



GluonCV: Image Classification

Tong He

Applied Scientist, Amazon Web Services.

2018.12.17

Introduction



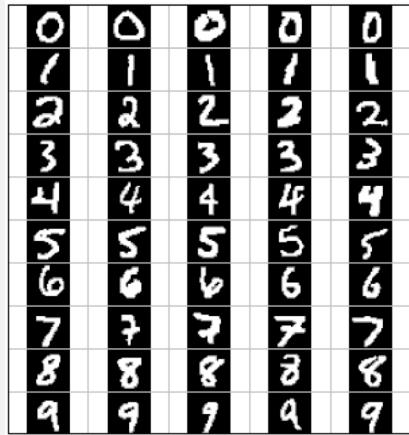
What is classification?

Tell you what is in the image

Dog



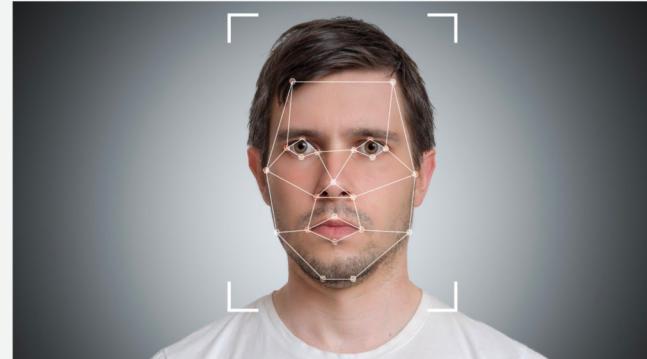
Applications



Digit Recognition

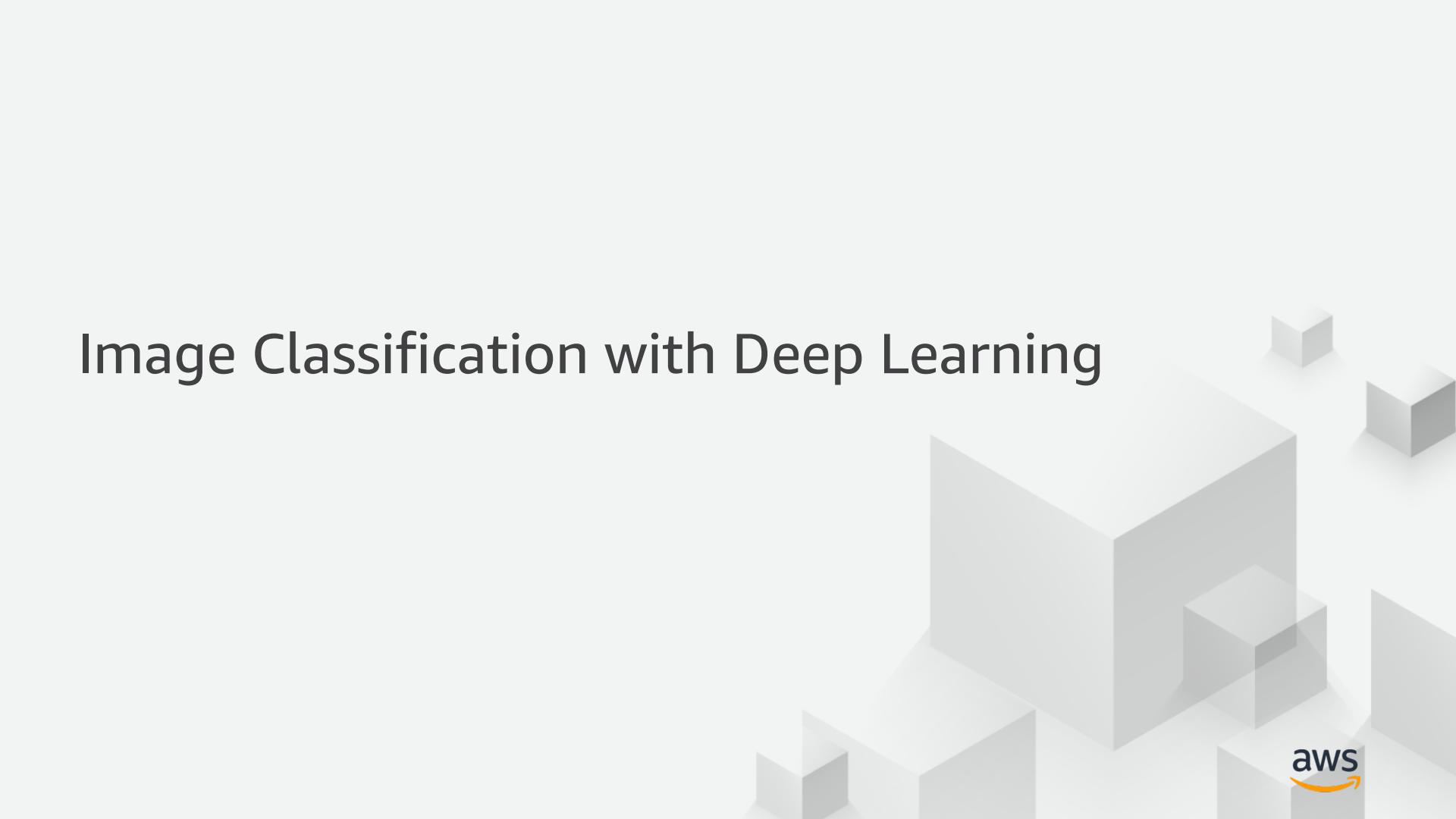


Car Model Detection



Facial Recognition

Image Classification with Deep Learning

The background features a minimalist design with several light gray, semi-transparent 3D cubes of varying sizes scattered across the white surface, suggesting a digital or data-oriented theme.

Classification with Deep Learning

MNIST: The “Hello World” dataset

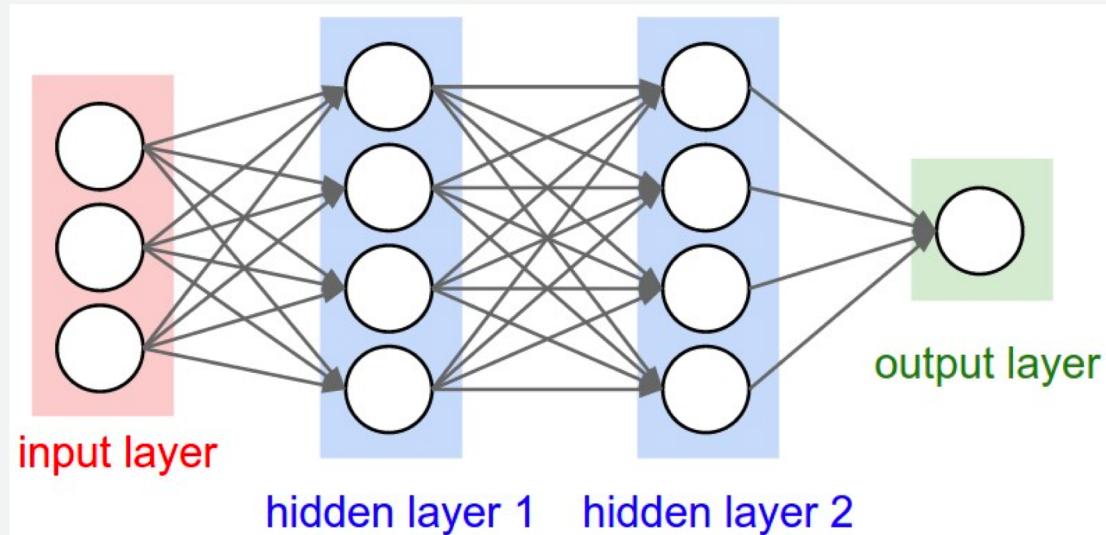
- 28 x 28 input
- Grayscale
- 50000 training
- 10000 test



Classification with Deep Learning

Fully-connected Layer

- Expensive
- 2D insensitive

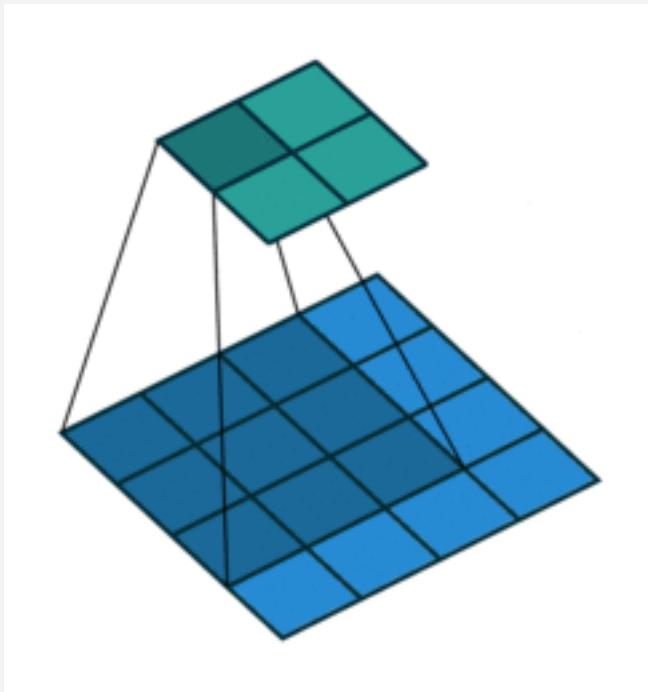


Classification with Deep Learning

Convolution Layer

Input			Kernel			Output		
0	1	2	*	0	1	=	19	25
3	4	5		2	3		37	43
6	7	8						

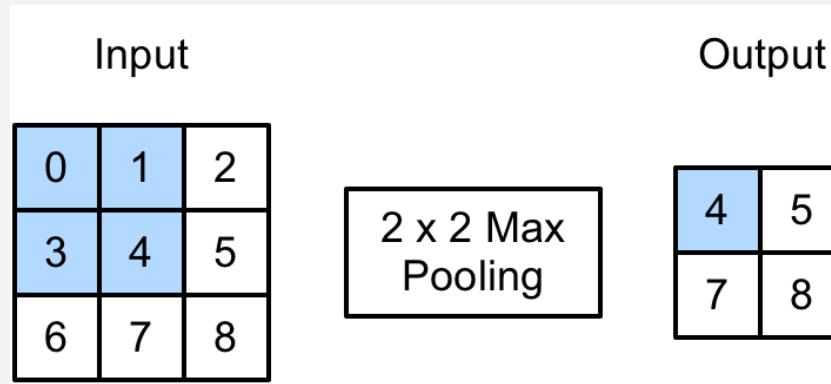
$$0*0 + 1*1 + 3*2 + 4*3 = 19$$



Classification with Deep Learning

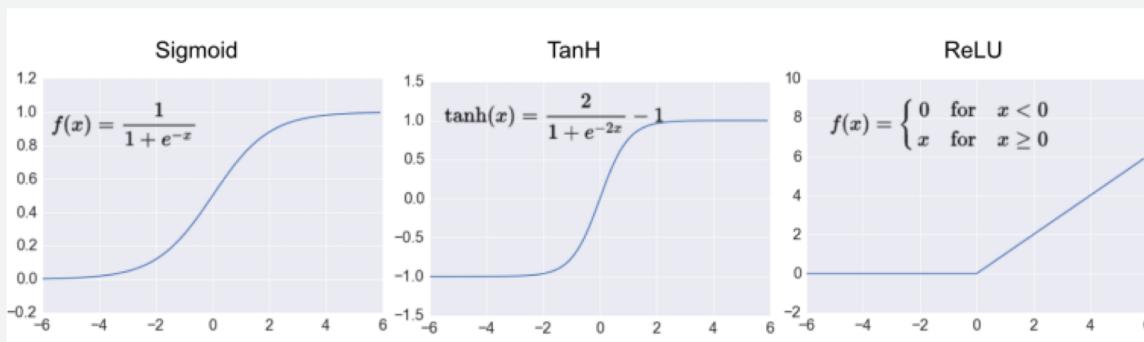
Pooling Layer

- Demonstration Video



Classification with Deep Learning

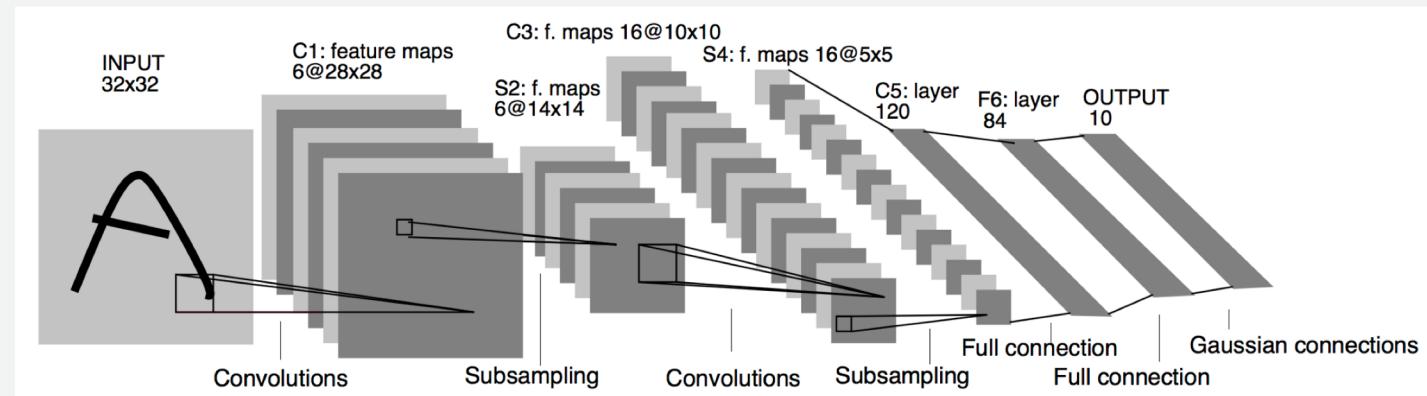
Activation Layer



Classification with Deep Learning

LeNet

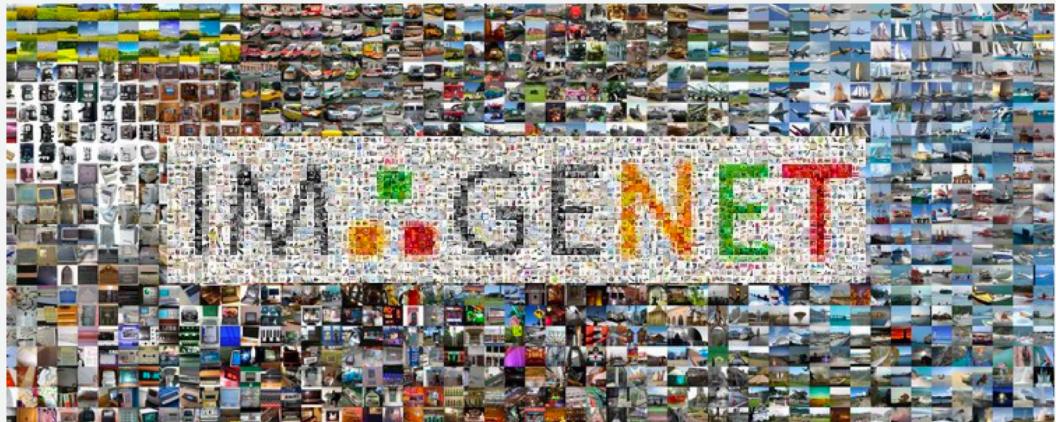
- [Demonstration WebSite](#)



Classification with Deep Learning

ImageNet Challenge

- Natural images
- 1000 classes
- 1.2 million images



Classification with Deep Learning

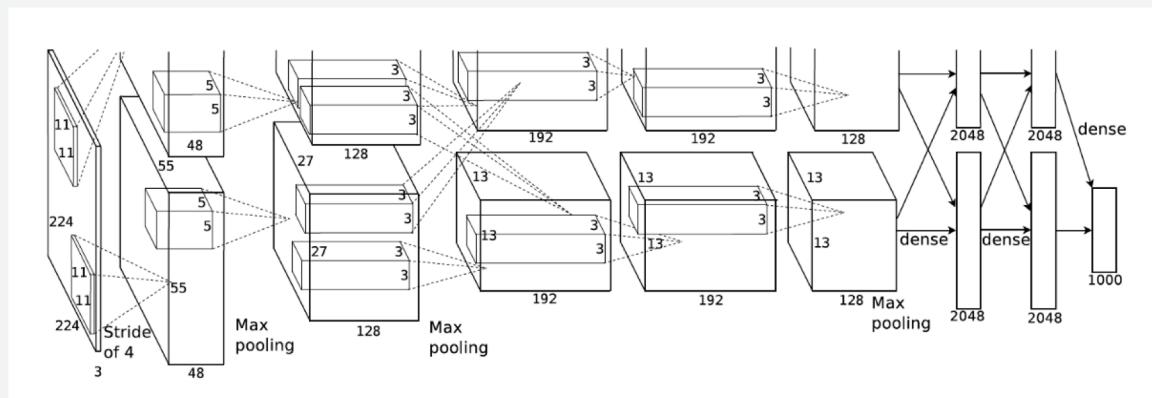
Difficulties

- LeNet is too small
- Hard to process huge amount of data

Classification with Deep Learning

AlexNet

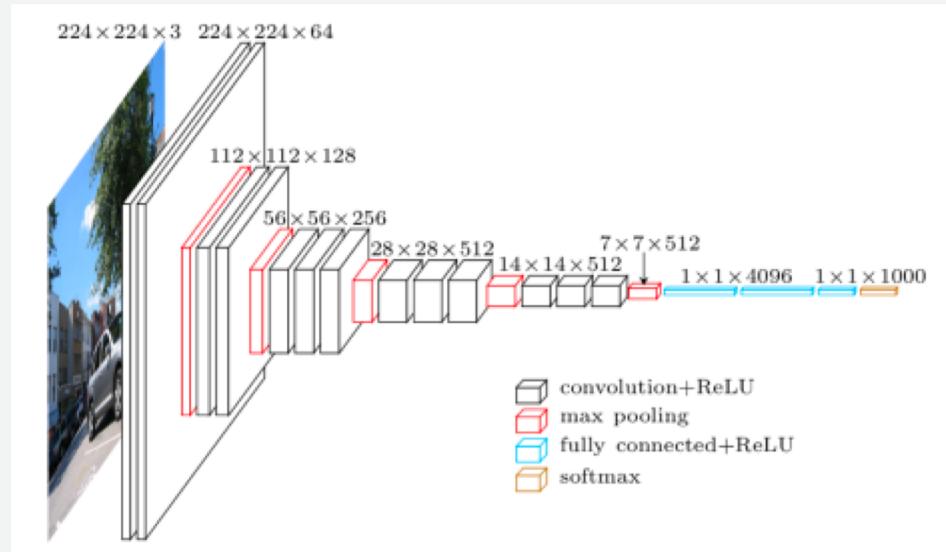
- GPU Accelerated
- Deep (8 layers)



Classification with Deep Learning

VGG

- Deeper
- Better architecture



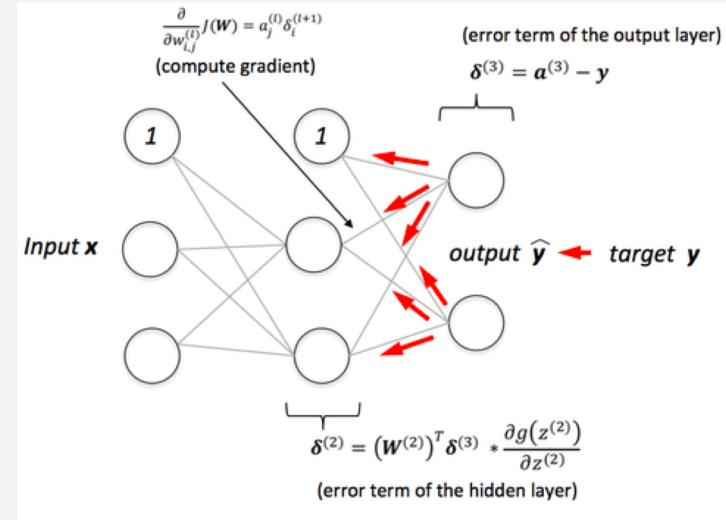
Classification with Deep Learning

Can we go deeper?

- Vanishing/Exploding Gradient

- Chain rule: $\frac{\partial f}{\partial w} = \frac{\partial f}{\partial g} \cdot \frac{\partial g}{\partial w}$

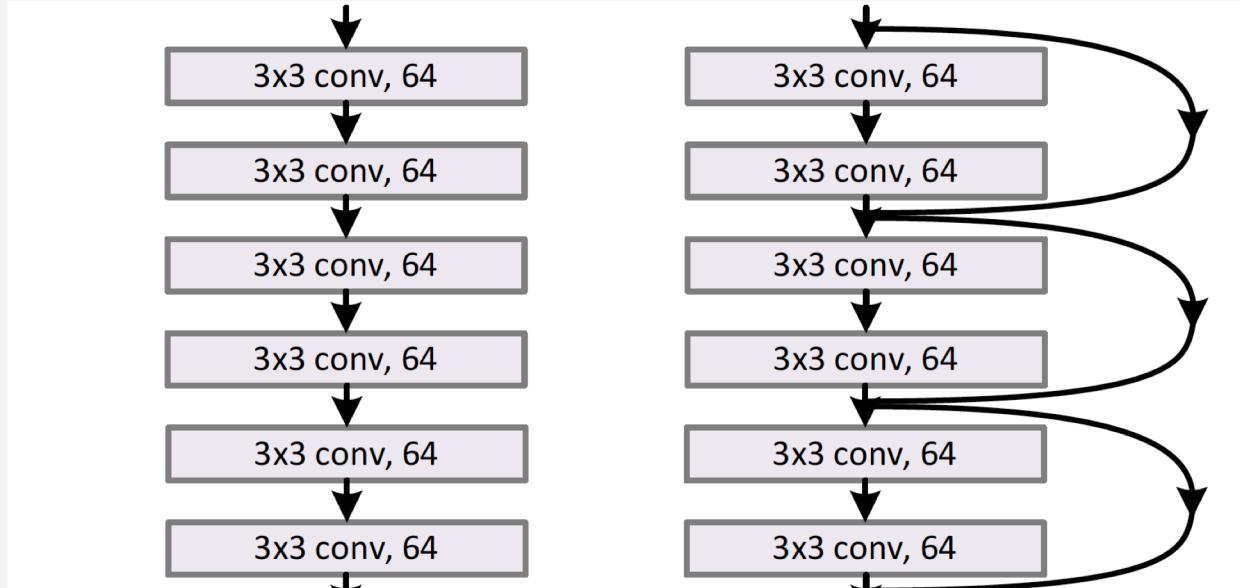
$$\left. \begin{aligned} \frac{\partial f}{\partial g} &= 0.1 \\ \frac{\partial g}{\partial w} &= 0.1 \end{aligned} \right\} \quad \frac{\partial f}{\partial w} = 0.01$$



Classification with Deep Learning

ResNet

- Residual
 - 152 layers!
- Modularized



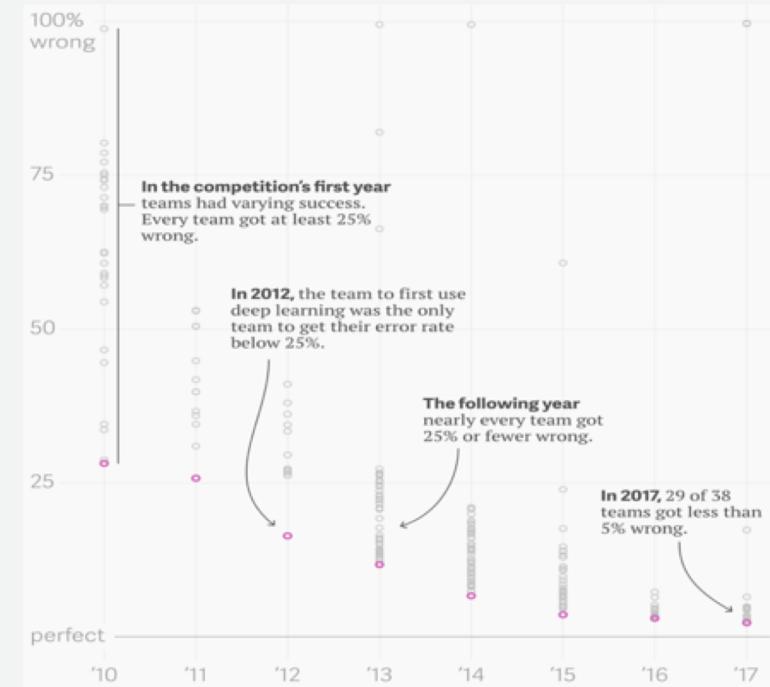
Plain network

Network with residual

Classification with Deep Learning

ImageNet Results

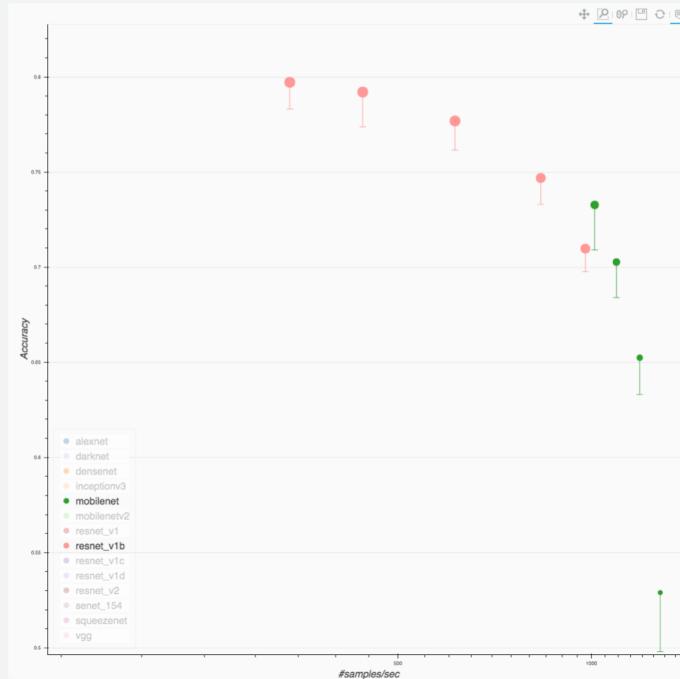
- Top-5 error: 25% to 5%
- Hundreds of new models



Classification with Deep Learning

MobileNet

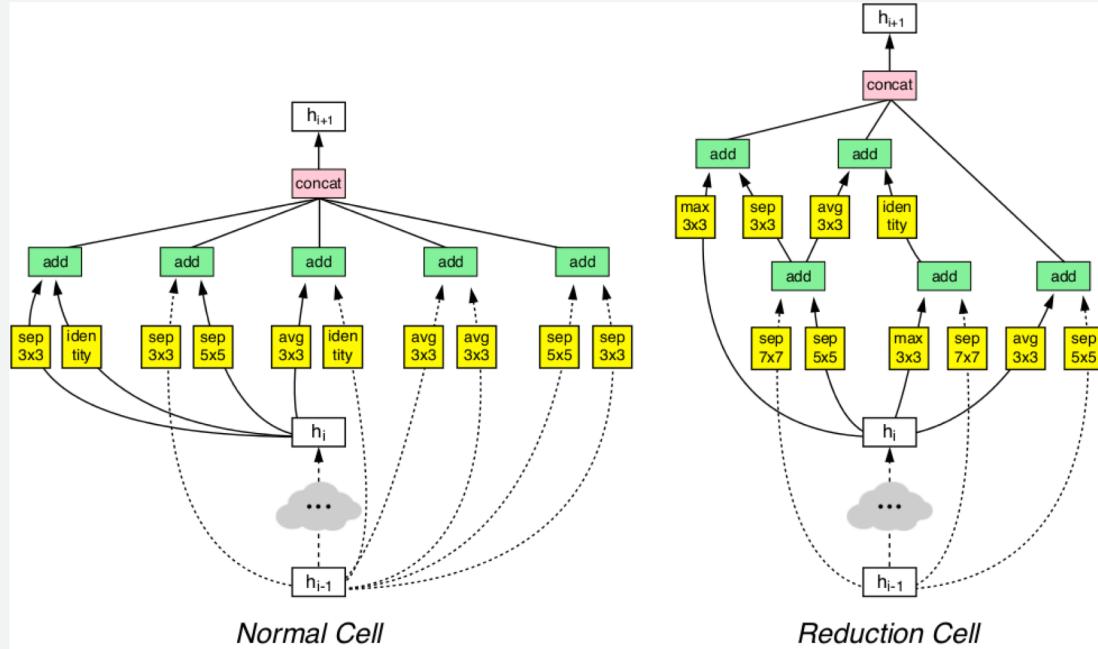
- Depth-wise Convolution
- Parameterized Size
- Fast Inference on devices



Classification with Deep Learning

NASNet

- Automatic Search
- Accurate and heavy



Model Summary

ResNet:

- Balanced accuracy and speed
- Well modularized

MobileNet:

- Small size
- Fast inference

NAS:

- Ongoing research topic

Classification with GluonCV

Classification with GluonCV

Model Zoo

- Pre-trained models
- Can be transferred or directly applied

Classification with GluonCV

GluonCV Model Zoo

- Comprehensive selection
 - AlexNet
 - VGG
 - ResNet
 - MobileNet
 - NASNet
 - ...
- One of the most accurate open-sourced libraries
- Reproducible

Classification with GluonCV

Training Essentials

- Data Preprocessing
- Network architecture definition
- Optimizer
- Loss
- Metric
- GPU Acceleration

Classification with GluonCV

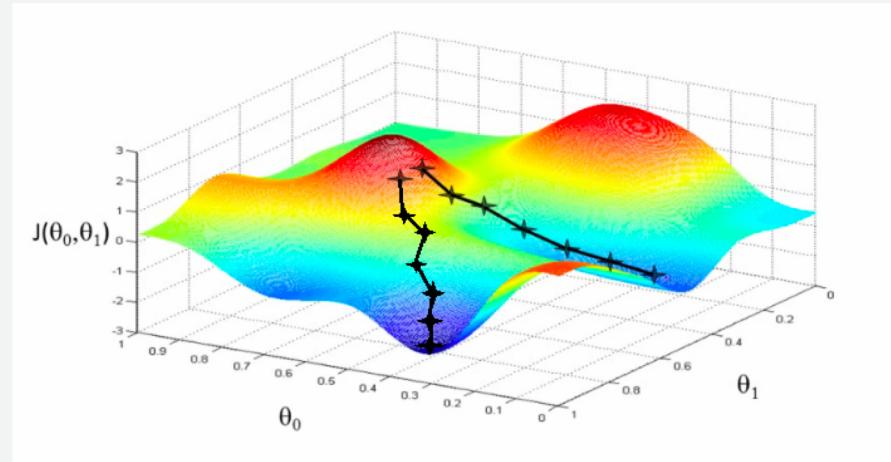
Data Preprocessing



Classification with GluonCV

Optimizers

- SGD
- Adam
- RMSProp
- ...



Classification with GluonCV

Advanced Tricks

- Label smoothing
- Learning rate schedule
- Mix-Up
- Knowledge Distillation

Classification with GluonCV

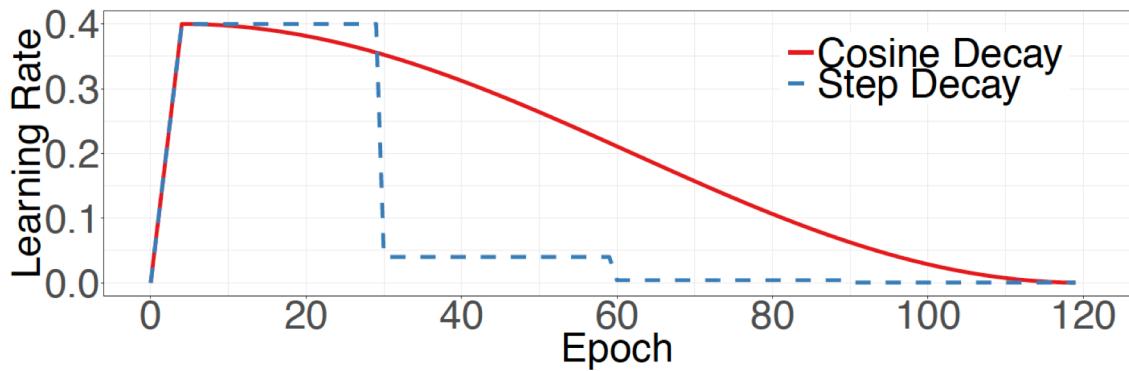
Label Smoothing

- One hot: $(0, 1, 0, 0, 0)$
- Smoothed: $(0.01, 0.96, 0.01, 0.01, 0.01)$
- Prevent overfitting!

Classification with GluonCV

Learning Rate Schedule

- Step
- Cosine
- Poly

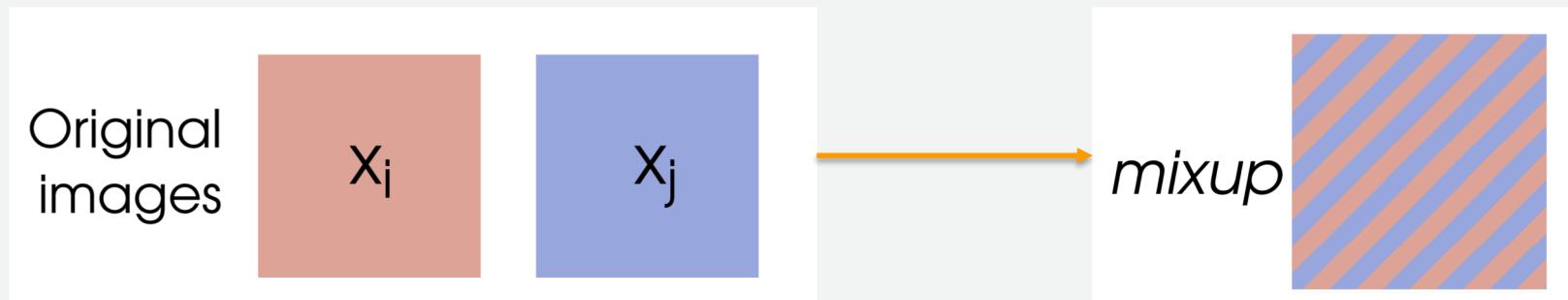


(a) Learning Rate Schedule

Classification with GluonCV

Mix-Up

- Linear mapping
- $f(ax_i + bx_j) = af(x_i) + bf(x_j)$



Classification with GluonCV

Knowledge Distillation

- Dark Knowledge
 - Dog vs Cat
 - Dog vs Car

cow	dog	cat	car
0	1	0	0

original hard targets

cow	dog	cat	car
10^{-6}	.9	.1	10^{-9}

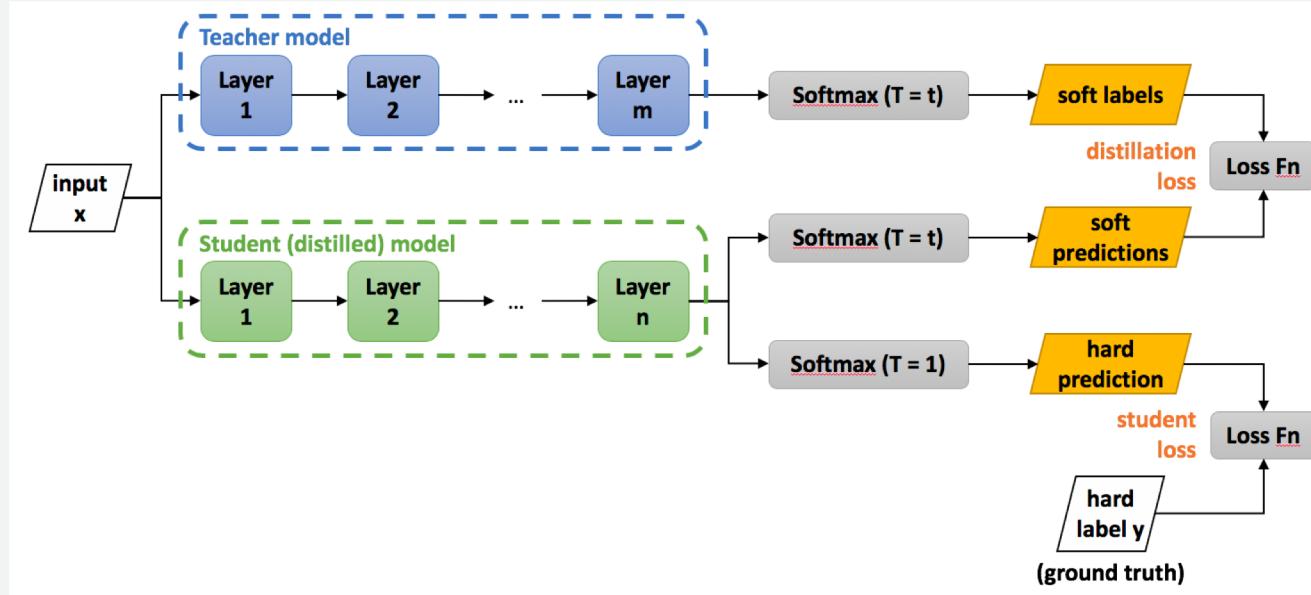
output of geometric ensemble

cow	dog	cat	car
.05	.3	.2	.005

softened output of ensemble

Classification with GluonCV

Knowledge Distillation



Classification with GluonCV

GluonCV Model Zoo

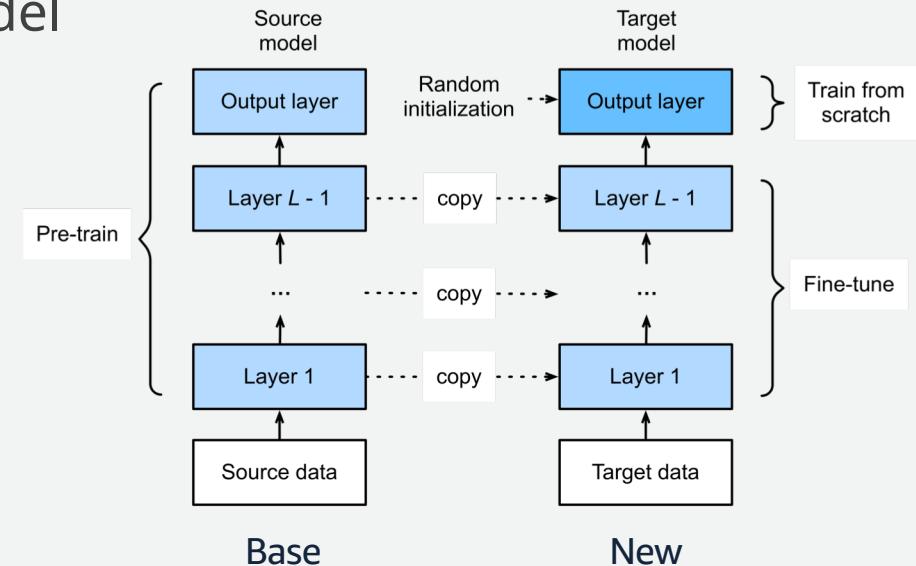
He, Tong, et al. "Bag of Tricks for Image Classification with Convolutional Neural Networks" arXiv preprint arXiv:1812.01187 (2018).



Classification with GluonCV

Transfer learning

- Based on a pre-trained model
- Re-define the output layer



Classification with GluonCV

Resources:

- Model Zoo: https://gluon-cv.mxnet.io/model_zoo/classification.html
- Tutorials: https://gluon-cv.mxnet.io/build/examples_classification/index.html
- Deep Learning Book: <http://diveintodeeplearning.org/>

Classification with GluonCV

Hands on!