Certifiable Robustness only formulas

Fabio Brau

March 1, 2023

Scuola Superiore Sant'Anna, Pisa.

Lp norms

$$\|\delta\|_p = \left(\sum_{i=1}^{n} |\delta_i|^p \right)^{\frac{1}{p}}$$

$$\|\delta\|_\infty = \max_i |\delta_i|$$

$$\|\delta\|_0 = \#\{i \,:\, \delta_i \neq 0\}$$

$$\mathcal{L} : \mathbb{R}^C \times \{1, \cdots, C\} \rightarrow \mathbb{R}$$

$$\mathcal{L}(y, c) = -\log\left(\frac{e^{y_c}}{\sum_{i=1}^{C} e^{y_i}}\right)$$

$$\begin{aligned} \max_{\delta \in \mathbb{R}^n} \quad & \mathcal{L}(f(x+\delta), l_{true}) \\ \text{s.t.} \quad & \|\delta\|_p \leq \varepsilon \\ & 0 \leq x + \delta \leq 1 \\ & Q * (x+\delta) \in \{0, \cdots, Q\} \end{aligned}$$

$$\begin{aligned} \min_{\delta \in \mathbb{R}^n} \quad & \|\delta\|_p \\ \text{s.t.} \quad & \mathcal{K}(x+\delta) \neq l_{true} \\ & 0 \leq x + \delta \leq 1 \\ & Q * (x+\delta) \in \{0, \cdots, Q\} \end{aligned}$$

$$\min_{\delta \in \mathbb{R}^n} \quad \mathcal{L}(f(x+\delta), l_{target})$$
$$\text{s.t.} \quad \|\delta\|_p \leq \varepsilon$$
$$0 \leq x+\delta \leq 1$$
$$Q*(x+\delta) \in \{0, \cdots, Q\}$$

$$\min_{\delta \in \mathbb{R}^n} \quad \|\delta\|_p$$
$$\text{s.t.} \quad \mathcal{K}(x+\delta) = l_{target}$$
$$0 \leq x+\delta \leq 1$$
$$Q*(x+\delta) \in \{0, \cdots, Q\}$$

$$\delta^* = \varepsilon \cdot \text{sign}\left(\nabla_x \mathcal{L}(f(x), l_{true})\right)$$

$$x^* = x + \delta^*, \quad \|\delta^*\|_\infty = \varepsilon$$

$$\tilde{\mathcal{L}}(f(x), l) = \alpha \cdot \mathcal{L}(f(x), l) \quad + \quad (1-\alpha) \cdot \mathcal{L}(f(x+\delta^*), l)$$

$$\begin{aligned} \min_{\delta \in \mathbb{R}^n} \quad & c\,\|\delta\|_p + \mathcal{L}(f(x+\delta), l_{target}) \\ \text{s.t.} \quad & 0 \le x + \delta \le 1 \\ & Q * (x+\delta) \in \{0, \cdots, Q\} \end{aligned}$$

$$\delta(c) \quad \mathcal{K}(x + \delta(c)) = l_{target}$$

$$c_{left} = 0,\ c_{right} = 100, \quad c_{test} = \frac{c_{left} + c_{out}}{2}$$

$$c_{left} = c_{left}, c_{right} = c_{test} \quad \text{if} \quad \mathcal{K}(x + \delta(c_{test})) = l$$

<++>

Thanks for the attention

Fabio Brau

Scuola Superiore Sant'Anna, Pisa

fabio.brau@santannapisa.it

retis.santannapisa.it/~f.brau

linkedin.com/in/fabio-brau

TELECOMMUNICATIONS, COMPUTER ENGINEERING, AND PHOTONICS INSTITUTE

Sant'Anna School of Advanced Studies – Pisa

ECCELLENZA MIUR 2018-2022

ROBOTICS & AI

Sant'Anna Scuola Universitaria Superiore Pisa

retis Real-Time Systems Laboratory