Help me identify tree species!

造林学 B4 小林　壱徳久

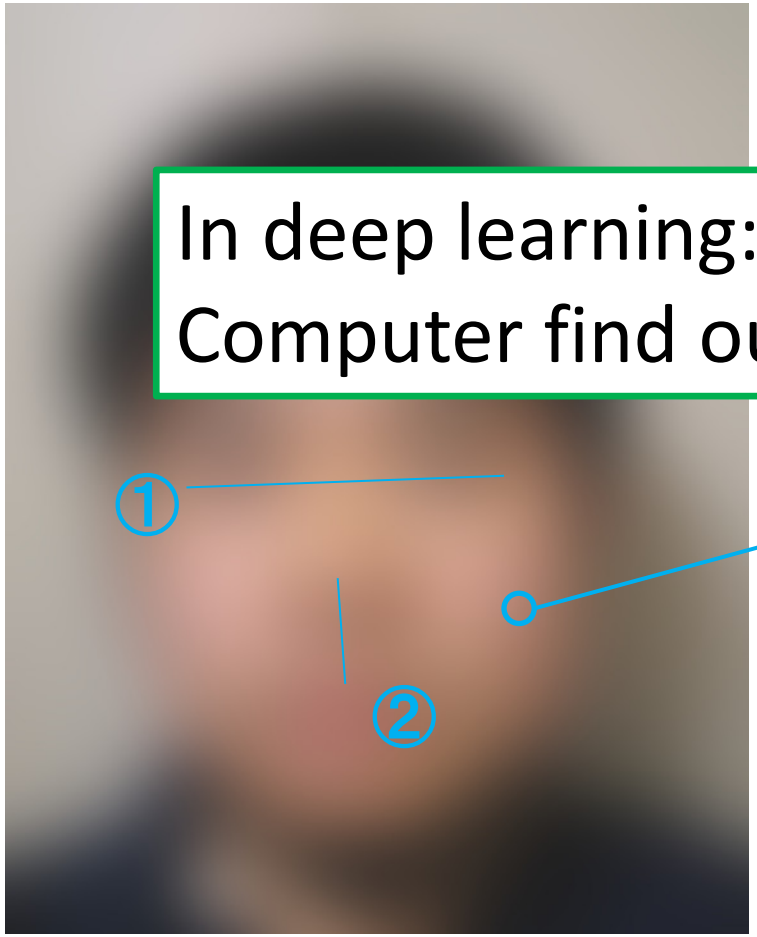
Play with Tensorflow

- Tensorflow is deep learning system Google developed, which is used in AlphaGo
- It is available for free

I tried to make computer identify tree species using this system

What is deep learning?

For example, in face recognition



In traditional method:
To identify face, human must

In deep learning:
Computer find out characteristics by itself

・ratio of distances between parts

etc…

(Red:177 Green:126 Blue:107)

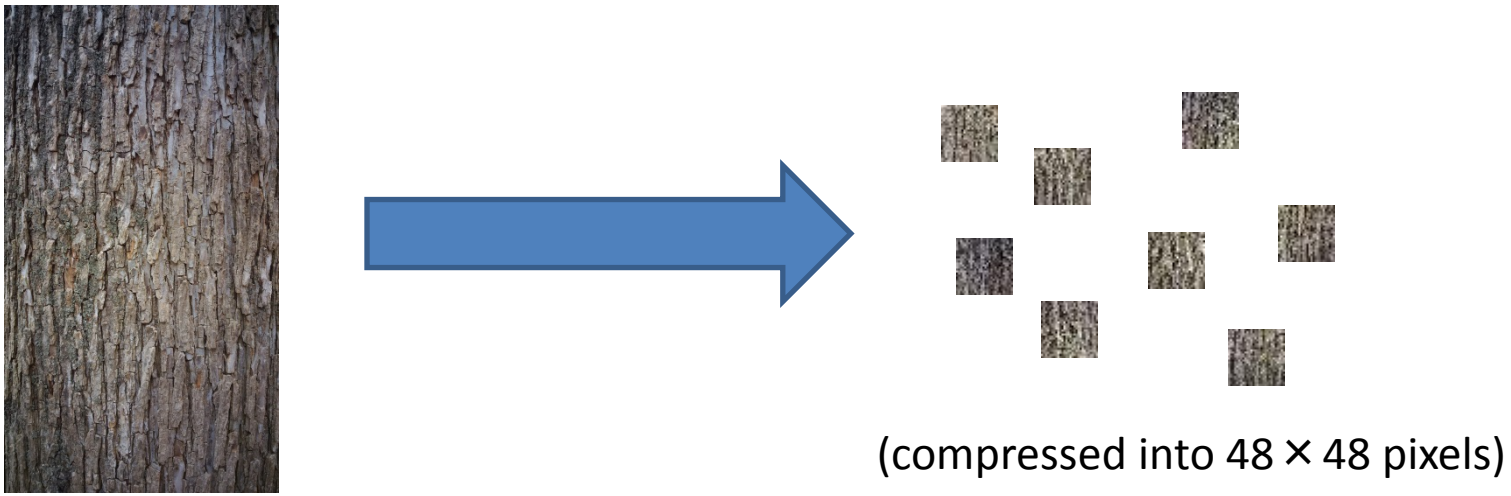
①/② = 2.97

Material&Method(1)

- I chose 6 species to identify
  ハルニレ (*Ulmus davidiana* var. *japonica*)
  シラカンバ (*Betula platyphylla*)
  カツラ (*Cercidiphyllum japonicum*)
  エゾヤマザクラ(*Cerasus sargentii*)
  サトザクラ (cherry blossoms for cultivar)
  モミジバスズカケノキ (*Platanus x acerifolia*)
- 4 samples set up per species
- All samples are in Sapporo campus

Material&Method(2)

- I took 4 pictures of bark per sample
- randomly trimmed a picture and got 10 patterns per picture
- So 160 patterns per species provided



(compressed into 48 × 48 pixels)

Material&Method(3)

- All patterns are divided into two data, TRAIN_DATA and TEST_DATA

All 6 ×160 pattens

TRAIN_DATA(768 pattens): Computer learns from this data

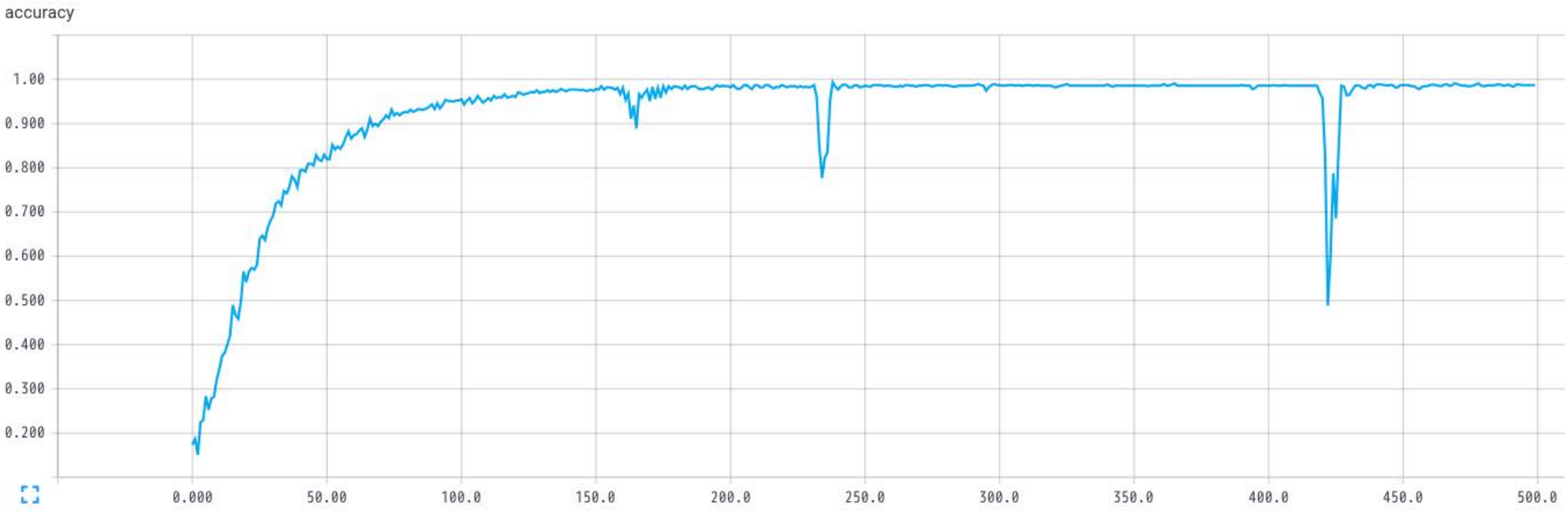
TEST_DATA(192 pattens): used to evaluate accuracy of model Computer produced

This step is very important to evaluate model exactly
(Questions in exam should not be the same ones in textbook)

Material&Method(4)

- When learning, computer randomly pick up 100 patterns from TRAIN_DATA and optimize parameters of model
- This step are repeated 500 times

Result(1)



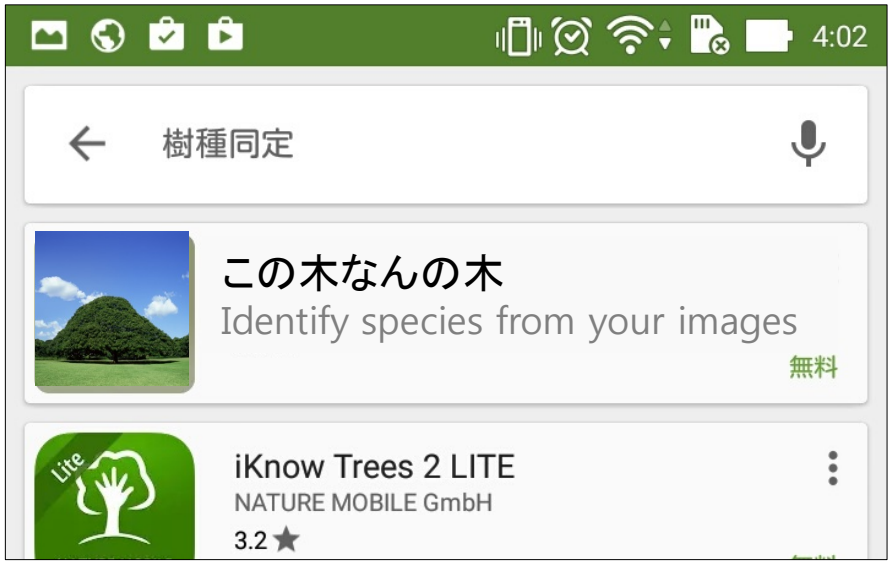
This graph shows accuracy of model applying to TRAIN_DATA

Result(2)

- Accuracy of model applying to TEST_DATA is **84.02%**
- Too low for practical use
- In each species:

87.50%	ハルニレ (*Ulmus davidiana* var. *japonica*)
68.75%	シラカンバ (*Betula platyphylla*)
94.11%	カツラ (*Cercidiphyllum japonicum*)
90.62%	エゾヤマザクラ(*Cerasus sargentii*)
93.75%	サトザクラ (cherry blossoms for cultivar)
68.75%	モミジバスズカケノキ (*Platanus x acerifolia*)

Application idea



- Help people to learn trees and forest

① take a picture

② send image to server

③ searching

④ return information:
name, property, usage …

For more accuracy

- Collect more samples
- Use bigger images
- Combine with other information
  -leaf shape
  -where a picture is taken (using GPS)
  → reduce possible species

Application in our field

- Assume distribution of species from aerial photographs

- Identify mycorrhizal fungi and automatically count them

- With drone?

What Tensorflow does?

(you can sleep now)

- Just matrix calculation
- A pattern is 48 × 48 pixels, so input data is represented by 2304-dimension vector(x_i)
- Output data is 6-dimension(6 species) vector(y_i), whose element shows probability of input data being each species.
- Simplest model as follow

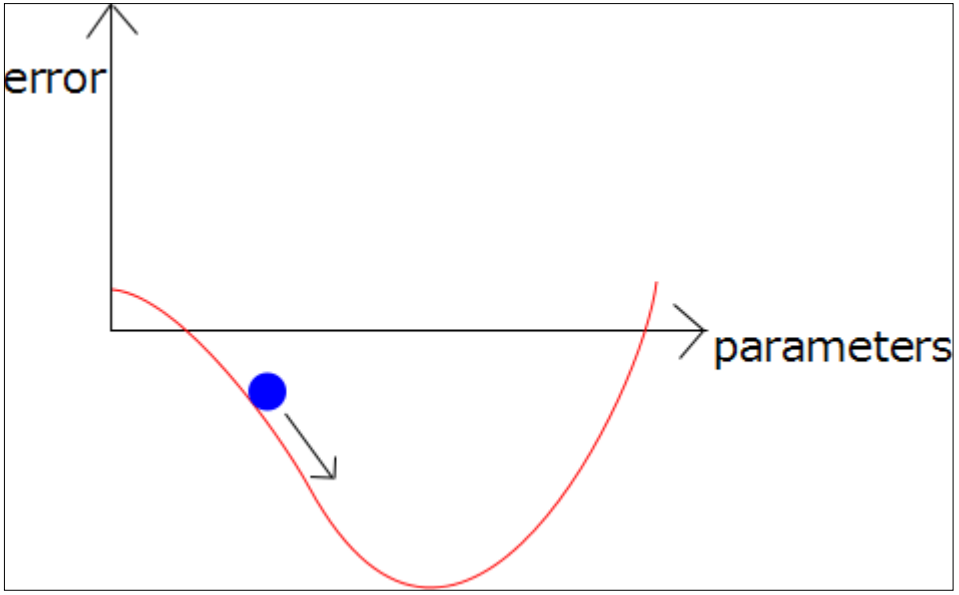
$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{pmatrix} = A \times \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{2303} \\ X_{2304} \end{pmatrix}$$

where A has 6 rows and 2304 columns

Tensorflow tries to find best matrix A by which inferred species match true species

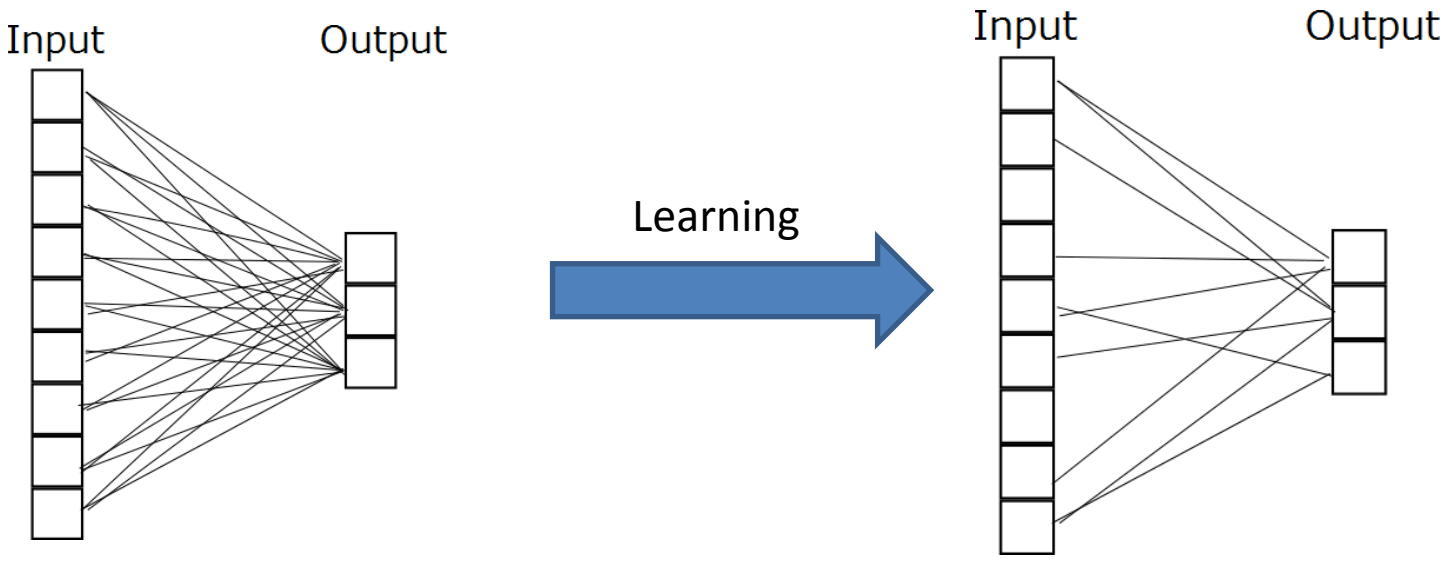
How to find best matrix A

- Just like regression analysis
- Defining **error function** between inferred and true species
- In each learning step, Tensorflow changes parameters of A slightly as reducing the value of error function



By the way

- If $y_1 = 2.1x_1 + 3.4x_2 + 0x_3 \cdots$ computer thinks x_3 is useless to identify y_1
- As computer learns patterns, computer is getting good at selecting useful information from input



Neuron network

- This phenomena is similar to Neurons in brain
- So deep learning is expected to be possible way to make Artificial Intelligence
- Don’t worry about losing our jobs!
  (there still are many problems to make AI)