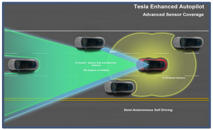
Towards Certifiable Data-Driven Systems

Vishaal Krishnan, Abed AlRahman Al Makdah, Fabio Pasqualetti
December 13, 2020

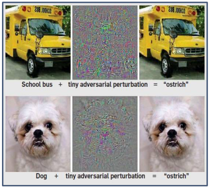
Department of Mechanical Engineering, University of California, Riverside

Data-driven systems in the real world

Success of data-driven decision/control systems



Documented failures



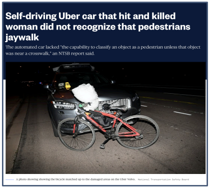

Tesla Model 3: Autopilot engaged during fatal crash

17 May 2019

Share

NTSB

The Tesla Model 3 after the crash


Self-driving Uber car that hit and killed woman did not recognize that pedestrians jaywalk

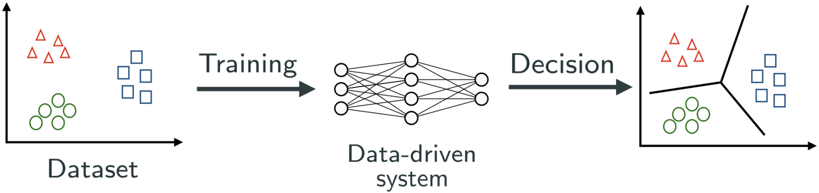
The automated car lacked "the capability to classify an object as a pedestrian unless that object was near a crosswalk," an NTSB report said.

A photo showing showing the bicycle matched up to the damaged areas on the Uber Volvo. National Transportation Safety Board

Certifiability remains a key challenge

A conceptual model of data-driven systems

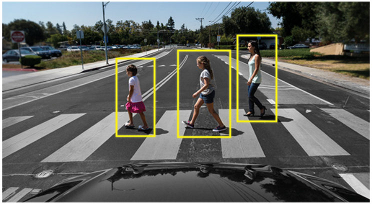
Utilizing dataset to optimize performance on given task



- System design tightly coupled to dataset
- **Goal:** generalizing to new datapoints

Why data-driven systems fail

A problem of generalization



Pedestrians on crosswalk recognized

Pedestrian jaywalking **not** recognized

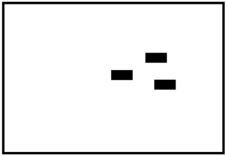
Operating conditions unseen in training

The problem is much worse

Adversarial examples

“Stop”

+



Perturbation

=

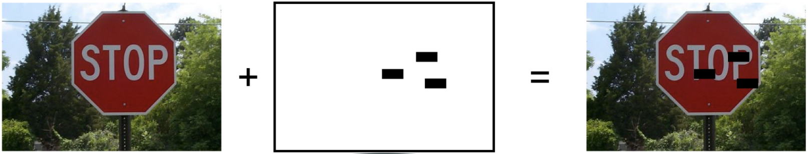
“Speed limit 45mph”

[Eykholt et al., 2017]

Sensitivity to “meaningless” perturbations

Sensitivity to perturbations

Meaning not intrinsic to dataset

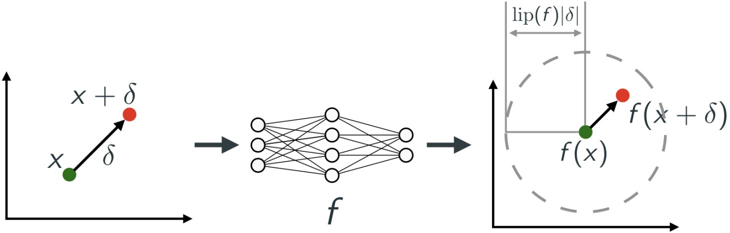


"Stop" + "Meaningless" ? perturbation = "Speed limit 45mph"

Approach: Tune sensitivity to *all* perturbations

Lipschitz constant: A sensitivity measure

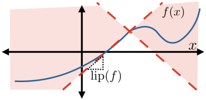
Lipschitz constant controls response to perturbations



Lipschitz constant as a **robustness certificate**

(Low Lipschitz constant $\Rightarrow$ Robust model)

Tuning sensitivity of data-driven models



Lipschitz-regularized learning

[Gouk et al., 2018] [Finlay et al., 2018]

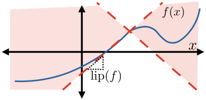
Lipschitz constant estimation

[Weng et al., 2018] [Fazlyab et al., 2019]

Robustness-constrained learning

[Wong et al., 2018] [Pauli et al., 2020]

Tuning sensitivity of data-driven models



Lipschitz-regularized learning

[Gouk et al., 2018] [Finlay et al., 2018]

Lipschitz constant estimation

[Weng et al., 2018] [Fazlyab et al., 2019]

Robustness-constrained learning

[Wong et al., 2018] [Pauli et al., 2020]

What's lacking?

- A formal theory of Lipschitz-robust learning
- An understanding of tradeoffs involved
- A unifying framework for design

Lipschitz-robust learning

A closer look

Lipschitz-robust learning problem:
Minimize (strictly convex) loss with Lipschitz constraint

$$\min_{f \in \mathrm{Lip}} L(f)$$
$$\text{s.t. } \mathrm{lip}(f) \leq \alpha$$

A closer look

Lipschitz-robust learning problem:
Minimize (strictly convex) loss with Lipschitz constraint

$$\min_{f \in \mathrm{Lip}} \; L(f)$$
$$\text{s.t. } \mathrm{lip}(f) \leq \alpha$$

Theorem (Saddle point)

A unique saddle point (with Lipschitz bound α) exists

A closer look: first-order conditions

1. Stationarity:

$$\nabla \cdot (\lambda \nabla f) + \mathbb{E}\left[\partial_f L\right] = 0$$

Key insight: Saddle point given by Poisson PDE

$\nabla \cdot (\lambda \nabla)$ – Laplace operator (λ – Lagrange multiplier)

A closer look: first-order conditions

1. Stationarity:

$$\nabla \cdot (\lambda \nabla f) + \mathbb{E}\left[\partial_f L\right] = 0$$

Key insight: Saddle point given by Poisson PDE

$\nabla \cdot (\lambda \nabla)$ – Laplace operator (λ – Lagrange multiplier)

2. Feasibility: $\mathrm{lip}(f) \leq \alpha \qquad \lambda \geq 0$

A closer look: first-order conditions

1. Stationarity:

$$\nabla \cdot (\lambda \nabla f) + \mathbb{E}\left[\partial_f L\right] = 0$$

Key insight: Saddle point given by Poisson PDE

$\nabla \cdot (\lambda \nabla)$ – Laplace operator (λ – Lagrange multiplier)

2. Feasibility: $\mathrm{lip}(f) \leq \alpha \qquad \lambda \geq 0$

3. Complementary slackness:

$$\lambda \left(|\nabla f| - \alpha\right) = 0 \quad \text{over domain}$$

Robustness via Laplacian smoothing

A heat flow analogy

$$\nabla \cdot (\lambda \nabla f) + \mathbb{E}\left[\partial_f L\right] = 0$$

(steady state temperature profile)

- Map f as temperature profile
- Multiplier λ as conductivity
- Derivative of loss $\partial_f L$ as heat source

Robustness via Laplacian smoothing

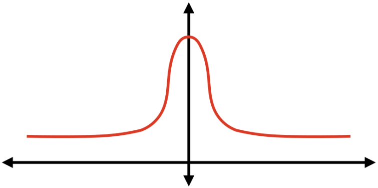
A heat flow analogy

$$\nabla \cdot (\lambda \nabla f) + \mathbb{E}\left[\partial_f L\right] = 0$$

(steady state temperature profile)

- Map f as temperature profile
- Multiplier λ as conductivity
- Derivative of loss $\partial_f L$ as heat source

Sensitivity tuned via Laplacian smoothing



Robustness via Laplacian smoothing

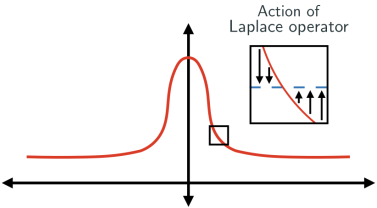
A heat flow analogy

$$\nabla \cdot (\lambda \nabla f) + \mathbb{E}\left[\partial_f L\right] = 0$$

(steady state temperature profile)

- Map f as temperature profile
- Multiplier λ as conductivity
- Derivative of loss $\partial_f L$ as heat source

Sensitivity tuned via Laplacian smoothing



Robustness via Laplacian smoothing

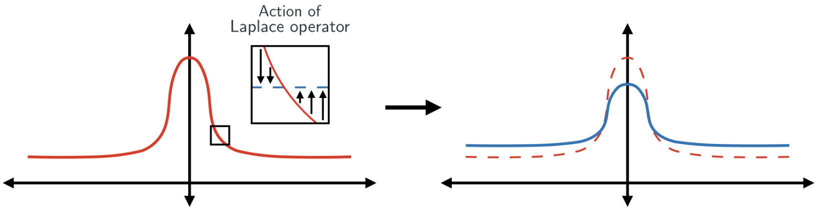
A heat flow analogy

$$\nabla \cdot (\lambda \nabla f) + \mathbb{E}\left[\partial_f L\right] = 0$$

(steady state temperature profile)

- Map f as temperature profile
- Multiplier λ as conductivity
- Derivative of loss $\partial_f L$ as heat source

Sensitivity tuned via Laplacian smoothing



Key takeaways

- Lipschitz-robust learning $\rightarrow$ Solutions of **Poisson-type**
- Robustness enforced by **Laplacian smoothing**
- Active constraint $\Rightarrow$ Tradeoff between accuracy and robustness
  (property of underlying dataset)

Key takeaways

- Lipschitz-robust learning $\rightarrow$ Solutions of **Poisson-type**
- Robustness enforced by **Laplacian smoothing**
- Active constraint $\Rightarrow$ Tradeoff between accuracy and robustness
  (property of underlying dataset)

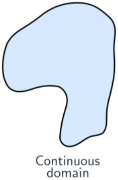
Heat flow-based training algorithms

1. Discretize function space to obtain model family
2. Heat flow to converge to Lipschitz-robust model

Algorithm design

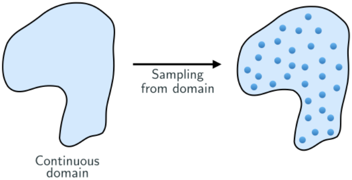
A graph-based learning framework

Graph-discretizing the (continuous) input domain



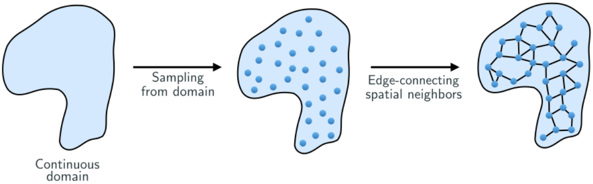
A graph-based learning framework

Graph-discretizing the (continuous) input domain



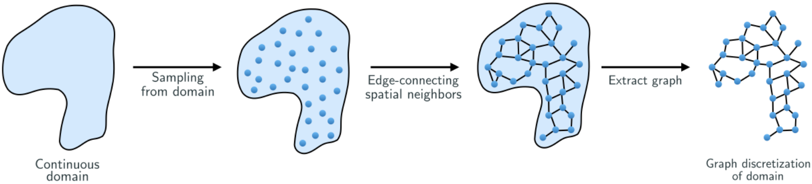
A graph-based learning framework

Graph-discretizing the (continuous) input domain



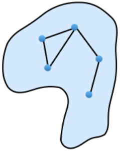
A graph-based learning framework

Graph-discretizing the (continuous) input domain



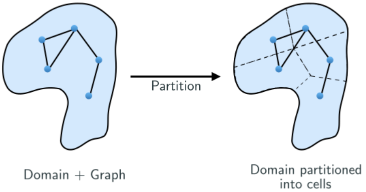
A graph-based learning framework

Discretizing the space of functions/maps


Domain + Graph

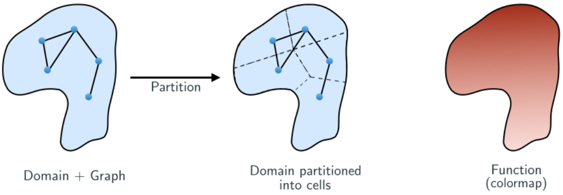
A graph-based learning framework

Discretizing the space of functions/maps



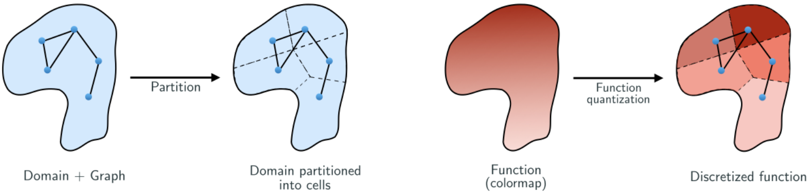
A graph-based learning framework

Discretizing the space of functions/maps



A graph-based learning framework

Discretizing the space of functions/maps



Lipschitz-robust learning on graph

Discrete formulation

$$\min_{(f_1,\dots,f_n)} \; L(f_1,\dots,f_n)$$
$$\text{s.t. } |f_i - f_j| \leq \alpha |x_i - x_j|$$

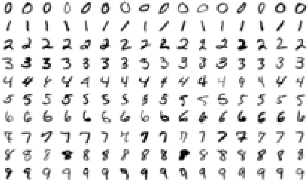
vertex i – position x_i, value f_i

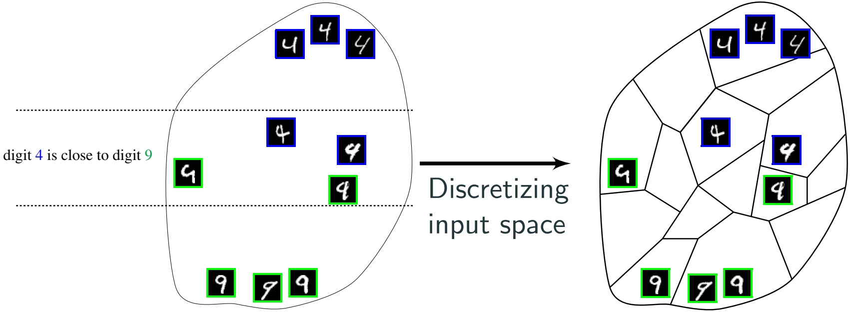
i, j – edge-connected vertices

- Lipschitz $\to$ Edge constraint
- Edge-Lipschitz bound α
- Smoothing by graph Laplacian

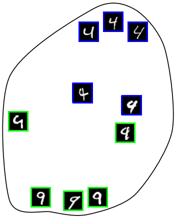
Handwritten digit classification

Applying framework to MNIST dataset

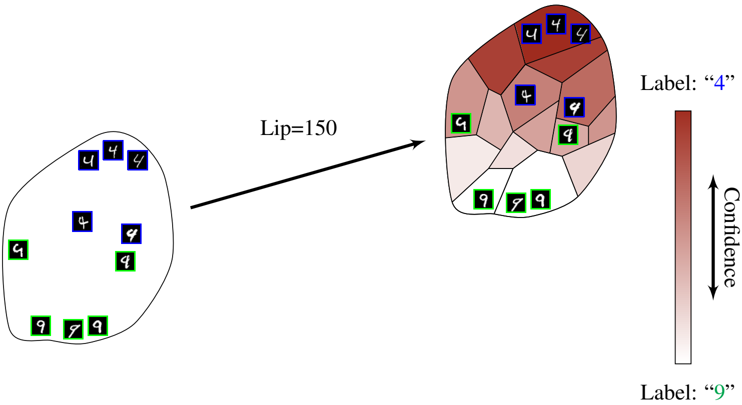




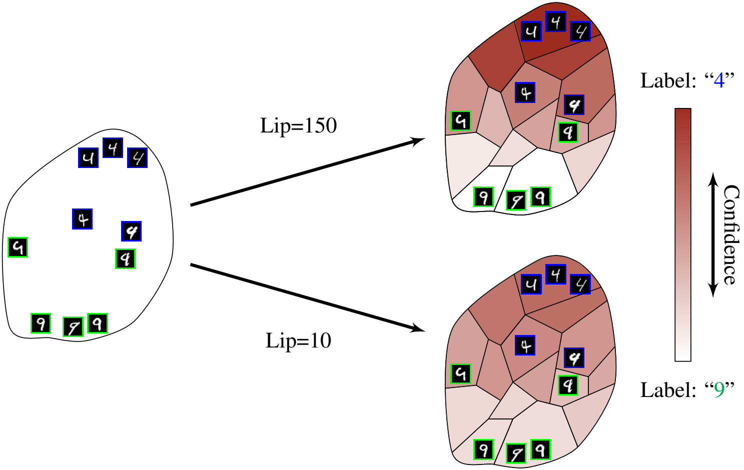
Handwritten digit classification



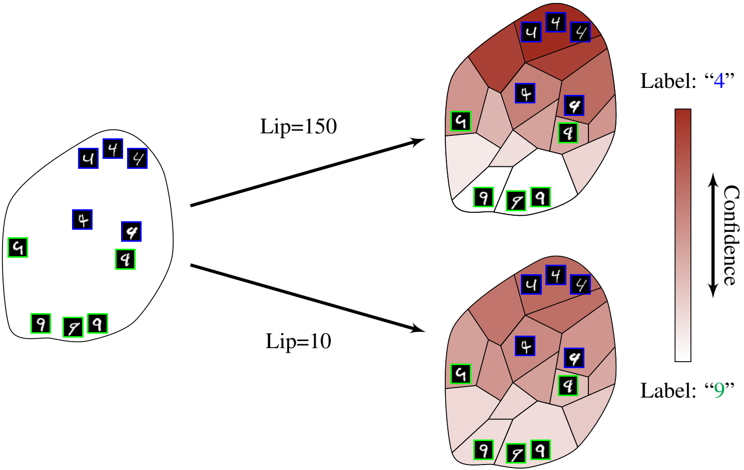
Handwritten digit classification



Handwritten digit classification

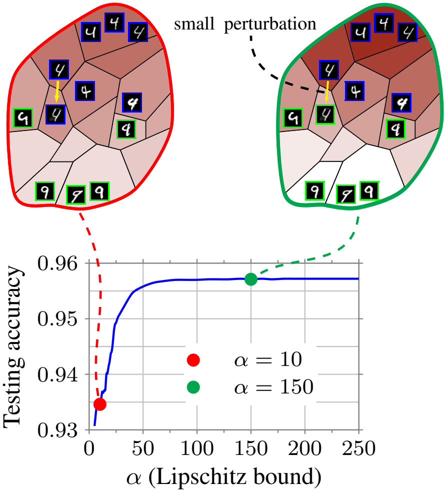


Handwritten digit classification



Learned map is smoothed by decreasing Lipschitz constant

Accuracy vs robustness



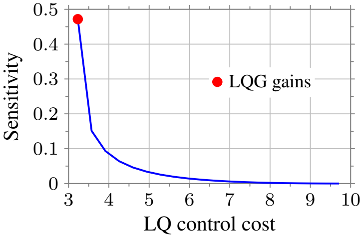
Summary

1. Lipschitz-robustness $\leftrightarrow$ Laplacian smoothing
2. Performance vs robustness tradeoff in learning
3. A graph-based robust learning framework

Ongoing work: Closed-loop setting

A preliminary diagnosis: tradeoff in learning-based control

[Makdah et al., ACC '20]

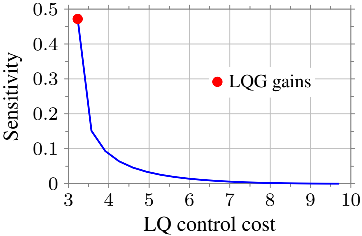


- Perception-based LQG control
- Uncertainty in sensor noise
- Robustness increases at the expenses of performance

Ongoing work: Closed-loop setting

A preliminary diagnosis: tradeoff in learning-based control

[Makdah et al., ACC '20]



- Perception-based LQG control
- Uncertainty in sensor noise
- Robustness increases at the expenses of performance

Ongoing:

- Lipschitz-robust learning for control
- Understanding performance vs robustness tradeoff
- Learning + control co-design (not separable)

References

- Krishnan, Makdah, Pasqualetti
  Lipschitz bounds and provably robust training by Laplacian smoothing
  NeurIPS 2020
- Makdah, Katewa, Pasqualetti
  Accuracy prevents robustness in perception-based control
  ACC 2020
- Makdah, Katewa, Pasqualetti
  A fundamental performance limitation for adversarial classification
  LCSS 2020