ROBOT INTELLIGENCE LAB

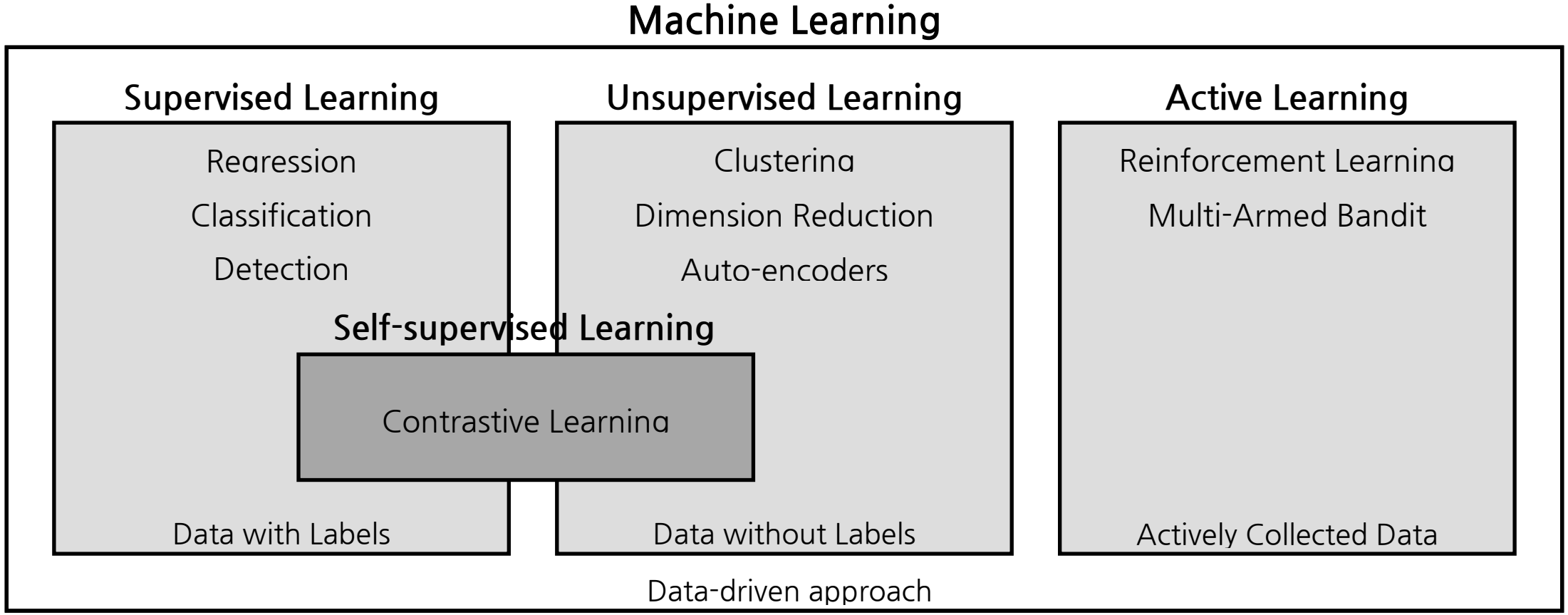
Title

Subtitle

Sungjoon Choi, Korea University

kakaobrain

Introduction



Introduction

ROBOT INTELLIGENCE LAB

- Key Components of Deep Learning
  - The data that we can learn from.
  - A model of how to transform the data.
  - A loss function that quantifies the badness of the model.
  - An algorithm to adjust the parameters to minimize the loss.

Data

ROBOT INTELLIGENCE LAB

Classification

Semantic Segmentation

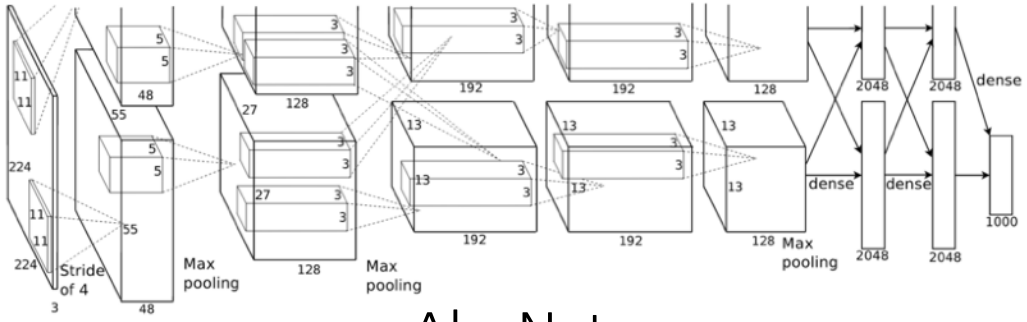
Detection

Pose Estimation

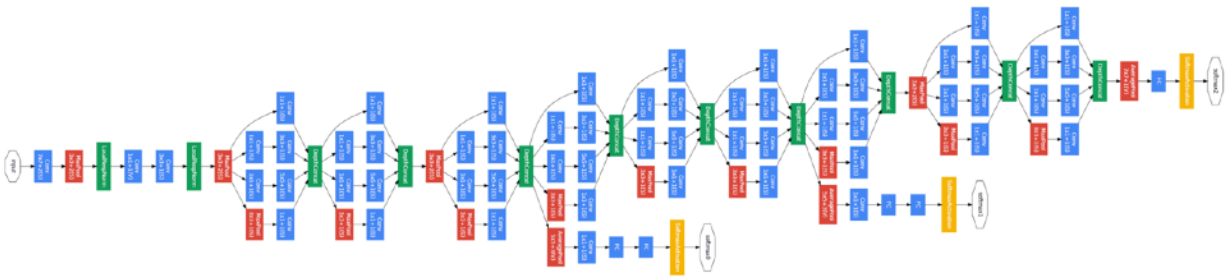
Visual QnA

Model Architectures

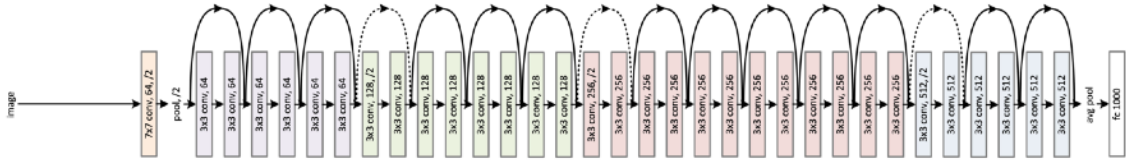
ROBOT INTELLIGENCE LAB



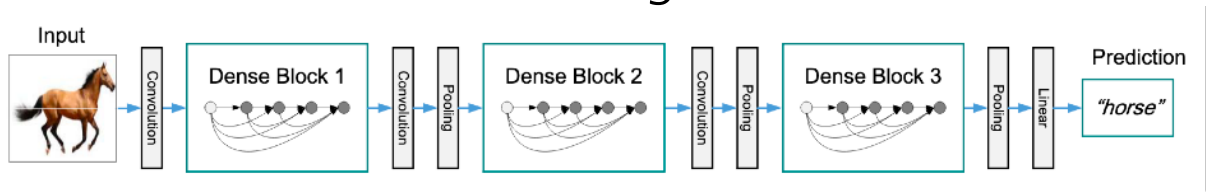
AlexNet



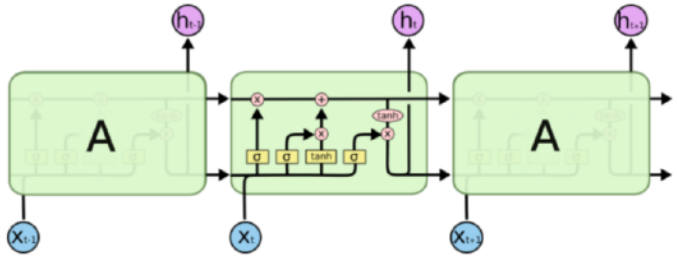
GoogLeNet



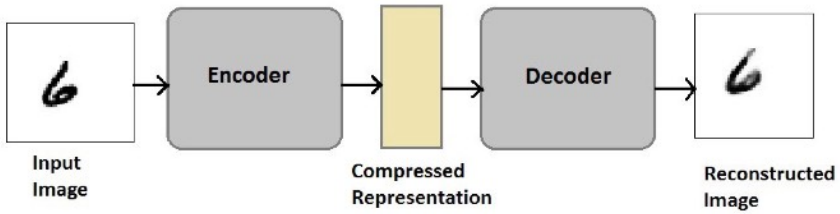
ResNet



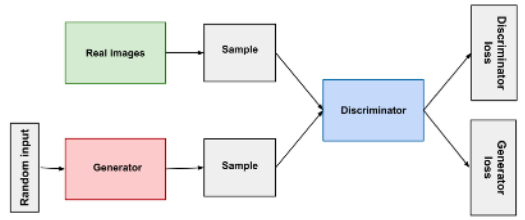
DenseNet



LSTM



Deep AutoEncoders



GAN

Loss

ROBOT INTELLIGENCE LAB

Regression Task

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}\sum_{d=1}^{D}(y_i^{(d)} - \hat{y}_i^{(d)})^2$$

Classification Task

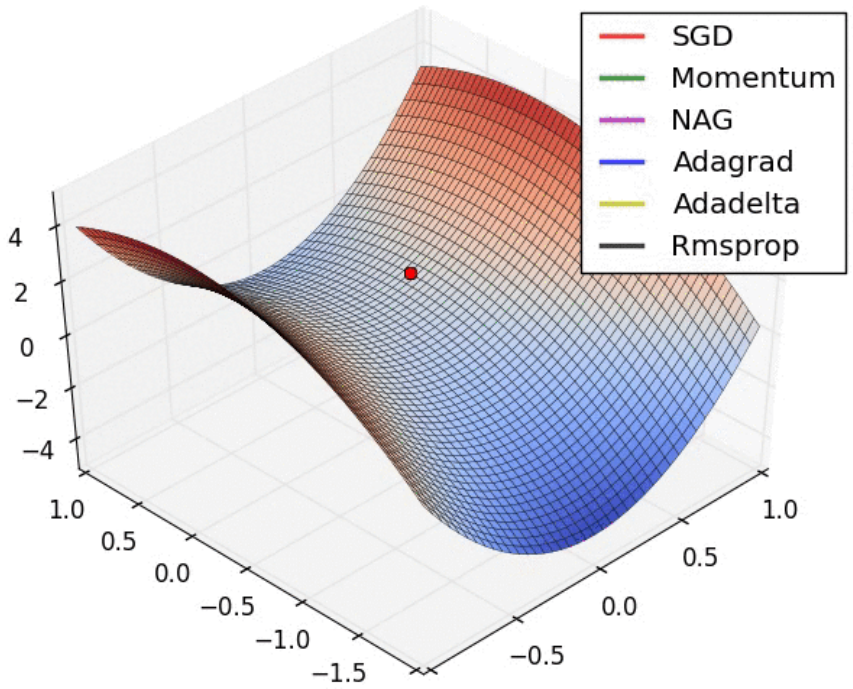
$$\mathrm{CE} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{d=1}^{D} y_i^{(d)} \log \hat{y}_i^{(d)}$$

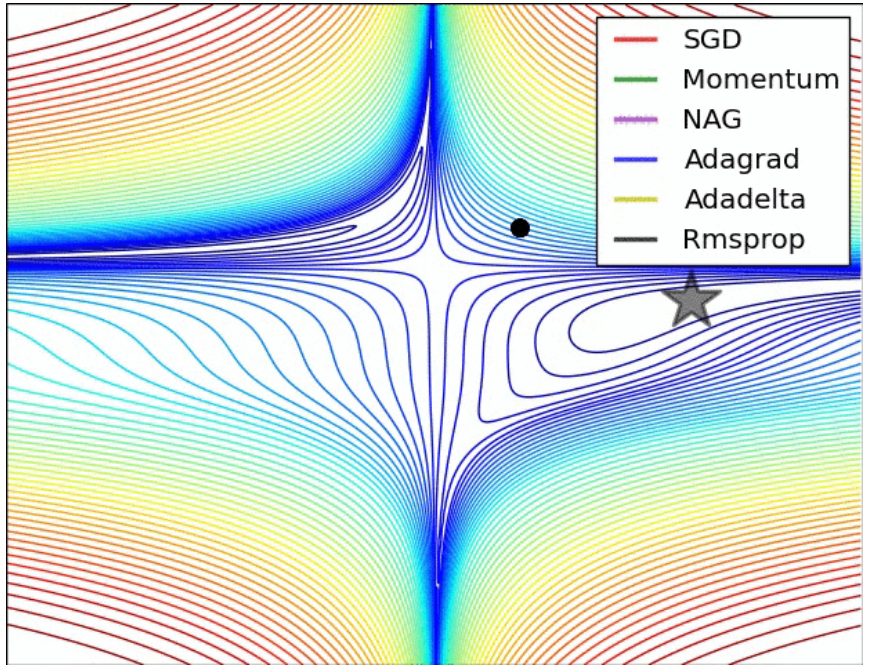
Probabilistic Task

$$\mathrm{MLE} = \frac{1}{N}\sum_{i=1}^{N}\sum_{d=1}^{D} \log \mathcal{N}(y_i^{(d)}; \hat{y}_i^{(d)}, 1)$$ (=MSE)

Algorithm

ROBOT INTELLIGENCE LAB





Dropout

Early stopping

k-fold validation

Weight decay

Batch normalization

MixUp

Ensemble

Bayesian Optimization

ROBOT INTELLIGENCE LAB

Historical Review

Deep Learning's Most Important Ideas - A Brief Historical Review

Denny Britz

2020-07-29

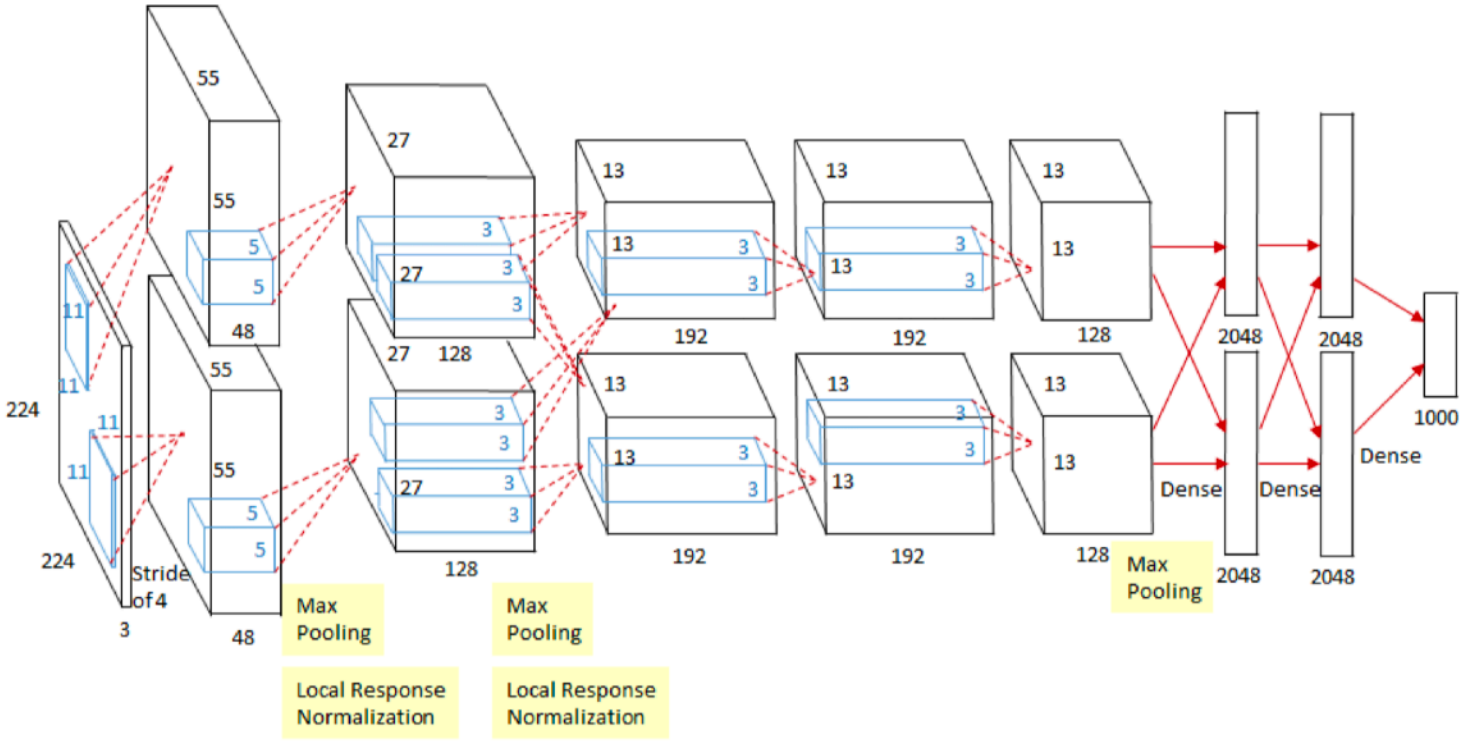
Historical Review

ROBOT INTELLIGENCE LAB

The goal of this post is to review those ideas that have stood the **test of time**.

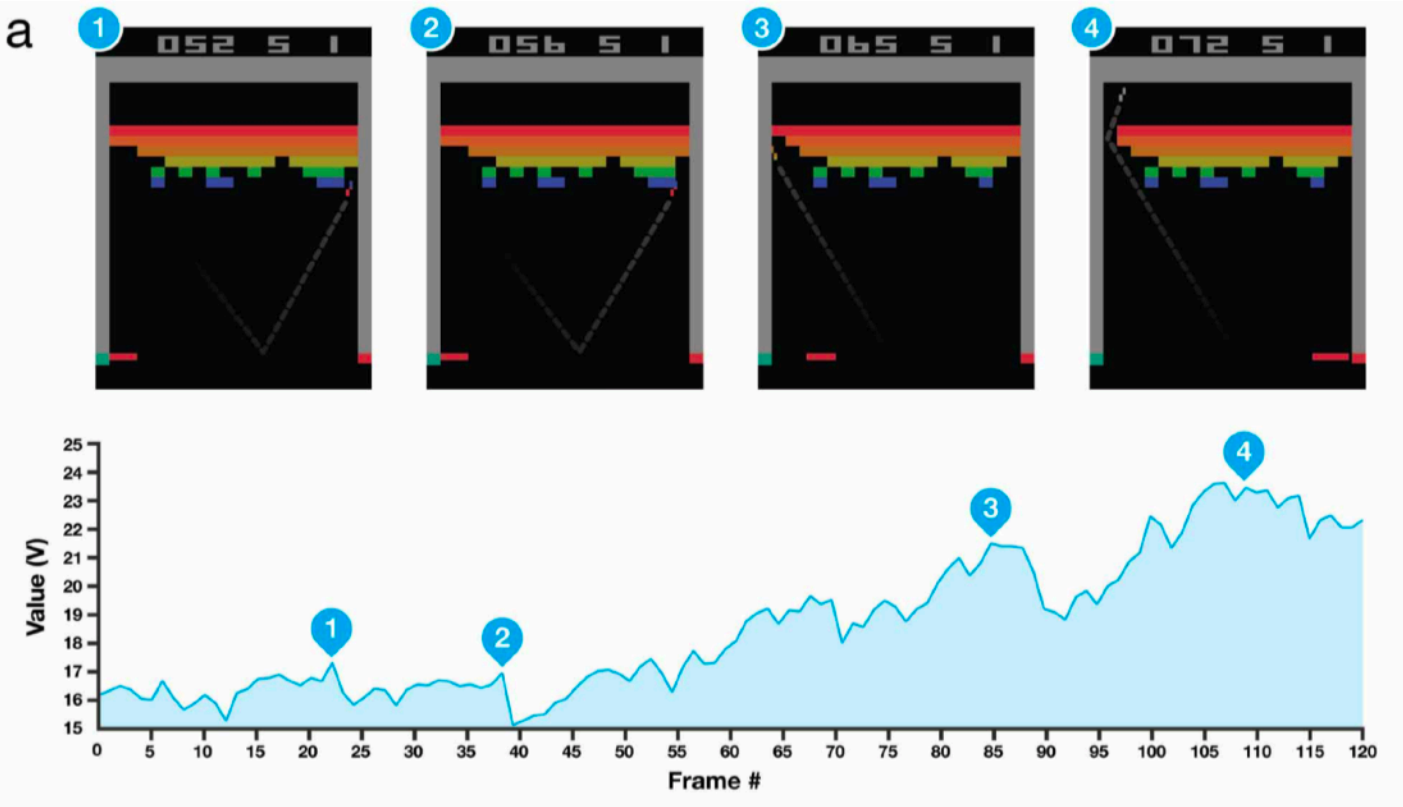
2012 - AlexNet

ROBOT INTELLIGENCE LAB



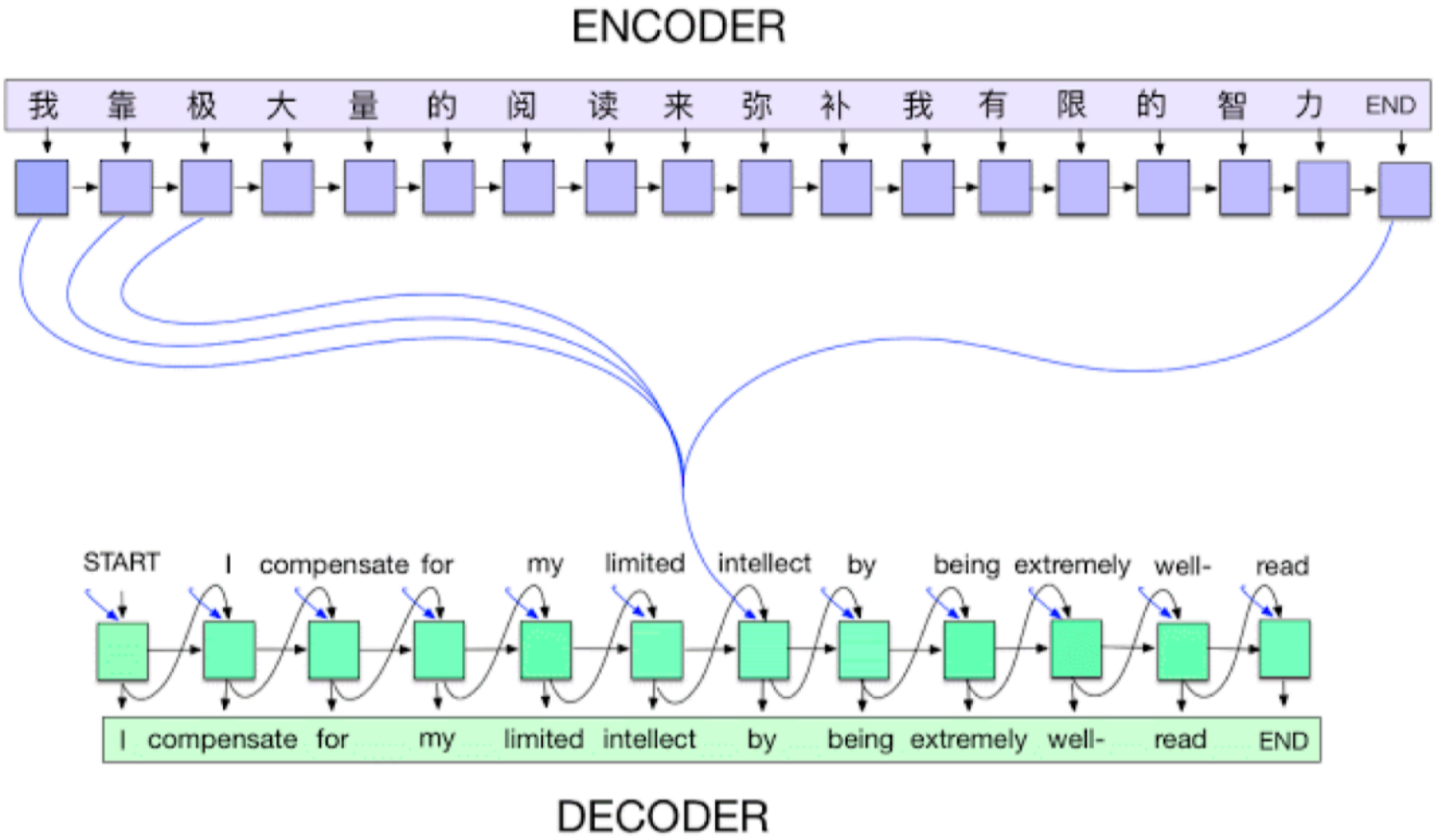
2013 - DQN

ROBOT INTELLIGENCE LAB

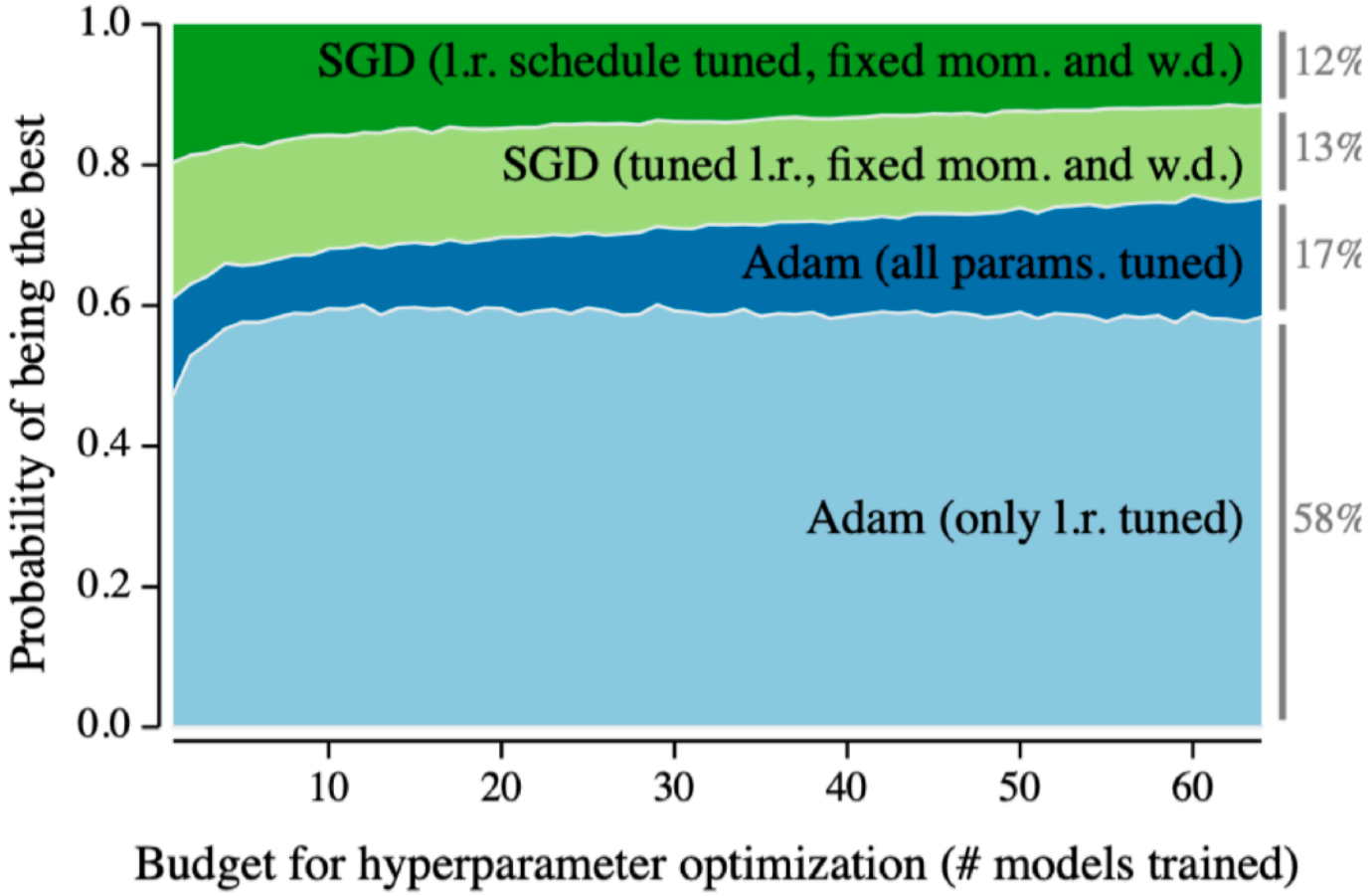


2014 - Encoder-Decoder with Attention

ROBOT INTELLIGENCE LAB

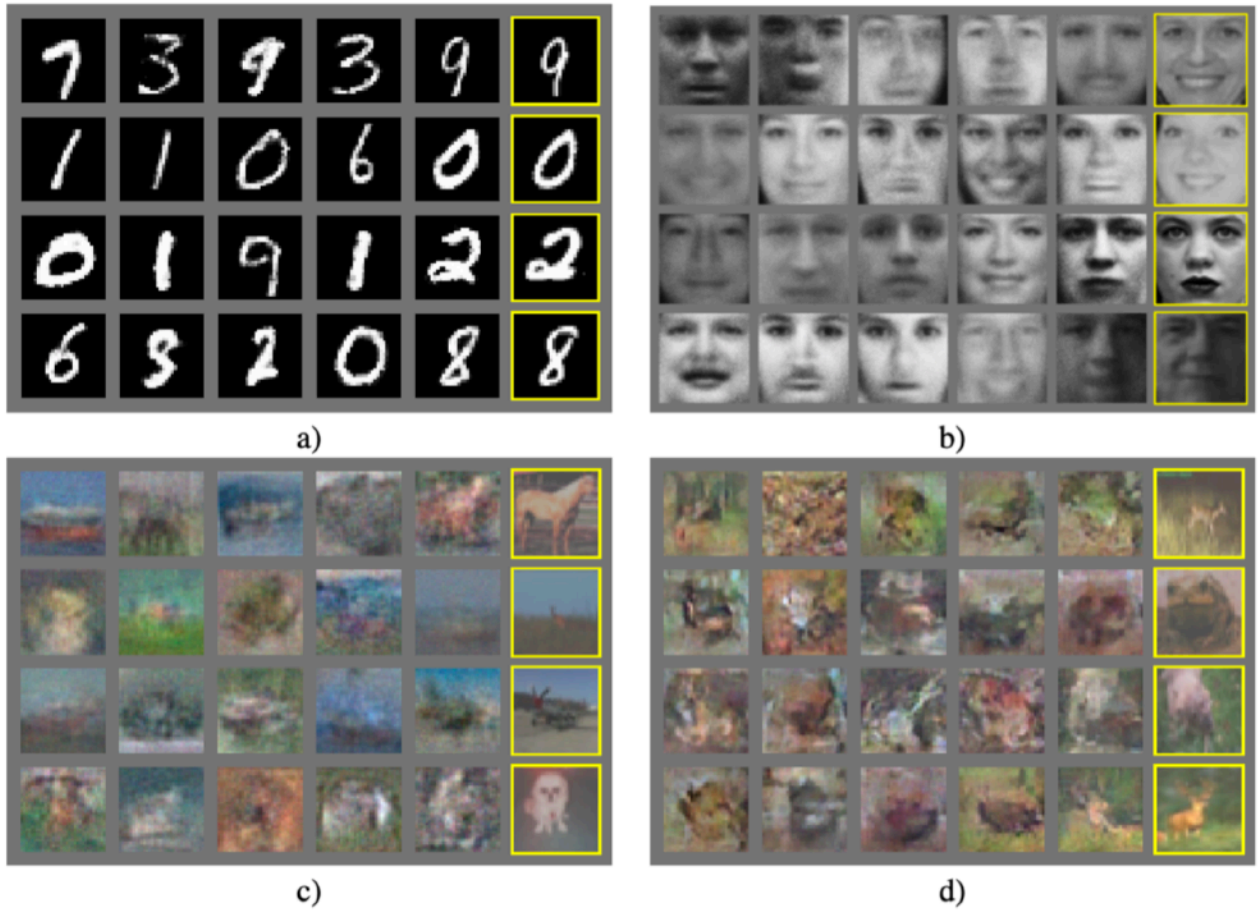


2014 - Adam


SGD (l.r. schedule tuned, fixed mom. and w.d.)
SGD (tuned l.r., fixed mom. and w.d.)
Adam (all params. tuned)
Adam (only l.r. tuned)
12%
13%
17%
58%
Probability of being the best
1.0
0.8
0.6
0.4
0.2
0.0
10
20
30
40
50
60
Budget for hyperparameter optimization (# models trained)

2015 - GAN

ROBOT INTELLIGENCE LAB



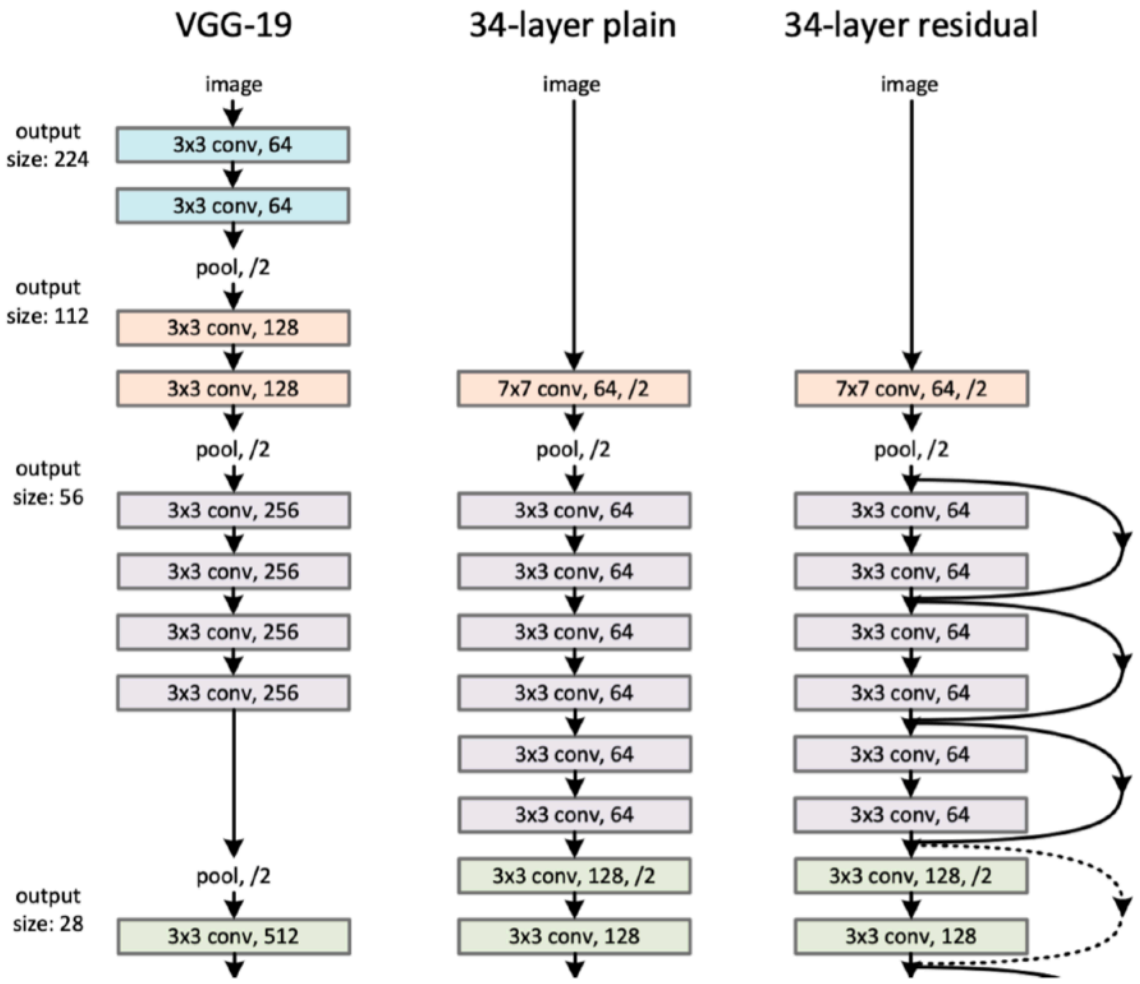
2015 - GAN

ROBOT INTELLIGENCE LAB

“Finally, we would like to thank Les Trois Brasseurs for stimulating our creativity.”

2015 - ResNet

ROBOT INTELLIGENCE LAB



2017 - Transformer

ROBOT INTELLIGENCE LAB

Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com

Noam Shazeer* Google Brain noam@google.com

Niki Parmar* Google Research nikip@google.com

Jakob Uszkoreit* Google Research usz@google.com

Llion Jones* Google Research llion@google.com

Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu

Łukasz Kaiser* Google Brain lukaszkaiser@google.com

Illia Polosukhin* ‡ illia.polosukhin@gmail.com

2017 - Transformer

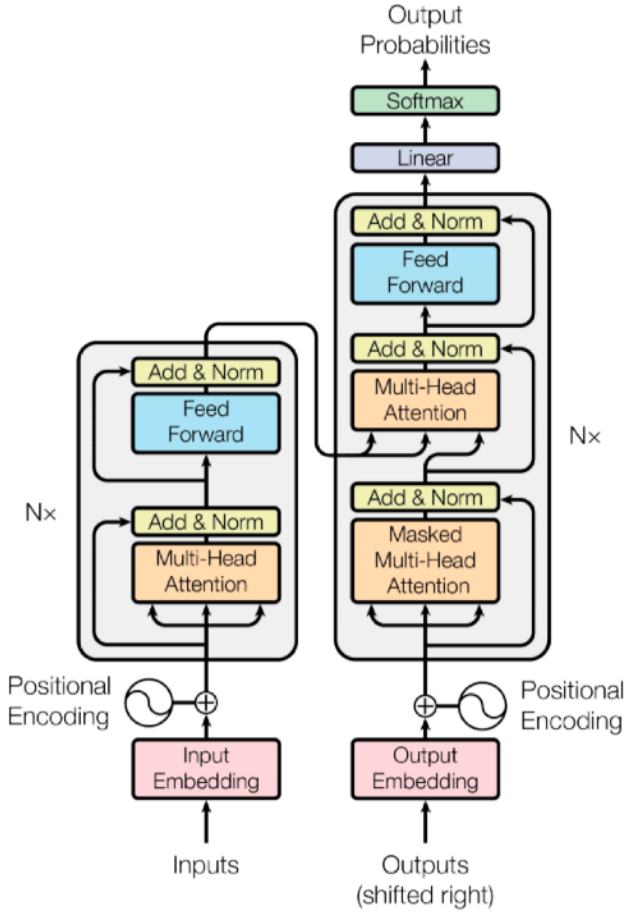


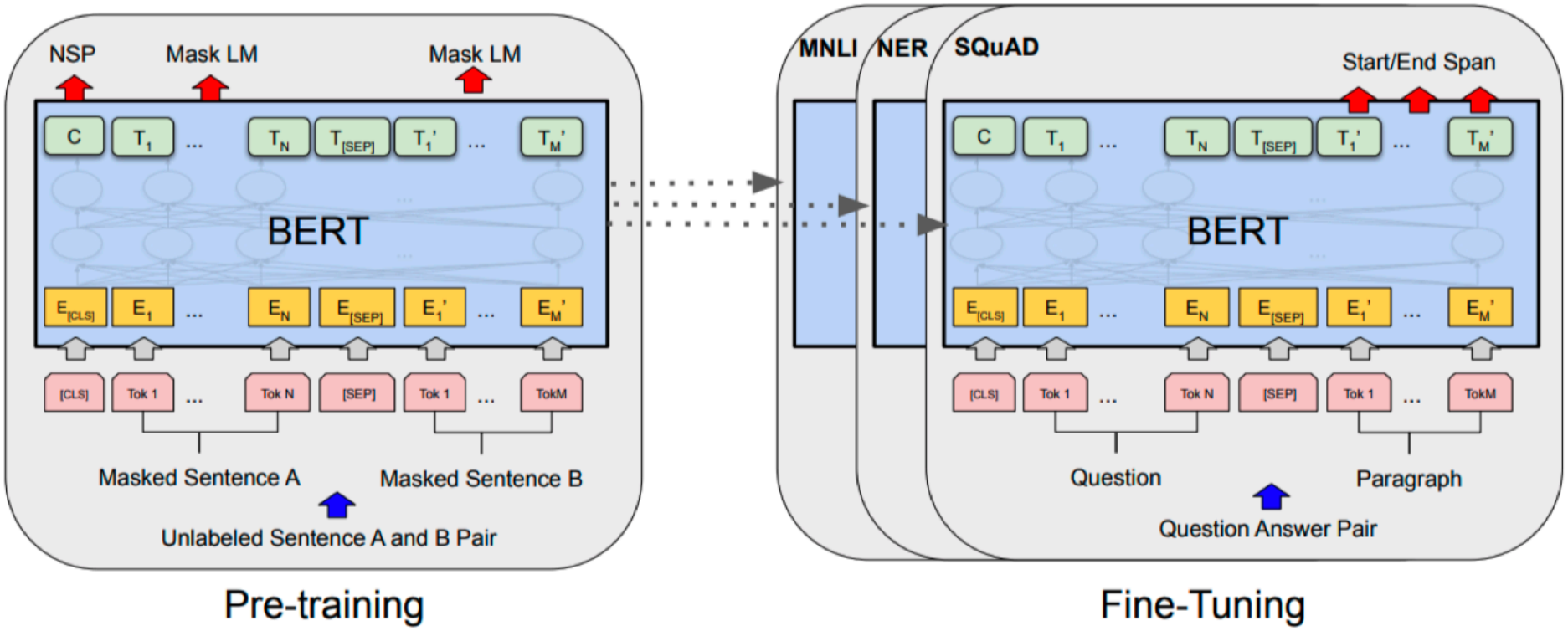
Figure 1: The Transformer - model architecture.

Single Head Self-Attention Calculation

$$\text{softmax}\left(\frac{Q \times K^T}{\sqrt{d_k}}\right) V = Z$$

2018 - BERT

ROBOT INTELLIGENCE LAB



2019 - Pre-trained Language Model

ROBOT INTELLIGENCE LAB

OpenAI

GPT-3, an autoregressive language model with 175 billion parameters

2020~ - Self Supervised Learning

ROBOT INTELLIGENCE LAB

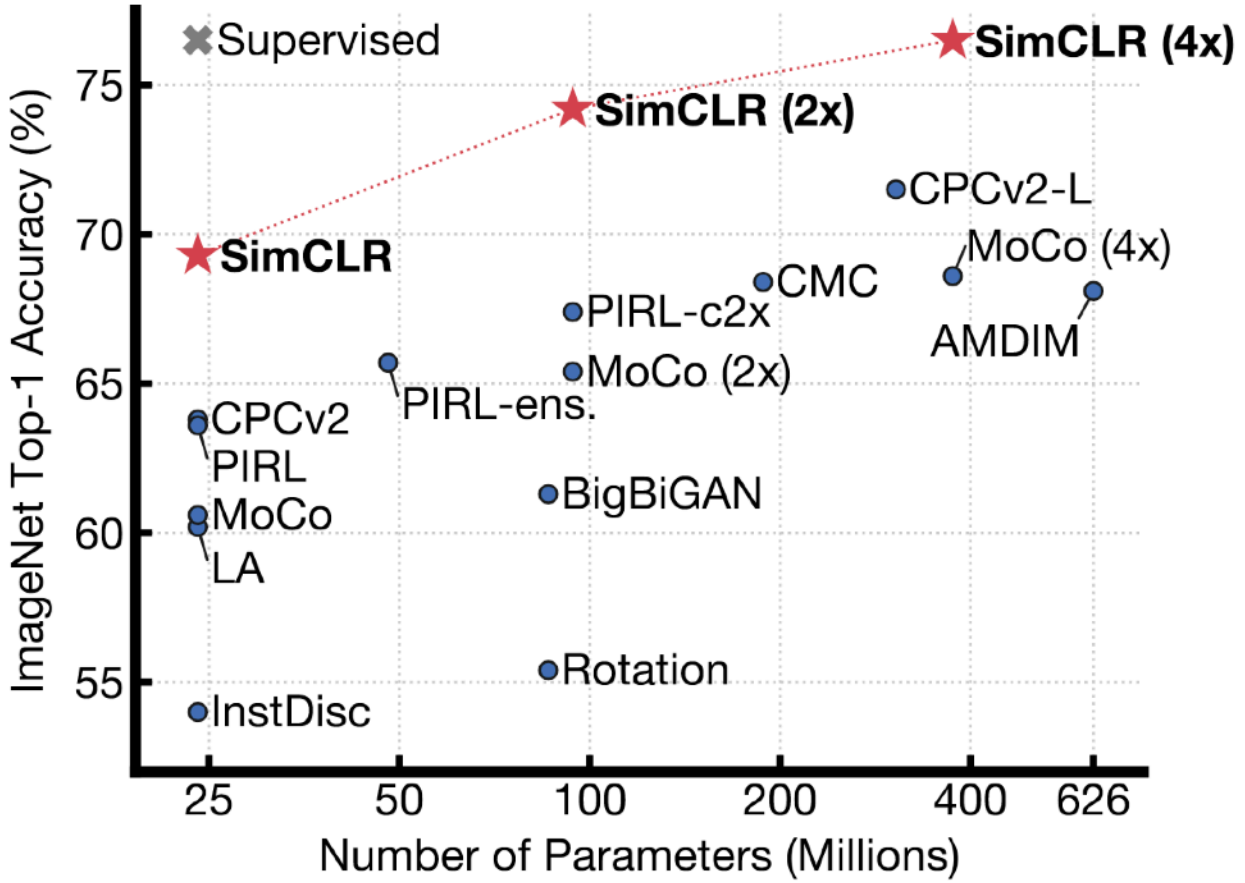
A Simple Framework for Contrastive Learning of Visual Representations

---

Ting Chen [1] **Simon Kornblith** [1] **Mohammad Norouzi** [1] **Geoffrey Hinton** [1]

2020~ - Self Supervised Learning

ROBOT INTELLIGENCE LAB



Thank You