AGENDA

i) GoogleNet
ii) ResNet
iii) Test cases for previous CNN app.

Top-1 error rate

actual class: Dog → prediction: Dog

Top-5 error rate

actual class: Dog → prediction: { i, j, k = Dog, l }

Google Net

top 5 error below 7%

o/p (softmax) → 30% Cat, 60% Dog, 0.2% Car, i

→ [60% Dog, 30% Cat, 5% bike, ...]

T

10 images → actual

93% 93/10 = 93%

7%

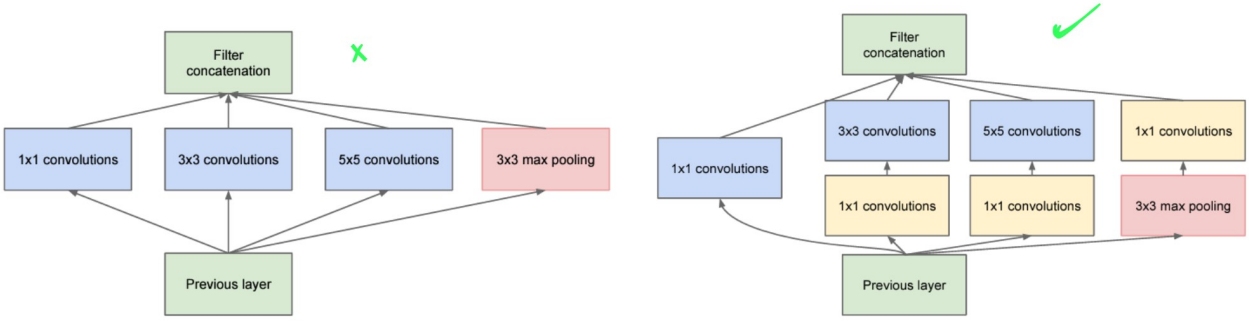
GoogleNet :-

↳ Christian Szegedy et. al. Google research team

→ Aim :- i) Deeper CNN layers ii) Reduce the no. of parameters

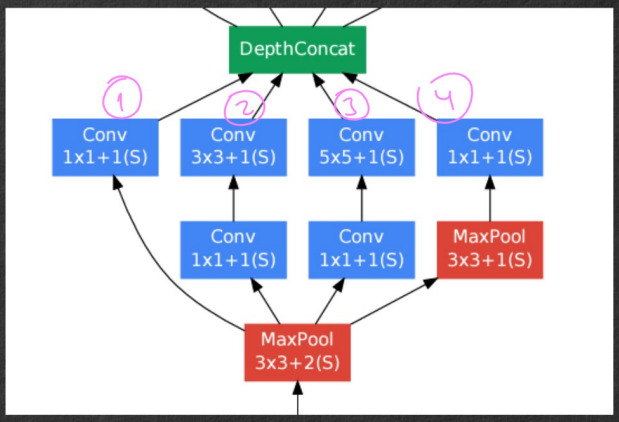
10 times less parameters than AlexNet

No of layers → 22 layers



(a) Inception module, naïve version (b) Inception module with dimension reductions

Figure 2: Inception module



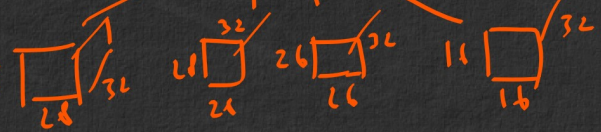
Conv
1X1 + 1(S)
1X1 kernel size
1 stride

MaxPool
3X3 + 1(S)
Overlapping pooling

padding = "same"

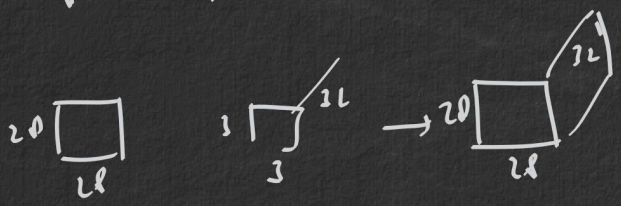
Depth Concatination

Depth wise concatination



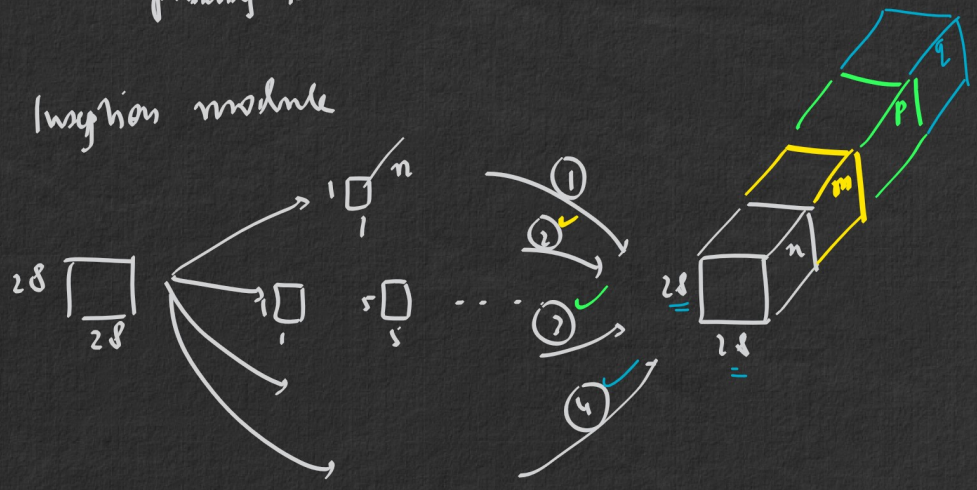
Ques:- if size from all branches are not same. then is it possible to concatenate?
NO.
possible way → padding = "same"

Previously



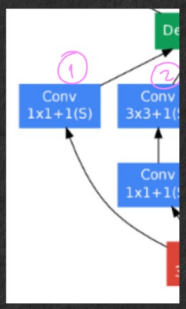
padding : same

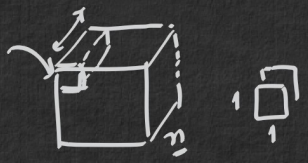
Inception module



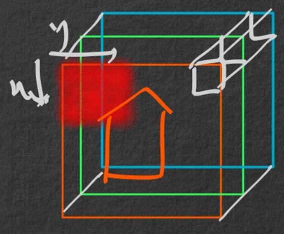
① branch of Inception module :-

kernel = 1X1
stride = 1.





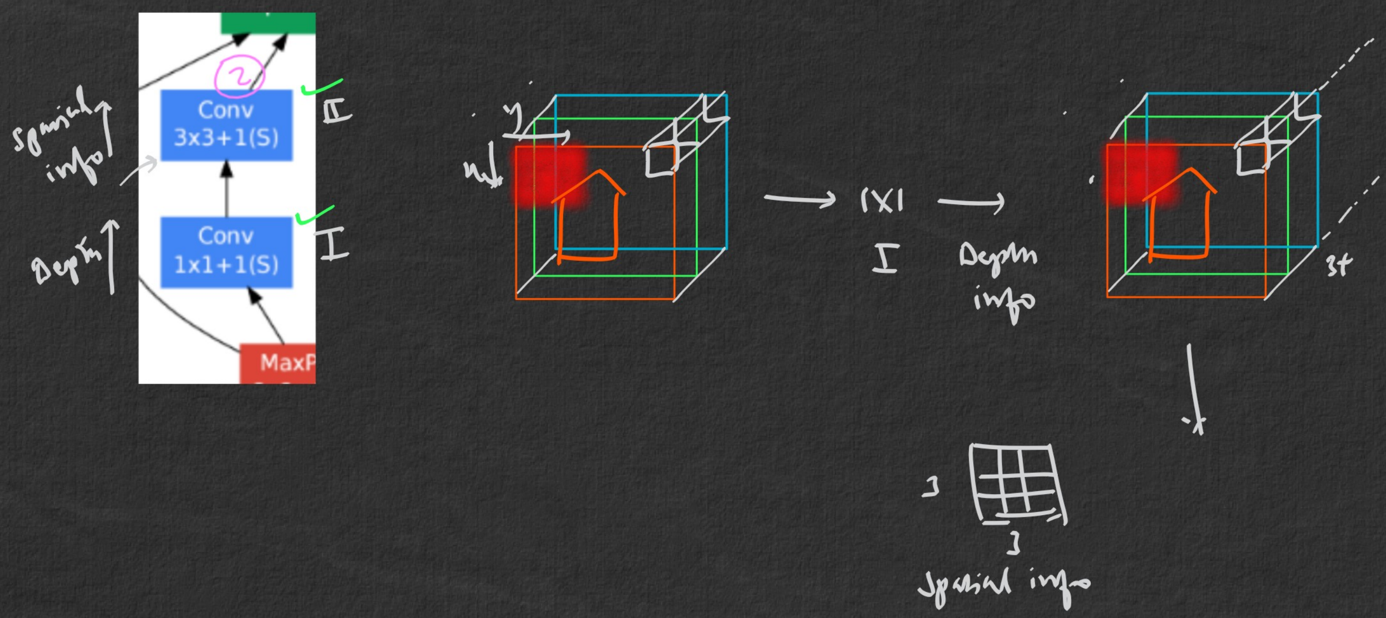
i) operating pixel by pixel
ii) It helps to get depth wise information.
iii) It will not capture the spatial info.



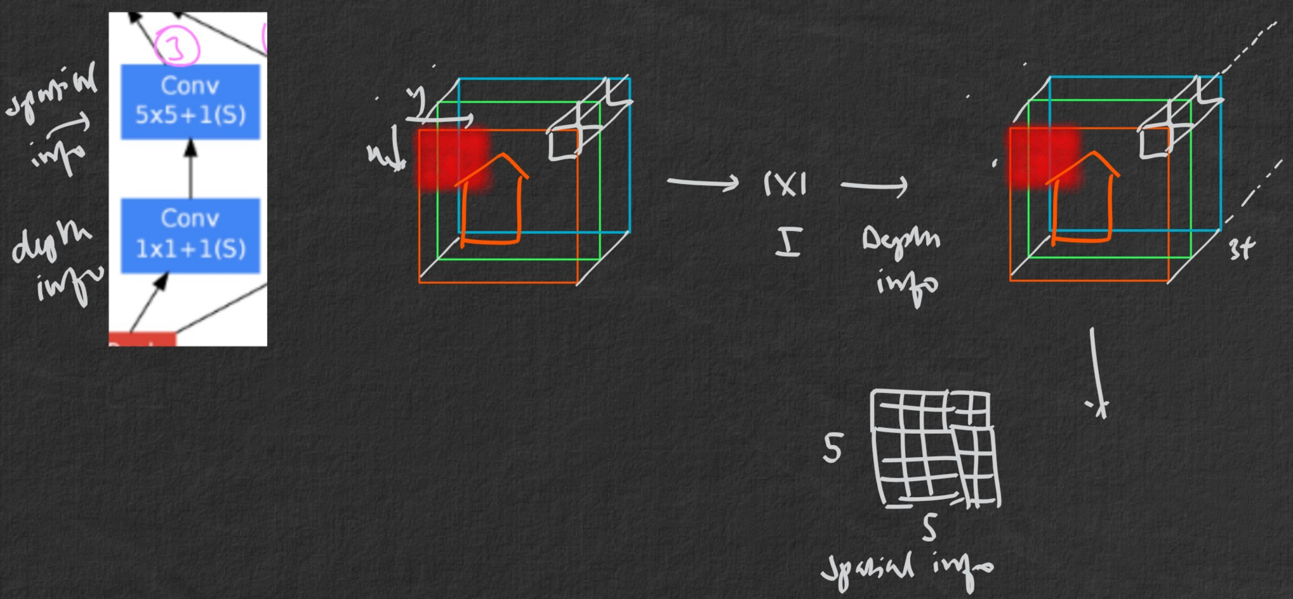
→ Spatial info : → shape of the object , ○

→ depth info: → colours → shades if → ⊖

② Branch of inception module:-



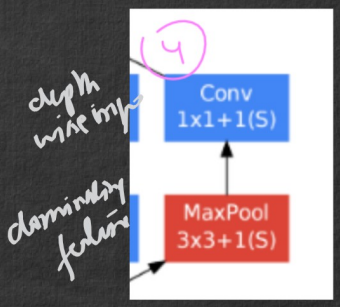
③ branch of inception module:-



Increased receptive fields

25 pixel at a time

(4) branch of inception module :-



MaxPool → get the dominant features w.r.t to pixel values.

Stride 1 →

i> get dominating feature
ii> get to preserve most of info as stride = 1.

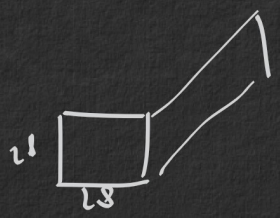
Pooling formula

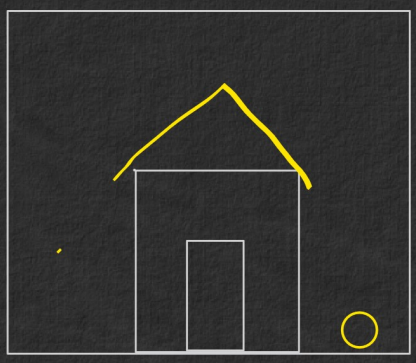
$$\frac{W-f}{S}+1$$

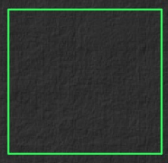
28x28 , 3x3 $\quad \frac{28-3}{1}+1 = 26\times 26$

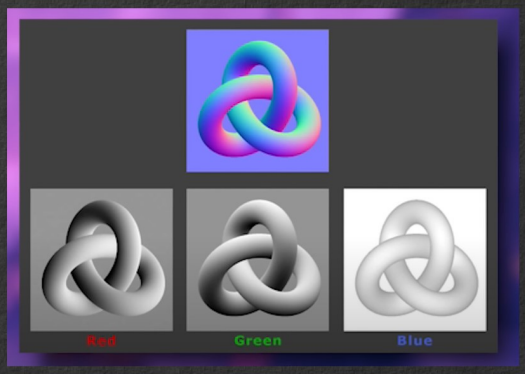
From all 4 branches

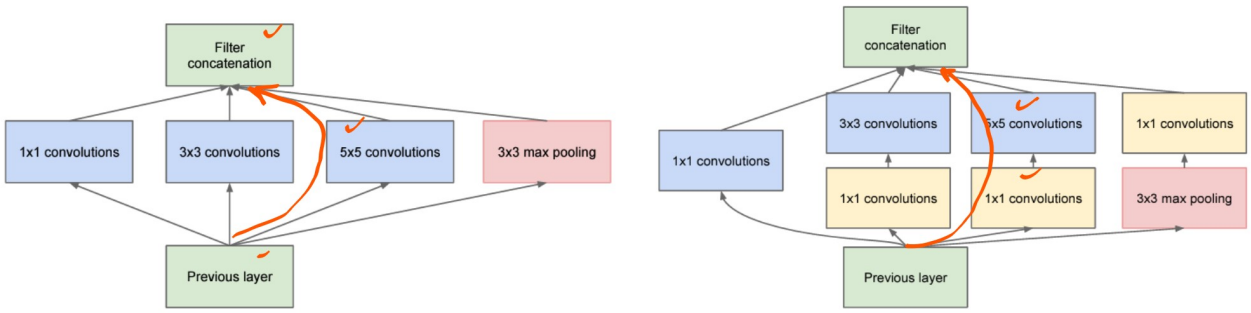
- (1) → Depth wise info only.
- (2) → Depth + spatial info formed in 3x3 pixel values
- (3) → Depth + spatial info formed in 5x5 pixel values
- (4) → Depth + Dominating feature







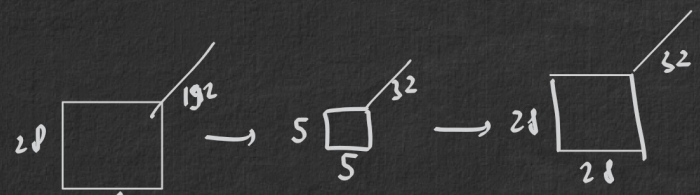




(a) Inception module, naïve version (b) Inception module with dimension reductions

Figure 2: Inception module

(a)

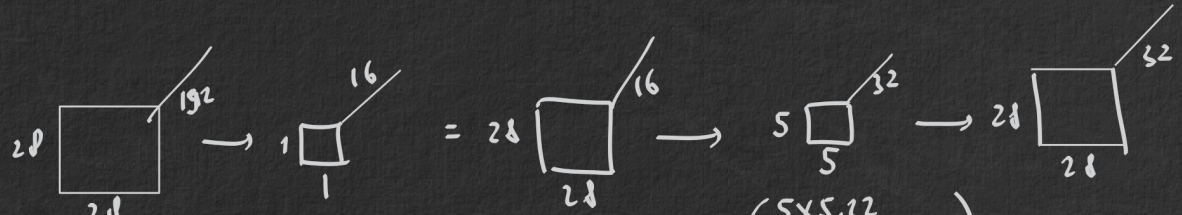


(5×5, 32
padding: same
stride: 1)

No. of operations $(5\times5 \times 28\times28\times192)\times32 = 120,422,400 \simeq 120M$ operations

(5×5 kernel size; 28×28×192 w h depth filters)

(b)



28×28×192 (1×1) 16 (5×5, 32 padding: same stride: 1)

$(1\times1 \times 28\times28\times192)\times16 + (5\times5\times28\times28\times16)\times32 =$

$2,408,448 + 10,035,200$

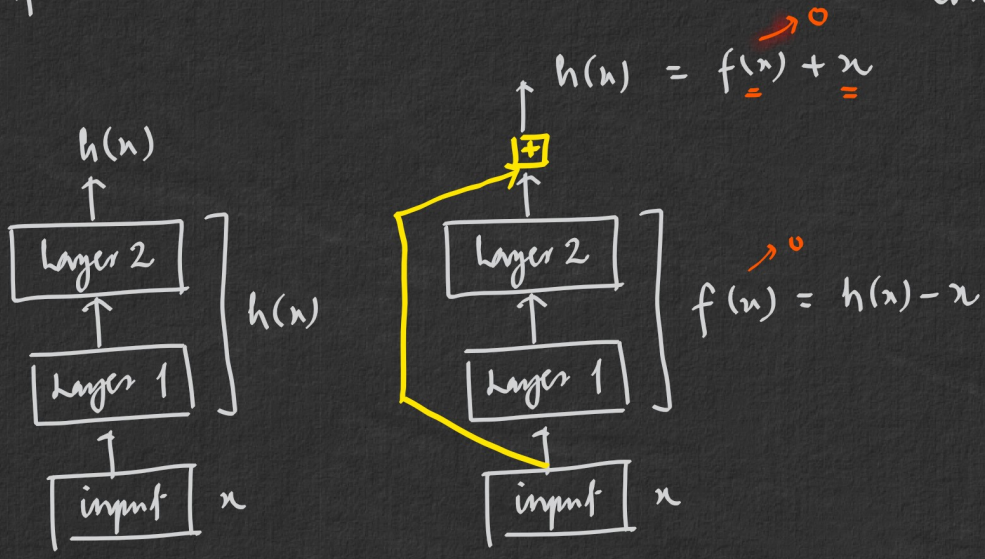
$= 12,443,648 \simeq 12.41M$ operations

(a) → 120M (b) → 12.4M

10 times less operations done in (b) case with more information

RESNET

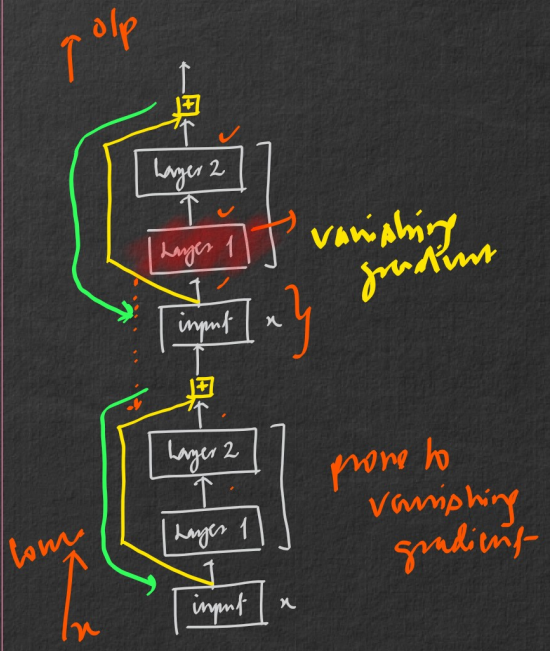
- by Kaiming He et. al.
- 2015
- ILSVRC 2015 challenge winner
- top 5 error rate < 3.6%
- Unique about RESNET → skip connection {aka shortcut connection}



i) Whenever we initialize NN
- → weights are close to zero.
- → NN will output closer to zero.
- → If we add skip connection :-
  - ↳ NN will result a copy of input initially
    - ↳ Model identity function

$$f(x) = x$$

Speed up the training because we are not starting from zero.



Vanishing gradient ?

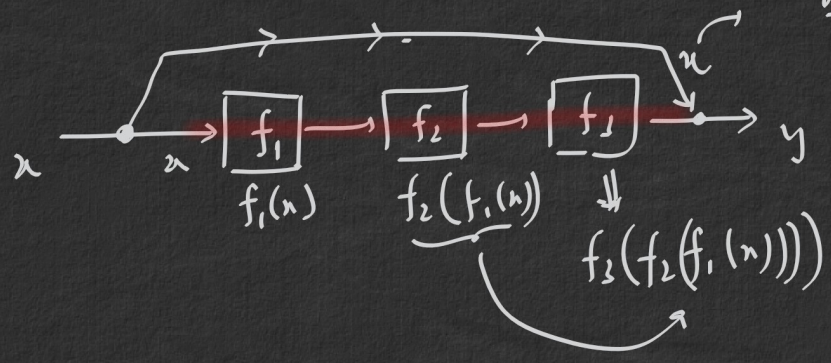
gradient → ΔW

$\Delta W \to 0$ ⇒ Vanishing gradient.

$W = W - \eta \Delta W$ (0)

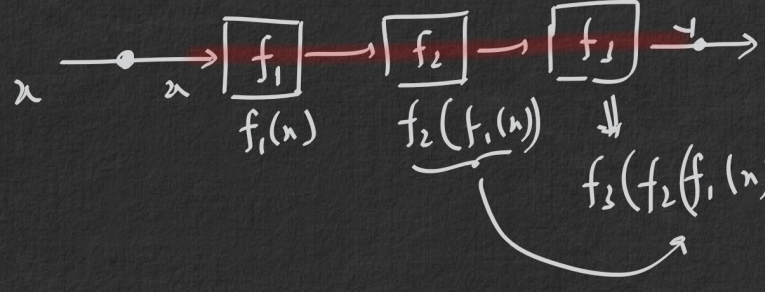
$W \simeq W$ weight update is negligible

ResNet 34, 50, 101, 152



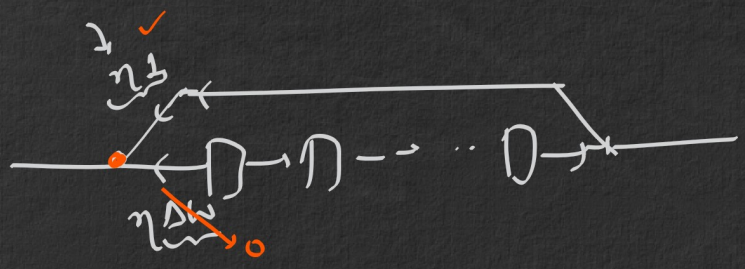
$\frac{dx}{dx} = 1$

$\frac{d}{dx} \cdot \cdot \cdot$



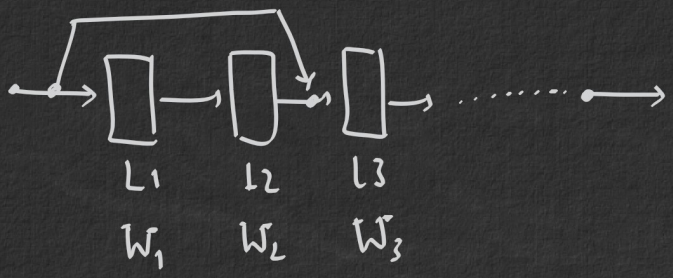
$\frac{d}{dx} \cdot \cdot \cdot$

→ ① Skip connection will help the gradient flow in backpropagation easily.



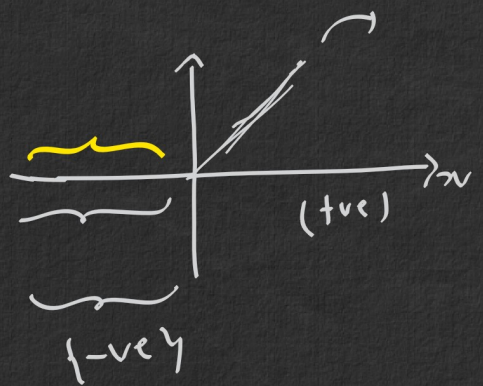
$$W = W - \eta \Delta w$$

$$(1 + \Delta w')$$

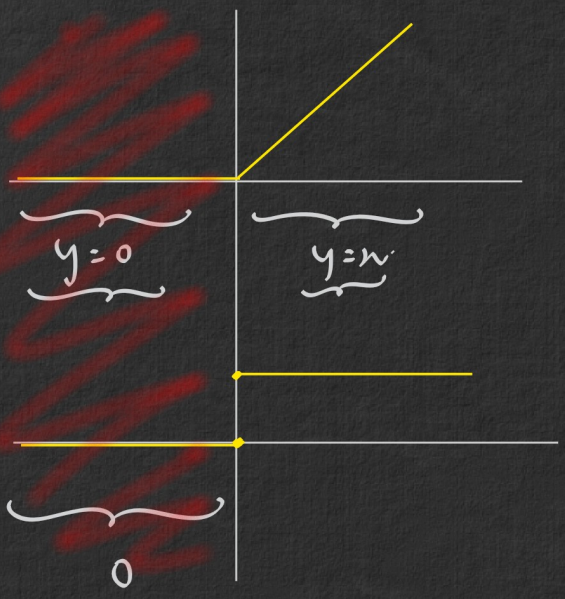


$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$A \cdot I = A$$



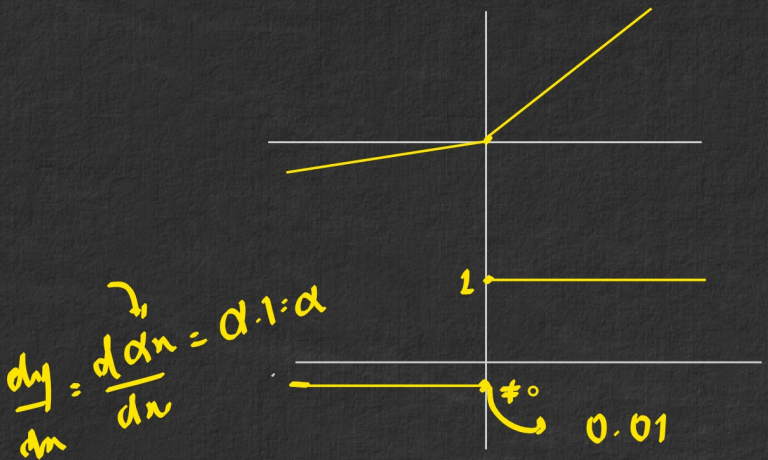
$$\text{relu} = \begin{cases} x & x > 0 \\ 0 & x \leq 0 \end{cases}$$



forward pass

$\frac{dy}{dx} = \frac{dx}{dx} = 1$

backward pass



$\frac{dy}{dx} = \frac{d\alpha x}{dx} = \alpha \cdot 1 = \alpha$

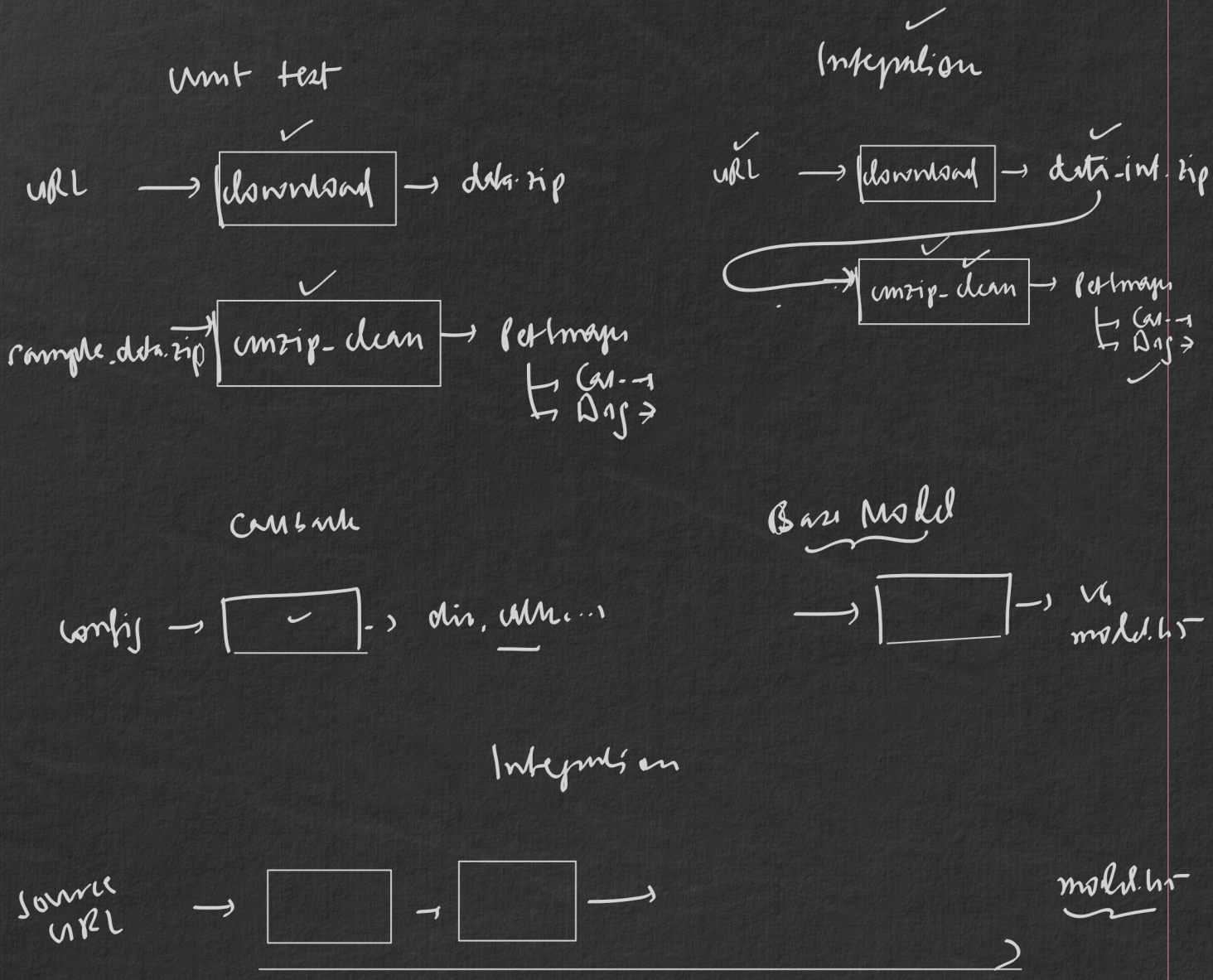
LeakyReLU: $\begin{cases} x & x > 0 \\ \alpha x & x \leq 0 \end{cases}$

$\alpha = 0.01$

→ Sample-data ⟶ 16 MB
  ↳ .txt
  .jpg = size = 0kb
  ↳ Cat ⟶ img, 0kb, .db
  ↳ Dog , .db

→ sample.data.zip

→ PetImages
  ↳ Cat ⟶ 3 img
  ↳ Dog ⟶ 3 imgs


Unit test
URL
download
data.zip
sample_data.zip
unzip_clean
PetImages
Cat
Dog
Integration
URL
download
data_int.zip
unzip_clean
PetImages
Cat
Dog
Callback
config
dirs, callb...
Base Model
VG
model.h5
Integration
Source URL
model.h5