Motivation

- Is Neural Network itself the best identifier of the confidence?
- Can we really trust the confidence rate reported by Neural Networks for their predictions?
- How confident predictions really are if they have a confidence rate of 95% reported by Neural Network?
- Can we have another solution to assess the reliability of Neural Network predictions?

Critical Applications

Trust on the confidence level of the network can result to drastic and irreparable damages

Misprediction Analysis of AlexNet

> 20% of the mispredictions have high confidence

Scenarios Behind Misprediction - AlexNet

- Multiple Objects in a Single Frame
- Similarity Between Different Classes
- Training Set Representativity
- Image Detail (Obstruction, Blur, Small Size, …)

Observations

- We can use reported low confidence rate as a unreliability sign, but we also need to assure the reliability of high confidence ones
- A significant portion of the mispredictions occur for high confidence results
- Network itself is not the best identifier of confidence
- Proposed Solution → Use specialist and accurate Neural Networks and decide on the reliability by combining the results

Proposed Solution

- Partition the dataset to different super-classes
- Train a specialist classifier for each super-class
- Train a broad classifier to choose the main classifier
- Train a confidence predictor to score the reliability according to the result of main and other classifiers

System of Specialists

Specialist Approach for CIFAR-10

General Classifier → 74%

Animal Specialist → 74%

Object Specialist → 88%