Resnet ⇒ Res ↓ Residual

Net ↓ network

⇒ Microsoft Research ⇒ 2015

LSTM ← Transformer

↳ Skip Connection

↳ identity network

Resnet 18, Resnet 34

3x Resnet 50, Resnet-101, Resnet 152

Resnet-x1, Resnet-D, wide Residual network (WRN)

Googlenet | Inception-v1 ⇒ 2014 (ILSVRC) ⇒ Imagenet

Parallel convolution filter

Googlenet

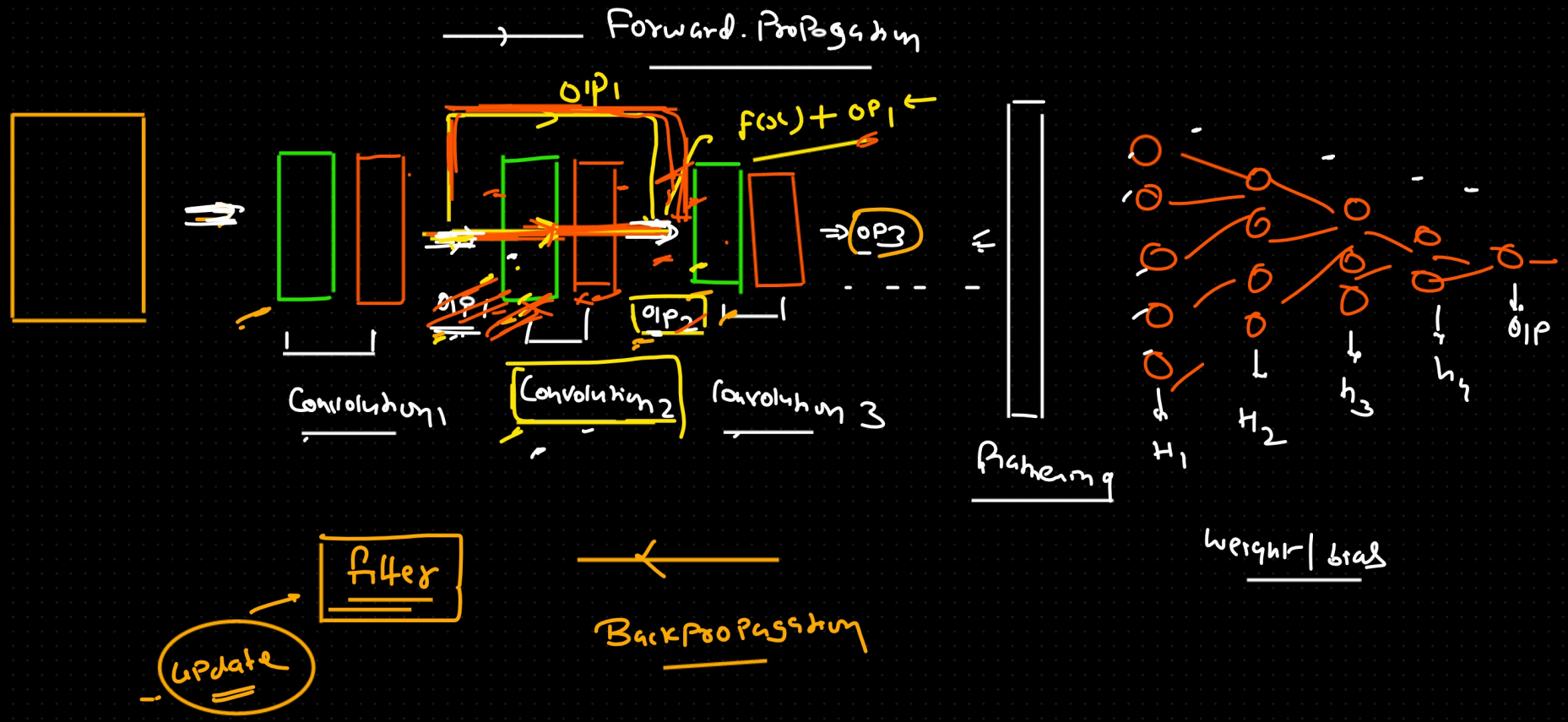
- Inception-v1
- Inception-v2
- Inception-v3

Inception-v4

→ Inception-Resnet ⇒ Parallel conv + Residual block

→ Xception (Extreme Inception)

→ Mobilenet, Nasnet, Efficient



Classical V/S Residual Neural

VGG16 | VGG19 ⇒ 6-7 L. → 3-4 L.

Microsoft

Depth of the Network → overfitting ?

→ underfitting

↳ vanishing GD

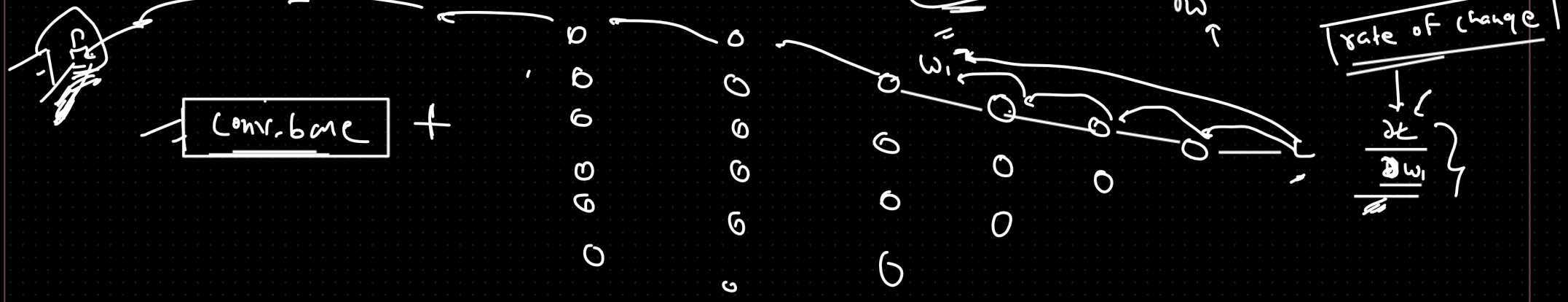
↳

Resnet

$$W_{new} = W_{old} - \eta \frac{\partial L}{\partial W}$$

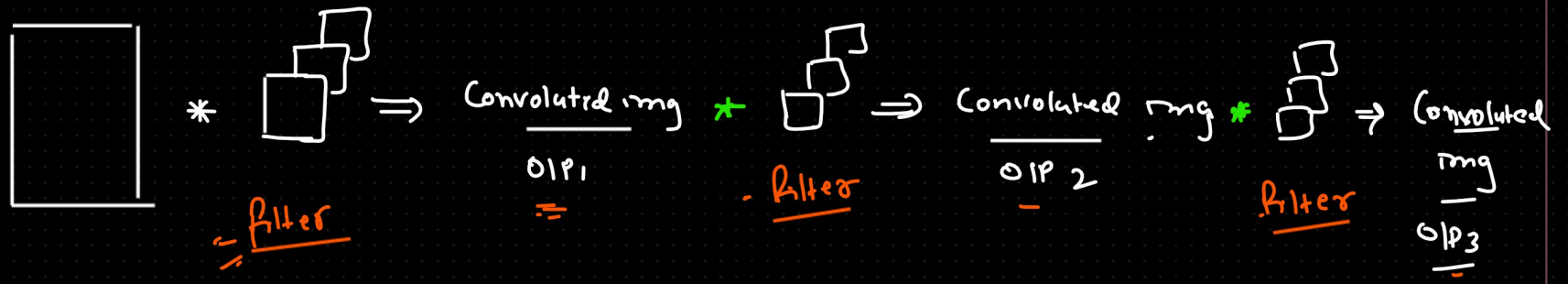
rate of change

$\frac{\partial L}{\partial W_1}$

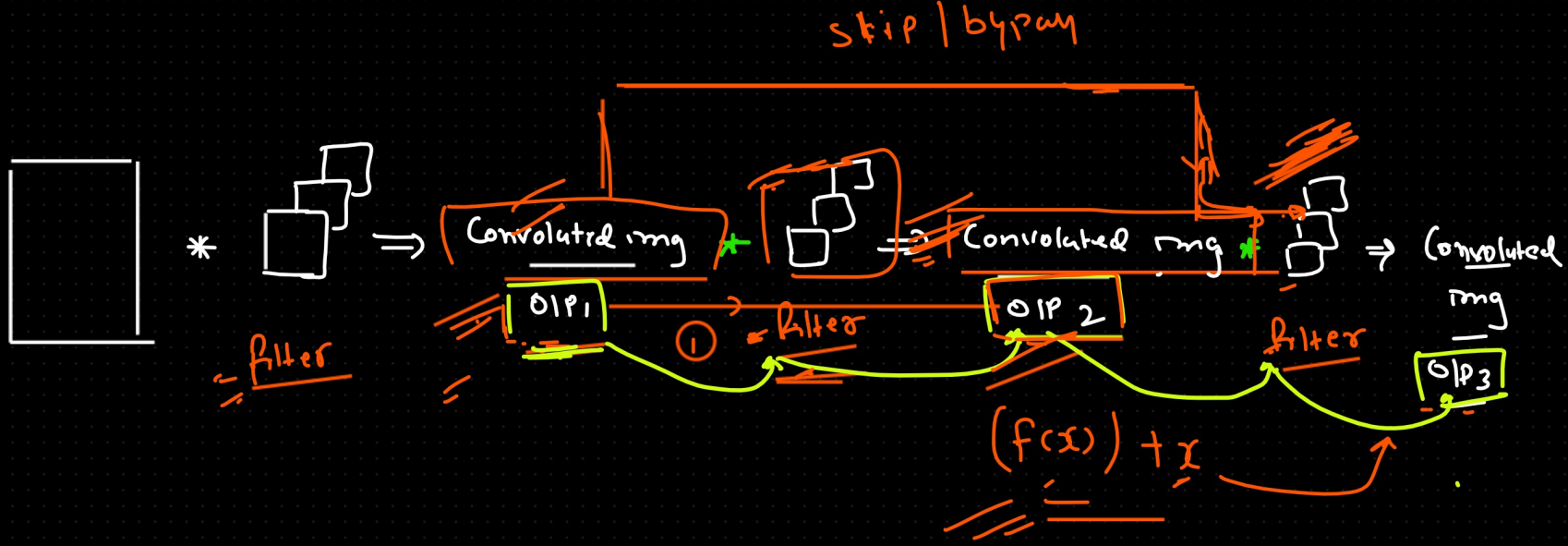


Residual network

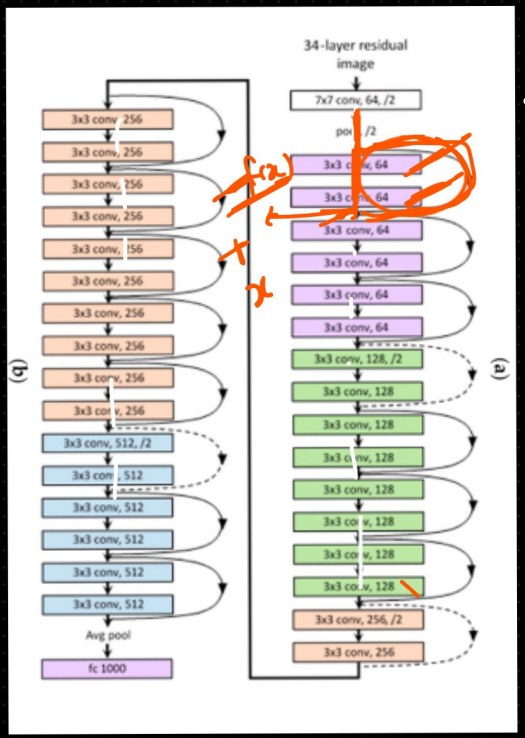
Residual network



Classical network



Residual Network



1. Skip Connection → residual conn work during FP and BP
2. train for the Deep CNN.
3. skip connection provide shortcut path for the gradient to flow through the nets.

Forward Pass

1. During the fp input is passing through a series of layers
2. In a network with skip conn, the input is also directly added to o/p of one or more layer deeper in the network

$= f(x) + x$

Calculate the loss

Backpropagation (updating weight & bias)

Gradient value

1. During the BP the gradient of loss with respect to the network weight will be calculated.

2. GD / Momentum / Adagrad / RmsProp / Adam

Gradient Flow through the skip Connection

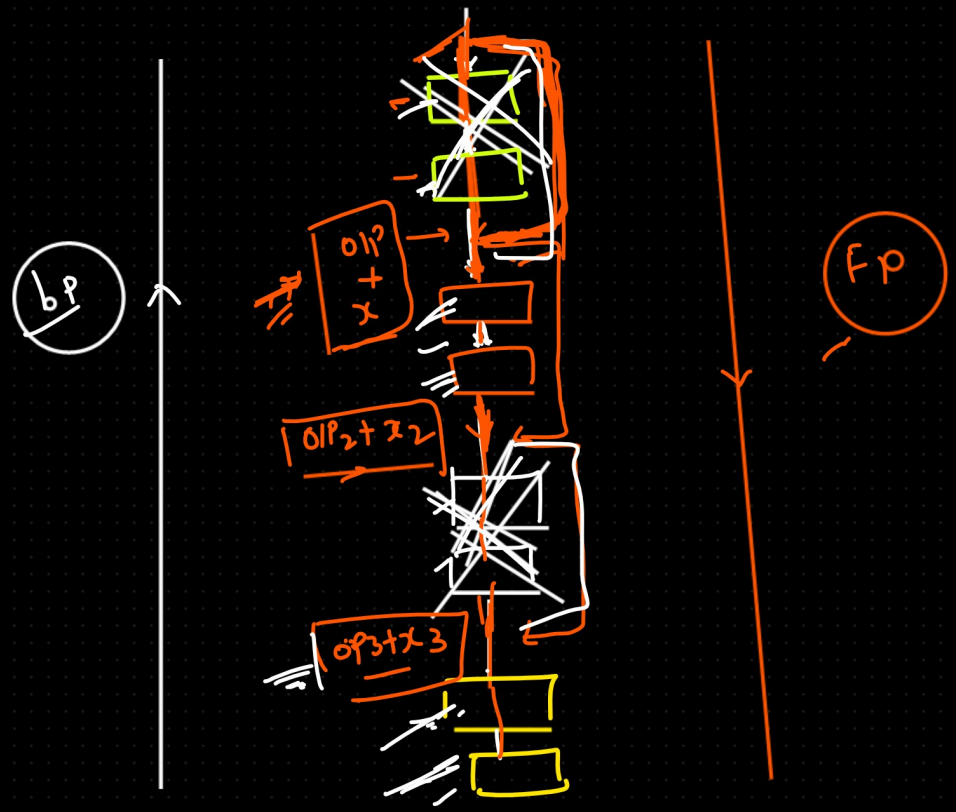
Key role of the skip Conn

→ In the tradional net- gradient is flowing to the entire

→ small / Dispear

→ slow or ineffective

Skip conn provide shortcut path for gradient to flow through the network



Topic	Microsoft — Resnet (2015)	google — inception (2014)
Basic Architecture	Deep Arc. with Residual block	Complex Arc. with Parallal conv.
Skip conn	Using Skip Conn and Jumping over the layes	not having skip conn. Pallal conv.
Depth	will find out till extreme Depth (18 → 152)	instead of deep it is a "wider"
No. of Params	Larger no. of Params	Compare to resnet, less Params
ILSVRC	Winner of 2015	Winner in 2014

Key - innovation

Skip - Connection } fast training with deep nets

- Parallal conv.
- multi scale feature extraction

Wider nets

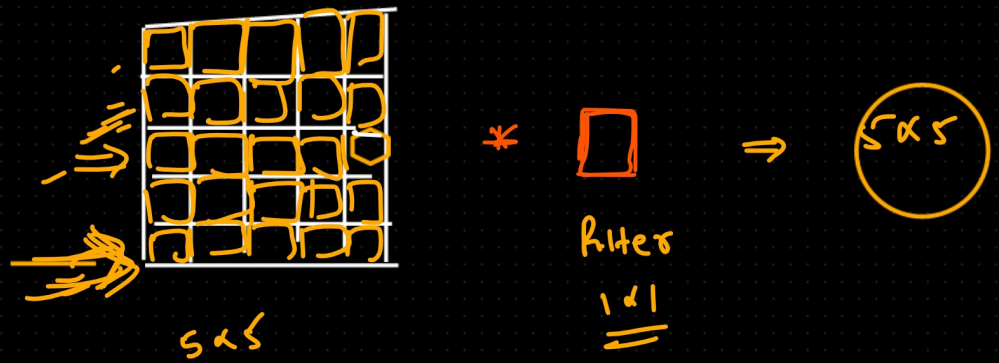
2023

Resnet | Inception

2015 ⇒ 3.5 d.

96.5%

Pointwise Convolution



$$\frac{n+2p-f}{s}+1 \Rightarrow \frac{5+2(0)-1}{1}+1 \Rightarrow 5+0+1+1 = \boxed{5 \times 5}$$