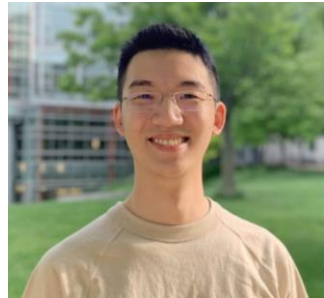


Provable Defense Against Geometric Transformations

Rem Yang, Jacob Laurel, Sasa Misailovic, Gagandeep Singh



UNIVERSITY OF
ILLINOIS
URBANA-CHAMPAIGN

vmware®

Vulnerability of Deep Neural Networks

Vulnerability of Deep Neural Networks

Correctly classified



x

Vulnerability of Deep Neural Networks

Perturbed: misclassified

Correctly classified

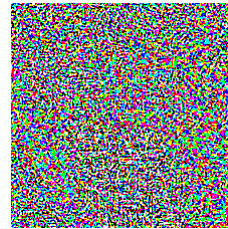


x



ℓ_p perturbations

+



=



δ where $\|\delta\|_p < \epsilon$

$x + \delta$

Vulnerability of Deep Neural Networks

Perturbed: misclassified

Correctly classified



x



ℓ_p perturbations

$$+ \begin{array}{c} \text{[Noise Image]} \\ \delta \text{ where } \|\delta\|_p < \epsilon \end{array} = \begin{array}{c} \text{[Kitten Image]} \\ x + \delta \end{array}$$

Geometric transformations



Rotation



Scaling



Shearing



Contrast



Brightness

Vulnerability of Deep Neural Networks

Perturbed: misclassified

Correctly classified



x



ℓ_p perturbations

$$+ \begin{array}{c} \text{[Noise Image]} \\ \delta \text{ where } \|\delta\|_p < \epsilon \end{array} = \begin{array}{c} \text{[Kitten Image]} \\ x + \delta \end{array}$$

Geometric transformations



Rotation



Scaling



Shearing



Contrast



Brightness

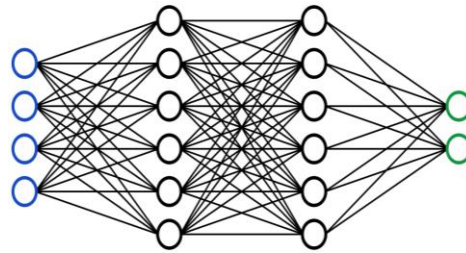
Certified Robustness

Certified Robustness

Set of perturbed images X'



Classifier f

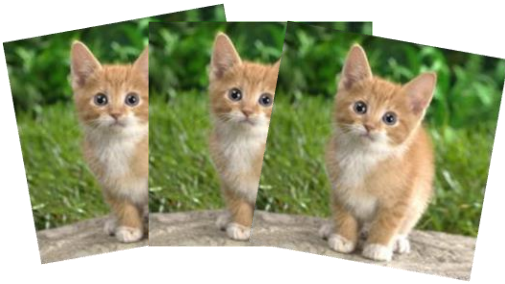


Provably correct?

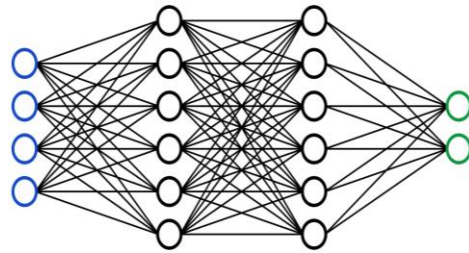
$$y = \operatorname{argmax}_i f_i(x') \quad \forall x' \in X'$$

Certified Robustness

Set of perturbed images X'



Classifier f



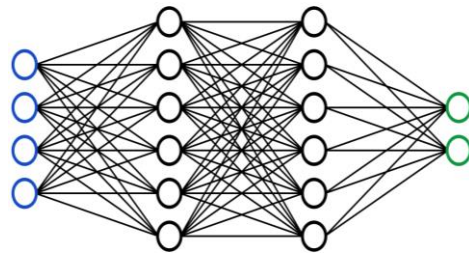
Provably correct?

$$y = \operatorname{argmax}_i f_i(x') \quad \forall x' \in X'$$

Set of perturbed images X'



Regression net f



Certified output bounds

$$\underline{y} \leq \min_{x' \in X'} f(x') \leq \max_{x' \in X'} f(x') \leq \bar{y}$$

Geometric Robustness Verification

Probabilistic (Fischer et al., 2020; Hao et al., 2022; Li et al., 2021)	Deterministic (Balunovic et al., 2019; Mohapatra et al., 2020)
Scales to larger datasets	Only scaled up to CIFAR-10
Large inference overhead	No inference overhead

Geometric Robustness Verification

Probabilistic (Fischer et al., 2020; Hao et al., 2022; Li et al., 2021)	Deterministic (Balunovic et al., 2019; Mohapatra et al., 2020)
Scales to larger datasets	Only scaled up to CIFAR-10
Large inference overhead	No inference overhead

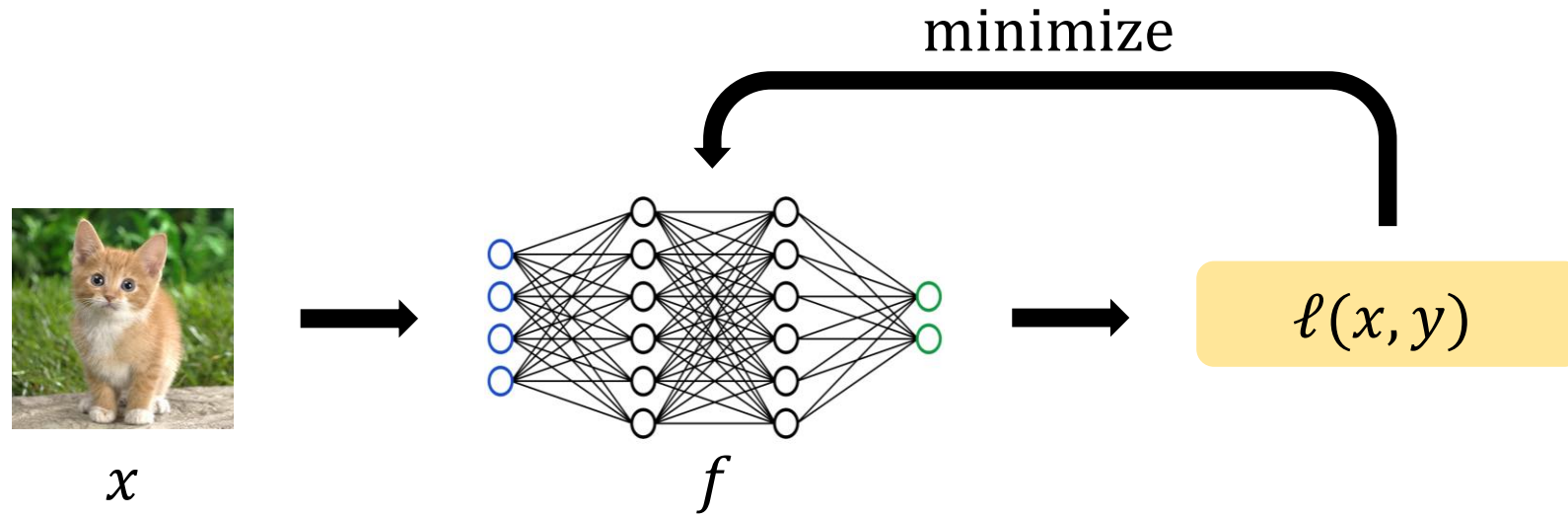
Geometric Robustness Verification

Probabilistic (Fischer et al., 2020; Hao et al., 2022; Li et al., 2021)	Deterministic (Balunovic et al., 2019; Mohapatra et al., 2020)
Scales to larger datasets	Only scaled up to CIFAR-10
Large inference overhead	No inference overhead

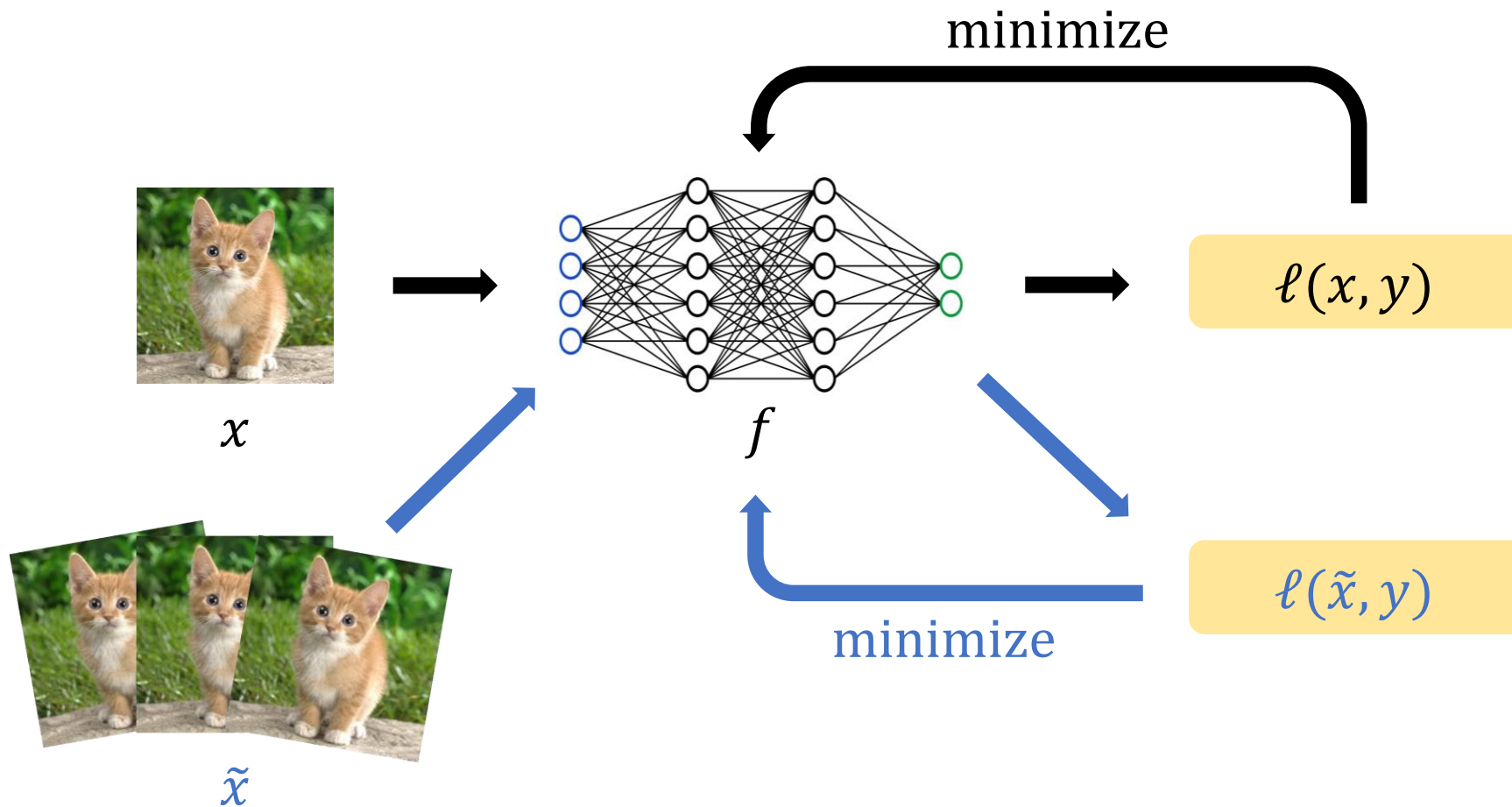
These works only verify networks not explicitly trained to be provably robust

Provable Defense

Provable Defense



Provable Defense



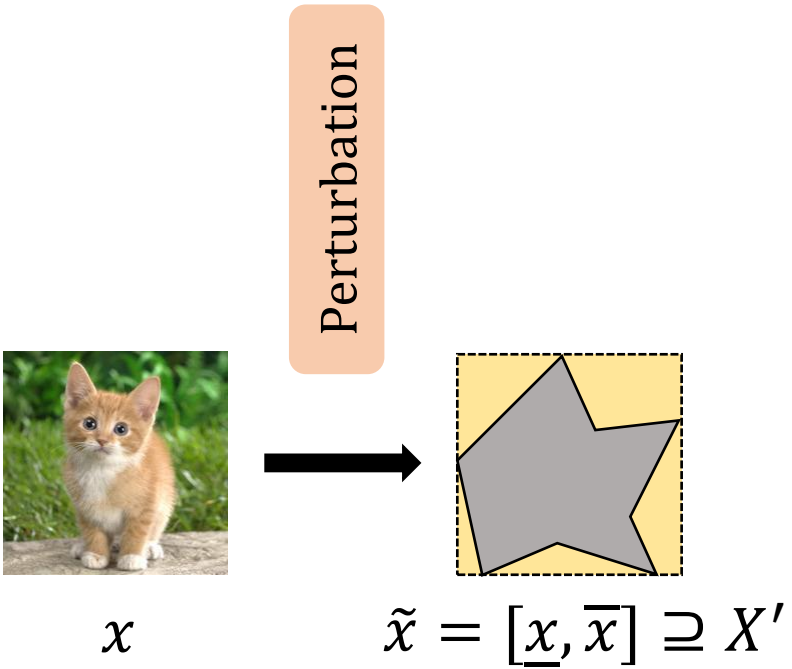
Interval Bound Propagation

Interval Bound Propagation

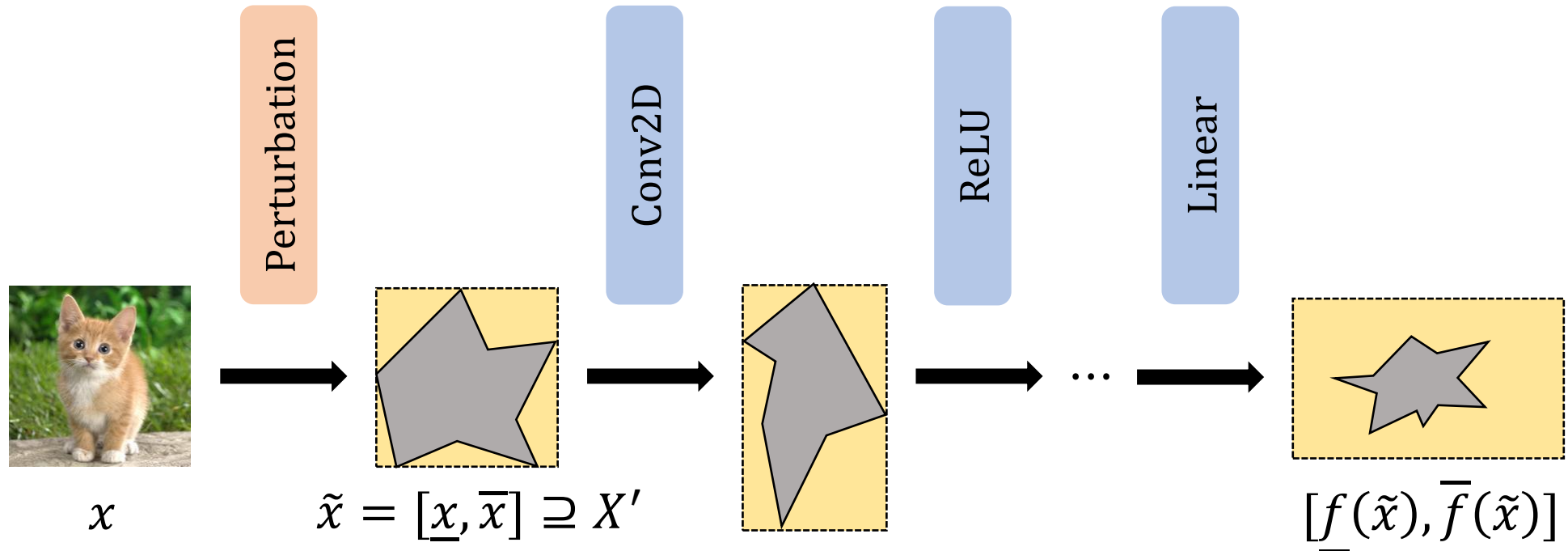


x

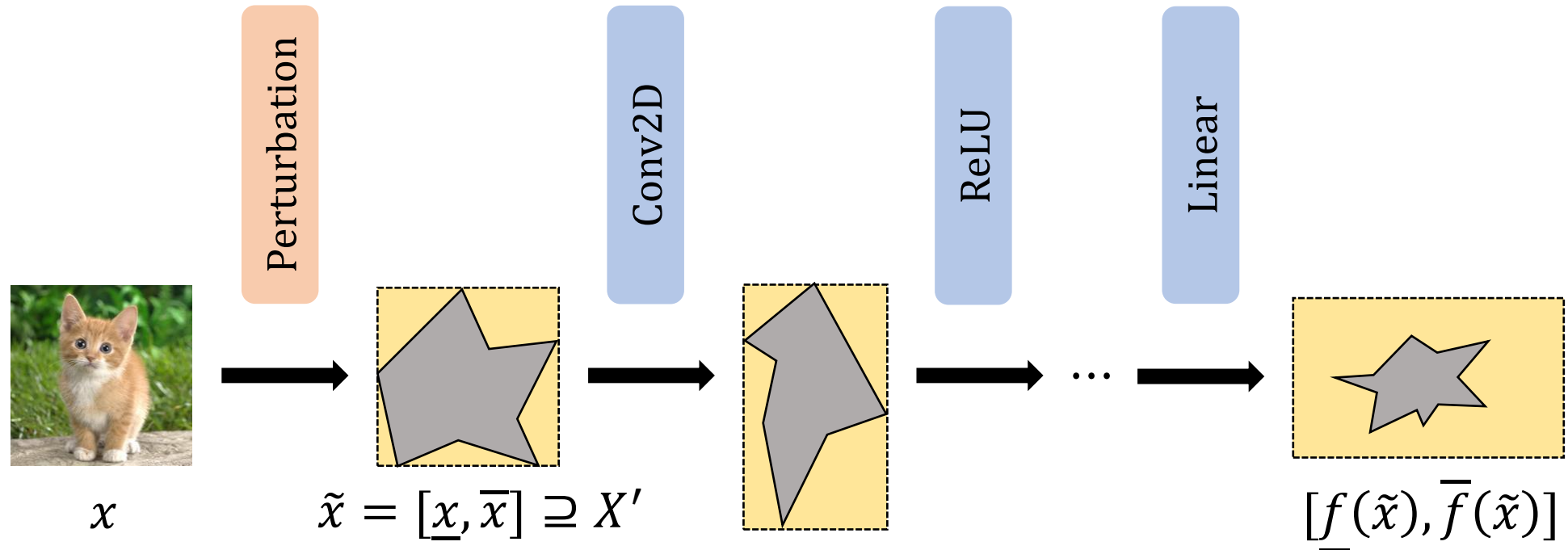
Interval Bound Propagation



Interval Bound Propagation

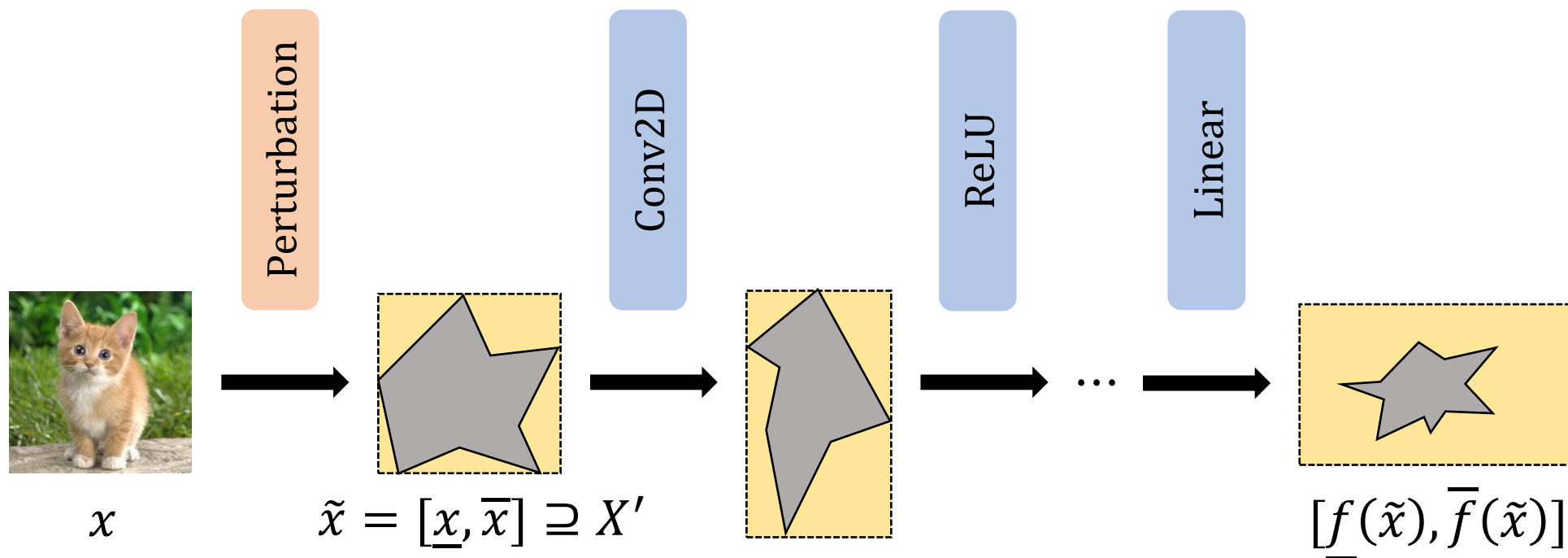


Interval Bound Propagation



Classification: $\underline{f}_y(\tilde{x}) > \bar{f}_j(\tilde{x}) \forall j \neq y$

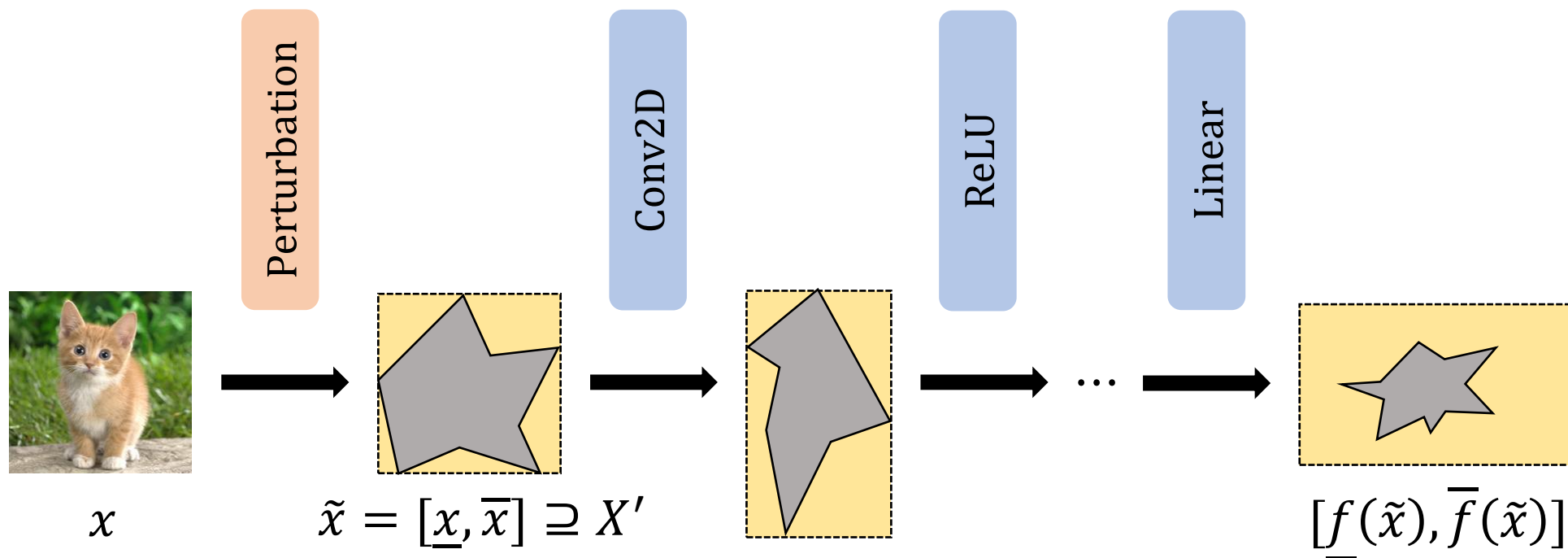
Interval Bound Propagation



Classification: $f_y(\tilde{x}) > \bar{f}_j(\tilde{x}) \forall j \neq y$

Regression: directly use obtained bounds

Interval Bound Propagation



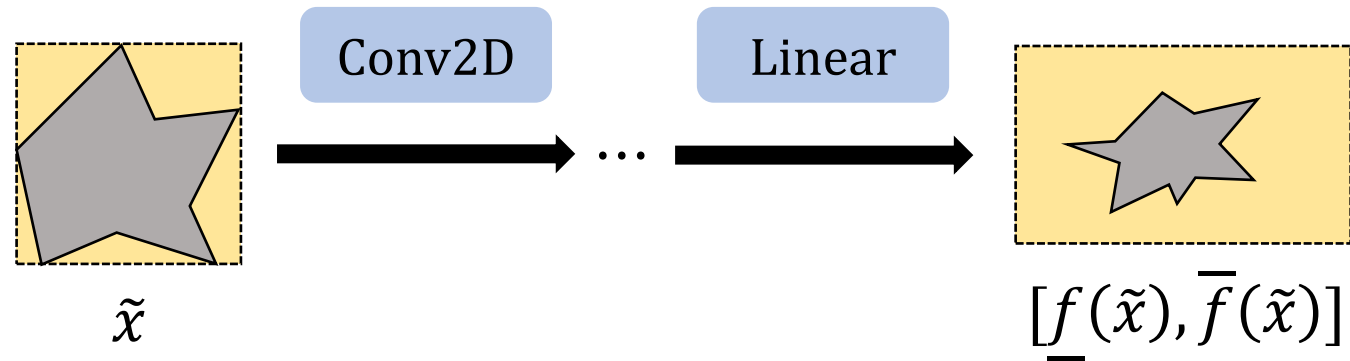
Classification: $f_y(\tilde{x}) > \bar{f}_j(\tilde{x}) \forall j \neq y$

Regression: directly use obtained bounds

Existing works* only handle perturbations with simple formulas, e.g., $\tilde{x} = [x - \epsilon, x + \epsilon]$

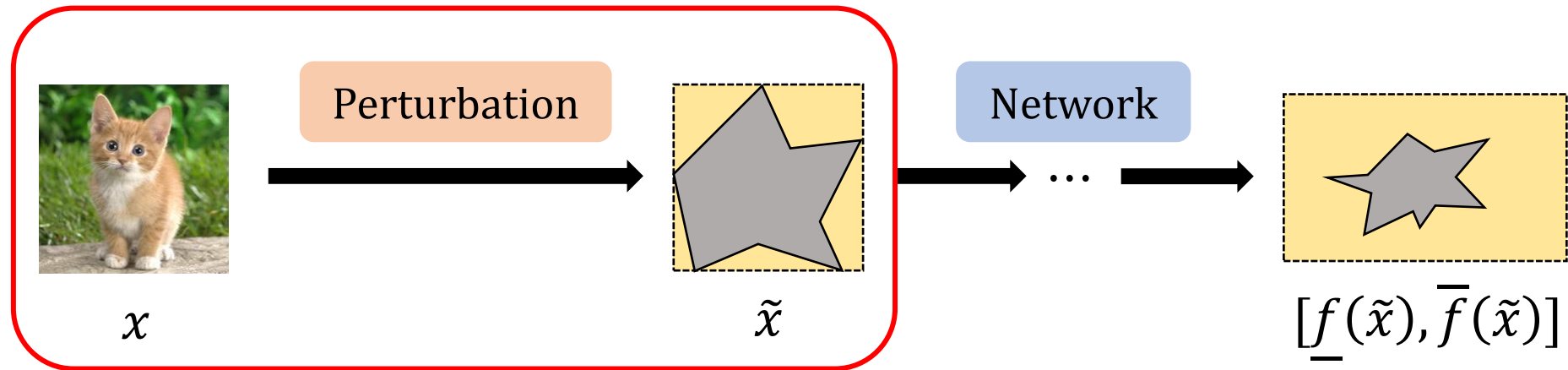
* (Gowal et al., 2019; Mirman et al., 2018; Xu et al., 2020; Zhang et al., 2020)

Interval Bound Propagation



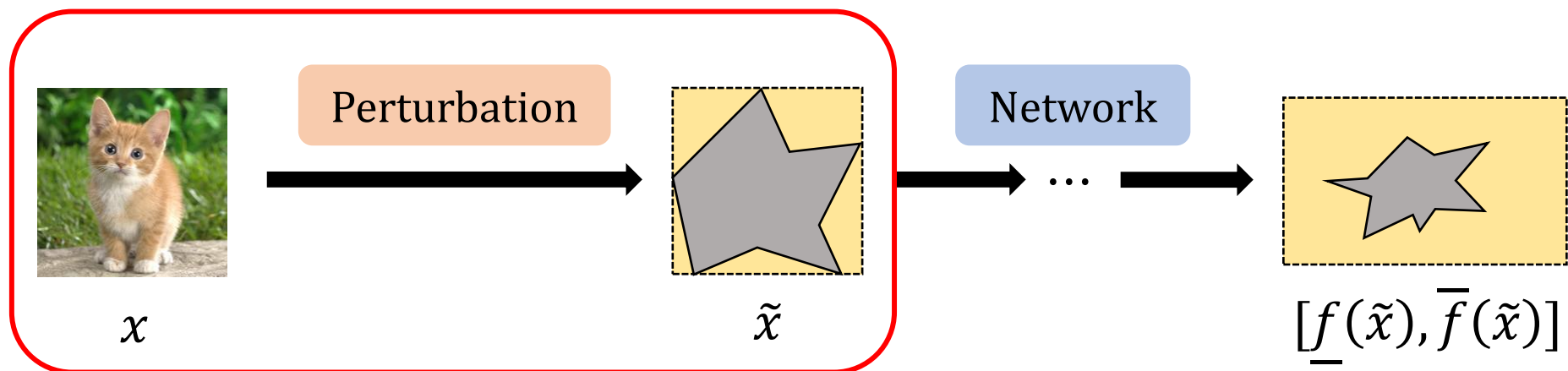
Dataset	Time to Propagate Bounds (s)
CIFAR-10	0.004
Tiny ImageNet	0.018

Interval Bound Propagation



Dataset	Time to Compute Bounds (s)	Time to Propagate Bounds (s)
CIFAR-10	22.81	0.004
Tiny ImageNet	62.83	0.018

Interval Bound Propagation



Dataset	Time to Compute Bounds (s)	Time to Propagate Bounds (s)
CIFAR-10	22.81	0.004
Tiny ImageNet	62.83	0.018

Need faster way to compute geometric perturbation bounds on GPU

Geometric Certification with Splitting

Geometric Certification with Splitting

Define geometric transformation $P: \mathbb{R}^{C \times H \times W} \times \mathbb{R}^{|\theta|} \rightarrow \mathbb{R}^{C \times H \times W}$ and interval range of parameters $\tilde{\theta} = [\underline{\theta}, \bar{\theta}]$

Geometric Certification with Splitting

Define geometric transformation $P: \mathbb{R}^{C \times H \times W} \times \mathbb{R}^{|\theta|} \rightarrow \mathbb{R}^{C \times H \times W}$ and interval range of parameters $\tilde{\theta} = [\underline{\theta}, \bar{\theta}]$



\underline{x}

\bar{x}

$$\tilde{x} = P(x, \tilde{\theta})$$

Geometric Certification with Splitting

Define geometric transformation $P: \mathbb{R}^{C \times H \times W} \times \mathbb{R}^{|\theta|} \rightarrow \mathbb{R}^{C \times H \times W}$ and interval range of parameters $\tilde{\theta} = [\underline{\theta}, \bar{\theta}]$



\underline{x}



\bar{x}

$$\tilde{x} = P(x, \tilde{\theta})$$



\underline{x}_1



\bar{x}_1

$$\tilde{x}_1 = P(x, \tilde{\theta}_1)$$

...



\underline{x}_K



\bar{x}_K

$$\tilde{x}_K = P(x, \tilde{\theta}_K)$$

Geometric Certification with Splitting

Define geometric transformation $P: \mathbb{R}^{C \times H \times W} \times \mathbb{R}^{|\theta|} \rightarrow \mathbb{R}^{C \times H \times W}$ and interval range of parameters $\tilde{\theta} = [\underline{\theta}, \overline{\theta}]$



\underline{x}



\overline{x}

$$\tilde{x} = P(x, \tilde{\theta})$$



\underline{x}_1



\overline{x}_1

$$\tilde{x}_1 = P(x, \tilde{\theta}_1)$$

...



\underline{x}_K



\overline{x}_K

$$\tilde{x}_K = P(x, \tilde{\theta}_K)$$

Classification: $\underline{f}_y(\tilde{x}_k) > \overline{f}_j(\tilde{x}_k) \forall j \neq y \quad \forall k \in \{1, 2, \dots, K\}$

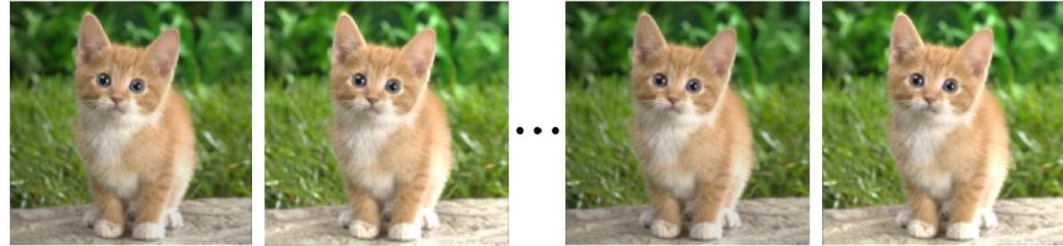
Geometric Certification with Splitting

Define geometric transformation $P: \mathbb{R}^{C \times H \times W} \times \mathbb{R}^{|\theta|} \rightarrow \mathbb{R}^{C \times H \times W}$ and interval range of parameters $\tilde{\theta} = [\underline{\theta}, \bar{\theta}]$



\underline{x} \bar{x}

$$\tilde{x} = P(x, \tilde{\theta})$$



\underline{x}_1 \bar{x}_1 ... \underline{x}_K \bar{x}_K

$$\tilde{x}_1 = P(x, \tilde{\theta}_1)$$

$$\tilde{x}_K = P(x, \tilde{\theta}_K)$$

Classification: $f_{\underline{y}}(\tilde{x}_k) > \overline{f_j}(\tilde{x}_k) \forall j \neq y \quad \forall k \in \{1, 2, \dots, K\}$

Regression: $\cup_{k=1}^K \{ f(\tilde{x}_k) \}$

Geometric Certification with Splitting

Define geometric transformation $P: \mathbb{R}^{C \times H \times W} \times \mathbb{R}^{|\theta|} \rightarrow \mathbb{R}^{C \times H \times W}$ and interval range of parameters $\tilde{\theta} = [\underline{\theta}, \overline{\theta}]$



$$\underline{x} \quad \overline{x}$$
$$\tilde{x} = P(x, \tilde{\theta})$$



$$\underline{x}_1 \quad \overline{x}_1$$
$$\tilde{x}_1 = P(x, \tilde{\theta}_1)$$



$$\underline{x}_K \quad \overline{x}_K$$
$$\tilde{x}_K = P(x, \tilde{\theta}_K)$$

Classification: $f_{\underline{y}}(\tilde{x}_k) > \overline{f_j}(\tilde{x}_k) \forall j \neq y \quad \forall k \in \{1, 2, \dots, K\}$

Regression: $\cup_{k=1}^K \{ f(\tilde{x}_k) \}$

Need to account for splitting in training formulation

Main Contributions

Main Contributions

Certified Geometric Training (CGT)

Main Contributions

Certified Geometric Training (CGT)

- Fast Geometric Verifier (FGV)

Main Contributions

Certified Geometric Training (CGT)

- Fast Geometric Verifier (FGV)
- First provable training formulation for deterministic geometric robustness

Main Contributions

Certified Geometric Training (CGT)

- Fast Geometric Verifier (FGV)
- First provable training formulation for deterministic geometric robustness

Key Empirical Results

Main Contributions

Certified Geometric Training (CGT)

- Fast Geometric Verifier (FGV)
- First provable training formulation for deterministic geometric robustness

Key Empirical Results

- Deterministic geometric certification **60 – 42,000 × faster** than the SOTA

Main Contributions

Certified Geometric Training (CGT)

- Fast Geometric Verifier (FGV)
- First provable training formulation for deterministic geometric robustness

Key Empirical Results

- Deterministic geometric certification **60 – 42,000 × faster** than the SOTA
- Scales beyond CIFAR-10 to **Tiny ImageNet** and **Udacity Self-Driving** datasets

Computing Coordinate in Original Image



Scale up
→



Computing Coordinate in Original Image

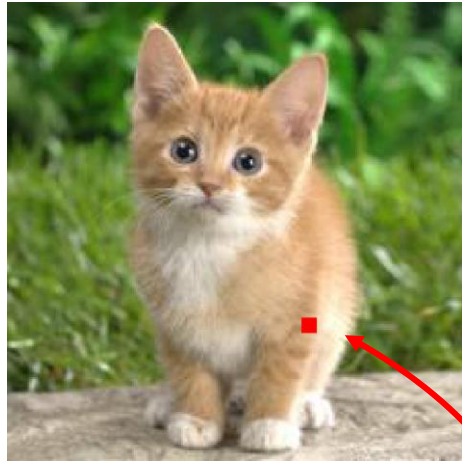


Scale up
→



(i, j)

Computing Coordinate in Original Image



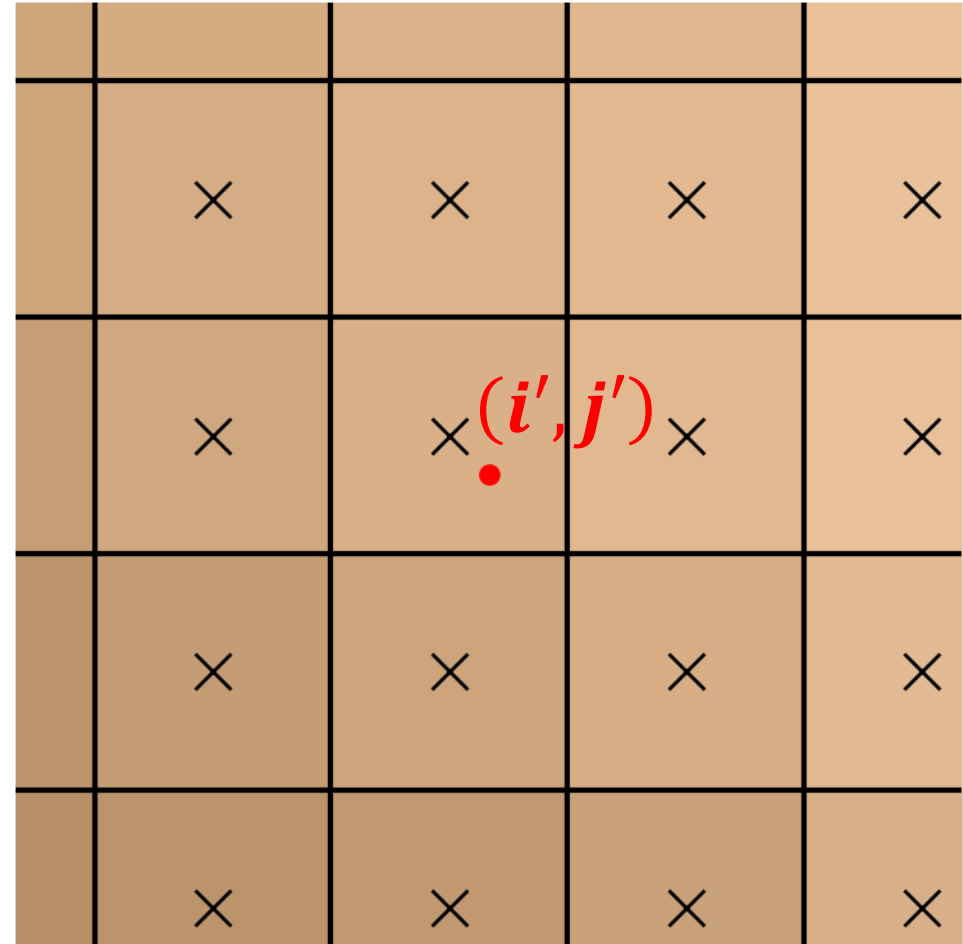
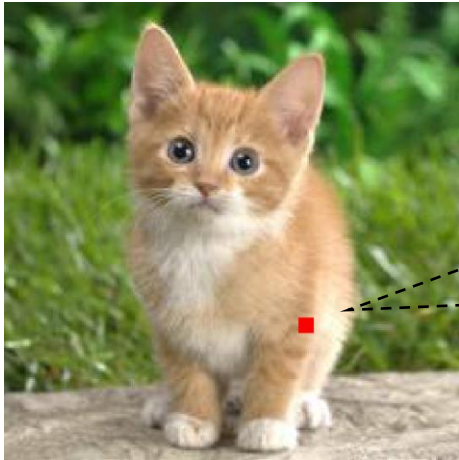
$$(i', j') = T^{-1}_{\theta}(i, j)$$

Scale up
→

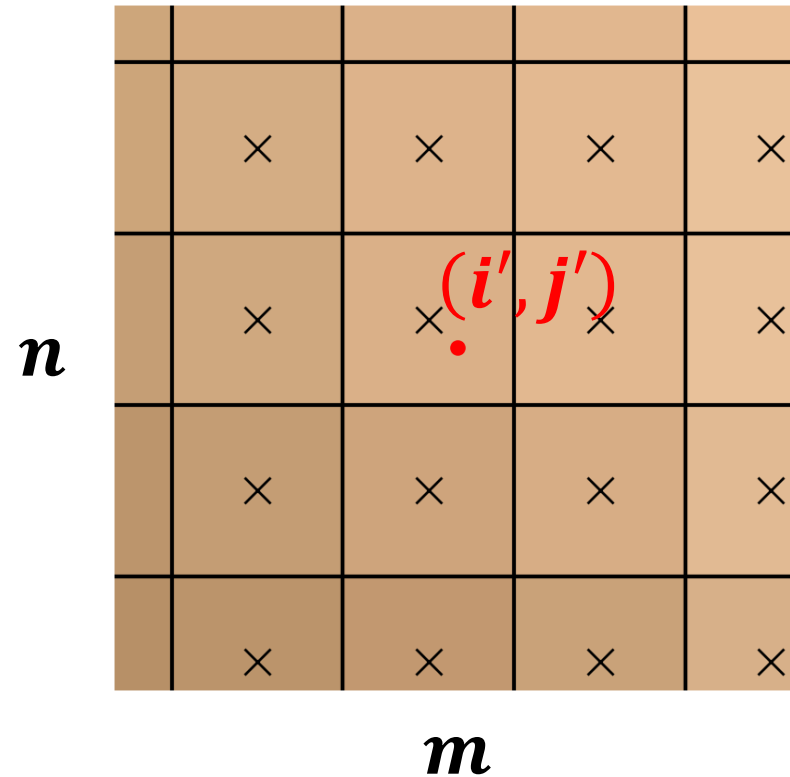


$$(i, j)$$

Coordinates Are Not Integers

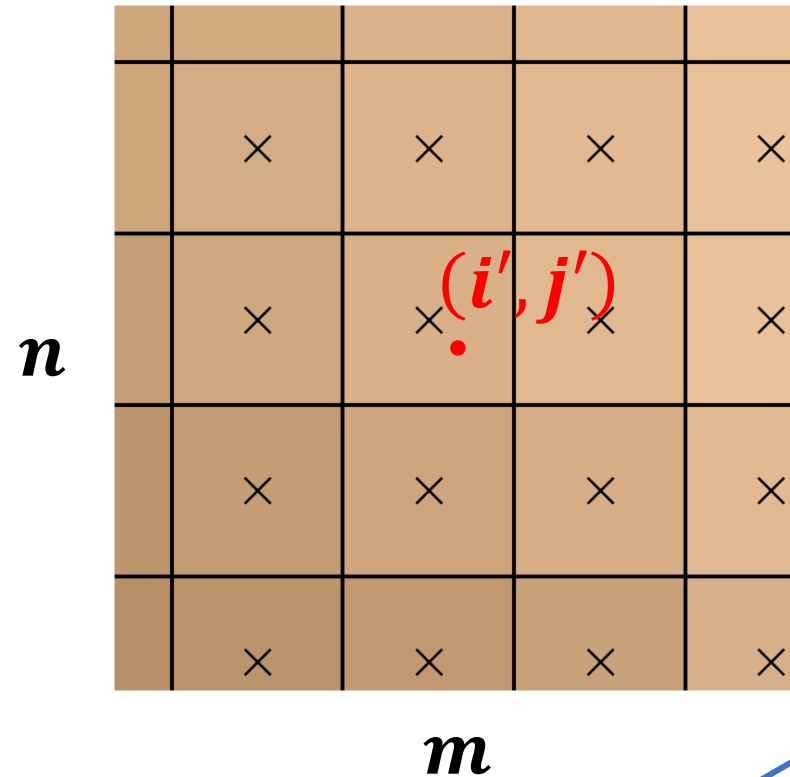


Interpolation



$$x'_{i,j} = \sum_{n=0}^{H-1} \sum_{m=0}^{W-1} x_{n,m} \cdot \max(0, 1 - |i' - n|) \cdot \max(0, 1 - |j' - m|)$$

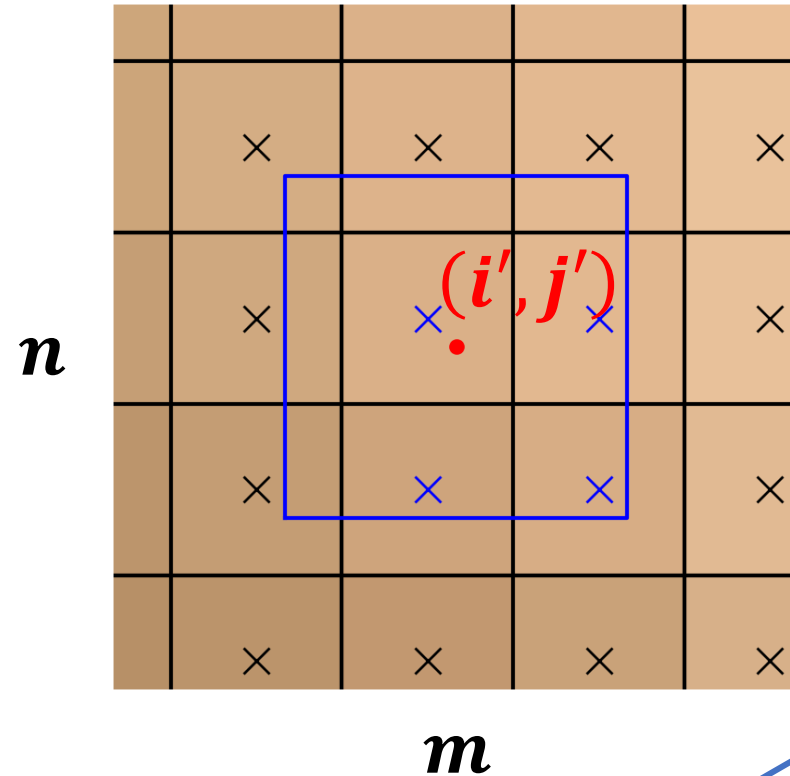
Interpolation



$$x'_{i,j} = \sum_{n=0}^{H-1} \sum_{m=0}^{W-1} x_{n,m} \cdot \max(0, 1 - \underbrace{|i' - n|}_{\text{vertical distance}}) \cdot \max(0, 1 - \underbrace{|j' - m|}_{\text{horizontal distance}})$$

< 1

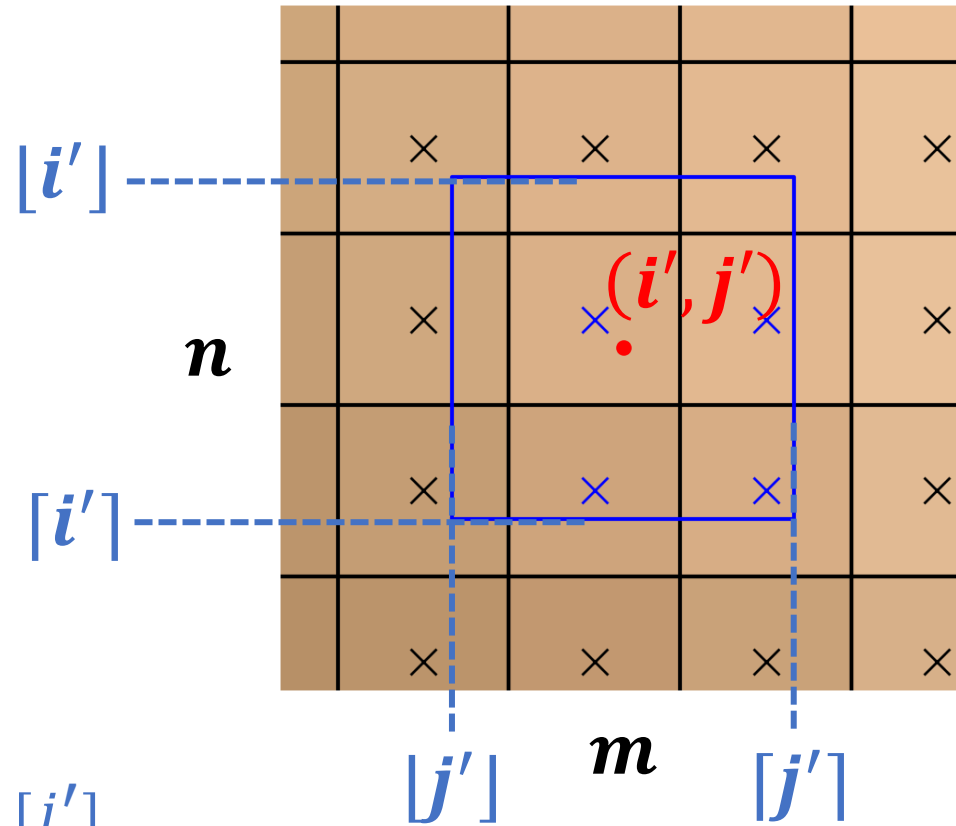
Interpolation



$$x'_{i,j} = \sum_{n=0}^{H-1} \sum_{m=0}^{W-1} x_{n,m} \cdot \max(0, 1 - \underbrace{|i' - n|}_{\text{vertical distance}}) \cdot \max(0, 1 - \underbrace{|j' - m|}_{\text{horizontal distance}})$$

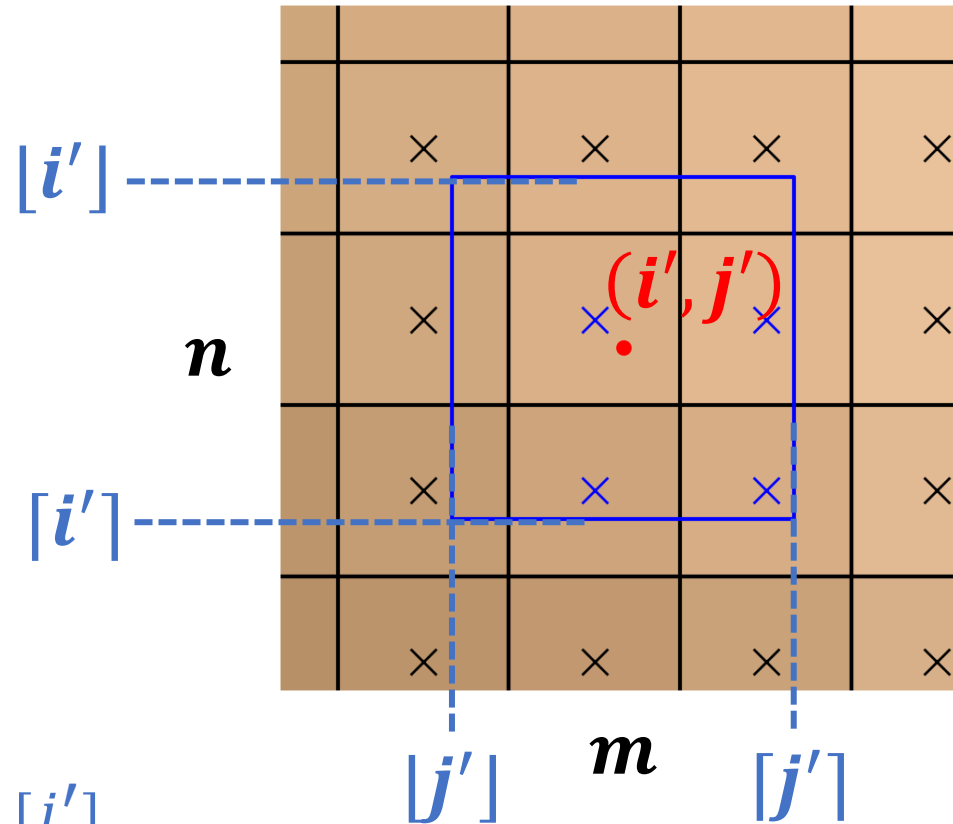
< 1

Interpolation



$$x'_{i,j} = \sum_{n=[i']}^{[i']} \sum_{m=[j']}^{[j']} x_{n,m} \cdot \max(0, 1 - |i' - n|) \cdot \max(0, 1 - |j' - m|)$$

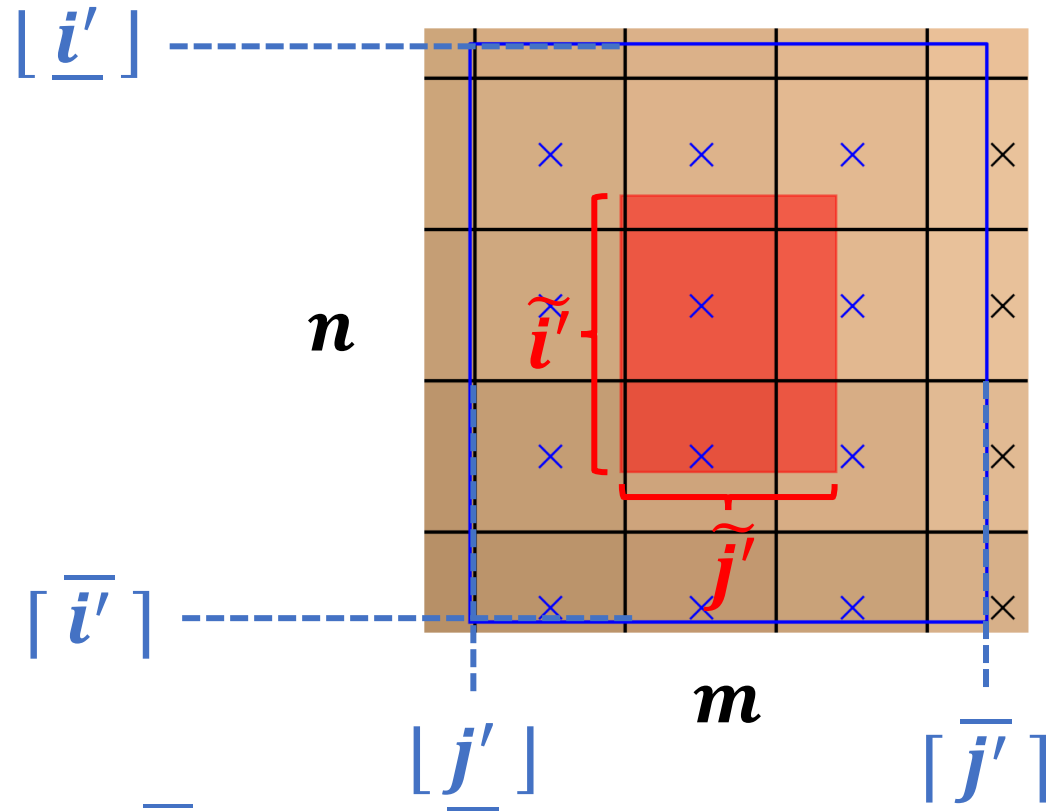
Interpolation



What if we consider an **interval** range of parameters $\tilde{\theta}$?

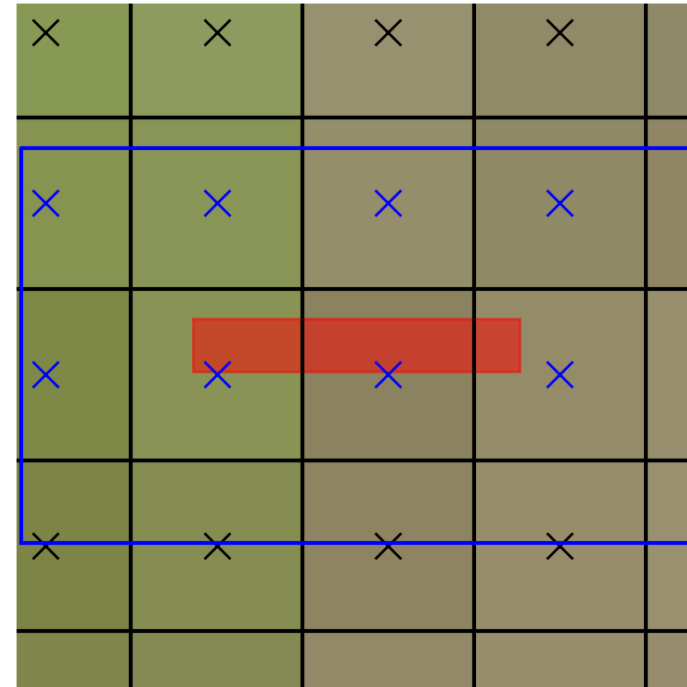
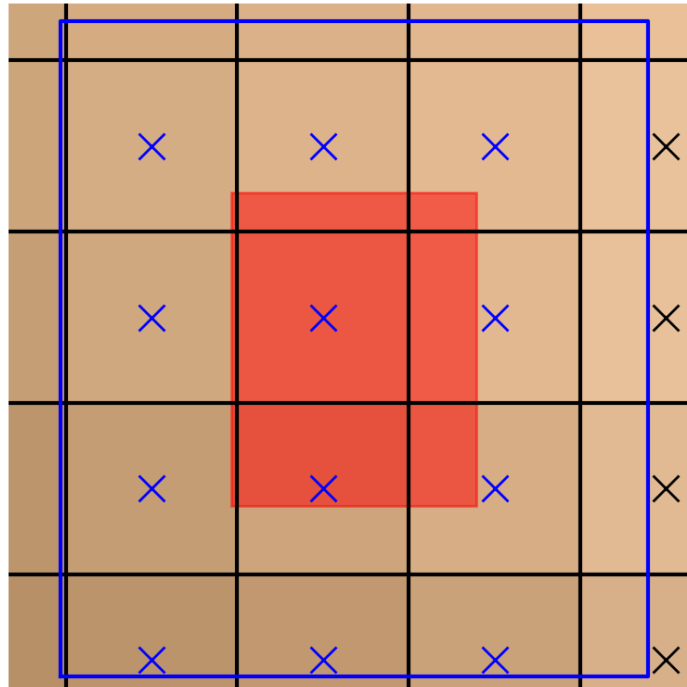
$$x'_{i,j} = \sum_{n=[i']}^{[i']} \sum_{m=[j']}^{[j']} x_{n,m} \cdot \max(0, 1 - |i' - n|) \cdot \max(0, 1 - |j' - m|)$$

Interval Interpolation



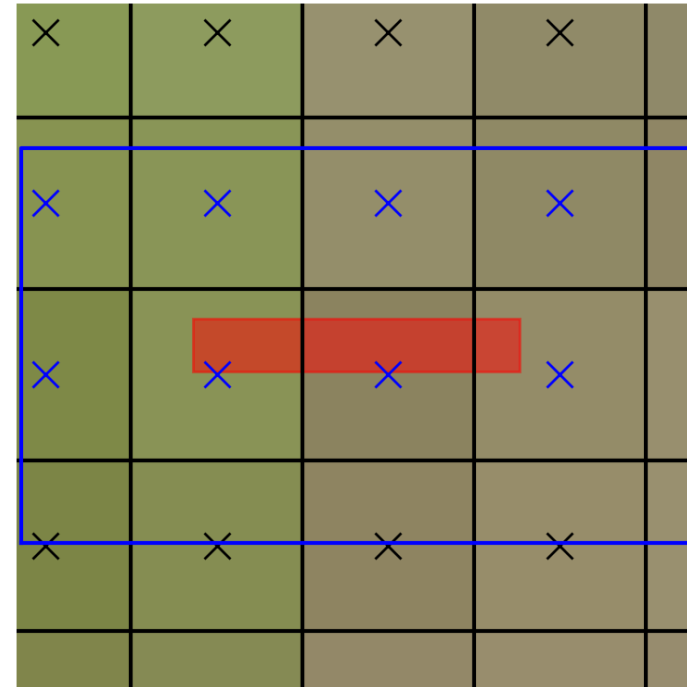
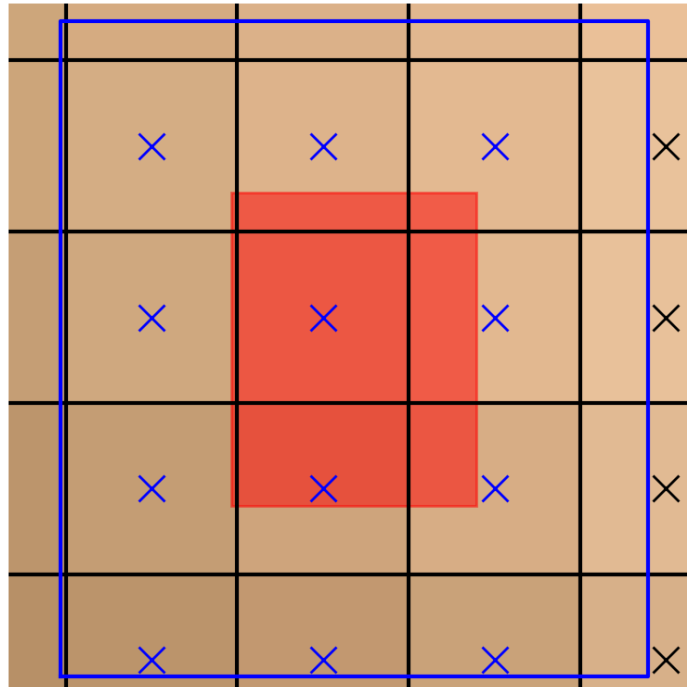
$$\tilde{x}'_{i,j} = \sum_{n=\lfloor \underline{i}' \rfloor}^{\lceil \bar{i}' \rceil} \sum_{m=\lfloor \underline{j}' \rfloor}^{\lceil \bar{j}' \rceil} x_{n,m} \cdot \max(0, 1 - |\tilde{i}' - n|) \cdot \max(0, 1 - |\tilde{j}' - m|)$$

Interval Interpolation



Interpolation region for a different (\tilde{i}', \tilde{j}')

Interval Interpolation



Interpolation region for a different (\tilde{i}', \tilde{j}')

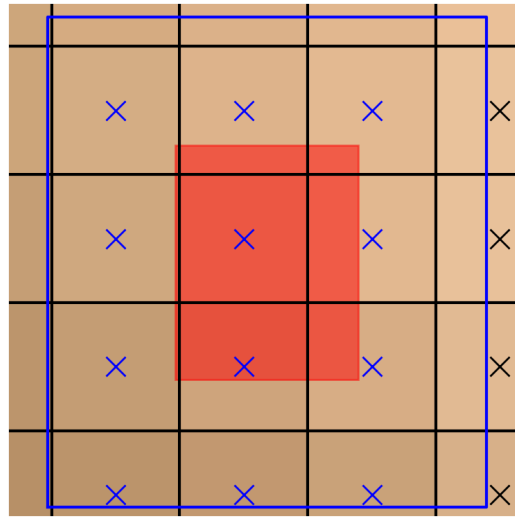
Each pixel's interpolation bounds differ; not GPU-parallelizable!

How do we GPU-parallelize interval interpolation?

How do we GPU-parallelize interval interpolation?

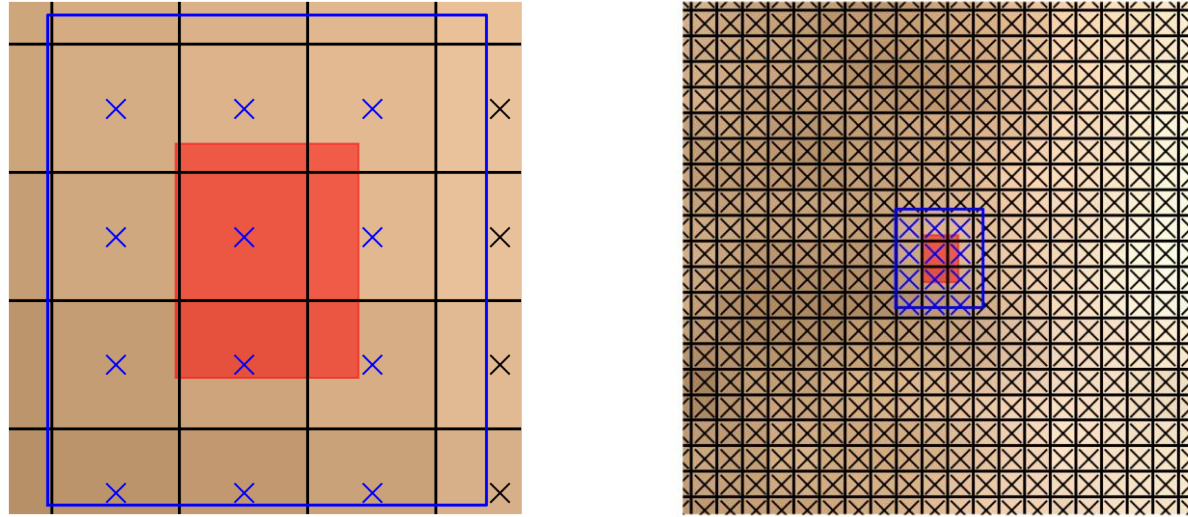
Fast Geometric Verifier

Fast Geometric Verifier



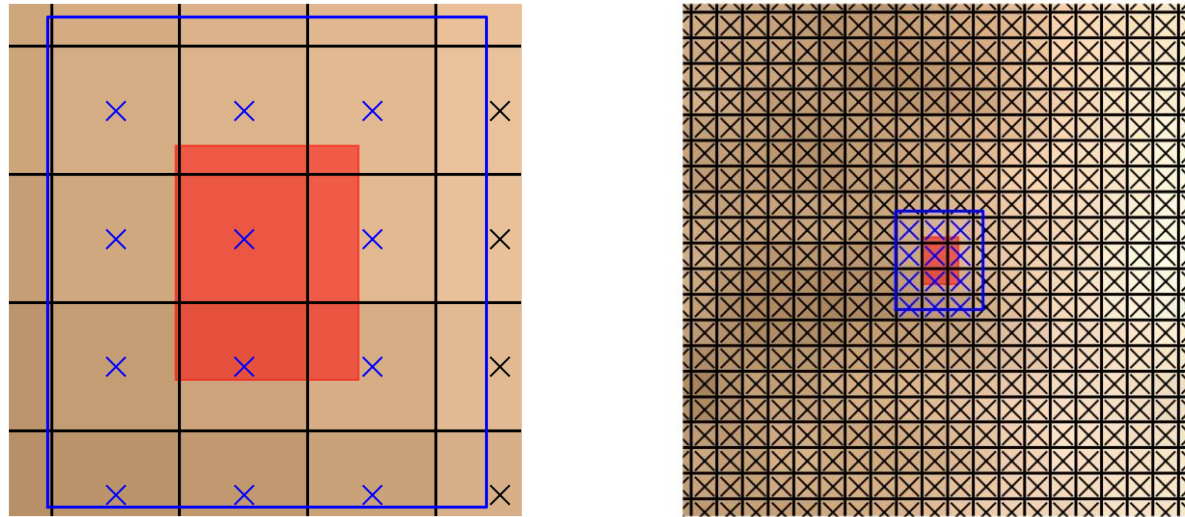
$$\tilde{x}'_{i,j} = \sum_{n=\lfloor \underline{i}' \rfloor}^{\lceil \bar{i}' \rceil} \sum_{m=\lfloor \underline{j}' \rfloor}^{\lceil \bar{j}' \rceil} x_{n,m} \cdot \max(0, 1 - |\tilde{i}' - n|) \cdot \max(0, 1 - |\tilde{j}' - m|)$$

Fast Geometric Verifier



$$\tilde{x}'_{i,j} = \sum_{n=0}^{H-1} \sum_{m=0}^{W-1} x_{n,m} \cdot \max(0, 1 - |\tilde{i}' - n|) \cdot \max(0, 1 - |\tilde{j}' - m|)$$

Fast Geometric Verifier



$$\tilde{x}'_{i,j} = \sum_{n=0}^{H-1} \sum_{m=0}^{W-1} x_{n,m} \cdot \max(0, 1 - |\tilde{i}' - n|) \cdot \max(0, 1 - |\tilde{j}' - m|)$$

Lots of multiplication of pixel values with zero distances!

Fast Geometric Verifier

Decompose expression and precompute interpolation distances

Fast Geometric Verifier

Decompose expression and precompute interpolation distances

$$\tilde{x}'_{i,j} = \sum_{n=0}^{H-1} \sum_{m=0}^{W-1} x_{n,m} \cdot \underbrace{\max(0, 1 - |\tilde{i}' - n|) \cdot \max(0, 1 - |\tilde{j}' - m|)}_{\tilde{d}_{n,m}^{i,j}}$$

Fast Geometric Verifier

Decompose expression and precompute interpolation distances

$$\tilde{x}'_{i,j} = \sum_{n=0}^{H-1} \sum_{m=0}^{W-1} x_{n,m} \cdot \underbrace{\max(0, 1 - |\tilde{i}' - n|) \cdot \max(0, 1 - |\tilde{j}' - m|)}_{\tilde{d}_{n,m}^{i,j}}$$

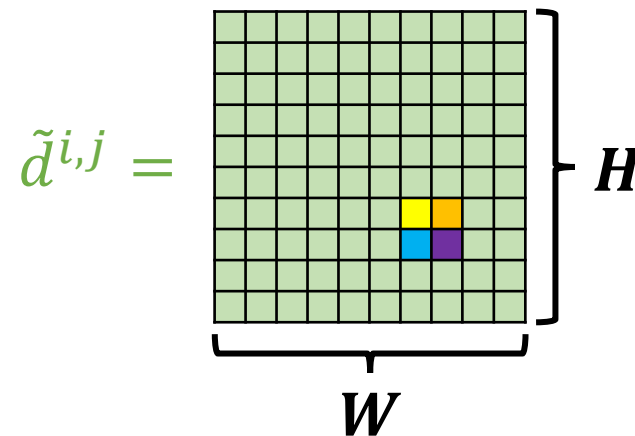
For a given (i, j) :

Fast Geometric Verifier

Decompose expression and precompute interpolation distances

$$\tilde{x}'_{i,j} = \sum_{n=0}^{H-1} \sum_{m=0}^{W-1} x_{n,m} \cdot \underbrace{\max(0, 1 - |\tilde{i}' - n|) \cdot \max(0, 1 - |\tilde{j}' - m|)}_{\tilde{d}_{n,m}^{i,j}}$$

For a given (i, j) :

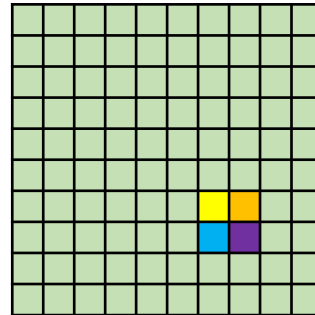


Fast Geometric Verifier

Decompose expression and precompute interpolation distances

$$\tilde{x}'_{i,j} = \sum_{n=0}^{H-1} \sum_{m=0}^{W-1} x_{n,m} \cdot \tilde{d}_{n,m}^{i,j}$$

For a given (i, j) :

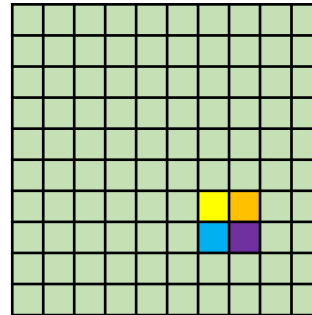


Fast Geometric Verifier

Decompose expression and precompute interpolation distances

$$\tilde{x}'_{i,j} = \sum_{n=0}^{H-1} \sum_{m=0}^{W-1} x_{n,m} \cdot \tilde{d}_{n,m}^{i,j}$$

For a given (i, j) :



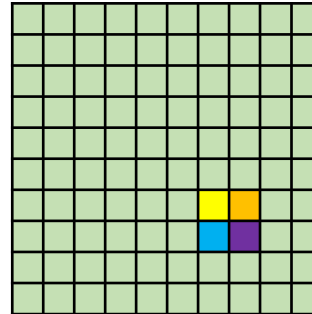
Fast Geometric Verifier

Decompose expression and precompute interpolation distances

$$\tilde{x}'_{i,j} = \sum_{n=0}^{H-1} \sum_{m=0}^{W-1} x_{n,m} \cdot \tilde{d}_{n,m}^{i,j}$$

For a given (i, j) :

$$\tilde{x}'_{i,j} = \Sigma$$



Fast Geometric Verifier

Decompose expression and precompute interpolation distances

$$\tilde{x}'_{i,j} = \sum_{n=0}^{H-1} \sum_{m=0}^{W-1} x_{n,m} \cdot \tilde{d}_{n,m}^{i,j}$$

For a given (i, j) :

$$\tilde{x}'_{i,j} = \sum \text{Image} \odot \text{Grid}$$

HW *HW*

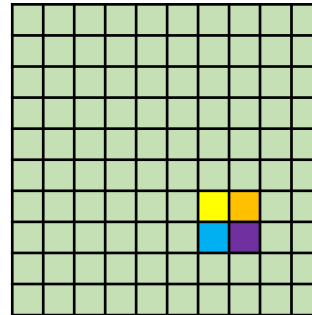
Fast Geometric Verifier

Decompose expression and precompute interpolation distances

$$\tilde{x}'_{i,j} = \sum_{n=0}^{H-1} \sum_{m=0}^{W-1} x_{n,m} \cdot \tilde{d}_{n,m}^{i,j}$$

For a given (i, j) :

$$\tilde{x}'_{i,j} = \sum$$



• *CHW* pixels

HW

HW

Fast Geometric Verifier

Decompose expression and precompute interpolation distances

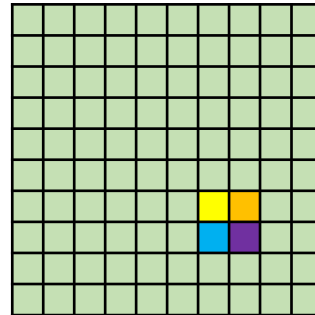
$$\tilde{x}'_{i,j} = \sum_{n=0}^{H-1} \sum_{m=0}^{W-1} x_{n,m} \cdot \tilde{d}_{n,m}^{i,j}$$

For a given (i, j) :

$$\tilde{x}'_{i,j} = \sum$$



HW



HW

- *CHW* pixels
- *B* images

Prohibitive memory overhead!

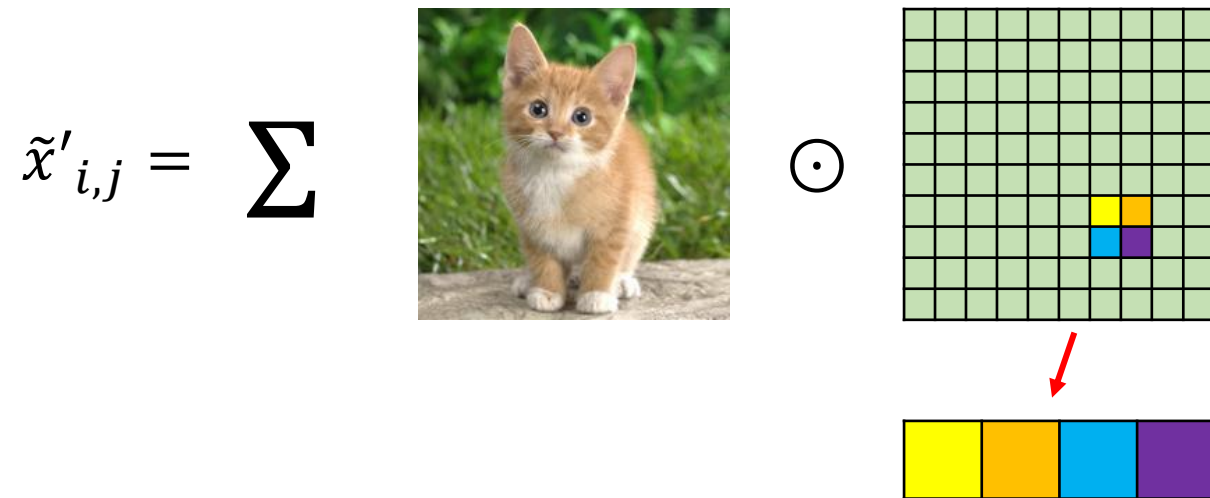
Fast Geometric Verifier

Leverage custom sparse tensor representation

$$\tilde{x}'_{i,j} = \Sigma \text{ (Kitten Image) } \odot \text{ (Grid Mask) }$$

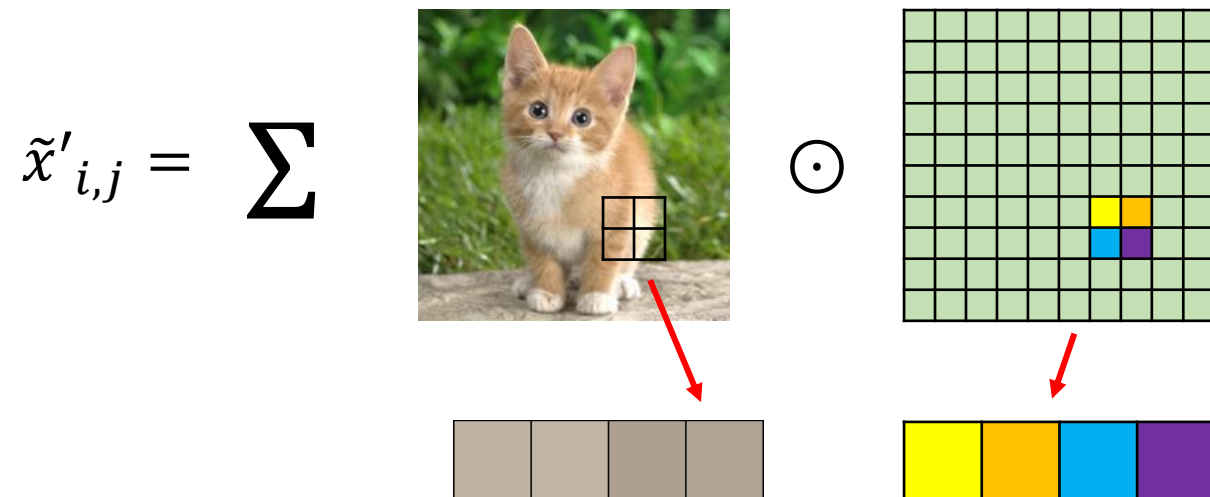

Fast Geometric Verifier

Leverage custom sparse tensor representation



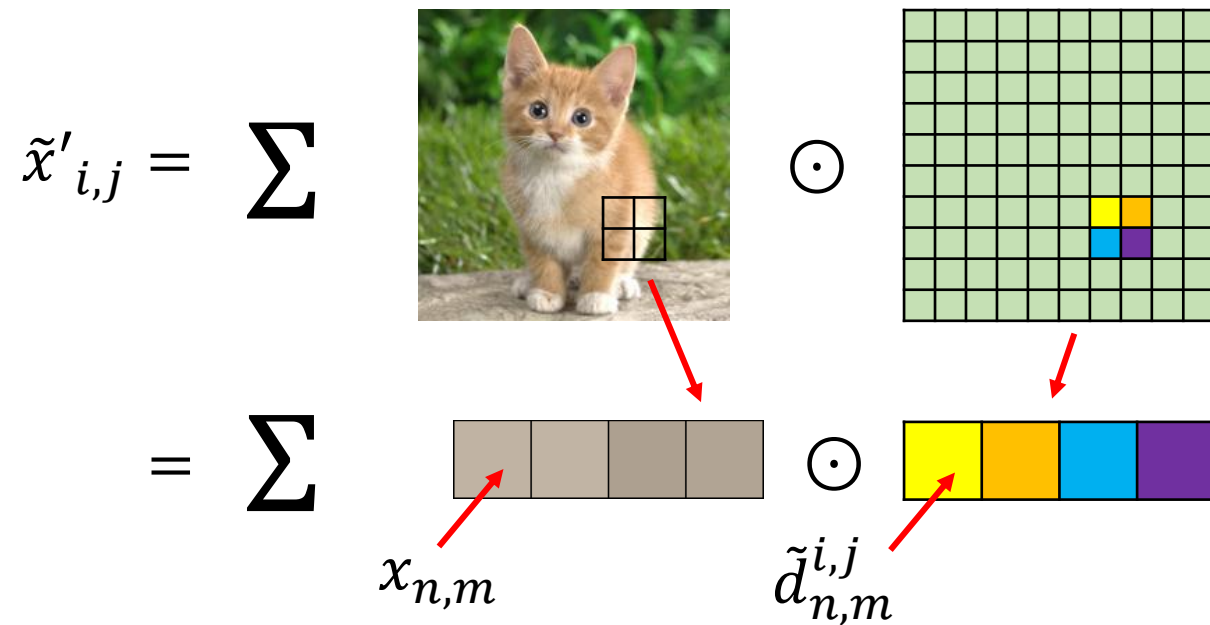
Fast Geometric Verifier

Leverage custom sparse tensor representation



Fast Geometric Verifier

Leverage custom sparse tensor representation



Fast Geometric Verifier

Leverage custom sparse tensor representation

$$\tilde{x}'_{i,j} = \sum \text{Image} \odot \text{Grid} = \sum x_{n,m} \odot \tilde{d}_{n,m}^{i,j}$$

The diagram illustrates the fast geometric verifier process. It shows the summation of a 4x4 grid of pixel values ($x_{n,m}$) and a 4x4 grid of distances ($\tilde{d}_{n,m}^{i,j}$) to produce a single value. The image shows a kitten with a 4x4 grid overlaid. The grid is then flattened into a 1x4 vector of colored blocks (yellow, orange, blue, purple). The distance grid is also flattened into a 1x4 vector of colored blocks (yellow, orange, blue, purple). The final result is a single value.

No multiplication of pixel values with zero distances

Robust Geometric Loss

Robust Geometric Loss

Define worst-case output vector \hat{f} where $\hat{f}_y = \underline{f}_y$ and $\hat{f}_j = \overline{f}_j \quad \forall j \neq y$

Robust Geometric Loss

Define worst-case output vector \hat{f} where $\hat{f}_y = \underline{f}_y$ and $\hat{f}_j = \overline{f}_j \quad \forall j \neq y$

Existing formulations [worst-case loss over *entire* perturbation region]

$$\ell(\hat{f}(\tilde{x}), y)$$

Robust Geometric Loss

Define worst-case output vector \hat{f} where $\hat{f}_y = \underline{f}_y$ and $\hat{f}_j = \overline{f}_j \quad \forall j \neq y$

Existing formulations [worst-case loss over *entire* perturbation region]

$$\ell(\hat{f}(\tilde{x}), y)$$

Our formulation [worst-case loss over *small, sampled* regions]

Robust Geometric Loss

Define worst-case output vector \hat{f} where $\hat{f}_y = \underline{f}_y$ and $\hat{f}_j = \overline{f}_j \quad \forall j \neq y$

Existing formulations [worst-case loss over *entire* perturbation region]

$$\ell(\hat{f}(\tilde{x}), y)$$

Our formulation [worst-case loss over *small, sampled* regions]

To enforce robustness to P across entire parameter range $\tilde{\theta} = [\underline{\theta}, \overline{\theta}]$:

Robust Geometric Loss

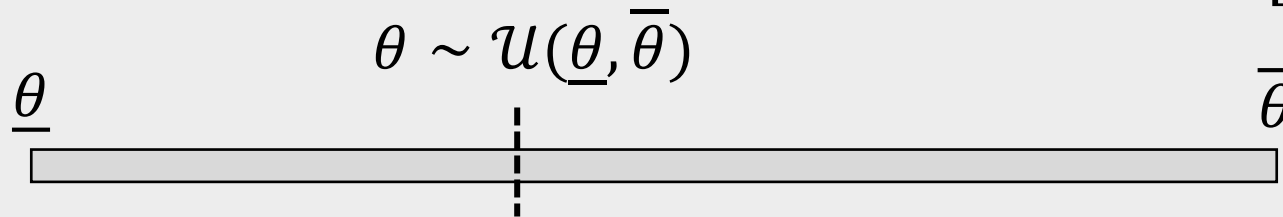
Define worst-case output vector \hat{f} where $\hat{f}_y = \underline{f}_y$ and $\hat{f}_j = \overline{f}_j \quad \forall j \neq y$

Existing formulations [worst-case loss over *entire* perturbation region]

$$\ell(\hat{f}(\tilde{x}), y)$$

Our formulation [worst-case loss over *small, sampled* regions]

To enforce robustness to P across entire parameter range $\tilde{\theta} = [\underline{\theta}, \overline{\theta}]$:



Robust Geometric Loss

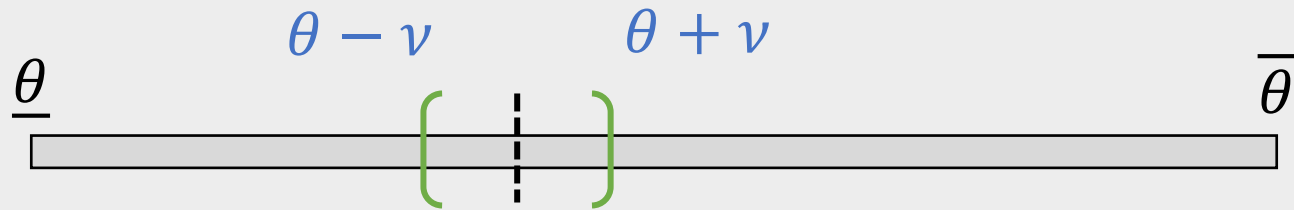
Define worst-case output vector \hat{f} where $\hat{f}_y = \underline{f}_y$ and $\hat{f}_j = \overline{f}_j \quad \forall j \neq y$

Existing formulations [worst-case loss over *entire* perturbation region]

$$\ell(\hat{f}(\tilde{x}), y)$$

Our formulation [worst-case loss over *small, sampled* regions]

To enforce robustness to P across entire parameter range $\tilde{\theta} = [\underline{\theta}, \overline{\theta}]$:



Robust Geometric Loss

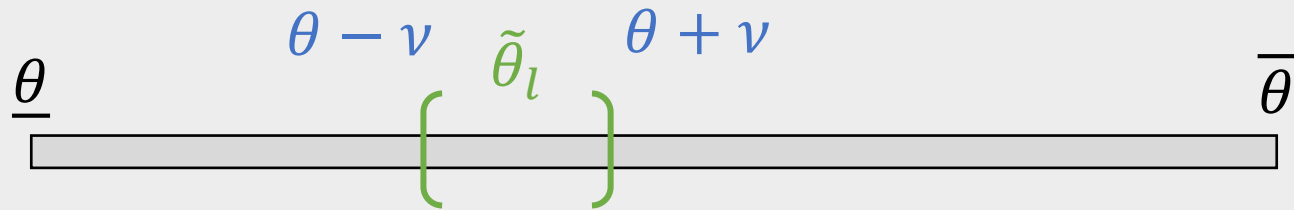
Define worst-case output vector \hat{f} where $\hat{f}_y = \underline{f}_y$ and $\hat{f}_j = \overline{f}_j \quad \forall j \neq y$

Existing formulations [worst-case loss over *entire* perturbation region]

$$\ell(\hat{f}(\tilde{x}), y)$$

Our formulation [worst-case loss over *small, sampled* regions]

To enforce robustness to P across entire parameter range $\tilde{\theta} = [\underline{\theta}, \overline{\theta}]$:



Robust Geometric Loss

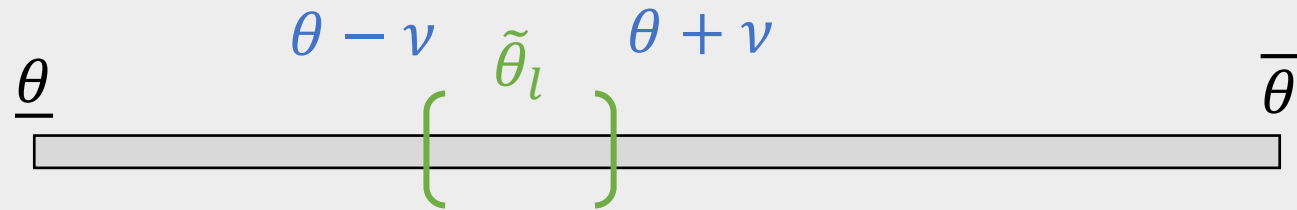
Define worst-case output vector \hat{f} where $\hat{f}_y = \underline{f}_y$ and $\hat{f}_j = \overline{f}_j \quad \forall j \neq y$

Existing formulations [worst-case loss over *entire* perturbation region]

$$\ell(\hat{f}(\tilde{x}), y)$$

Our formulation [worst-case loss over *small, sampled* regions]

To enforce robustness to P across entire parameter range $\tilde{\theta} = [\underline{\theta}, \overline{\theta}]$:



$$\ell(\hat{f}(P(x, \tilde{\theta}_l)), y)$$

Robust Geometric Loss

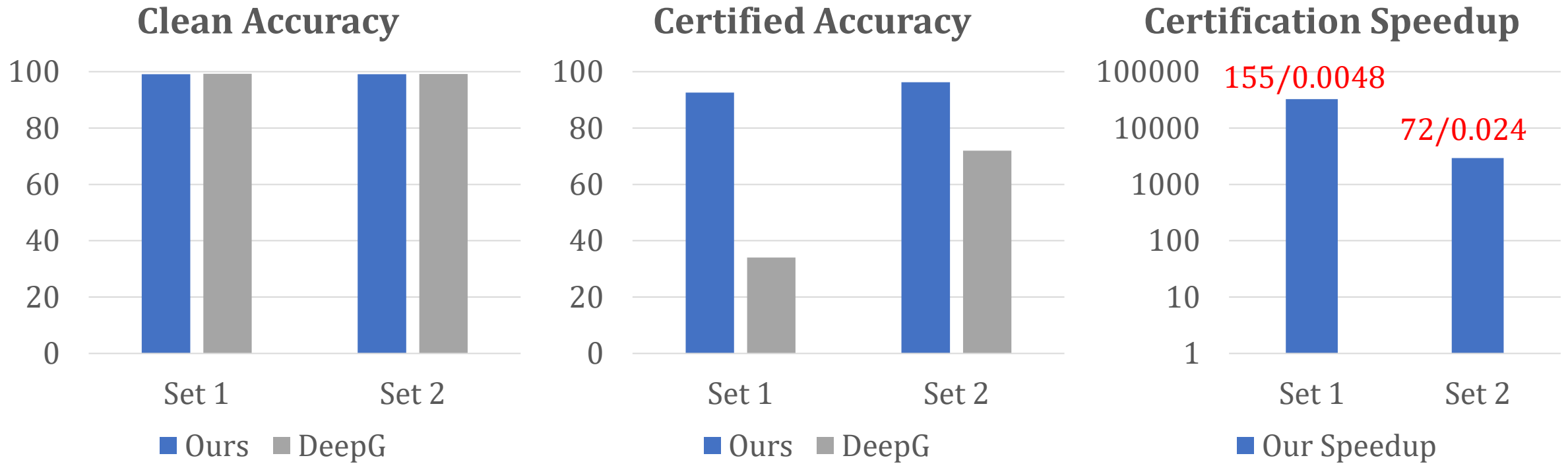
Regression

Minimize both the lower and upper bounds' distances to the ground truth

$$\frac{\ell\left(\underline{f}\left(P(x, \tilde{\theta}_l)\right), y\right) + \ell\left(\overline{f}\left(P(x, \tilde{\theta}_l)\right), y\right)}{2}$$

certified lower bound certified upper bound

MNIST

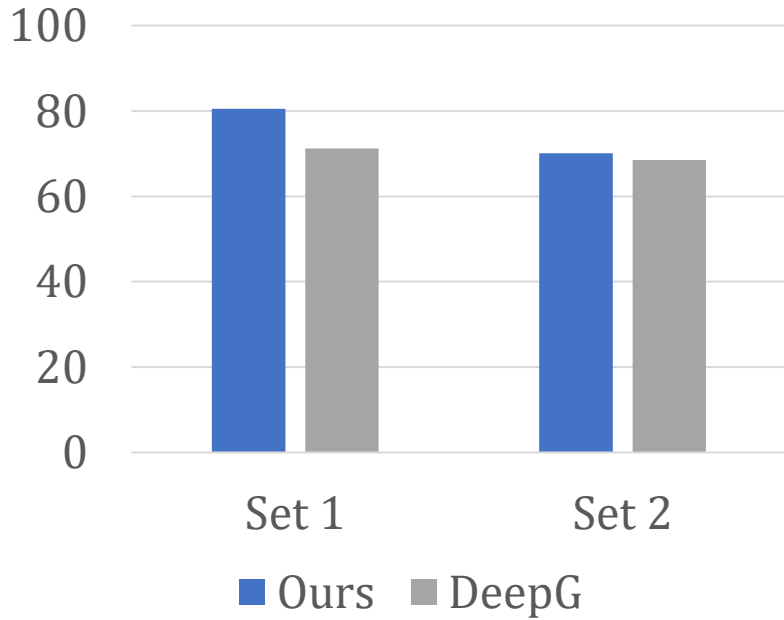


Set 1: Scale(5%), Rotate(5°), Contrast(5%), Brightness(0.01)

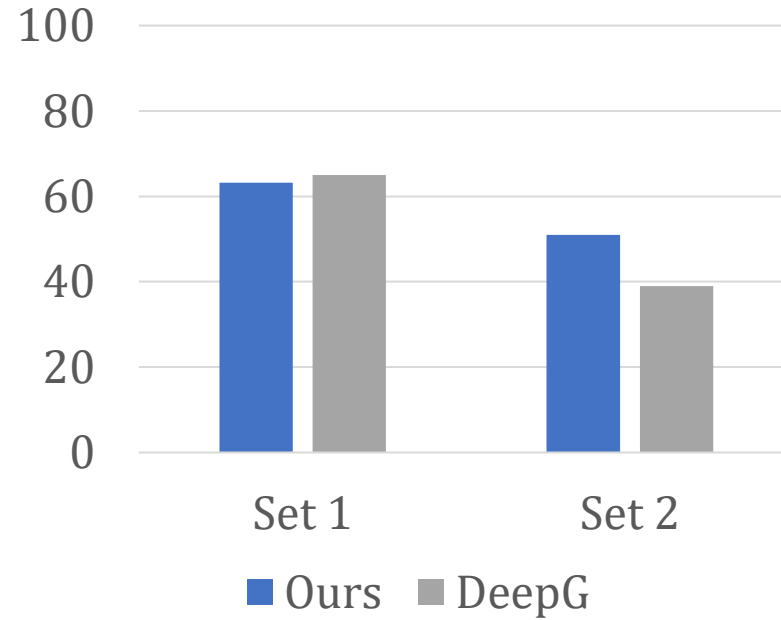
Set 2: Shear(2%), Rotate(2°), Scale(2%), Contrast(2%), Brightness(0.001)

CIFAR-10

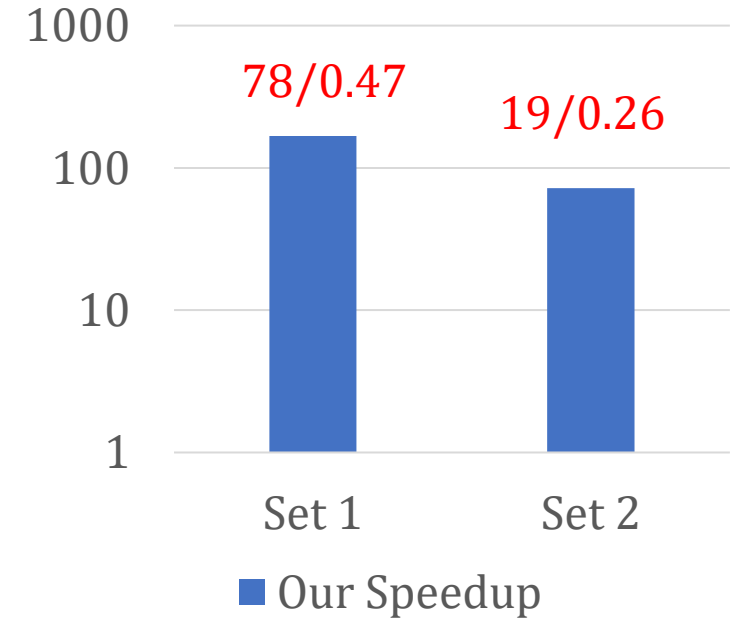
Clean Accuracy



Certified Accuracy



Certification Speedup



Set 1: Rotate(10°)

Set 2: Rotate(2°), Shear(2%)

Tiny ImageNet

First results for deterministic certified geometric robustness on CIFAR-10+

Transforms	Accuracy (%)	Certified (%)	Certification Time per Image (s)
Shear(2%)	35.5	25.7	0.214
Scale(2%)	33.1	21.3	0.205
Rotate(5°)	32.2	17.4	1.006

In ℓ_∞ -norm setting with $\epsilon = \frac{1}{255}$, 27.8% accuracy and 15.9% certified robustness in Xu et al., 2020

Udacity Self-Driving

Task: predict steering angle from $3 \times 66 \times 200$ driving scene image

Network: 9-layer convolutional network from Bojarski et al., 2016

Transformation: $\pm 2^\circ$ rotation

Udacity Self-Driving

Task: predict steering angle from $3 \times 66 \times 200$ driving scene image

Network: 9-layer convolutional network from Bojarski et al., 2016

Transformation: $\pm 2^\circ$ rotation



Green: ground truth label

Blue: network prediction

Red: certified bound

Udacity Self-Driving

Certified training can **help** network performance

Training Method	MAE	Certified MAE	Certification Time per Image (s)
Regular	6.07°	97.56°	0.11
Dropout	4.85°	96.65°	0.12
Ours	5.36°	8.05°	0.11

Conclusion

Robust geometric loss

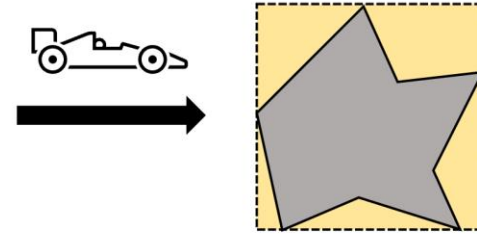
$$\ell \left(\hat{f} \left(P(x, \tilde{\theta}_l) \right), y \right)$$

$\theta \sim \mathcal{U}(\underline{\theta}, \bar{\theta})$ and $\tilde{\theta}_l = [\theta - \nu, \theta + \nu]$

Fast geometric verifier



x



$\tilde{x} = P(x, \tilde{\theta})$

Poster session 1

11:30 – 13:30

@ MH1-2-3-4 #155



Code