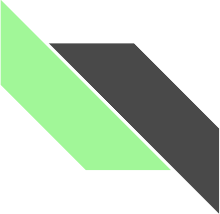

Uncertainty is the best form of defence

- Can intervalise the pixel data then retrain the neural network

- This effectively protects against both black and white box attacks

- Using uncertainty means that you do not have to be specific about which attack you are trying to prevent

19

https://github.com/MadryLab/mnist_challenge, Sadeghi et al. 2019

Uncertainty is the best form of defence

- Can intervalise the pixel data then retrain the neural network
- This effectively protects against both black and white box attacks
- Using uncertainty means that you do not have to be specific about which attack you are trying to prevent

https://github.com/MadryLab/mnist_challenge, Sadeghi et al. 2019