Residual Blocks

VGG-16

Layer	# filters	Filter size	Stride	Padding	Size of feature map	Activation function
Input	-	-			224 X 224 X 3	
Conv 1	64	3X3	1	1	224 X 224 X 64	ReLU
Conv 2	64	3X3	1	1	224 X 224 X 64	ReLU
Max Pooling 1	-	2X2	2		112 X 112 X 64	
Conv 3	128	3X3	1	1	112 X 112 X 128	ReLU
Conv 4	128	3X3	1	1	112 X 112 X 128	ReLU
Max Pooling 2	-	2X2	2		56 X 56 X 128	
Conv 5	256	3X3	1	1	56 X 56 X 256	ReLU
Conv 6	256	3X3	1	1	56 X 56 X 256	ReLU
Conv 7	256	3X3	1	1	56 X 56 X 256	ReLU
Max Pooling 3	-	2X2	2		28 X 28 X 256	
Conv 8	512	3X3	1	1	28 X 28 X 512	ReLU
Conv 9	512	3X3	1	1	28 X 28 X 512	ReLU
Conv 10	512	3X3	1	1	28 X 28 X 512	ReLU
Max Pooling 4	-	2X2	2		14 X 14 X 512	
Conv 11	512	3X3	1	1	14 X 14 X 512	ReLU
Conv 12	512	3X3	1	1	14 X 14 X 512	ReLU
Conv 13	512	3X3	1	1	14 X 14 X 512	ReLU
Max Pooling 5	-	2X2	2		7 X 7 X 512	
Flatten	-				25088	
Fully Connected 1					4096	ReLU
Fully Connected 2					4094	ReLU
Fully Connected 3					1000	Softmax

Number of layers = 16

Analytics Vidhya

GoogLeNet

type	patch size/ stride	output size	depth	#1×1	#3×3 reduce	#3×3	#5×5 reduce	#5×5	pool proj	params	ops
convolution	7×7/2	112×112×64	1							2.7K	34M
max pool	3×3/2	56×56×64	0								
convolution	3×3/1	56×56×192	2		64	192				112K	360M
max pool	3×3/2	28×28×192	0								
inception (3a)		28×28×256	2	64	96	128	16	32	32	159K	128M
inception (3b)		28×28×480	2	128	128	192	32	96	64	380K	304M
max pool	3×3/2	14×14×480	0								
inception (4a)		14×14×512	2	192	96	208	16	48	64	364K	73M
inception (4b)		14×14×512	2	160	112	224	24	64	64	437K	88M
inception (4c)		14×14×512	2	128	128	256	24	64	64	463K	100M
inception (4d)		14×14×528	2	112	144	288	32	64	64	580K	119M
inception (4e)		14×14×832	2	256	160	320	32	128	128	840K	170M
max pool	3×3/2	7×7×832	0								
inception (5a)		7×7×832	2	256	160	320	32	128	128	1072K	54M
inception (5b)		7×7×1024	2	384	192	384	48	128	128	1388K	71M
avg pool	7×7/1	1×1×1024	0								
dropout (40%)		1×1×1024	0								
linear		1×1×1000	1							1000K	1M
softmax		1×1×1000	0								

Number of layers = 22

Inception V3

type	patch size/stride or remarks	input size
conv	3×3/2	299×299×3
conv	3×3/1	149×149×32
conv padded	3×3/1	147×147×32
pool	3×3/2	147×147×64
conv	3×3/1	73×73×64
conv	3×3/2	71×71×80
conv	3×3/1	35×35×192
3×Inception	As in figure 5	35×35×288
5×Inception	As in figure 6	17×17×768
2×Inception	As in figure 7	8×8×1280
pool	8 × 8	8 × 8 × 2048
linear	logits	1 × 1 × 2048
softmax	classifier	1 × 1 × 1000

Number of layers = 42

Problem with Deep Networks

Problem with Deep Networks

1. Vanishing / Exploding gradients

Problem with Deep Networks

1. Vanishing / Exploding gradients

Solution:

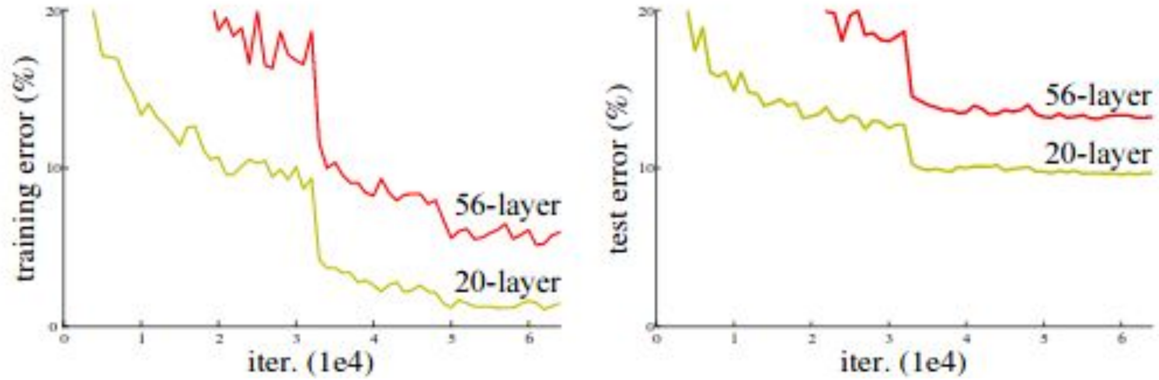
- Normalized Weight initialization
- Adding Batch Normalization

Problem with Deep Networks

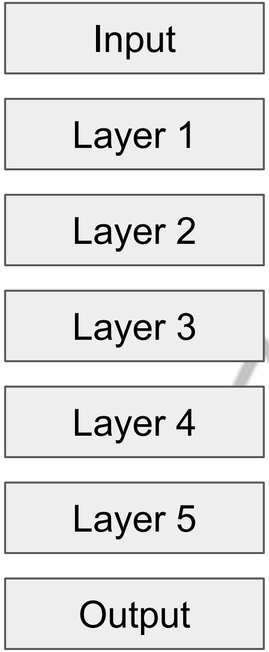
2. With increasing depth, performance starts degrading rapidly

Problem with Deep Networks

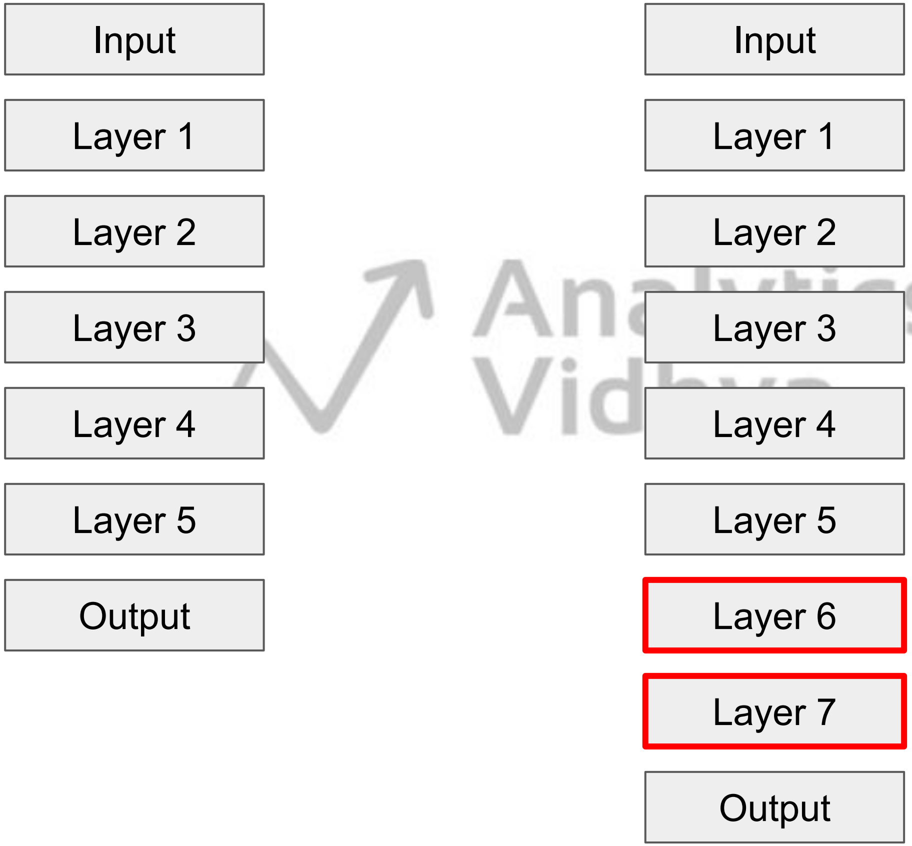
2. With increasing depth, performance starts degrading rapidly



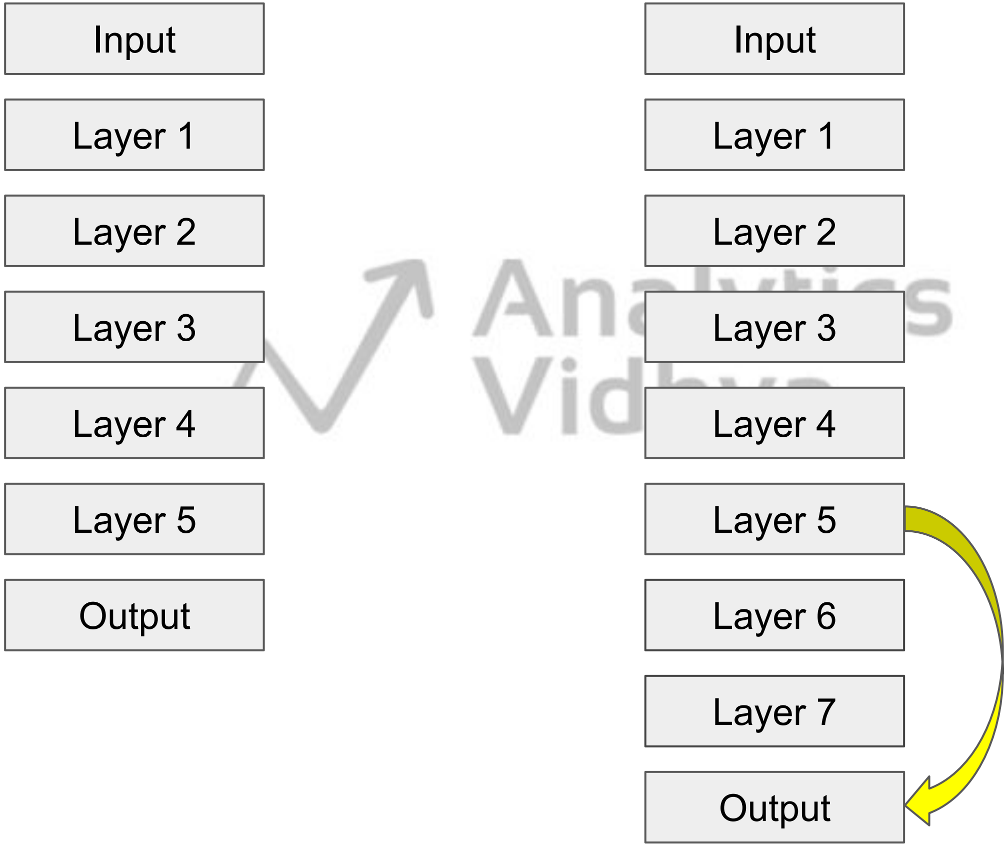
Problem with Deep Networks



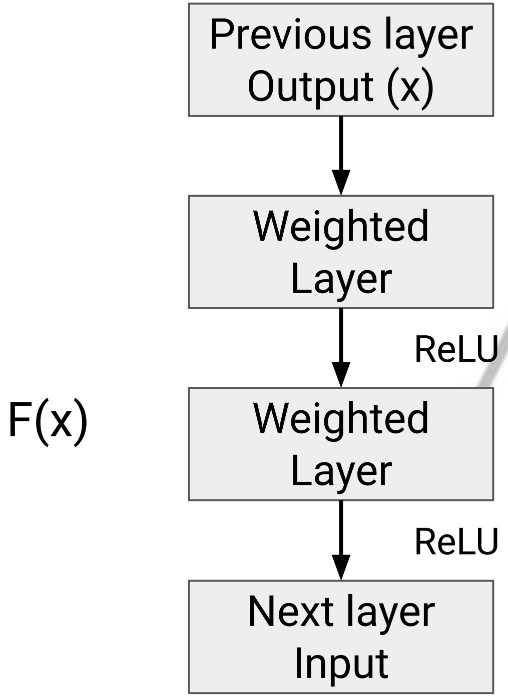
Problem with Deep Networks



Problem with Deep Networks

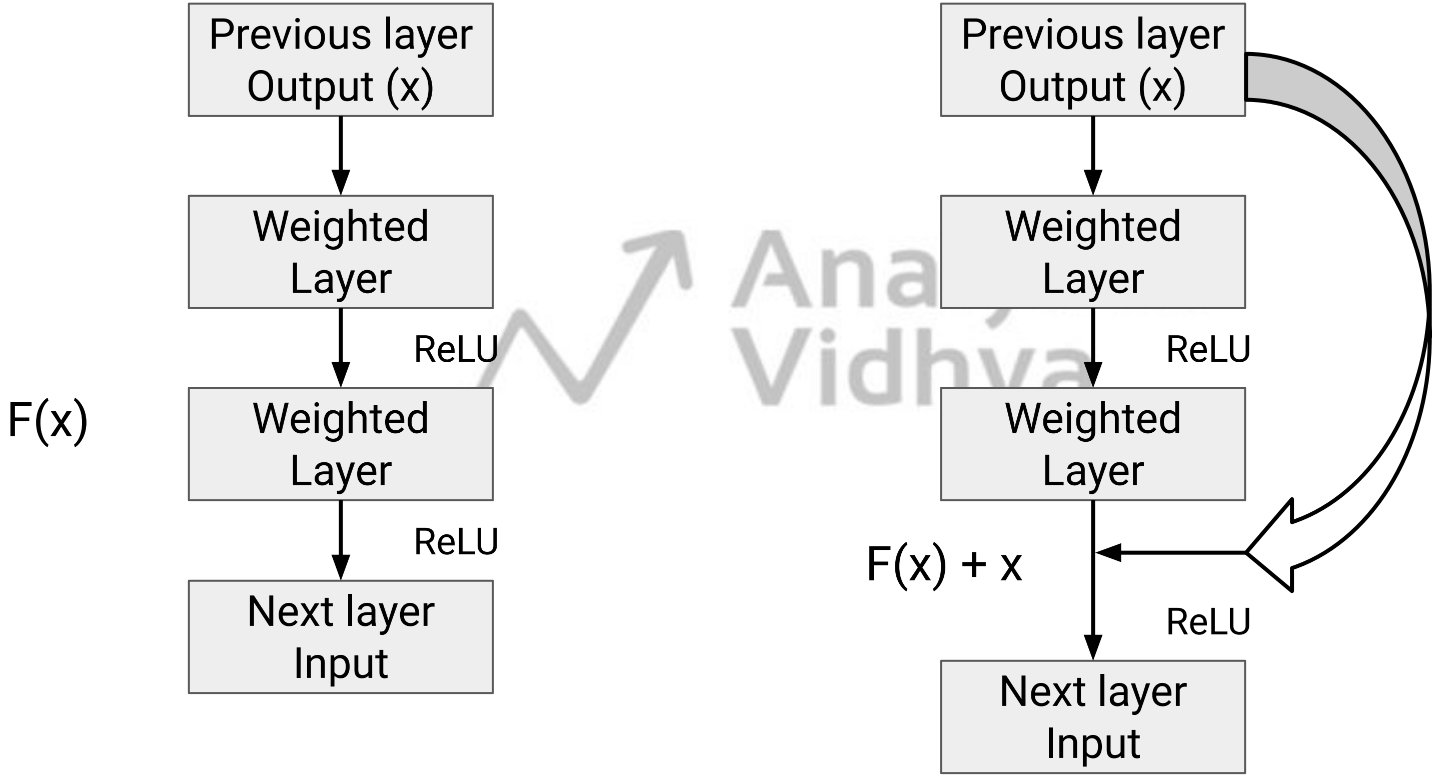


Residual Learning



Plain Block

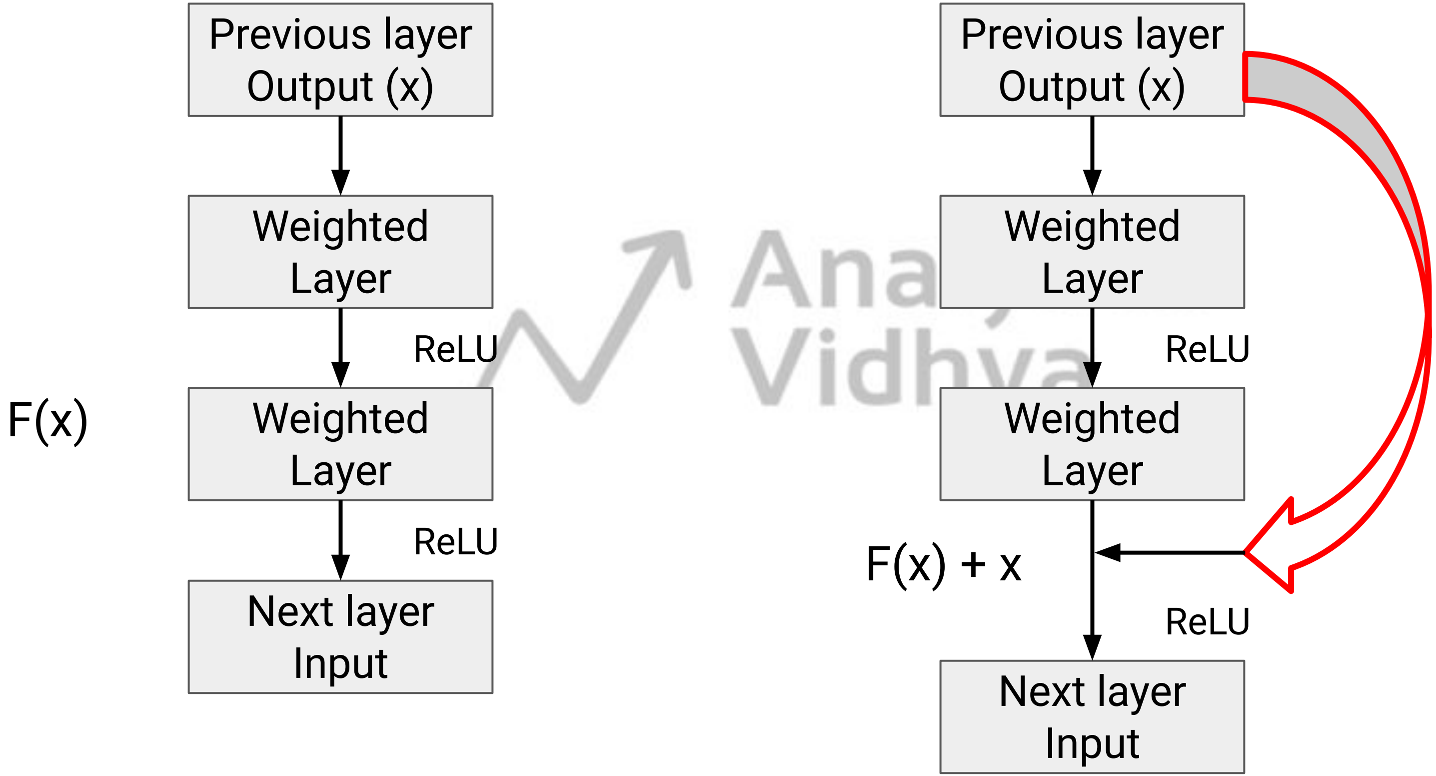
Residual Learning



Plain Block

Residual Block

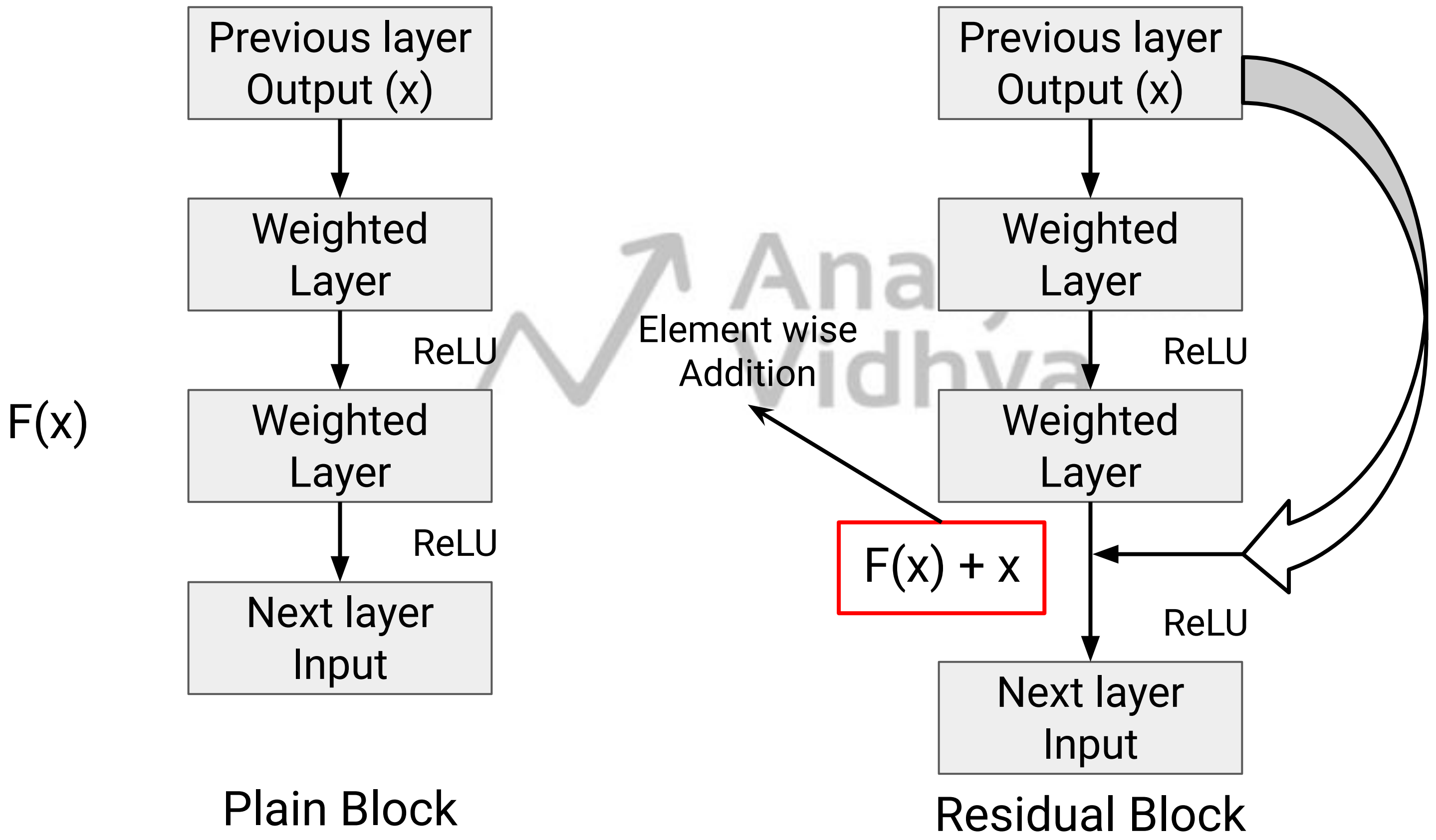
Residual Learning



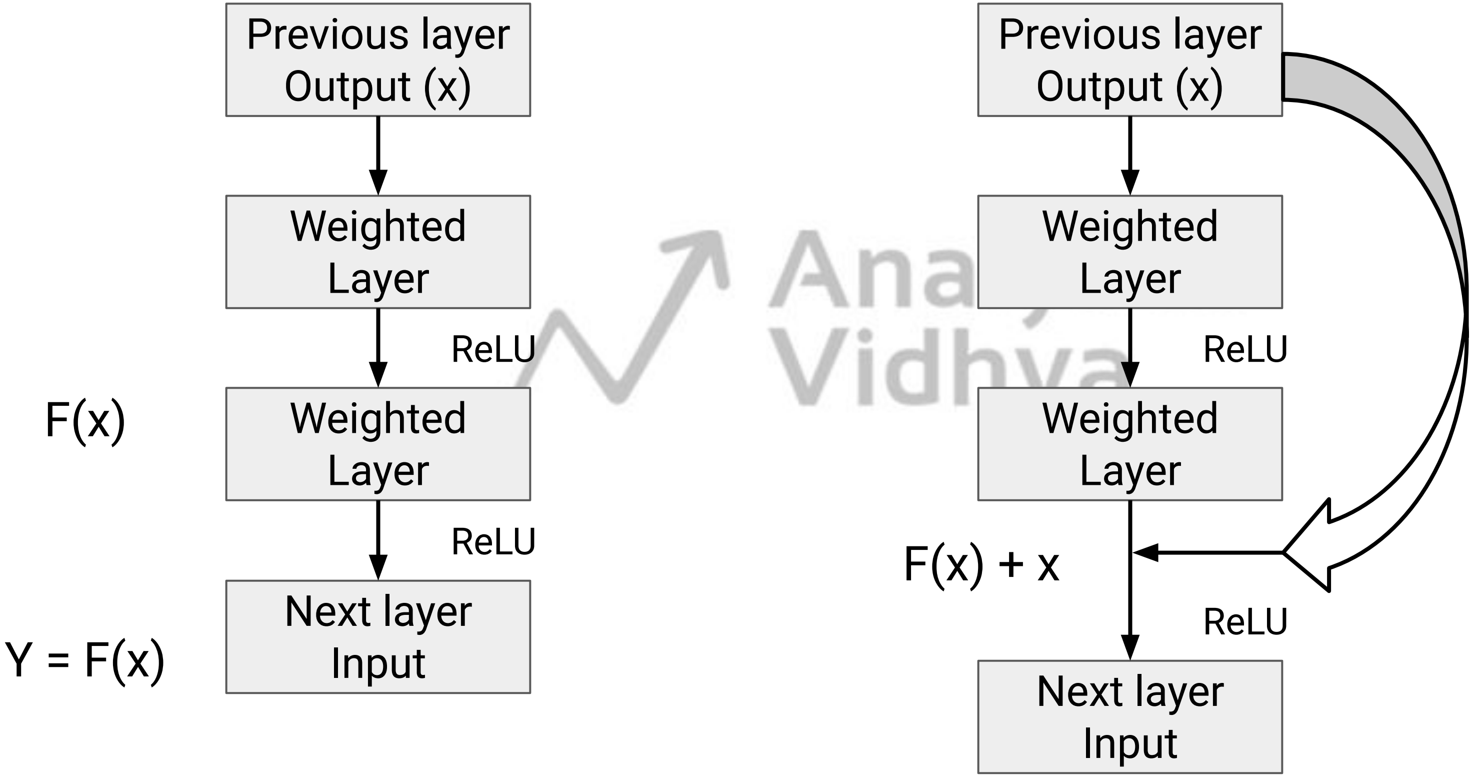
Plain Block

Residual Block

Residual Learning



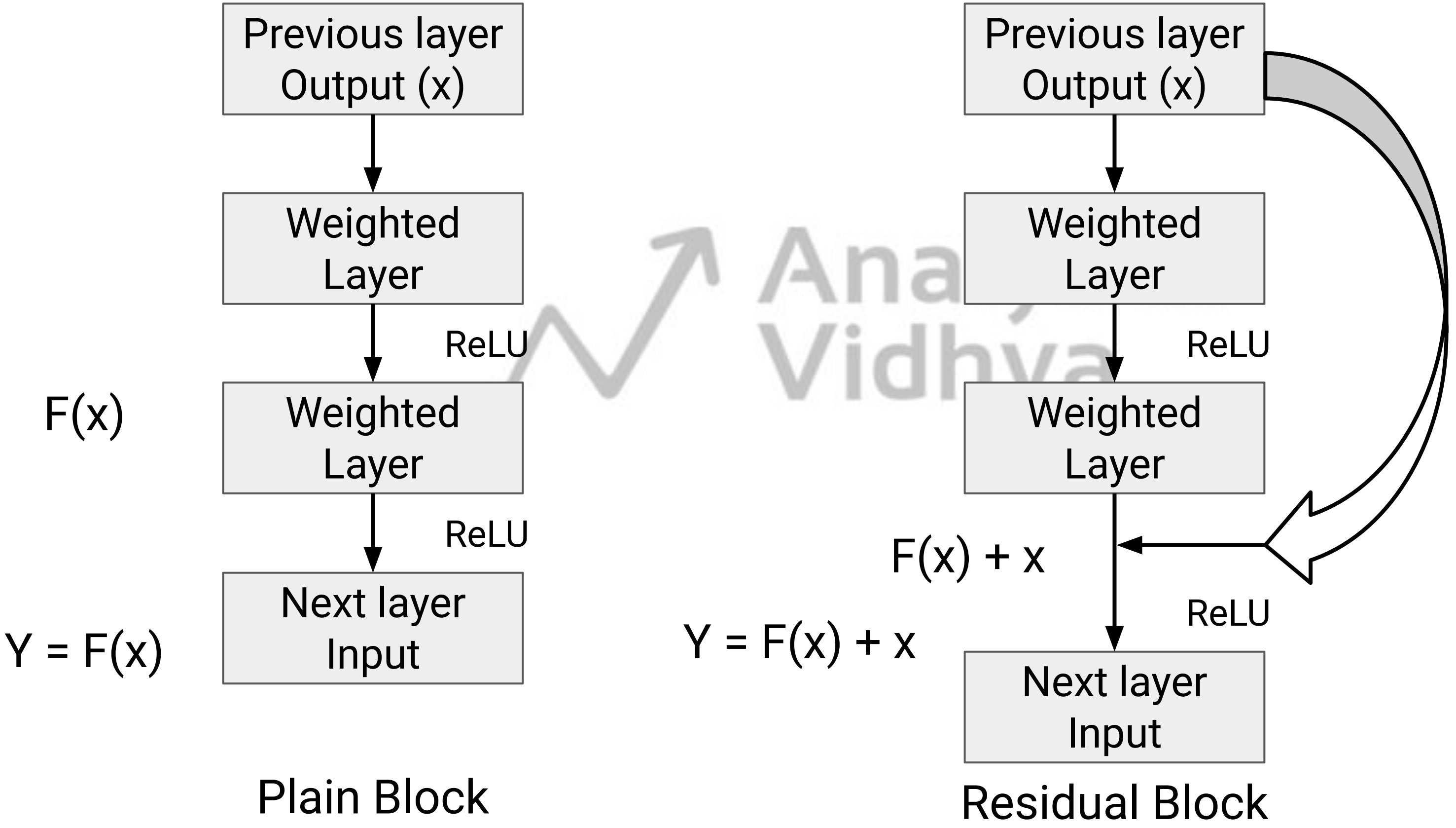
Residual Learning



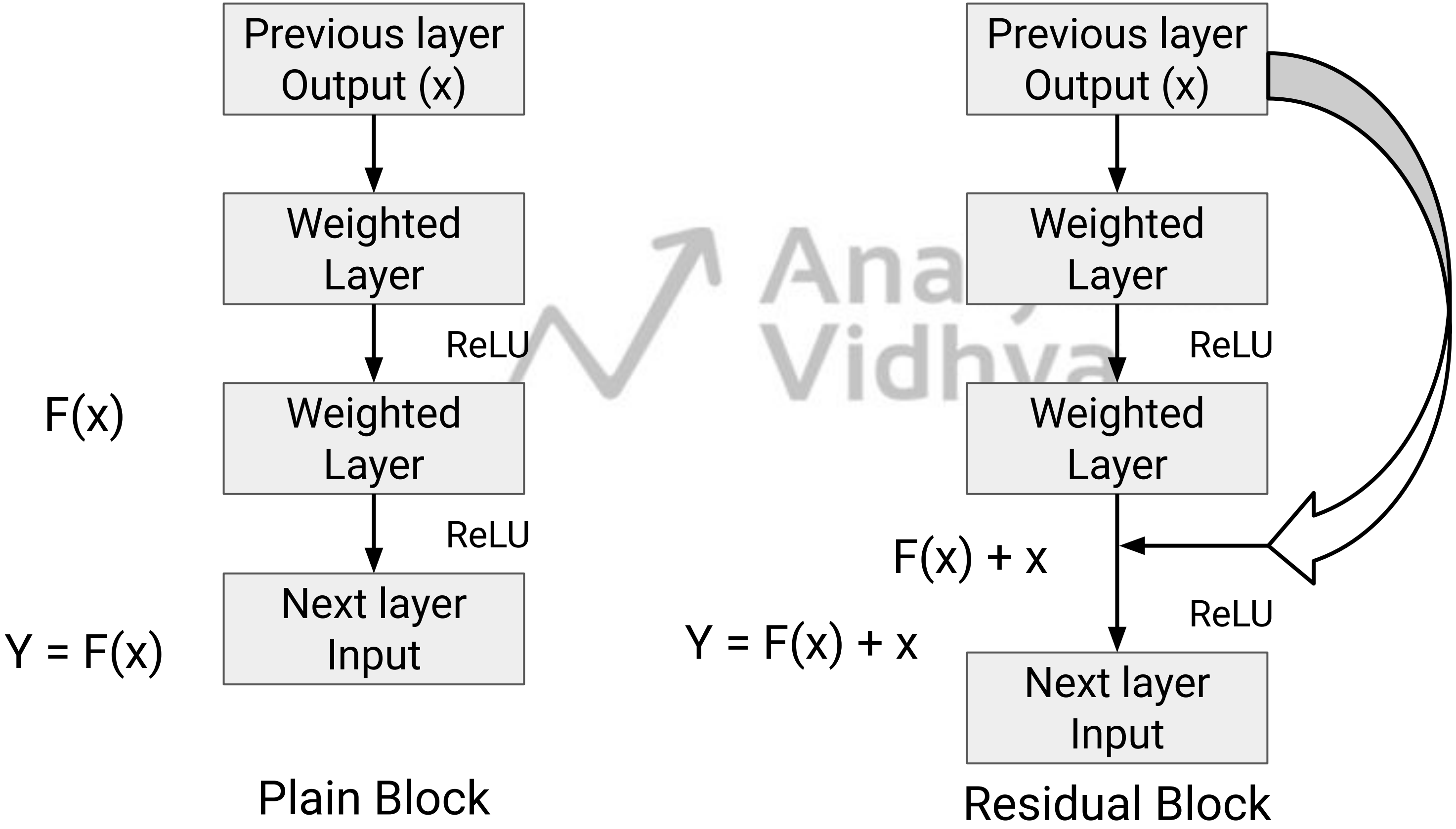
Plain Block

Residual Block

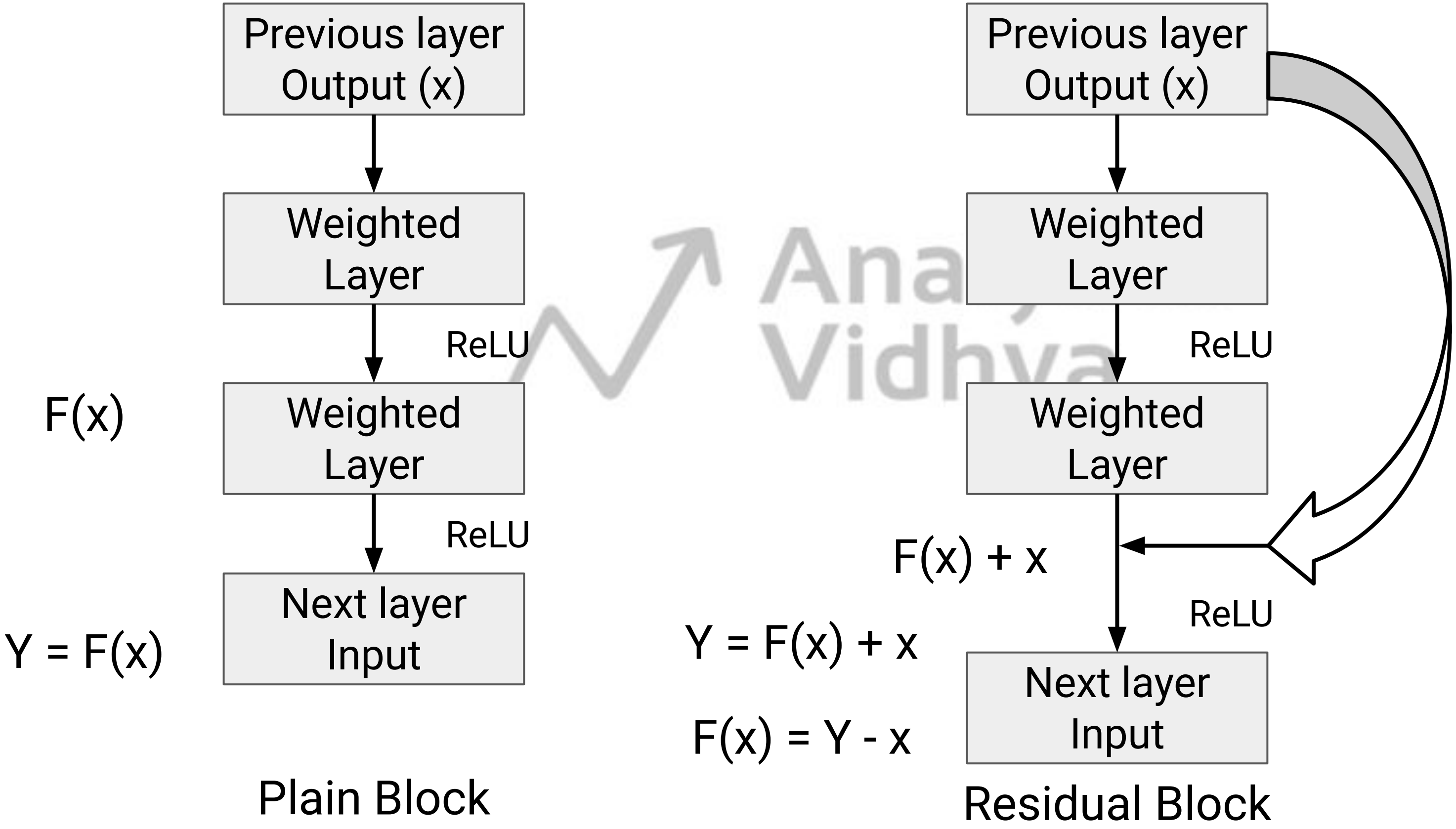
Residual Learning



Why Residual?



Why Residual?



Advantages of Residual Blocks

Advantages of Residual Blocks

- Can pass initial information to deeper layers

Advantages of Residual Blocks

- Can pass initial information to deeper layers
- $F(x) + x$ is an element wise addition, hence no extra parameters and no extra computational complexity

Advantages of Residual Blocks

- Can pass initial information to deeper layers
- $F(x) + x$ is an element wise addition, hence no extra parameters and no extra computational complexity
- Easier to optimize the residual mapping

Thank You