CS540 Introduction to Artificial Intelligence
Lecture 8

Young Wu
Based on lecture slides by Jerry Zhu, Yingyu Liang, and Charles Dyer

June 12, 2020

Review Session

Admin

Q1 (last)

- Midterm on June 30 and July 1.
- Review Sessions:
- A: June 23 and June 24 (~~Monday and~~ Tuesday).
- B: June 25 and June 26 (~~Wednesday and~~ Thursday).

and Wed

and Fri

SIFT and HOG Features

Motivation

SUM

- SIFT and HOG features are expensive to compute.
- Simpler features should be used for real time face detection tasks.

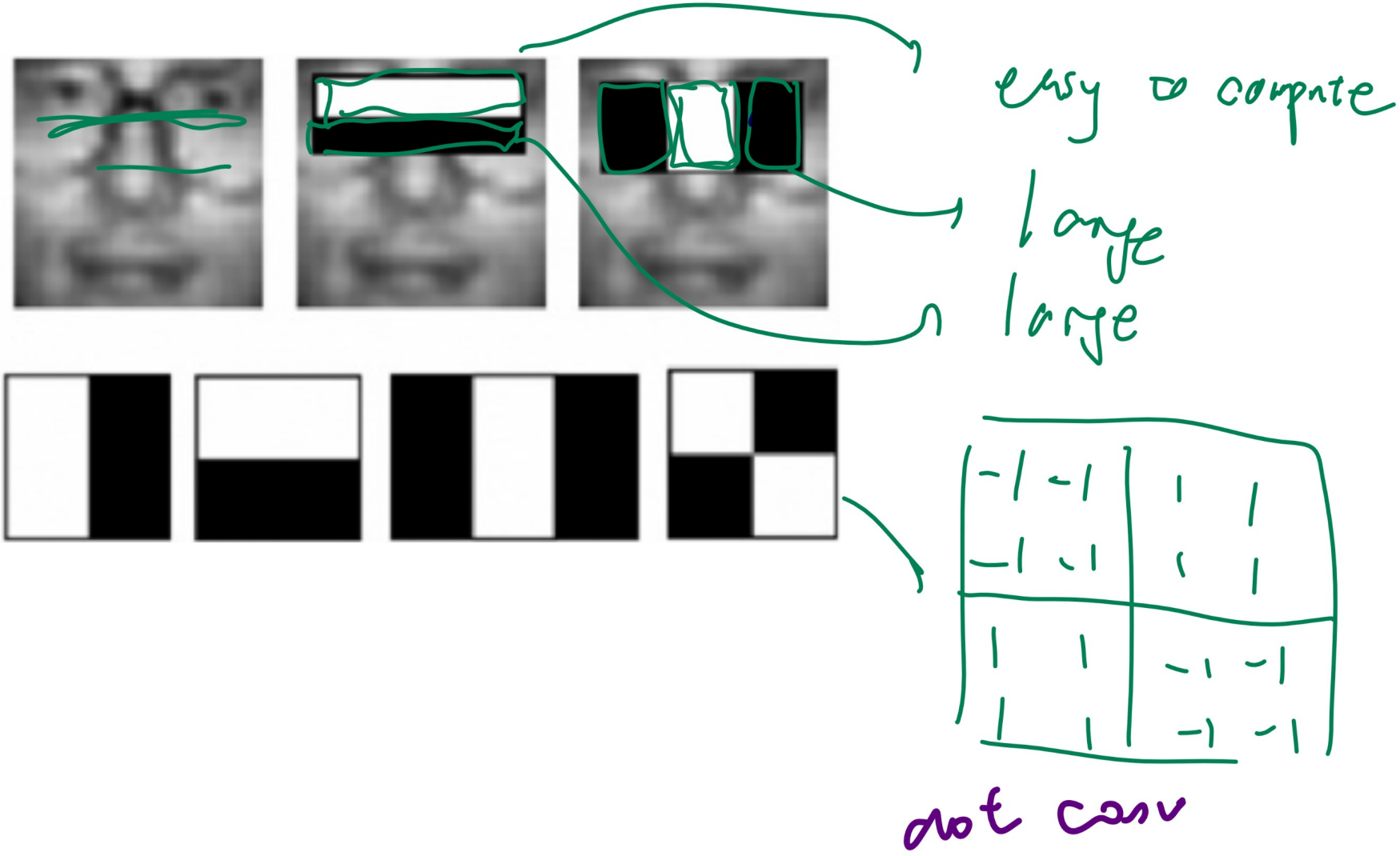
Real Time Face Detection

Motivation

- Each image contains 10000 to 500000 locations and scales.
- Faces occur in 0 to 50 per image.
- Want a very small number of false positives.

Haar Features Diagram

Motivation



Weak Classifiers

Definition

- Each weak classifier is a decision stump (decision tree with only one split) using one Haar feature x.

$$f(x) = \mathbb{1}_{\{x>\theta\}}$$

if $x > \theta \Rightarrow$ face
$x \leq \theta$ not

Info Gain

- Finding the threshold by comparing the information gain from all possible splits is too expensive, so θ is usually computed as the average of the mean values of the feature for each class.

$$\theta = \frac{1}{2}\left(\frac{1}{n_0}\sum_{i:y_i=0} x_i + \frac{1}{n_1}\sum_{i:y_i=1} x_i\right)$$

average value of non faces

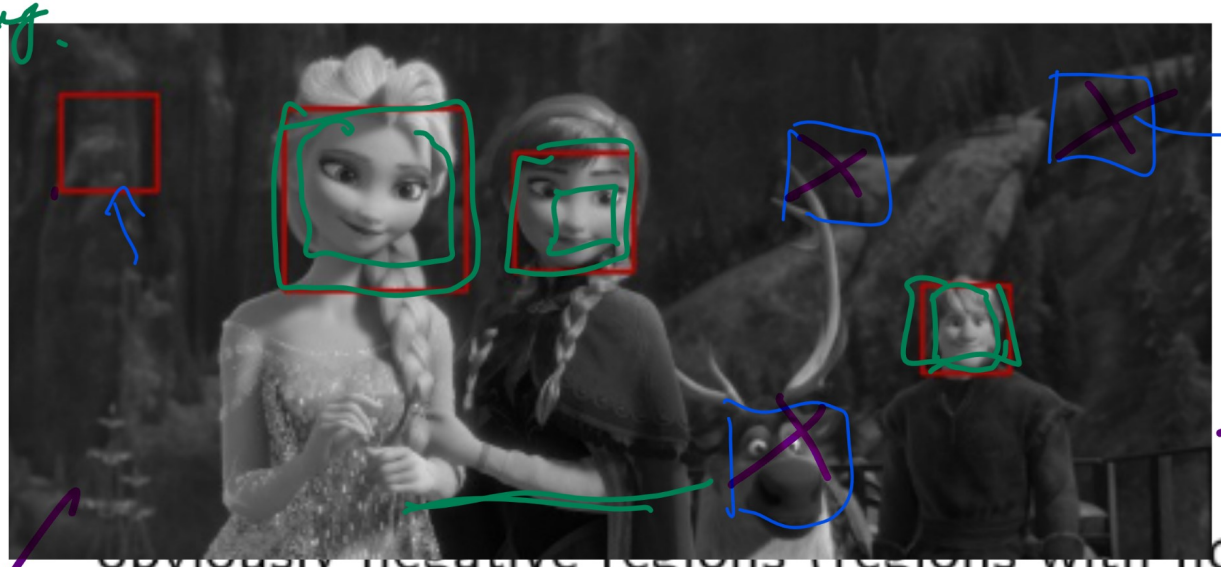
average value of faces

Strong Classifiers

Definition

- The weak classifiers are trained sequentially using ensemble methods such as AdaBoost.
- A sequence of T weak classifiers is called a T -strong classifier.
- Multiple T -strong classifiers can be trained for different values of T and combined into a cascaded classifier.

Cascaded Classifiers



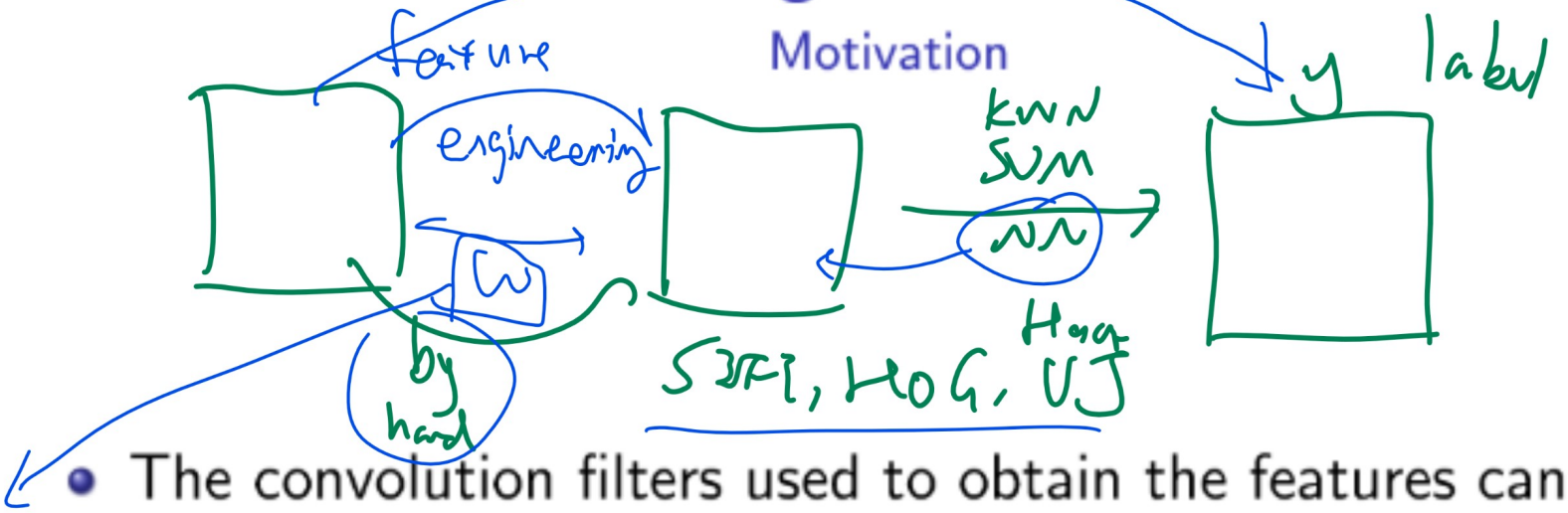
T, and use it reject obviously negative regions (regions with no faces).

- Train and use a T -strong classifier with larger T on only the regions that are not rejected.
- Repeat this process with stronger classifiers.

Viola Jones Diagram

Discussion

Learning Convolution



- The convolution filters used to obtain the features can be learned in a neural network. Such networks are called convolutional neural networks and they usually contain multiple convolutional layers with fully connected and softmax layers near the end.

Convolutional Layers

Definition

- In the (fully connected) neural networks discussed previously, each input unit is associated with a different weight.

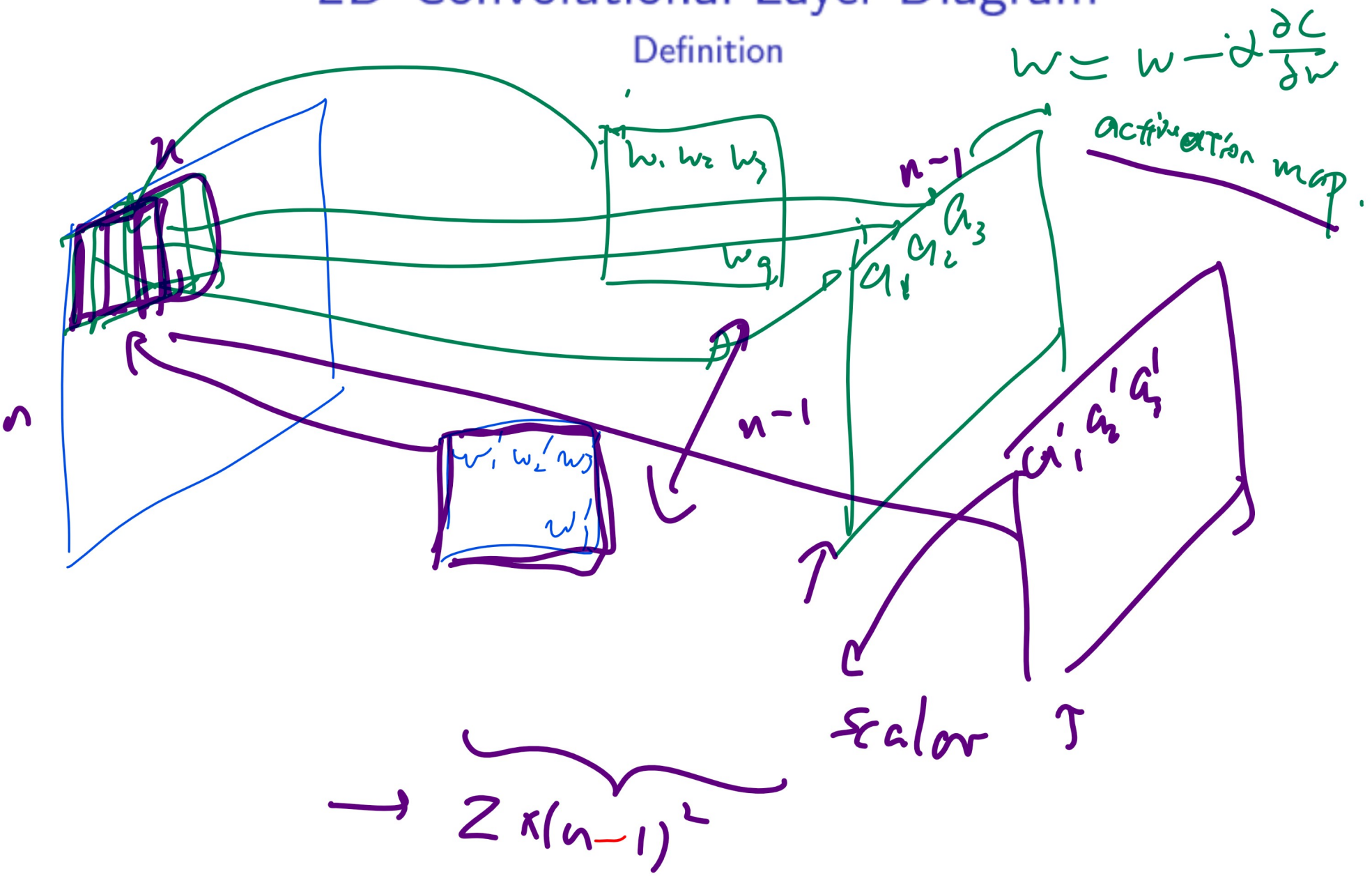
$$a = g\left(w^T x + b\right)$$

- In the convolutional layers, one single filter (a multi-dimensional array of weights) is used for all units (arranged in an array the same size as the filter).

$$A = g\left(W * X + b\right)$$

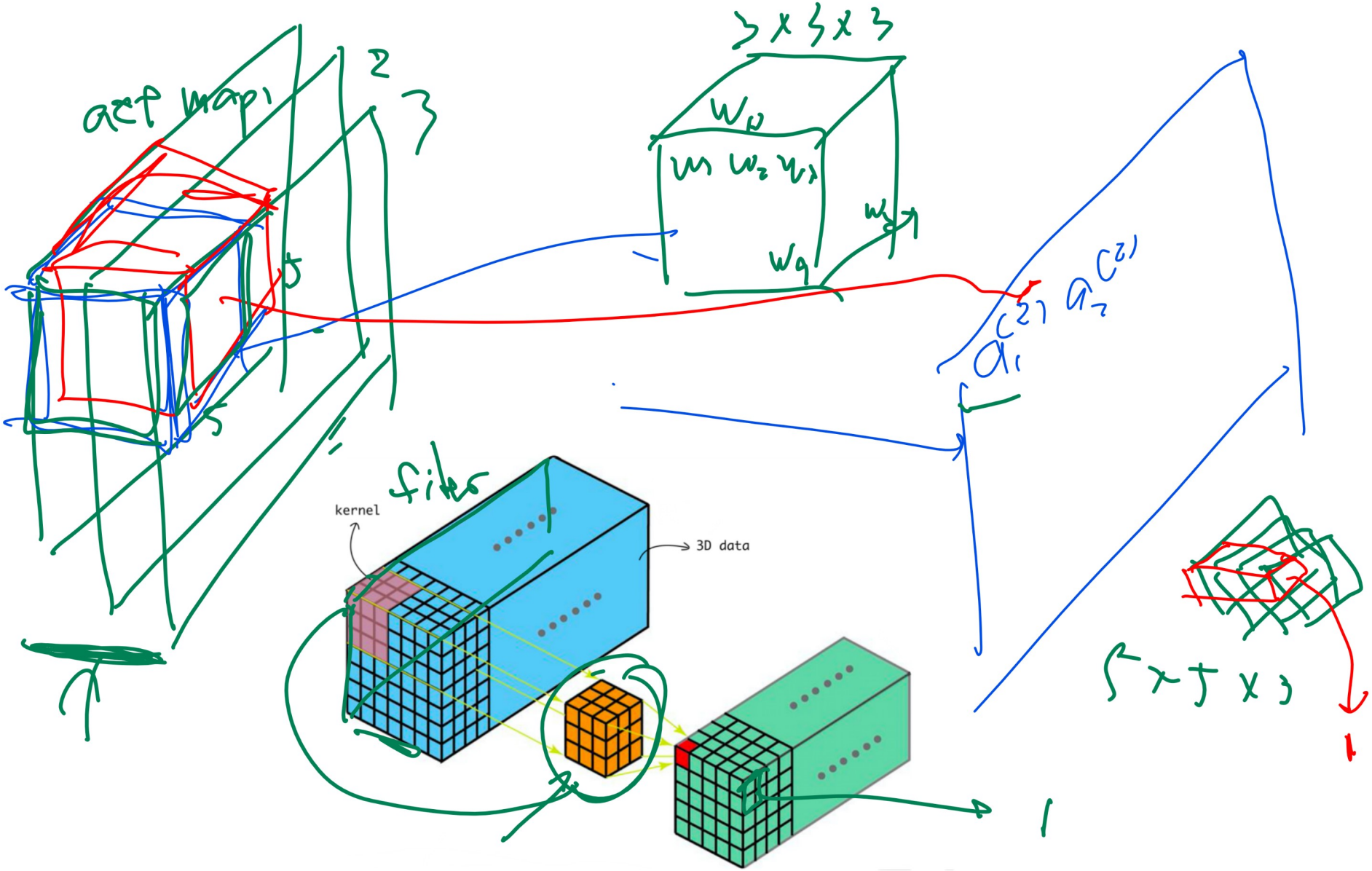
2D Convolutional Layer Diagram

Definition



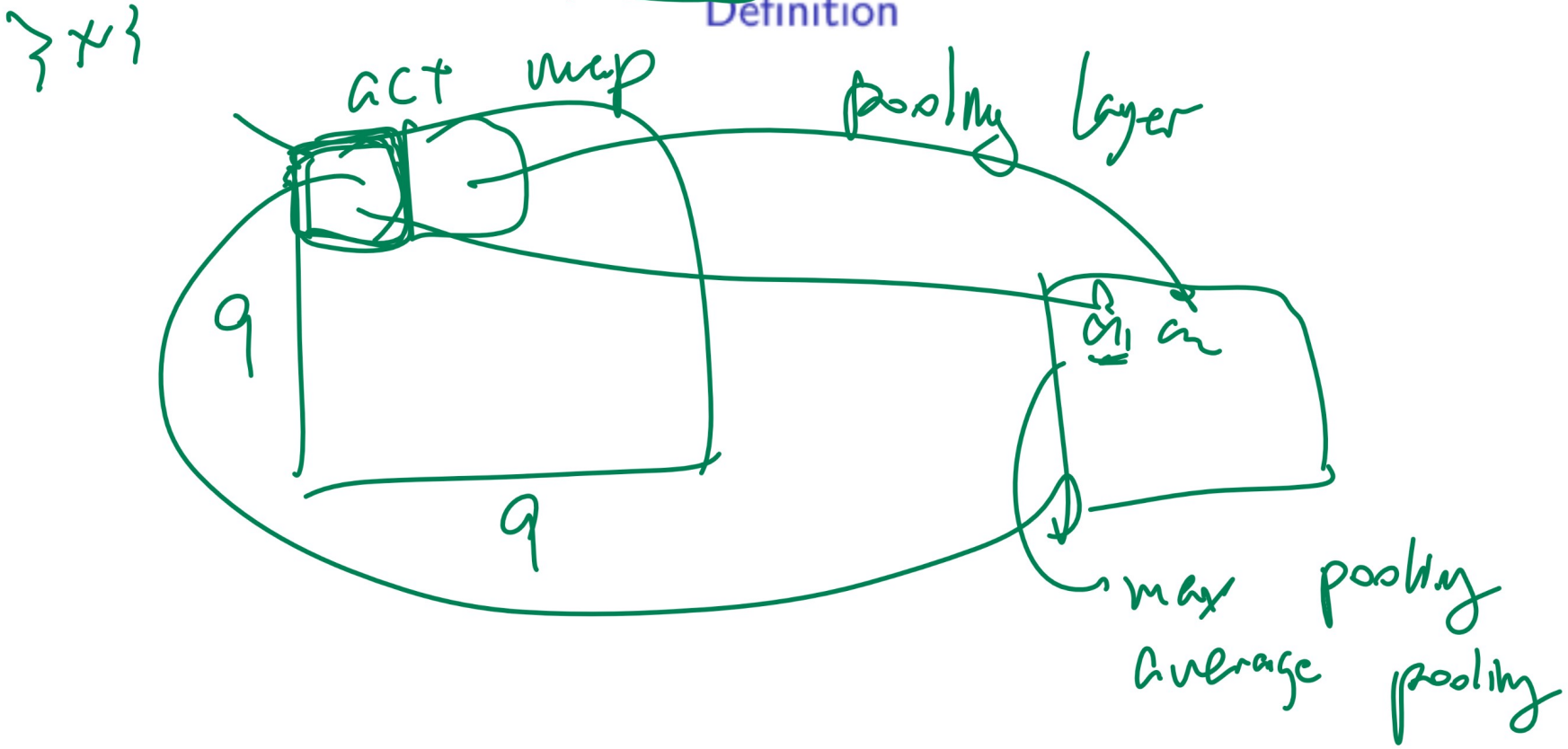
3D Convolutional Layer Diagram

Definition



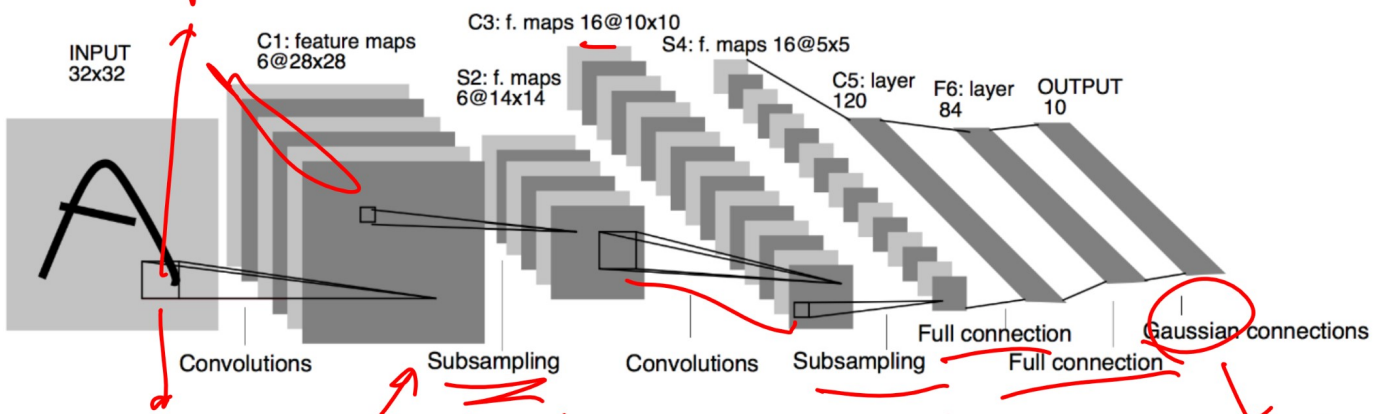
Pooling Diagram

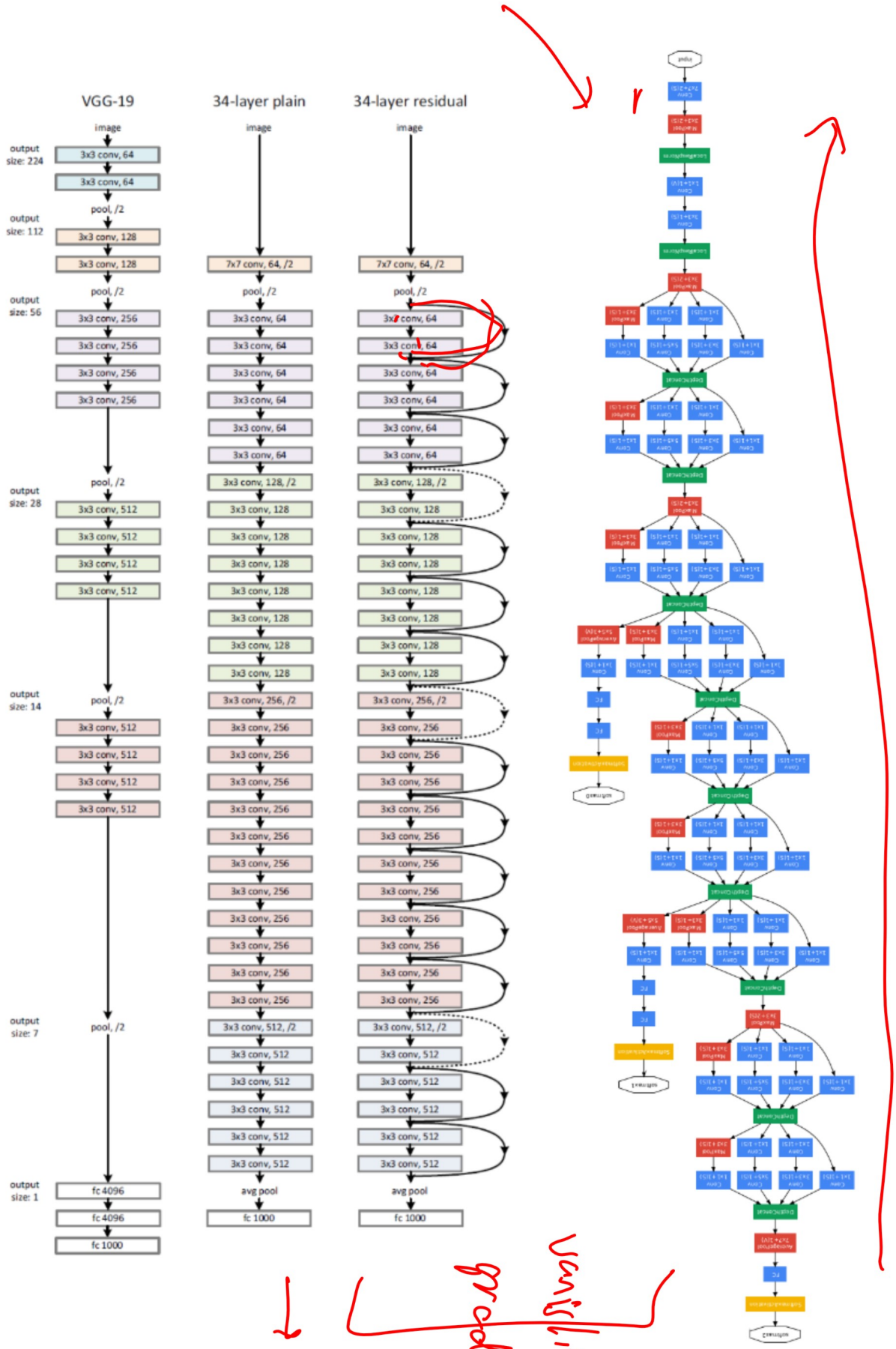
Definition



LeNet Diagram and Demo

Discussion




VGG-19
34-layer plain
34-layer residual
image
output size: 224
3x3 conv, 64
3x3 conv, 64
pool, /2
output size: 112
3x3 conv, 128
3x3 conv, 128
pool, /2
output size: 56
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
3x3 conv, 256
pool, /2
output size: 28
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
pool, /2
output size: 14
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
3x3 conv, 512
pool, /2
output size: 7
fc 4096
fc 4096
fc 1000
output size: 1
7x7 conv, 64, /2
pool, /2
3x3 conv, 64
3x3 conv, 128, /2
3x3 conv, 128
3x3 conv, 256, /2
3x3 conv, 256
3x3 conv, 512, /2
3x3 conv, 512
avg pool
fc 1000