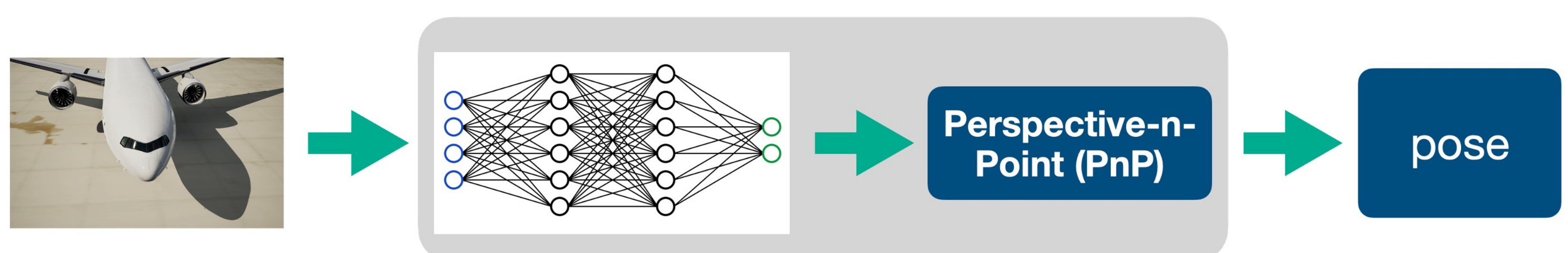
Certifying Robustness of Learning-based Pose Estimation Methods

Xusheng Luo[1], Tianhao Wei[1], Simin Liu[1], Ziwei Wang[1], Luis Mattei-Mendez[2], Taylor Loper[2], Joshua Neighbor[2], Casidhe Hutchison[1], Changliu Liu[1]

[1]Carnegie Mellon University, [2]The Boeing Company

Introduction

- Vision-based 6D pose estimation, i.e., 3D rotation and 3D translation of an object with respect to the camera, aims to identify the posture of objects through images.
- The adoption of neural networks have markedly surpassed techniques that depend on manually engineered features.
- The absence of performance assurances raises concerns about their integration into safety-critical applications.



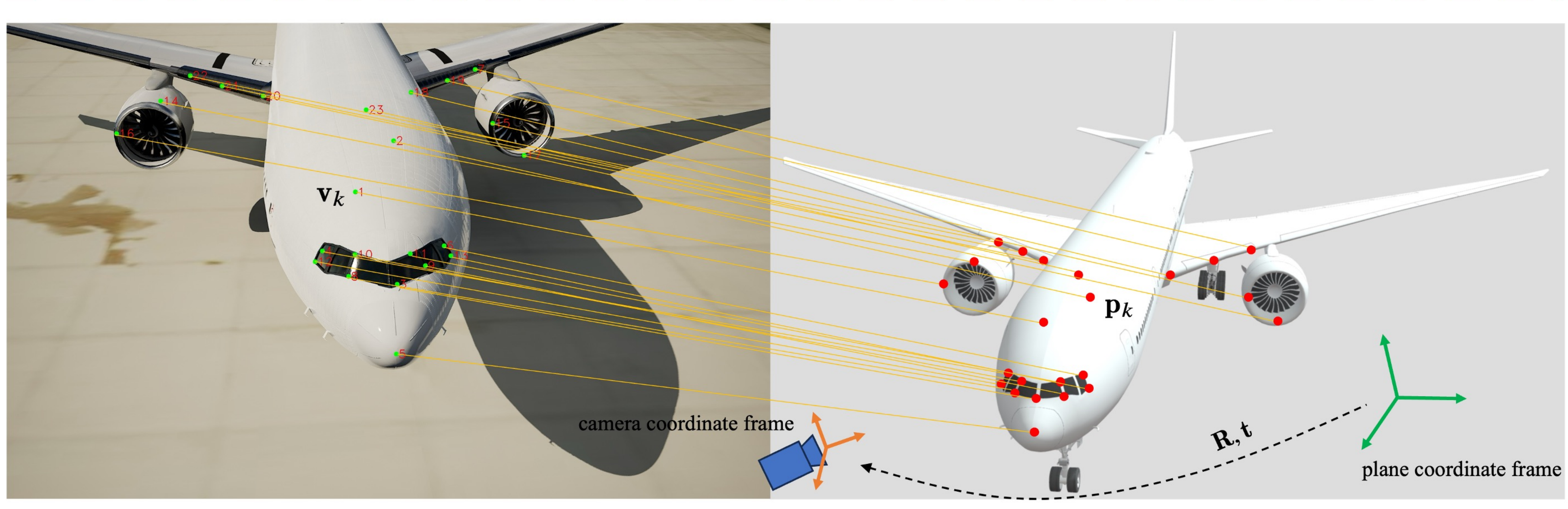
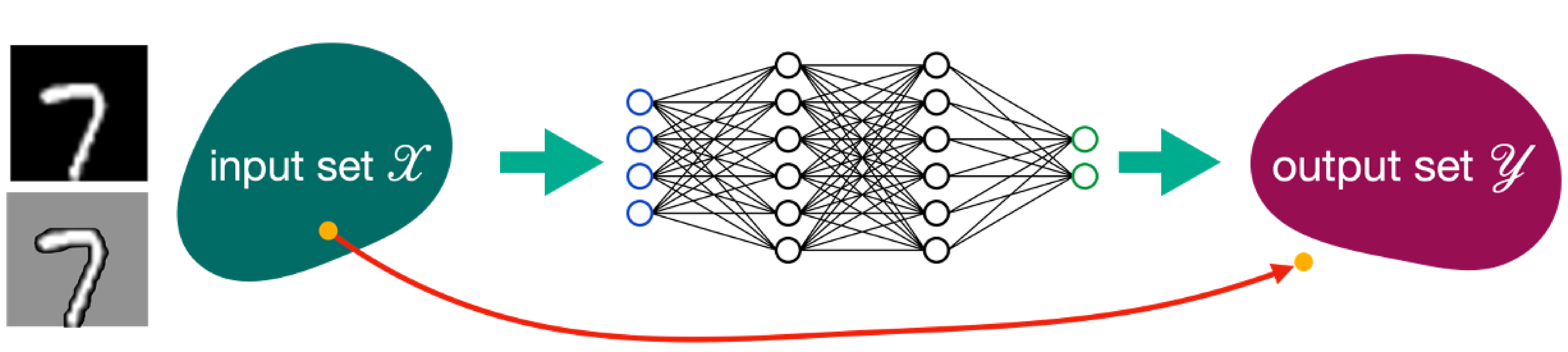


Figure 1: The pose of an airplane parked at an airport is estimated using a two-step approach based on keypoint detection.

Key contributions: To the best of our knowledge, this study is the first one to certify the robustness of large-scale, keypoint-based pose estimation problem encountered in the real world.

Problem Formulation

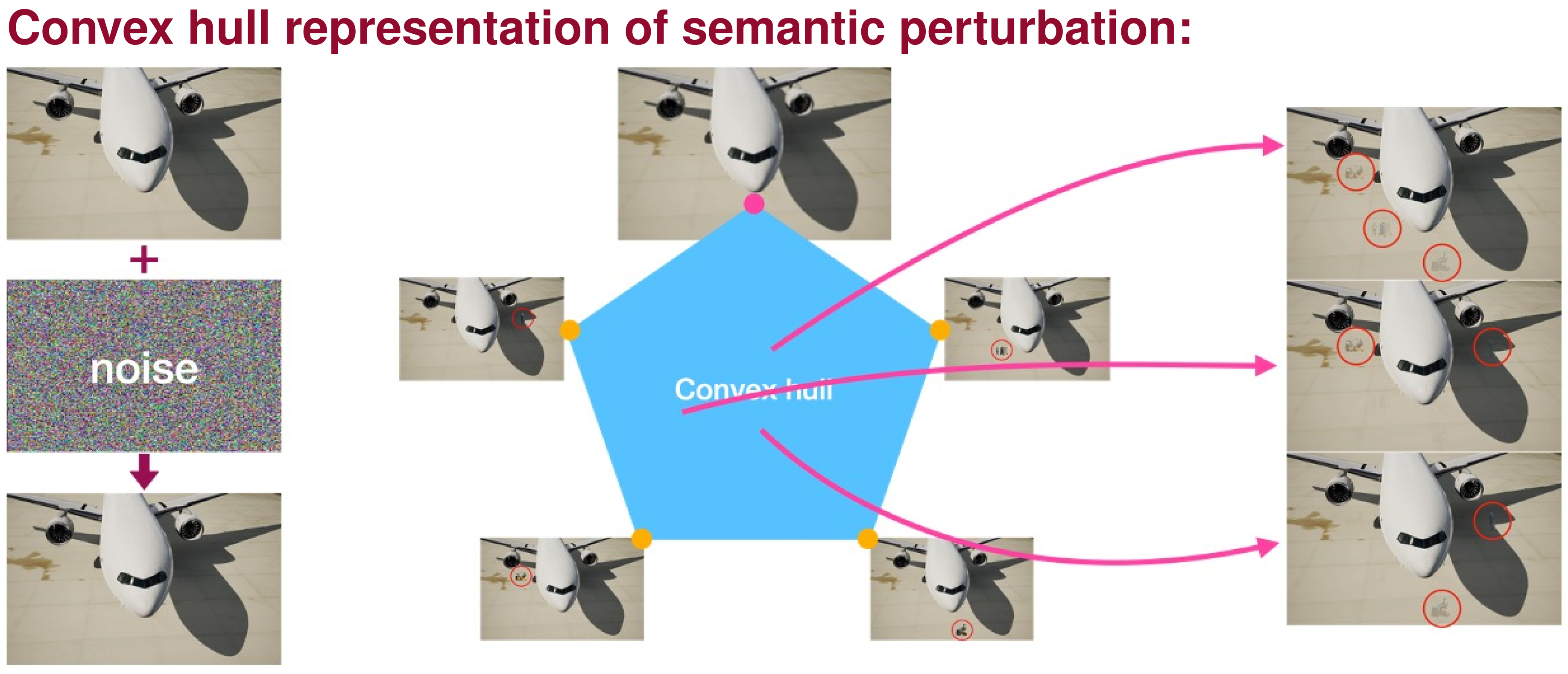
Robustness verification of standalone neural network:



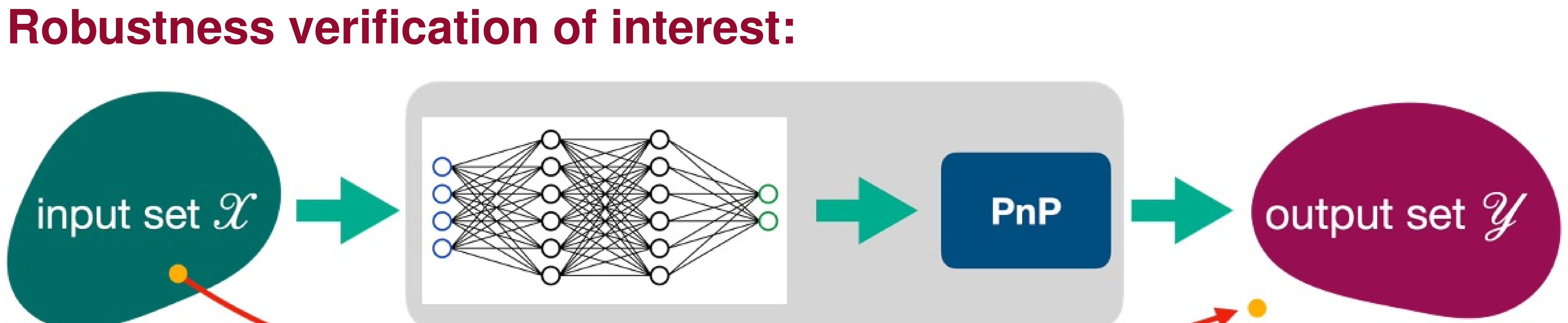
$\mathscr{X}$: Pixel-wise variation is at most 0.1 $\qquad$ $\mathscr{Y}$: prediction is always 7

Figure 2: The NN is considered robust if there exists one point in the input set such that the corresponding output falls outside the output set.

Convex hull representation of semantic perturbation:

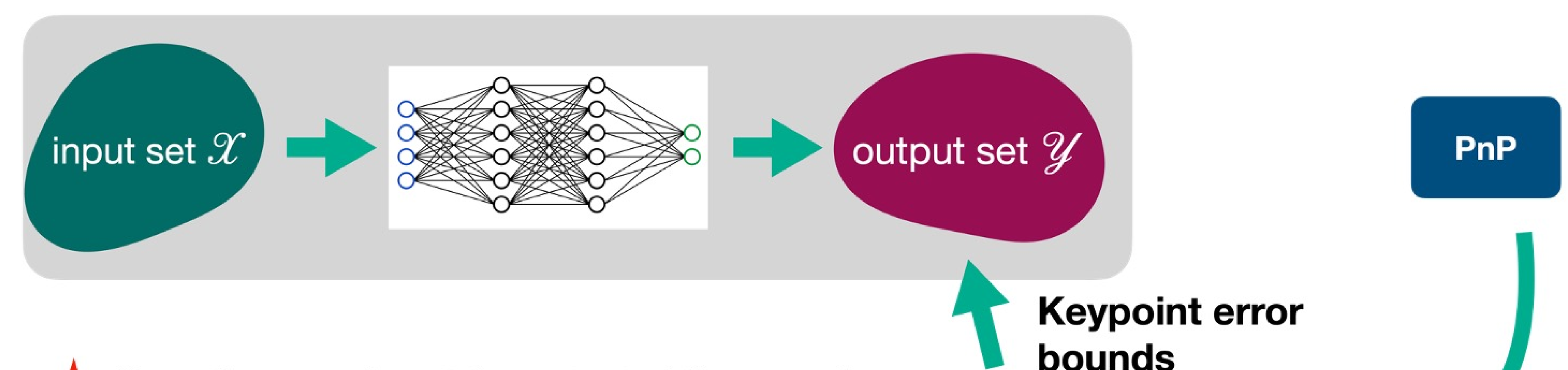


Robustness verification of interest:



Pose error bounds:
- Within **6** feet in the x and y directions
- Within **5** degrees in yaw
- Within **0.5** degrees in roll and pitch

Convert to Verification of Classification



★ Since the pose is not the output of the neural network, the main idea is to separate the PnP algorithm from the process, converting the problem into a standard robustness verification task for a classification network

Verification Results

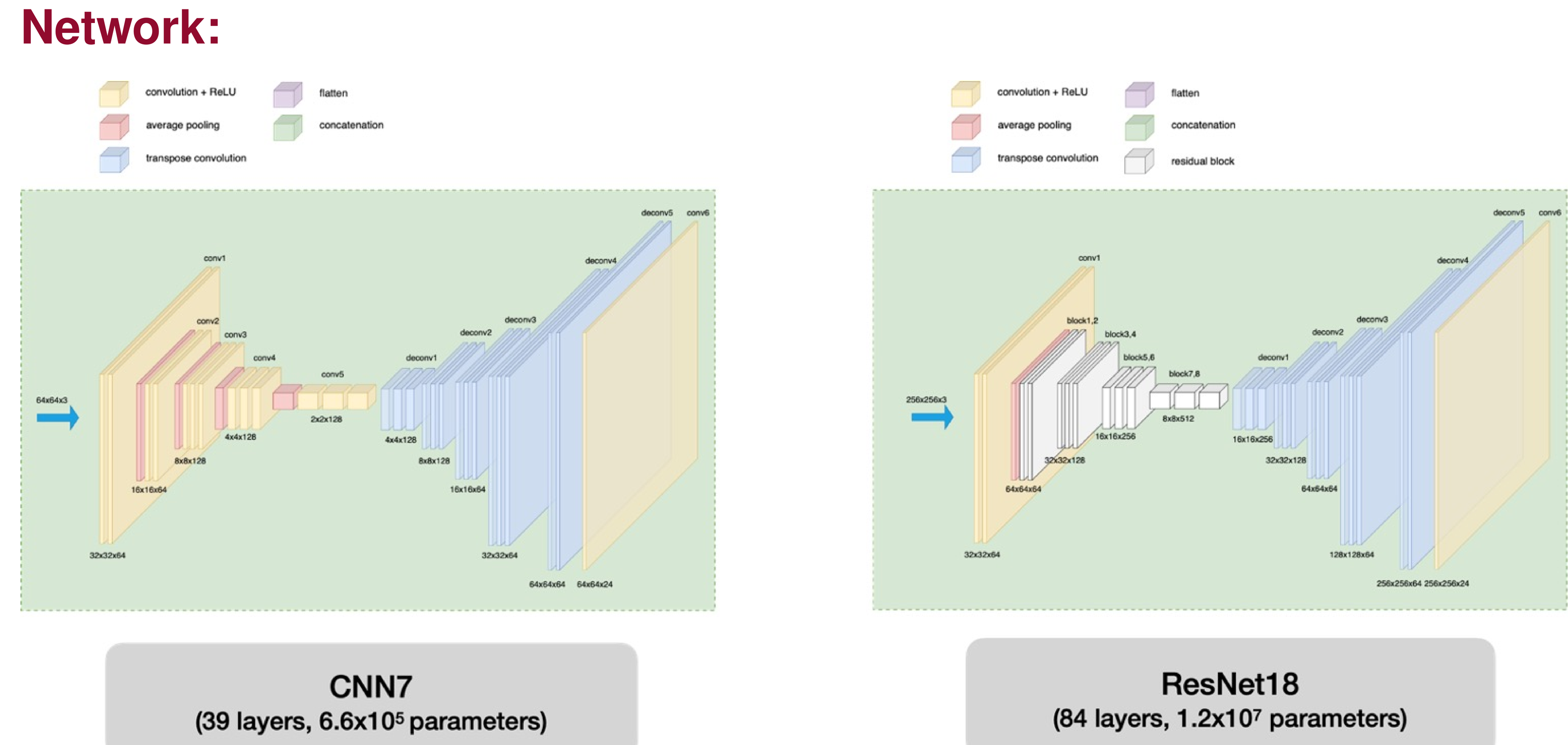
Questions:

- How computationally efficient is the certification method?
- How accurate is the certification method for pose estimation?

Metric:

- *Verification times*: statistics computed for each tested seed images with the specific perturbations.
- *Verified rate*: the proportion of cases where the verification algorithm confirms robustness at the keypoint level against those where seed images produce acceptable pose estimation errors.

Network:



Results:

- The neural network effectively minimizes the influence of background disturbances but remains more sensitive to perturbations that affect the airplane itself.

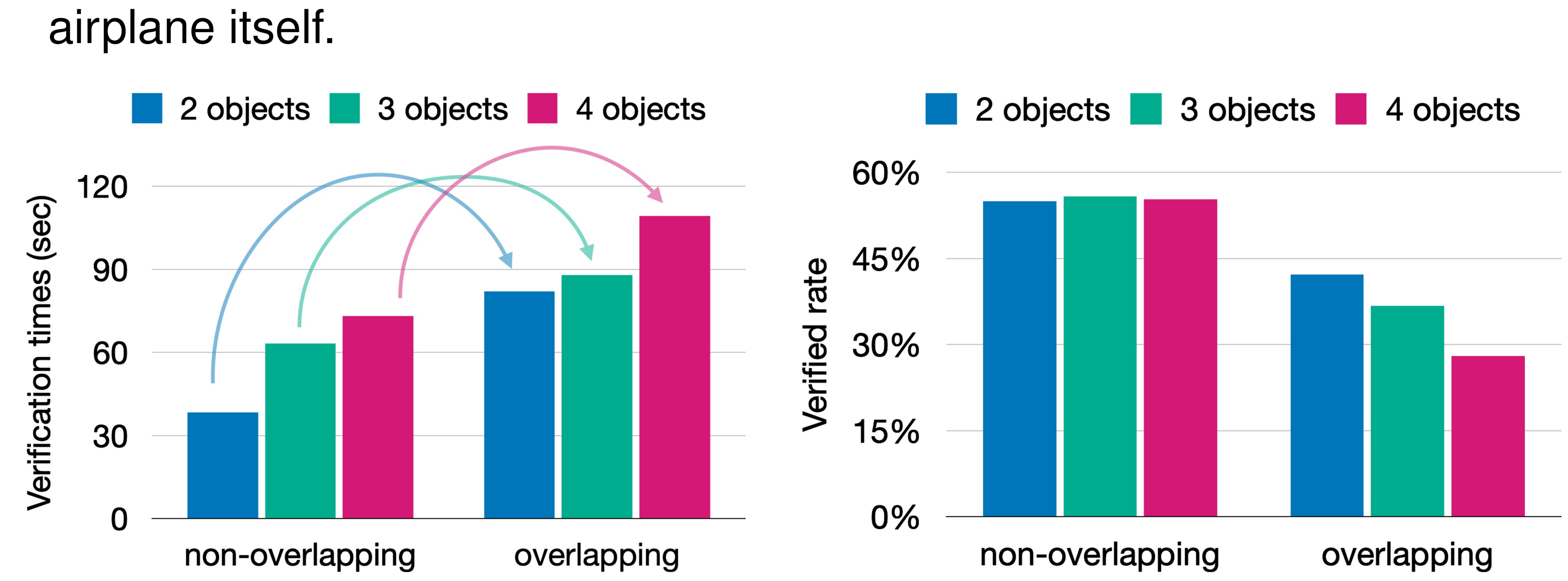
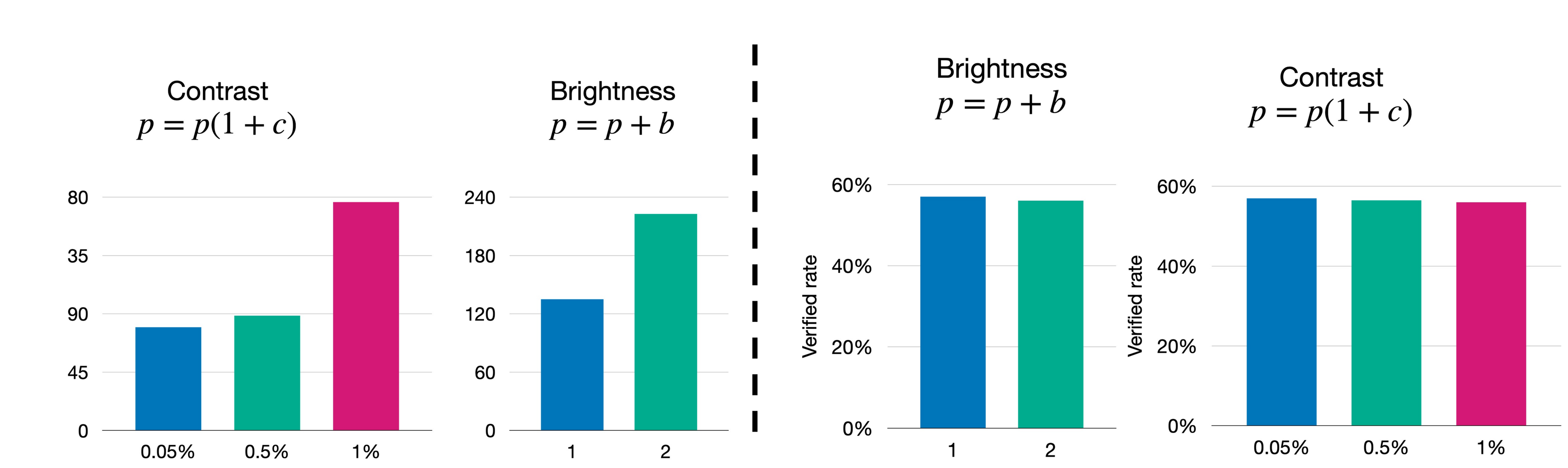


Figure 3: Local perturbation on CNN7.



(a) Verification times (sec) $\qquad$ (b) Verified rate

Figure 4: Global perturbation on ResNet18.