3

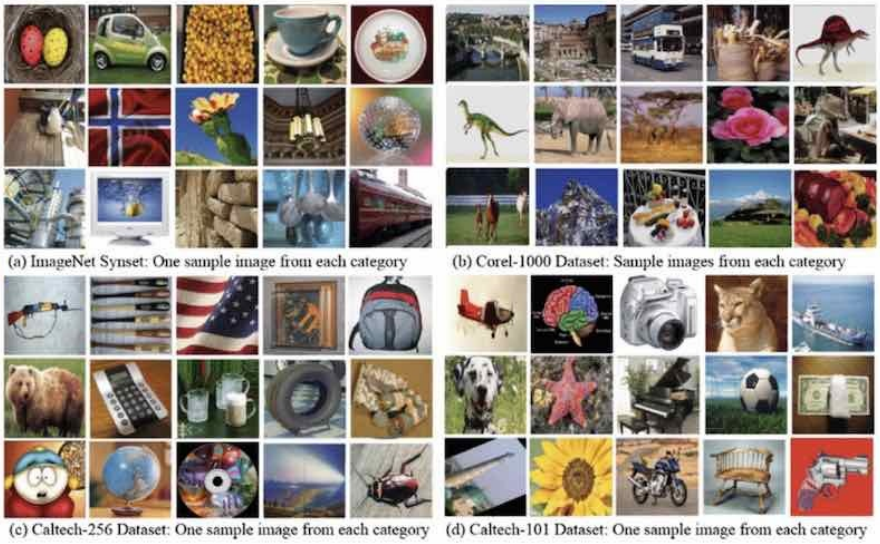
GitHub AI

AI “”

“”

“” ImageNet 120 Flickr

ImageNet paper



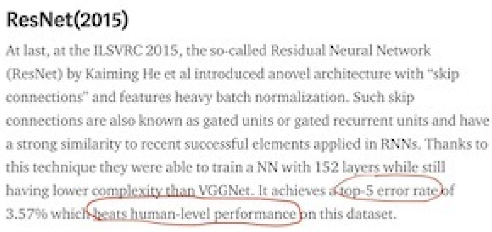
(a) ImageNet Synset: One sample image from each category (b) Corel-1000 Dataset: Sample images from each category

(c) Caltech-256 Dataset: One sample image from each category (d) Caltech-101 Dataset: One sample image from each category

“” benchmark “”

“Top-5 ”

“”“”AI “top-5 ” 5 top-5 “”



ResNet(2015)

At last, at the ILSVRC 2015, the so-called Residual Neural Network (ResNet) by Kaiming He et al introduced anovel architecture with "skip connections" and features heavy batch normalization. Such skip connections are also known as gated units or gated recurrent units and have a strong similarity to recent successful elements applied in RNNs. Thanks to this technique they were able to train a NN with 152 layers while still having lower complexity than VGGNet. It achieves a top-5 error rate of 3.57% which beats human-level performance on this dataset.

“top-5”“”“top-5”“”“top-5” top-5

“top-5” 5

1. “”
2. “”
3. “”
4. “”
5. “”

5 “”

5 ImageNet ILSVRCground truth

II: Object localization

The data for the classification and localization tasks will remain unchanged from ILSVRC 2012 . The validation and test data will consist of 150,000 photographs, collected from flickr and other search engines, hand labeled with the presence or absence of 1000 object categories. The 1000 object categories contain both internal nodes and leaf nodes of ImageNet, but do not overlap with each other. A random subset of 50,000 of the images with labels will be released as validation data included in the development kit along with a list of the 1000 categories. The remaining images will be used for evaluation and will be released without labels at test time. The training data, the subset of ImageNet containing the 1000 categories and 1.2 million images, will be packaged for easy downloading. The validation and test data for this competition are not contained in the ImageNet training data.

In this task, given an image an algorithm will produce 5 class labels $c_i, i = 1, \dots 5$ in decreasing order of confidence and 5 bounding boxes $b_i, i = 1, \dots 5$, one for each class label. The quality of a localization labeling will be evaluated based on the label that best matches the ground truth label for the image and also the bounding box that overlaps with the ground truth. The idea is to allow an algorithm to identify multiple objects in an image and not be penalized if one of the objects identified was in fact present, but not included in the ground truth.

The ground truth labels for the image are $C_k, k = 1, \dots n$ with n class labels. For each ground truth class label C_k, the ground truth bounding boxes are $B_{km}, m = 1 \dots M_k$, where M_k is the number of instances of the k^{th} object in the current image.

Let $d(c_i, C_k) = 0$ if $c_i = C_k$ and 1 otherwise. Let $f(b_i, B_k) = 0$ if b_i and B_k have more than 50% overlap, and 1 otherwise. The error of the algorithm on an individual image will be computed using:

$$e = \frac{1}{n} \cdot \sum_k min_i min_m max\{d(c_i, C_k), f(b_i, B_{km})\}$$

The winner of the object localization challenge will be the team which achieves the minimum average error across all test images.

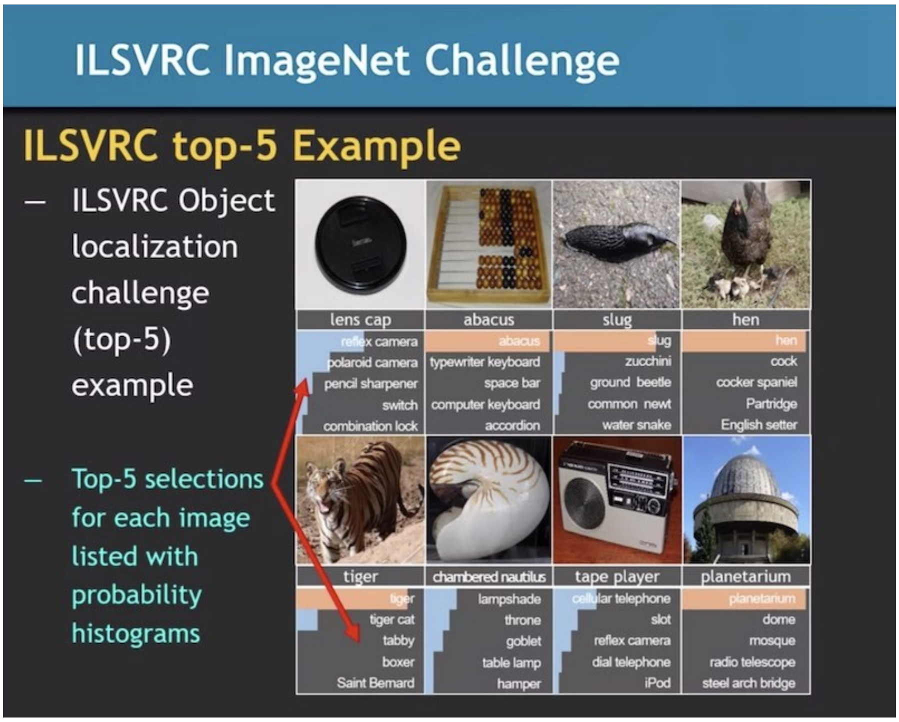
For each image, algorithms will produce a list of at most 5 scene categories in descending order of confidence. The quality of a labeling will be evaluated based on the label that best matches the ground truth label for the image. The idea is to allow an algorithm to identify multiple scene categories in an image given that many environments have multi-labels (e.g. a bar can also be a restaurant) and that humans often describe a place using different words (e.g. forest path, forest, woods).

For each image, an algorithm will produce 5 labels $l_j, j = 1, \dots, 5$. The ground truth labels for the image are $g_k, k = 1, \dots, n$ with n classes of scenes labeled. The error of the algorithm for that image would be

$$e = \frac{1}{n} \cdot \sum_k \min_j d(l_j, g_k).$$

$d(x, y) = 0$ if $x = y$ and 1 otherwise. The overall error score for an algorithm is the average error over all test images. Note that for this version of the competition, n=1, that is, one ground truth label per image.

5



Coursera top-5 top-5 “” top-5 computer keyboard accordion“” top-5 boxerSaint Bernard

“”

ILSVRC top-5 “” abacus computer keyboard accordion abacus top-5 top-5

“”“”

ILSVRC top-5AI top-5 “”

Top-5 top-1 ResNet-152 top-1 19.38% top-5 4.49%Top-1 “” top-5 “” top-5 5.1%

Table 3. Error rates (%, **10-crop** testing) on ImageNet validation. VGG-16 is based on our test. ResNet-50/101/152 are of option B that only uses projections for increasing dimensions.

method	top-1 err.	top-5 err.
VGG [41] (ILSVRC'14)	-	8.43†
GoogLeNet [44] (ILSVRC'14)	-	7.89
VGG [41] (v5)	24.4	7.1
PReLU-net [13]	21.59	5.71
BN-inception [16]	21.99	5.81
ResNet-34 B	21.84	5.71
ResNet-34 C	21.53	5.60
ResNet-50	20.74	5.25
ResNet-101	19.87	4.60
ResNet-152	**19.38**	**4.49**

Top-5

top-5 5 “top-5 ” 5 “top-5 ”

top-5 5.1% top-1 5.1% 5

ResNet-152 top-5 4.49% top-1 19.38% top-5

top-5 4.49% 5.1% 0.61%1% 3% “”“”

“ vs top-1 ”5.1% “ top-5 ” Andrej Karpathy ImageNet ImageNet “”

“” Andrej Karpathy Tesla AI ImageNet “”

Andrej Karpathy cs231n back-propagation “”

AI top-5 AI

“AI ”AI “”“90+% ”“top-5 ”“top-1 ”

top-1

Table 3. Error rates (%, **10-crop** testing) on ImageNet validation. VGG-16 is based on our test. ResNet-50/101/152 are of option B that only uses projections for increasing dimensions.

method	top-1 err.	top-5 err.
VGG [41] (ILSVRC'14)	-	8.43†
GoogLeNet [44] (ILSVRC'14)	-	7.89
VGG [41] (v5)	24.4	7.1
PReLU-net [13]	21.59	5.71
BN-inception [16]	21.99	5.81
ResNet-34 B	21.84	5.71
ResNet-34 C	21.53	5.60
ResNet-50	20.74	5.25
ResNet-101	19.87	4.60
ResNet-152	**19.38**	**4.49**

Table 1. Single crop validation error on ImageNet-1k (center 224x224 crop from resized image with shorter side = 256). The SENet-154 is one of our superior models used in ILSVRC 2017 Image Classification Challenge where we won the 1st place (Team name: WMW).

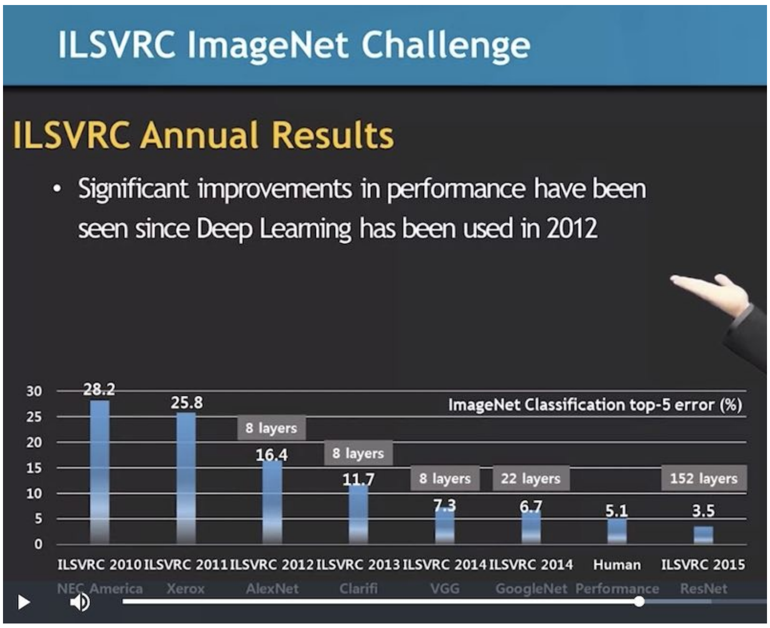
Model	Top-1	Top-5	Size	Caffe Model	Caffe Model
SE-BN-Inception	23.62	7.04	46 M	GoogleDrive	BaiduYun
SE-ResNet-50	22.37	6.36	107 M	GoogleDrive	BaiduYun
SE-ResNet-101	21.75	5.72	189 M	GoogleDrive	BaiduYun
SE-ResNet-152	21.34	5.54	256 M	GoogleDrive	BaiduYun
SE-ResNeXt-50 (32 x 4d)	20.97	5.54	105 M	GoogleDrive	BaiduYun
SE-ResNeXt-101 (32 x 4d)	19.81	4.96	187 M	GoogleDrive	BaiduYun
SENet-154	18.68	4.47	440 M	GoogleDrive	BaiduYun

top-1 ResNet-152 top-1 19.38%2017 ImageNet SENet-154top-1 18.68% ImageNet “”

top-1 80% SENet-154 ResNet-152 1.7 ResNet-152 ResNet-50 2.4 top-1

ResNet-50 ResNet-152 ResNet-152 ResNet-50 2.4 top-1 1.03% 22.37% 21.34% (22.37-21.24)/22.37 = 4.6% top-5 6.36% 5.54% 0.82% top-1 (6.36-5.54)/6.36 = 12.9%

AI top-5 “” top-5



top-1 top-1 top-1

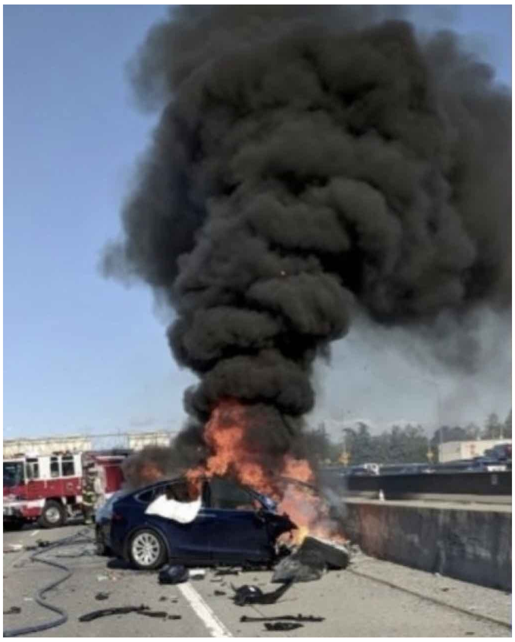
AI

AI

AI “”“AI ” Geoffrey Hinton “”

5 Tesla Autopilot top-5 “...... ” “...... ”

“”“” Tesla Autopilot Autopilot



“top-5 ”Autopilot

Autopilot

“”2018 3 Uber

Uber 6 “”“”“” 40 ......

Uber “” Uber

2018 12 Uber ......

“”“”

The letter is in response to a request for public comment by the NHTSA to a proposal it made last May to amend the Federal Motor Vehicle Safety Standards, a list of 75 rules that automakers must follow before selling cars to customers. Currently, those rules state that cars need to have controls such as a steering wheel and pedals.

BE QUICK ABOUT IT

But self-driving cars may not need these controls, proponents say, and the rules could be a hindrance to the technology being widely released at scale. Waymo and others like Cruise, the self-driving division of GM, and Ford hope to inevitably release tens of thousands of driverless cars without any human controls. Only by cutting the human completely out of the equation can an autonomous vehicle operate safely, these companies argue. And the NHTSA is considering rewriting the rules so self-driving car companies like Waymo can release cars without those features.

Waymo's letter is full of language like "promptly," "should move rapidly," and "urges NHTSA not to await" the completion of other third-party research into autonomous technology. The message it sends is one of urgency: the government needs to drop everything and change the damn rules already.

""" AI"

Tesla Uber AI

L1~L4 " L2 L3 L4 " " L2 "

""""

""""

AI AI

"""""" app……

"AI """ SiriAlexa"""""

""""""""

""""