

80 million images creation

- labels from web crawling
- 32×32 quality

2006



2009

CIFAR-10 creation:

- 10 classes and 60K images
- 50K train / 10k test
- Students paid to filter out mislabeled images

Problem

- votes have been lost
- errors have been found



2019

CIFAR-10H creation

- Switch original train/test sets
- Pay 2571 workers to label images
- $\simeq 50$ votes per image