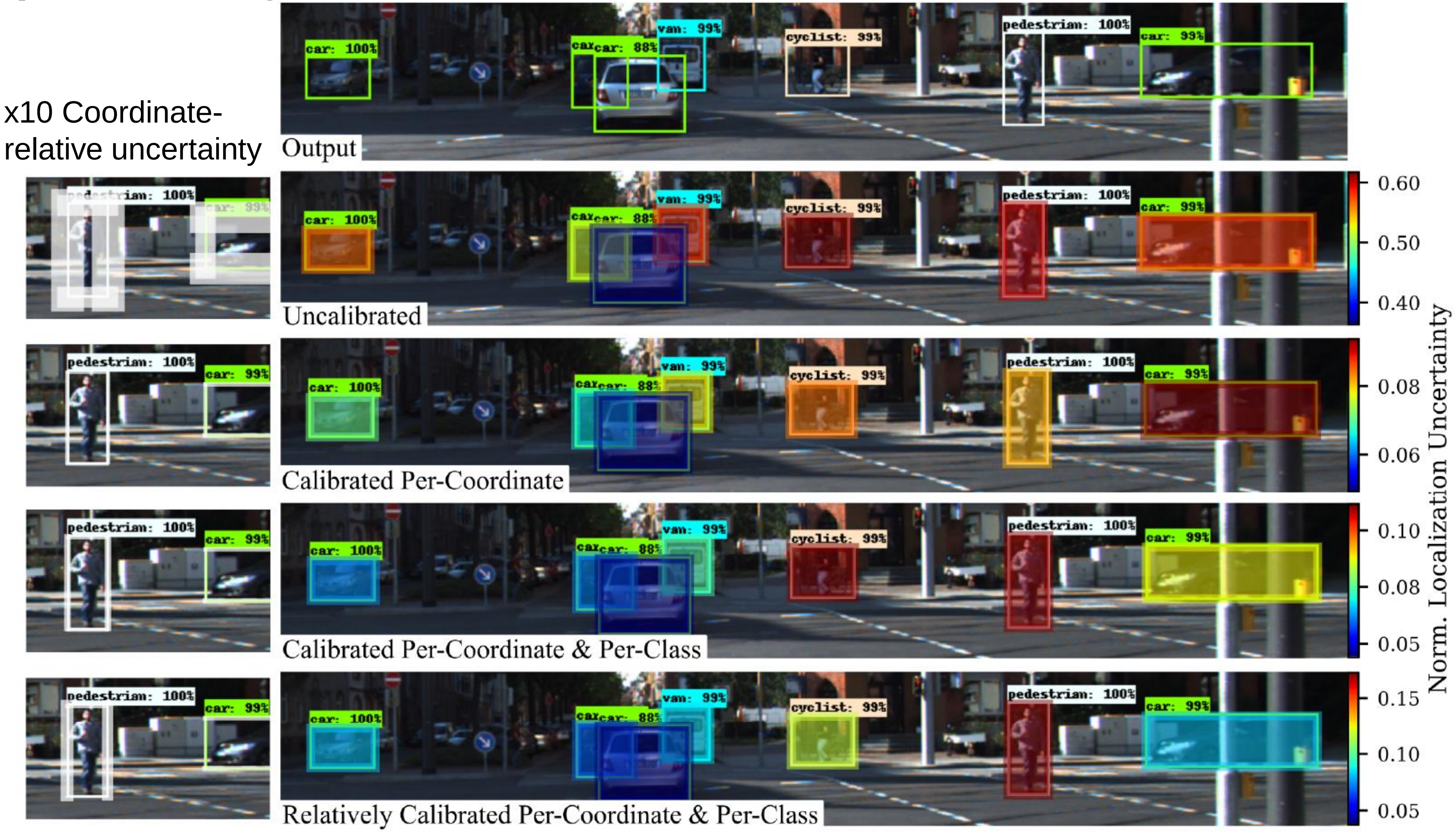
Appendix
Calibration Example Analysis

Calibration reduces uncertainty

Per-class calibration shifts uncertainty towards classes with lower performance

Relative calibration considers the effect of the different aspect ratios

x10 Coordinate-relative uncertainty



Appendix
BRISQUE Score

The Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) is a no-reference image quality assessment algorithm.

Original Image	JPEG2K Compression	Heavy Compression	Gaussian Noise	Median Blur
				
26.8286	30.7417	33.0692	79.8751	72.7044

BRISQUE:

- Is based on statistical models of natural image features, which are extracted from the image using a set of spatial and transform domain operators. The operators include discrete cosine transform (DCT), discrete wavelet transform (DWT), and local binary patterns (LBP), among others.
- Computes a score from the set of feature vectors that represents the degree of naturalness of the image, which correlates with the subjective quality of the image.
- Is a machine learning model, typically a Gaussian process or a support vector regression (SVR) model, trained on a large dataset of natural images TID2008, which were annotated with mean opinion scores (MOS) obtained from subjective experiments where human observers rated the perceived image quality.
- Ranges from 0 to 100, with higher scores indicating lower image quality and lower scores indicating higher image quality.

[15,16]

Appendix
Sources

1. Jiwoong Choi, Dayoung Chun, Hyun Kim, and Hyuk-Jae Lee. Gaussian yolov3: An accurate and fast object detector using localization uncertainty for autonomous driving. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 502–511, 2019.
2. Ali Harakeh, Michael Smart, and Steven L Waslander. Bayesod: A bayesian approach for uncertainty estimation in deep object detectors. In 2020 IEEE International Conference on Robotics and Automation (ICRA), pages 87–93. IEEE, 2020.
3. Yihui He, Chenchen Zhu, Jianren Wang, Marios Savvides, and Xiangyu Zhang. Bounding box regression with uncertainty for accurate object detection. In Proceedings of the ieee/cvf conference on computer vision and pattern recognition, pages 2888–2897, 2019.
4. Florian Kraus and Klaus Dietmayer. Uncertainty estimation in one-stage object detection. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC), pages 53– 60. IEEE, 2019.
5. Michael Truong Le, Frederik Diehl, Thomas Brunner, and Alois Knol. Uncertainty estimation for deep neural object detectors in safety-critical applications. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC), pages 3873–3878. IEEE, 2018.
6. Di Feng, Lars Rosenbaum, Claudius Glaeser, Fabian Timm, and Klaus Dietmayer. Can we trust you? on calibration of a probabilistic object detector for autonomous driving. arXiv preprint arXiv:1909.12358, 2019.
7. Max-Heinrich Laves, Sontje Ihler, Jacob F Fast, L¨uder A Kahrs, and Tobias Ortmaier. Recalibration of aleatoric and epistemic regression uncertainty in medical imaging. arXiv preprint arXiv:2104.12376, 2021.
8. Buu Phan, Rick Salay, Krzysztof Czarnecki, Vahdat Abdelzad, Taylor Denouden, and Sachin Vernekar. Calibrating uncertainties in object localization task. arXiv preprint arXiv:1811.11210, 2018.
9. Volodymyr Kuleshov, Nathan Fenner, and Stefano Ermon. Accurate uncertainties for deep learning using calibrated regression. In International conference on machine learning, pages 2796–2804. PMLR, 2018.
10. Max-Heinrich Laves, Sontje Ihler, Karl-Philipp Kortmann, and Tobias Ortmaier. Well-calibrated model uncertainty with temperature scaling for dropout variational inference. arXiv preprint arXiv:1909.13550, 2019.
11. Di Feng, Lars Rosenbaum, and Klaus Dietmayer. Towards safe autonomous driving: Capture uncertainty in the deep neural network for lidar 3d vehicle detection. In 2018 21st international conference on intelligent transportation systems (ITSC), pages 3266–3273. IEEE, 2018.
12. Di Feng, Lars Rosenbaum, Fabian Timm, and Klaus Dietmayer. Leveraging heteroscedastic aleatoric uncertainties for robust real-time lidar 3d object detection. In 2019 IEEE Intelligent Vehicles Symposium (IV), pages 1280–1287. IEEE, 2019.
13. Ivan Kobyzev, Simon JD Prince, and Marcus A Brubaker. Normalizing flows: An introduction and review of current methods. IEEE transactions on pattern analysis and machine intelligence, 43(11):3964–3979, 2020.
14. William W. S. Balakrishnan, N.and Chen. Lognormal Distributions and Properties, pages 5–6. Springer US, Boston, MA, 1999.
15. Mittal, Anish, Anush Krishna Moorthy, and Alan Conrad Bovik. No-reference image quality assessment in the spatial domain. IEEE Transactions on image processing 21.12 (2012): 4695-4708.
16. Kushashwa Ravi Shrimali. Image Quality Assessment: BRISQUE. URL: https://learnopencv.com/image-quality-assessment-brisque/. Accessed: 06.09.2023