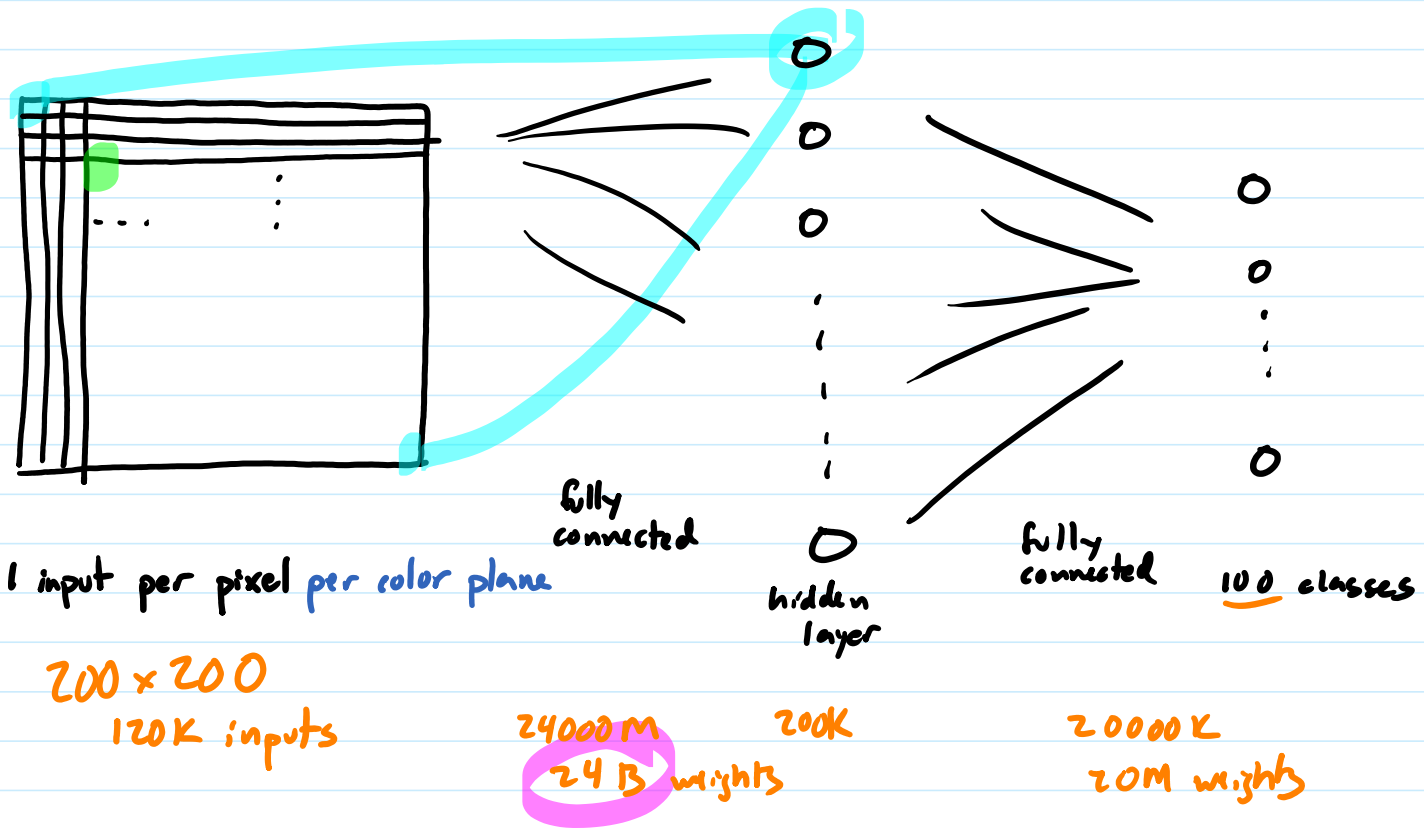
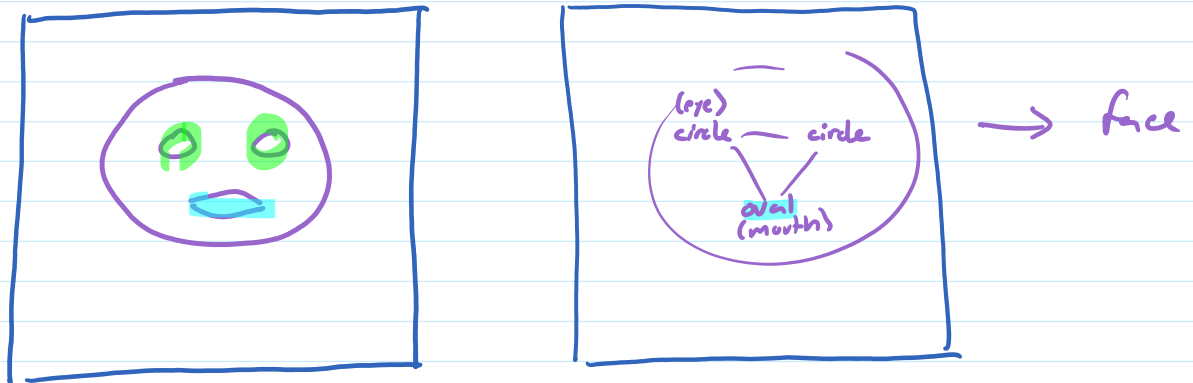


ANNs for Images

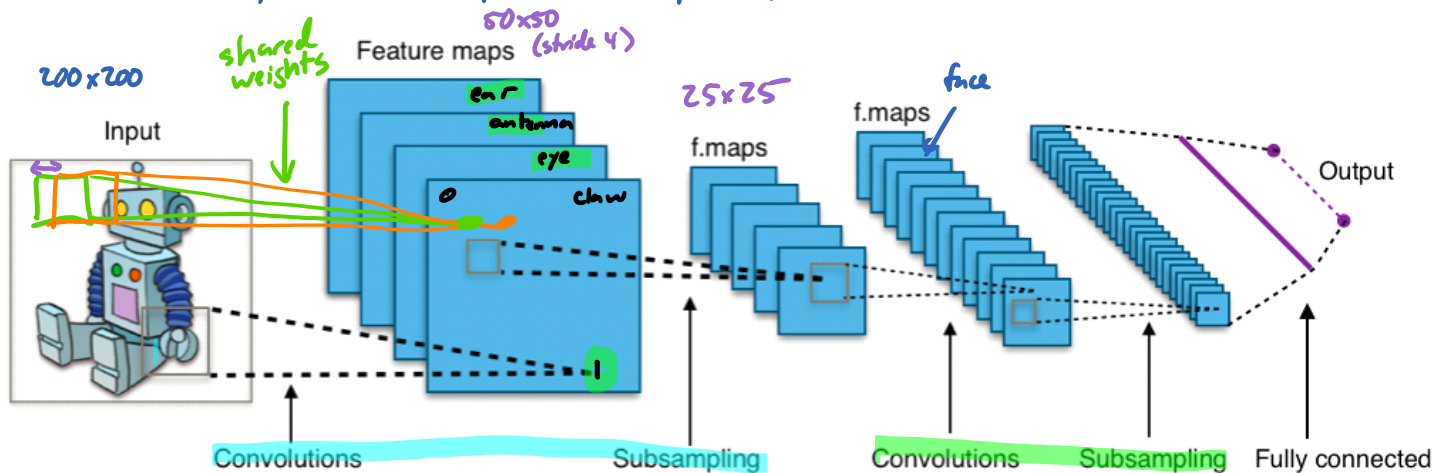


Convolutional Neural Networks

Deep Q network learning to play Pong



A much better picture from wikipedia user cphex34, who does not endorse these notes.



https://upload.wikimedia.org/wikipedia/commons/6/63/Typical_cnn.png

https://upload.wikimedia.org/wikipedia/commons/6/63/Typical_cnn.png

to reduce resolution

average over 2×2 regions

maximum over 2×2 regions



AlphaGo (2014-2017)

Step 1: supervised learning for convolutional deep neural network

3 weeks

using data from expert players
training examples

13 layers input $19 \times 19 \times 48$

output: a move $(19 \times 19 + 1)$

- matched expert's choice 55% of time
- + smaller network matching 25% of time
- so faster

hand-coded features

black
white
empty
opp captured
own captured
liberties
ladder capture
ladder escape

Step 2: reinforcement learning for convolutional deep neural network

1 day

beats SL network 80% of time

Step 3: reinforcement learning for value network

using data from RL step 2 network
playing itself 30M times
sample 1 position per game

output is estimate of value

+1 black win
0 draw
-1 white win

Step 4: MCTS

default policy uses fast SL network from step 1

initialize new nodes' values using network from step 3

tree policy uses PUCT

$$Q(s,a) + c \cdot P(s,a) \cdot \frac{\sqrt{\# \text{ times parent visited}}}{1 + \# \text{ times child visited}}$$

observed reward

from larger Step 1 network

Elo 3144 \rightarrow 3739 \rightarrow 5185
2015 (Fan Hui) 2016 (Lee Sedol) 2017 (retired)

Δ Elo 400 \rightarrow higher rated player has 92% chance of winning

AlphaGo Zero

zero human knowledge

input : $19 \times 19 \times 17$] \rightarrow current + last 7 positions + turn (all 1 = black, 0 = white)

output : move ($19 \times 19 + 1$) and value $[-1, 1]$

