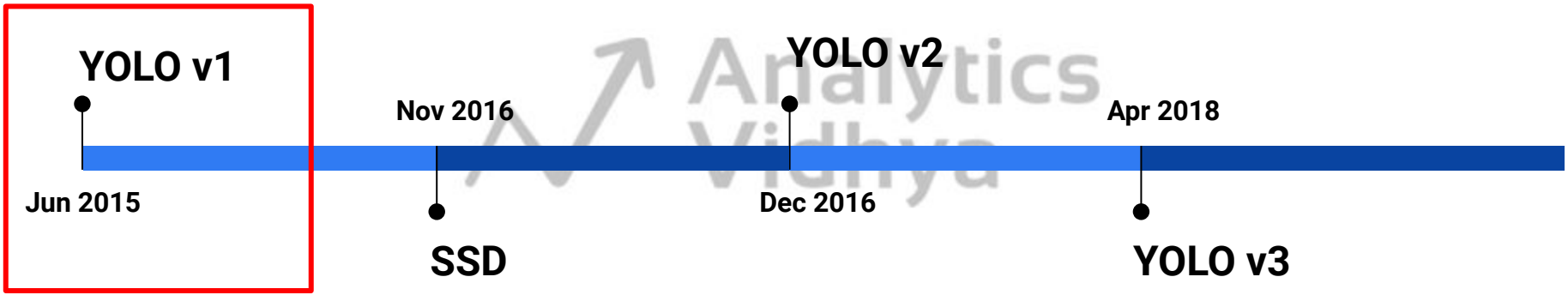
YOLO v1 - You Only Look Once

YOLO - You Only Look Once



YOLO - You Only Look Once

- YOLO has a single Neural Network

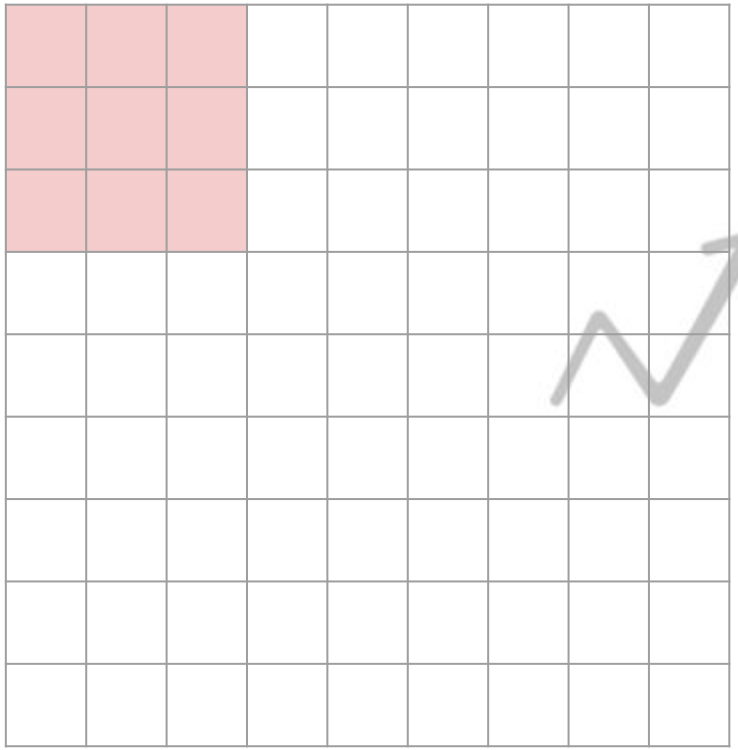
YOLO - You Only Look Once

- YOLO has a single Neural Network
- Lesser inference time than Faster R-CNN

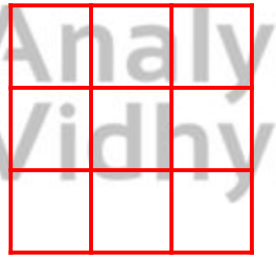
YOLO - You Only Look Once

- YOLO has a single Neural Network
- Lesser inference time than Faster R-CNN
- Network consists of Conv and Pooling layers

Intuition Behind YOLO

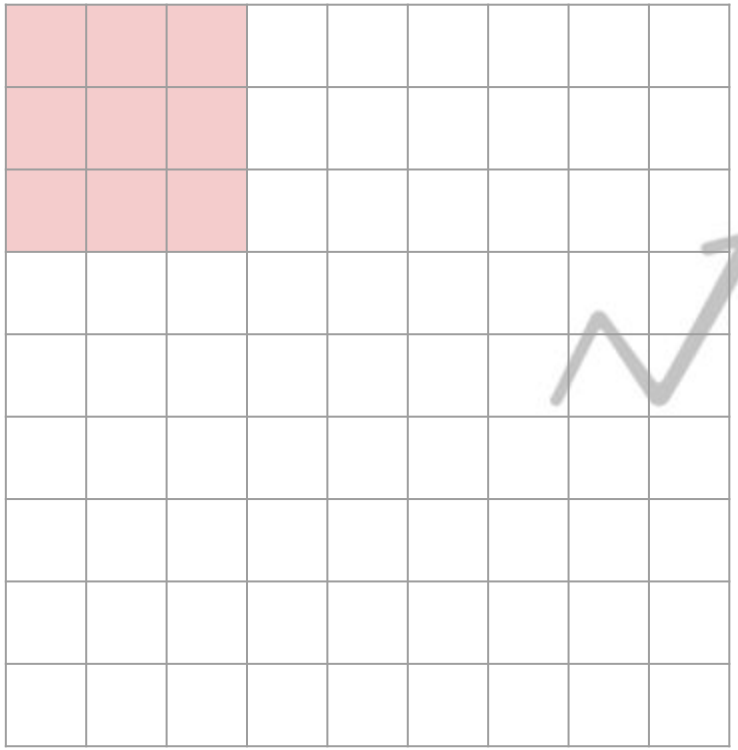


9 x 9

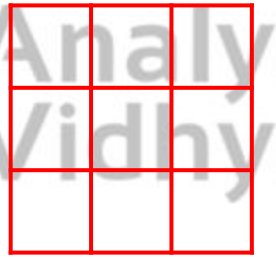


Filter size = 3 x 3
Stride = 3

Intuition Behind YOLO

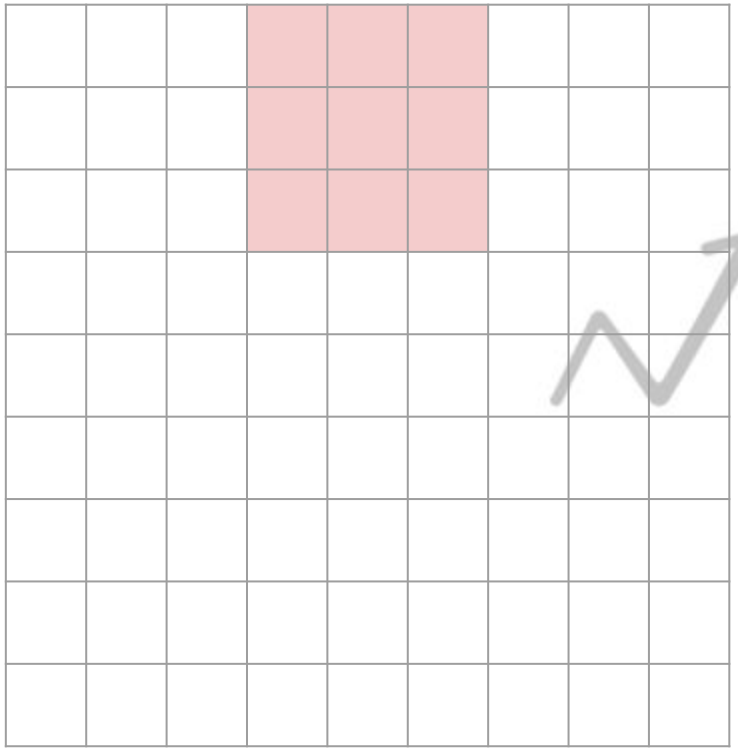


9 x 9

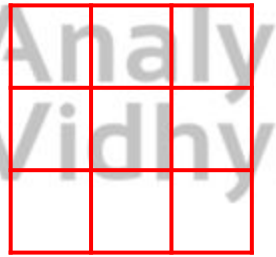


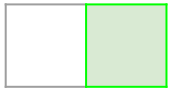
Filter size = 3 x 3
Stride = 3

Intuition Behind YOLO



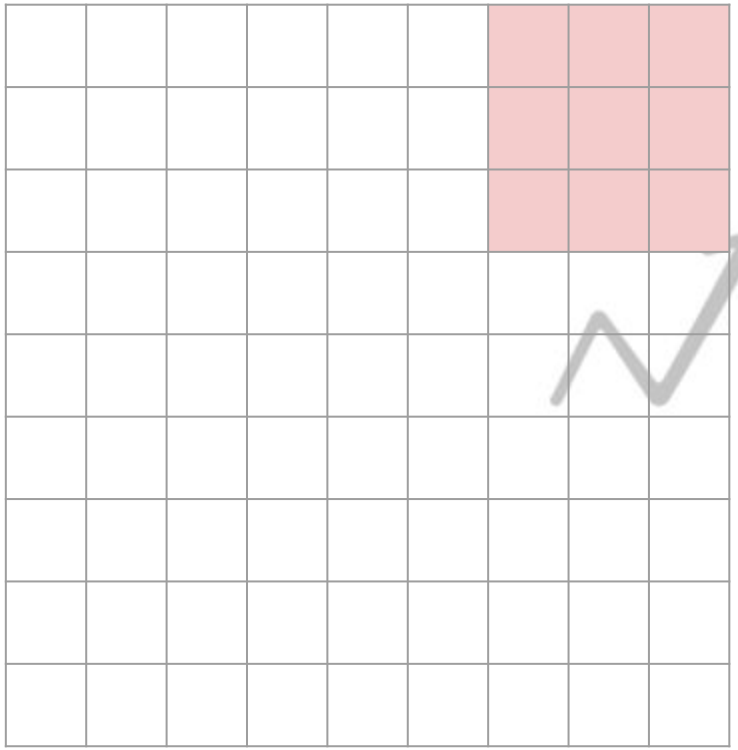
9 x 9



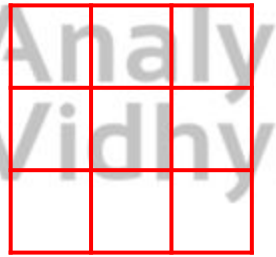


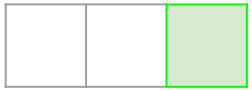
Filter size = 3 x 3
Stride = 3

Intuition Behind YOLO



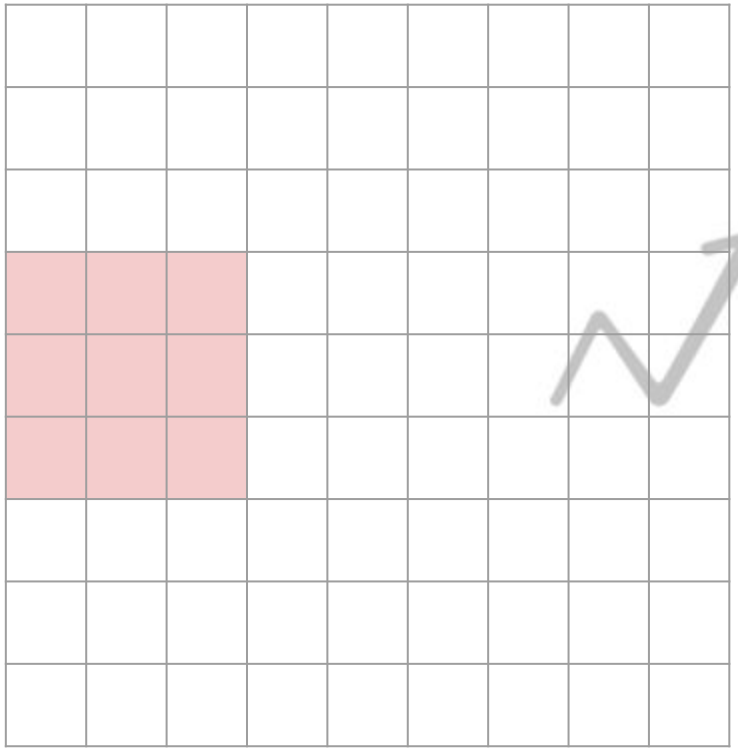
9 x 9



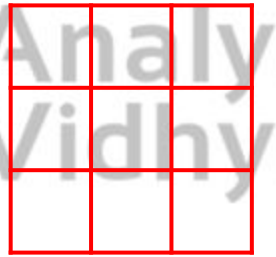


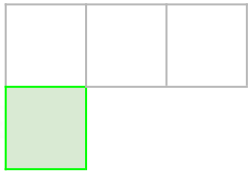
Filter size = 3 x 3
Stride = 3

Intuition Behind YOLO



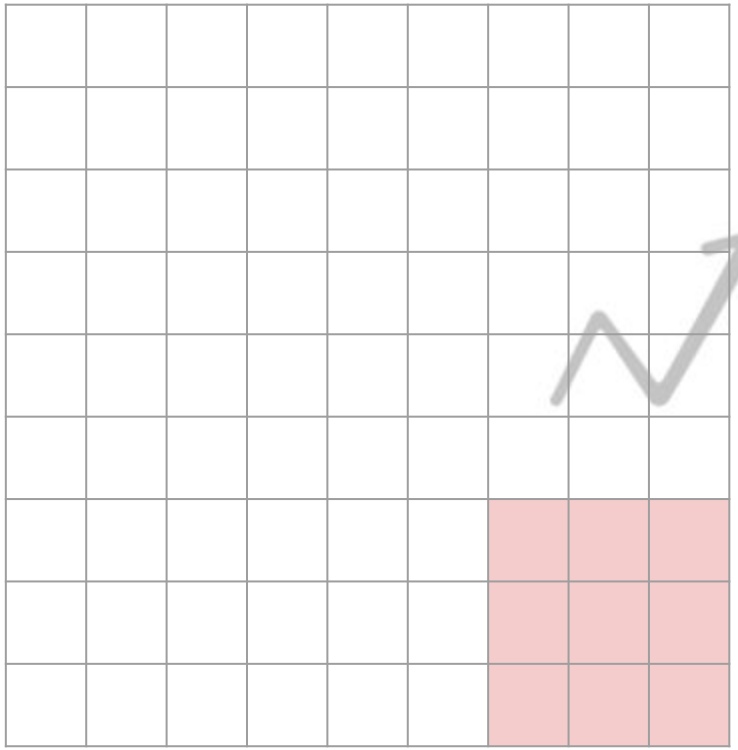
9 x 9



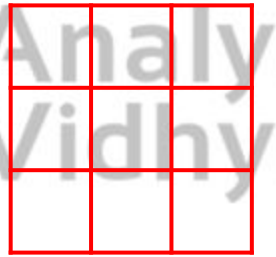


Filter size = 3 x 3
Stride = 3

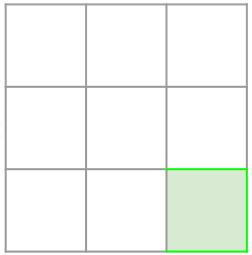
Intuition Behind YOLO



9 x 9

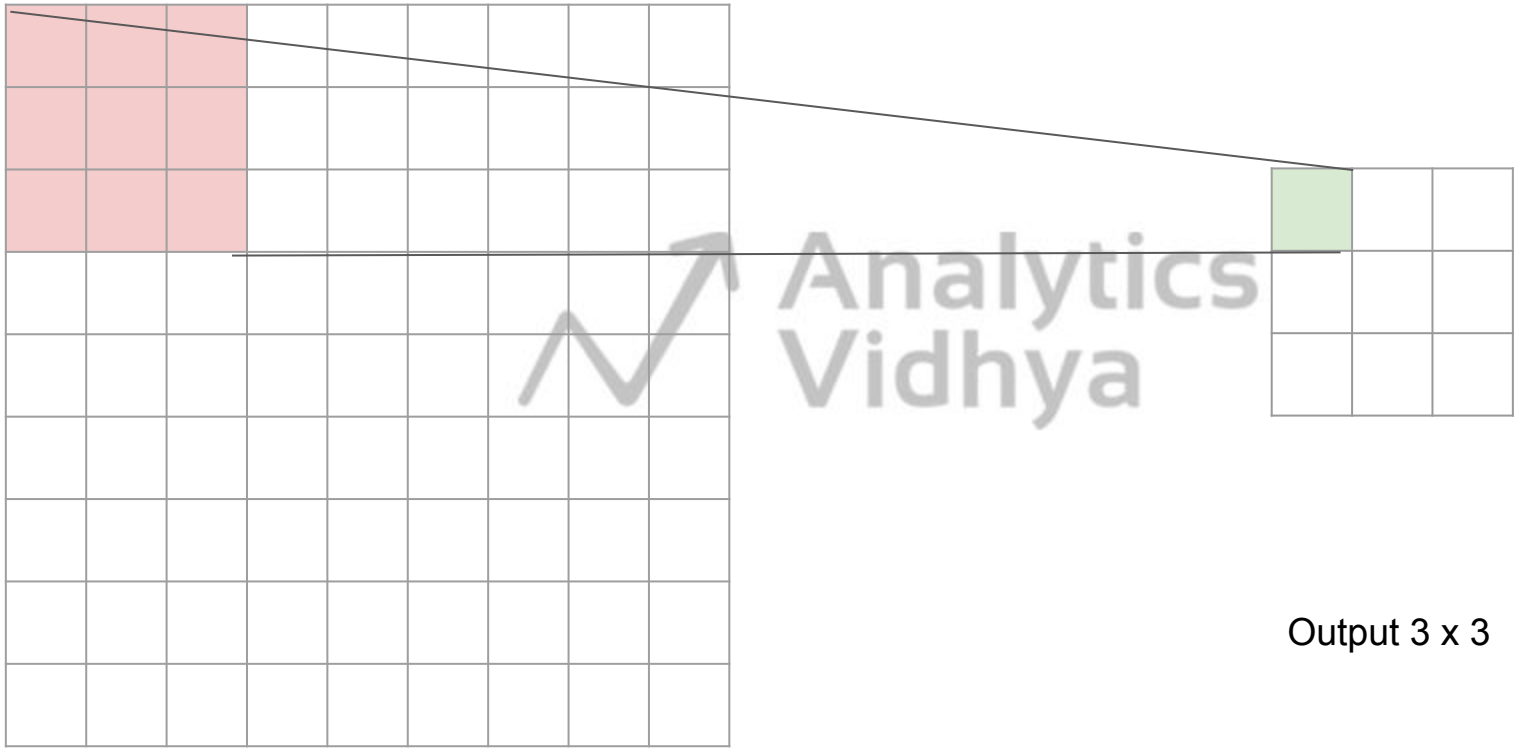


Filter size = 3 x 3
Stride = 3



3 x 3

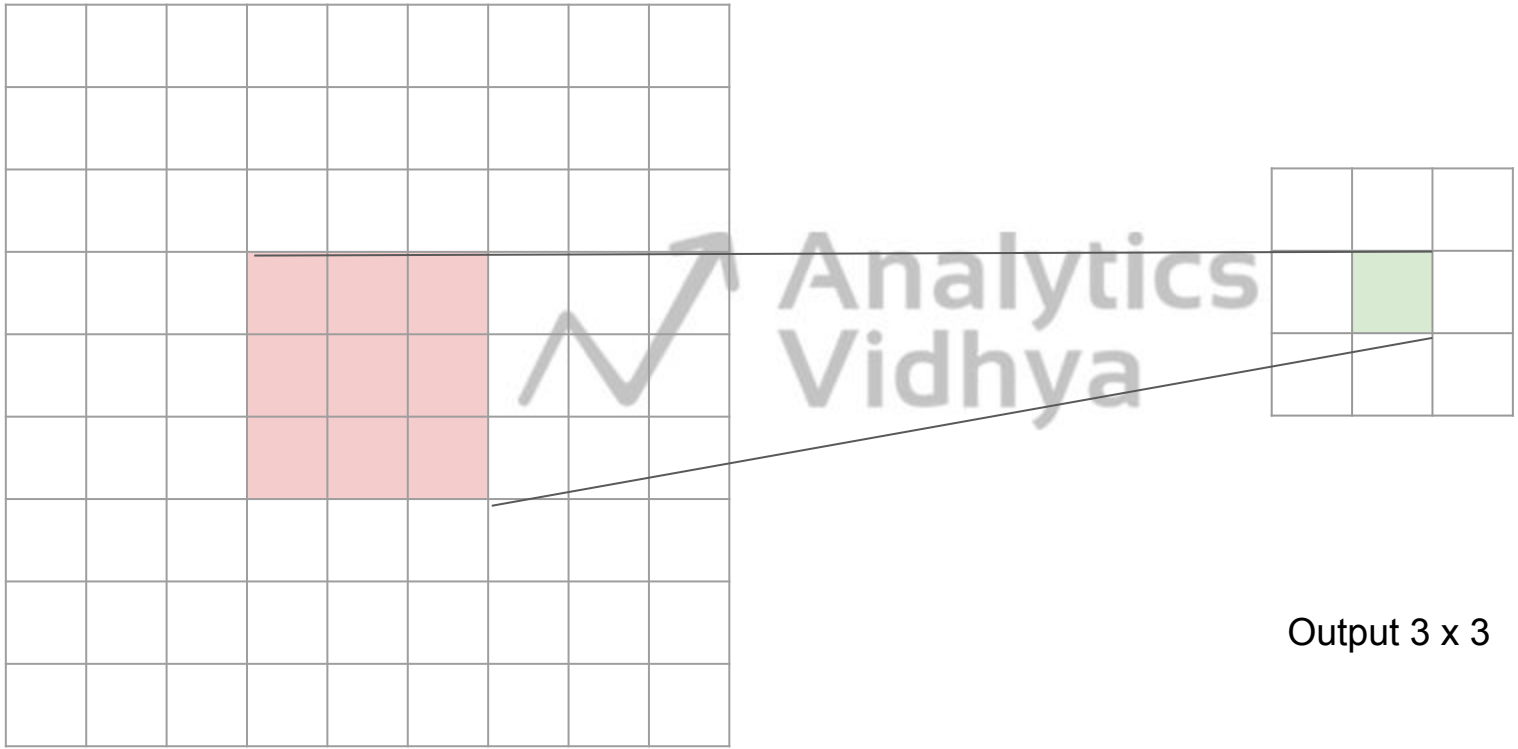
Intuition Behind YOLO



Output 3 x 3

Input Image - 9 x 9

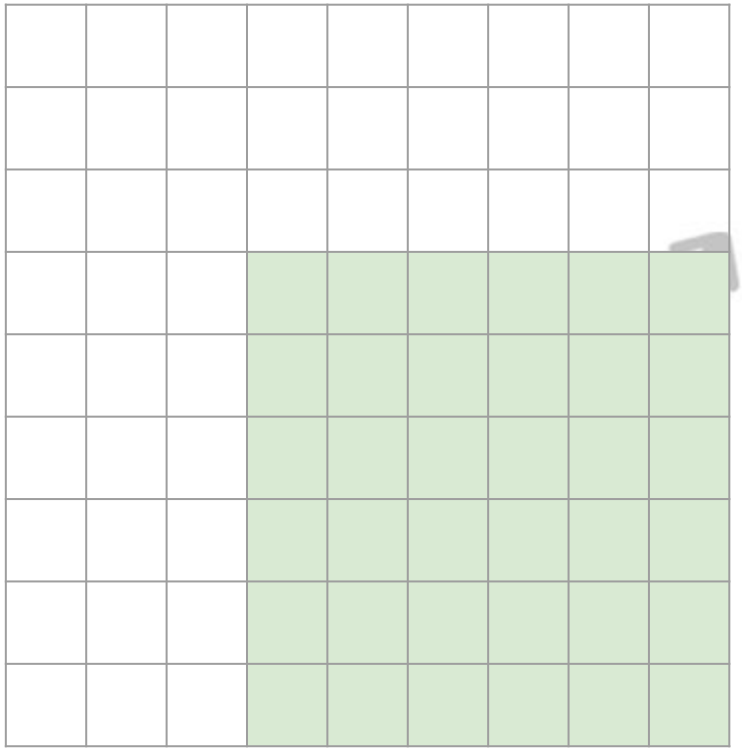
Intuition Behind YOLO



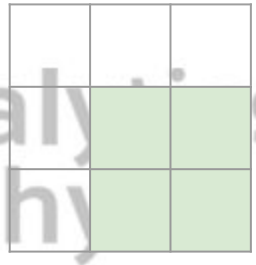
Output 3 x 3

Input Image - 9 x 9

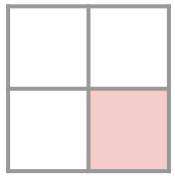
Intuition Behind YOLO



Input Image - 9 x 9



Output 3 x 3



Output 2 x 2

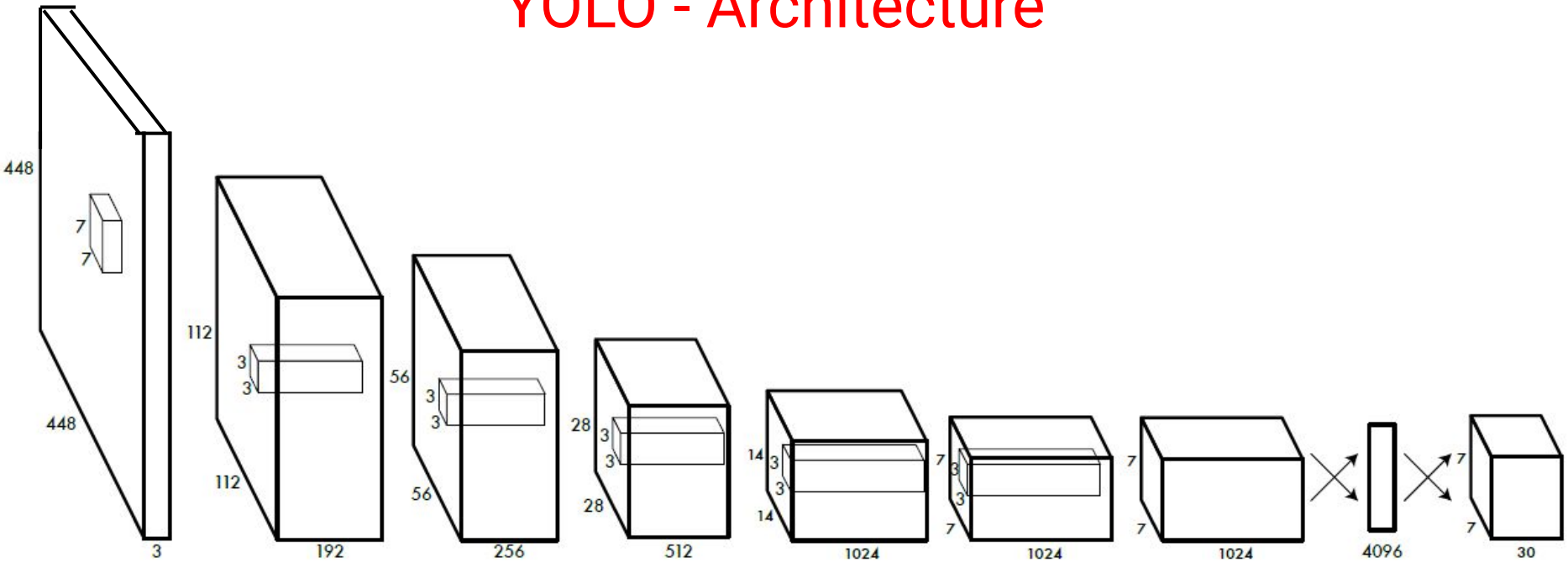
Sliding Window Approach

- Window slides through the complete image
- Generates predictions for the grids as output

YOLO - You Only Look Once

- YOLO has a single Neural Network
- Lesser inference time than Faster R-CNN
- Network consists of Conv and Pooling layers
- Uses the sliding window approach for object detection

YOLO - Architecture



Conv. Layer
7x7x64-s-2
Maxpool Layer
2x2-s-2

Conv. Layer
3x3x192
Maxpool Layer
2x2-s-2

Conv. Layers
1x1x128
3x3x256
1x1x256
3x3x512
Maxpool Layer
2x2-s-2

Conv. Layers
1x1x256
3x3x512
} ×4
1x1x512
3x3x1024
Maxpool Layer
2x2-s-2

Conv. Layers
1x1x512
3x3x1024
} ×2
3x3x1024
3x3x1024-s-2

Conv. Layers
3x3x1024
3x3x1024

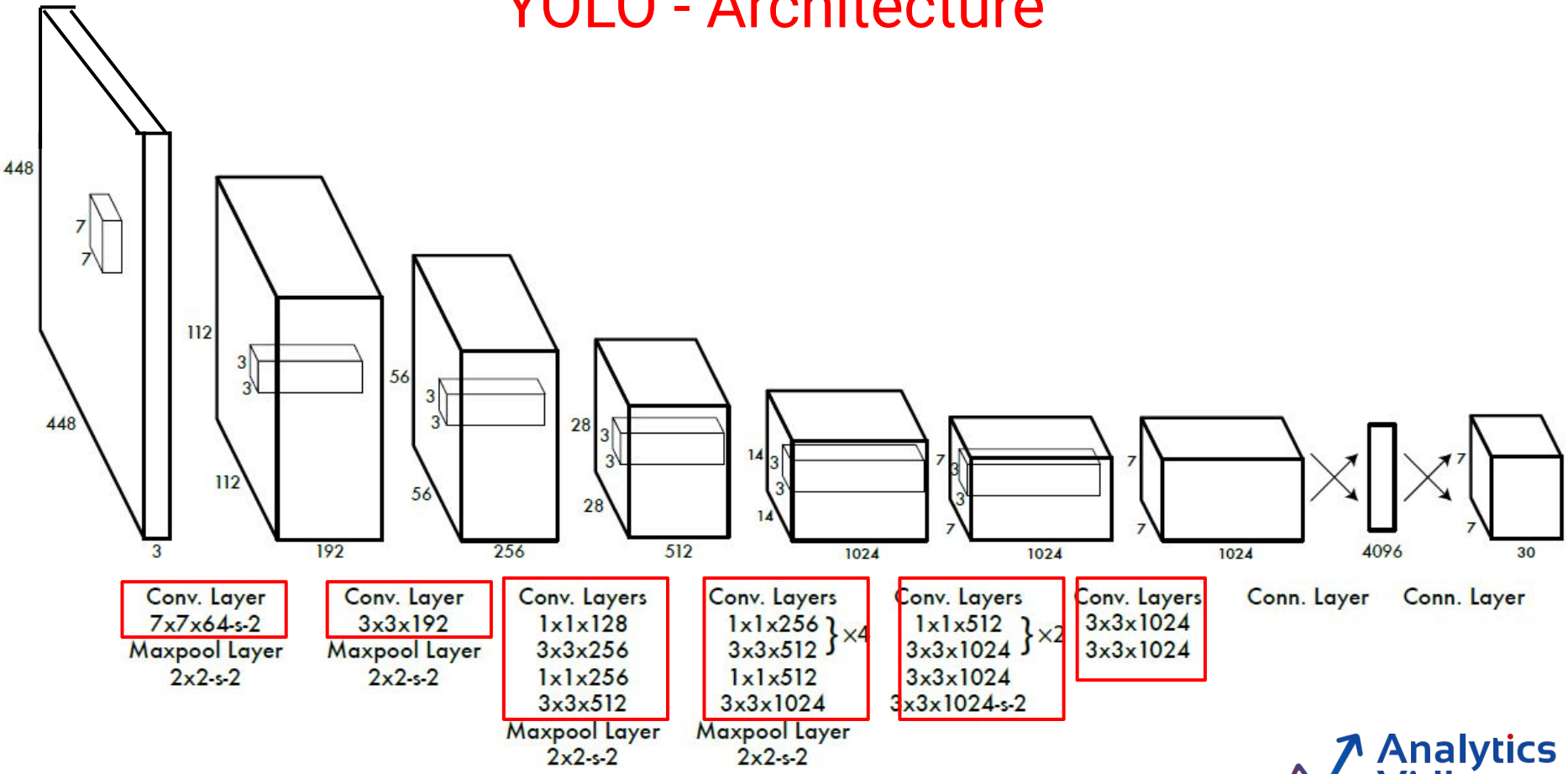
Conn. Layer

Conn. Layer

YOLO v1 architecture. Image source https://arxiv.org/pdf/1506.02640.pdf

Analytics Vidhya

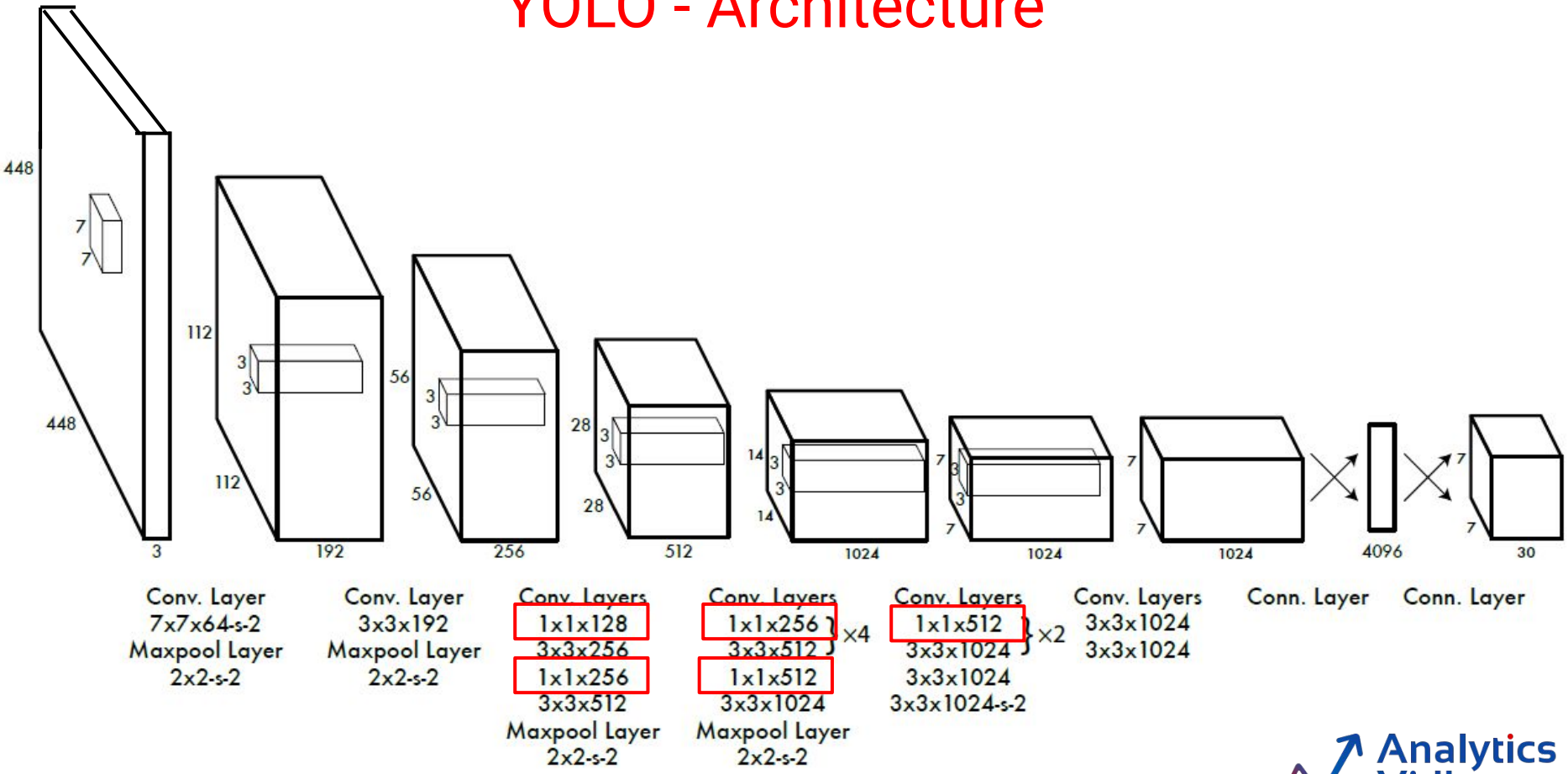
YOLO - Architecture



YOLO v1 architecture. Image source https://arxiv.org/pdf/1506.02640.pdf

Analytics Vidhya

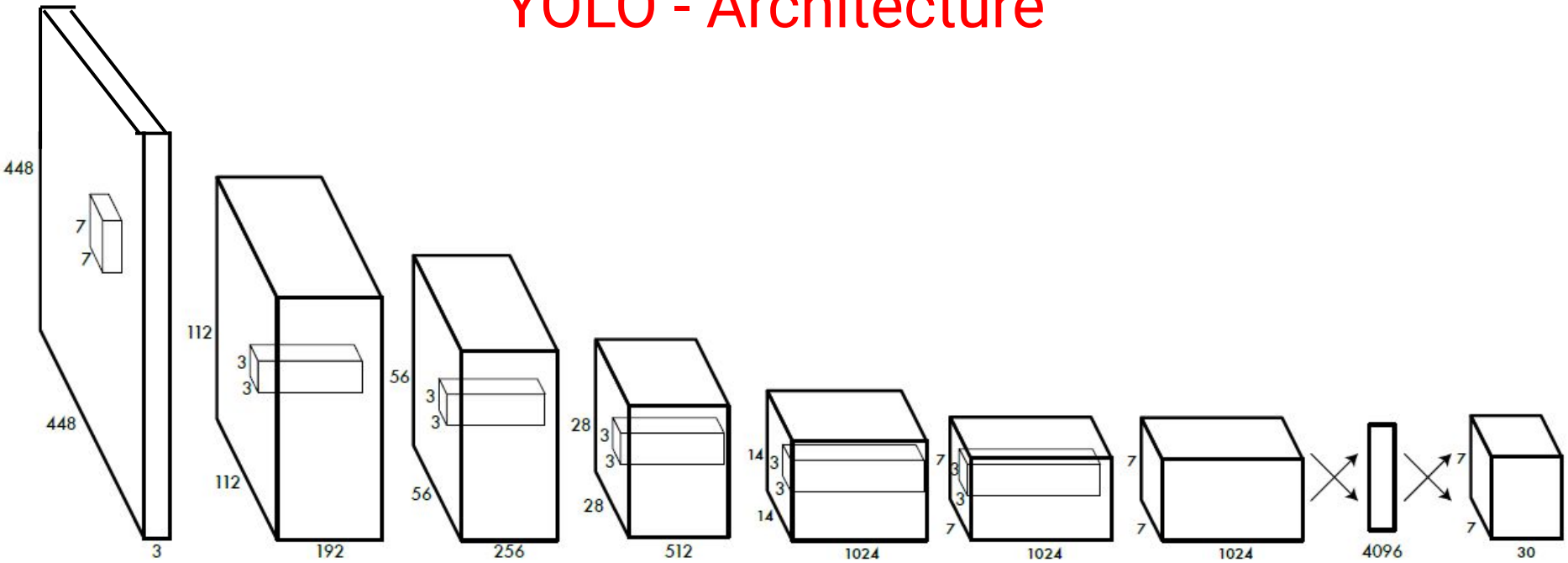
YOLO - Architecture



YOLO v1 architecture. Image source https://arxiv.org/pdf/1506.02640.pdf

Analytics Vidhya

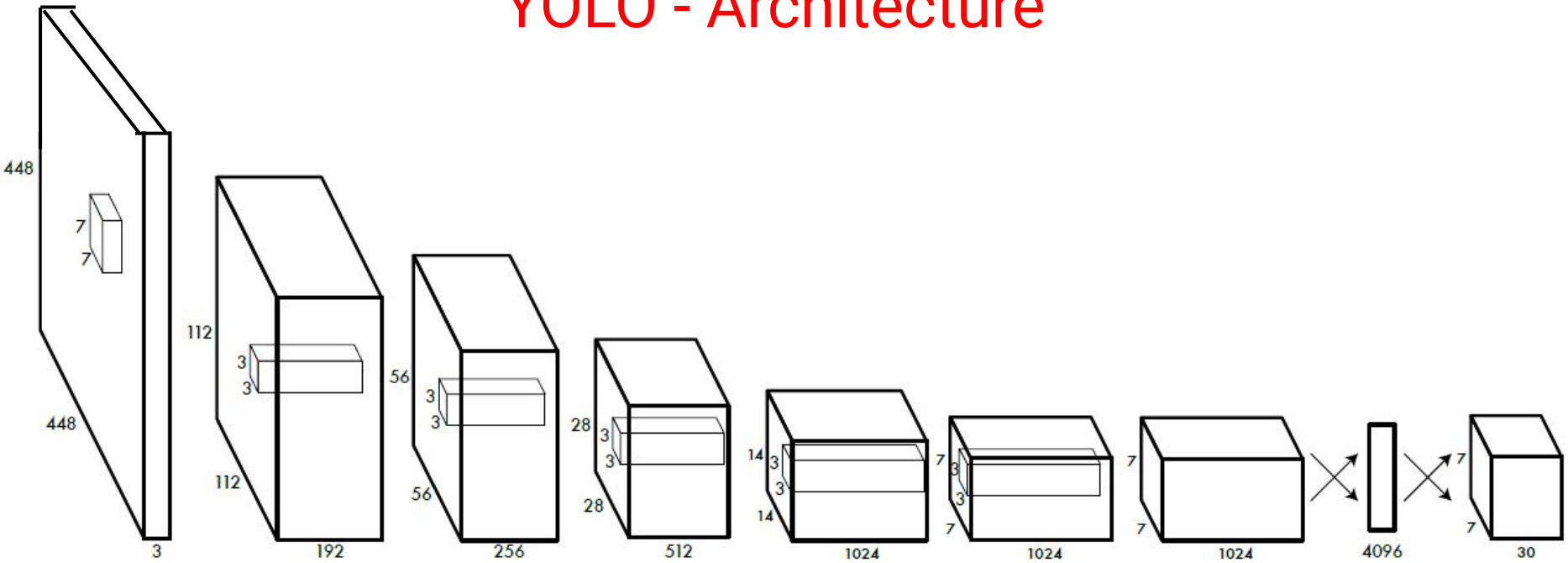
YOLO - Architecture



YOLO v1 architecture. Image source https://arxiv.org/pdf/1506.02640.pdf

Analytics Vidhya

YOLO - Architecture



Conv. Layer
7x7x64-s-2
Maxpool Layer
2x2-s-2

Conv. Layer
3x3x192
Maxpool Layer
2x2-s-2

Conv. Layers
1x1x128
3x3x256
1x1x256
3x3x512
Maxpool Layer
2x2-s-2

Conv. Layers
1x1x256
3x3x512
} ×4
1x1x512
3x3x1024
Maxpool Layer
2x2-s-2

Conv. Layers
1x1x512
3x3x1024
} ×2
3x3x1024
3x3x1024-s-2

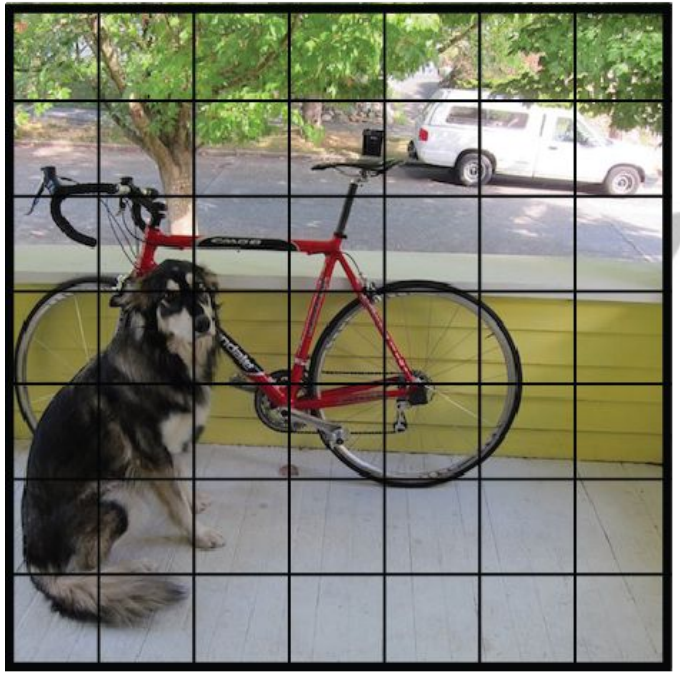
Conv. Layers
3x3x1024
3x3x1024

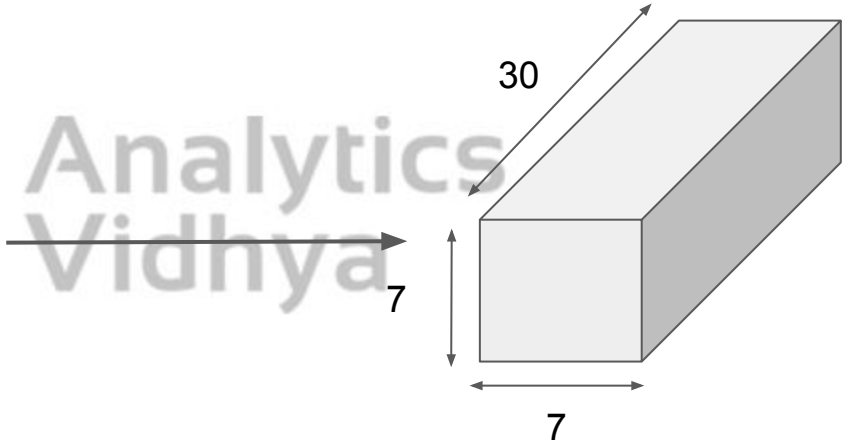
Conn. Layer Conn. Layer

YOLO v1 architecture. Image source https://arxiv.org/pdf/1506.02640.pdf

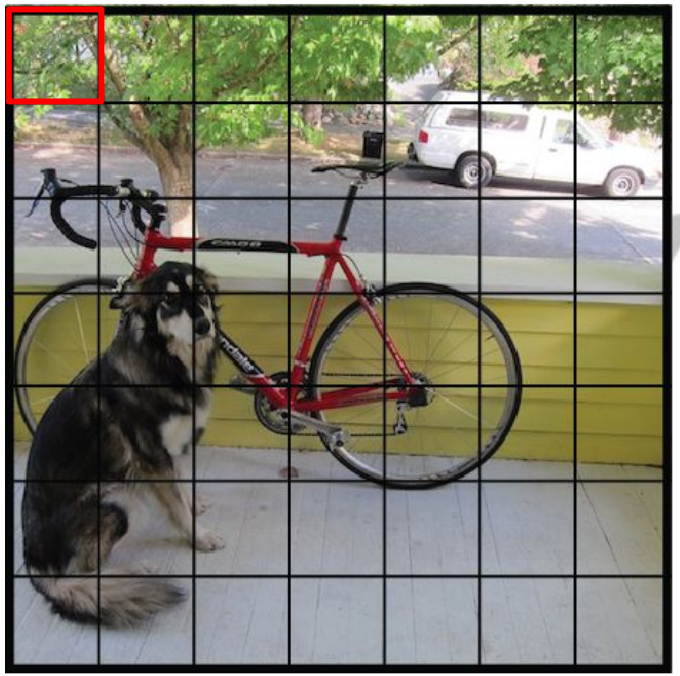
Analytics Vidhya

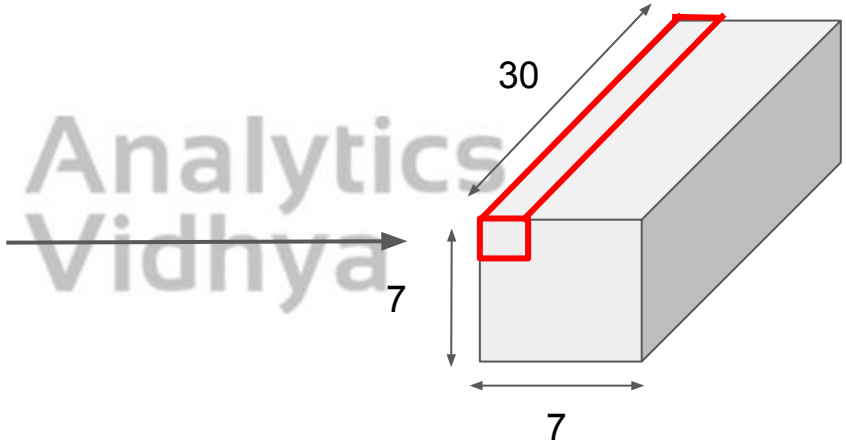
YOLO - Architecture



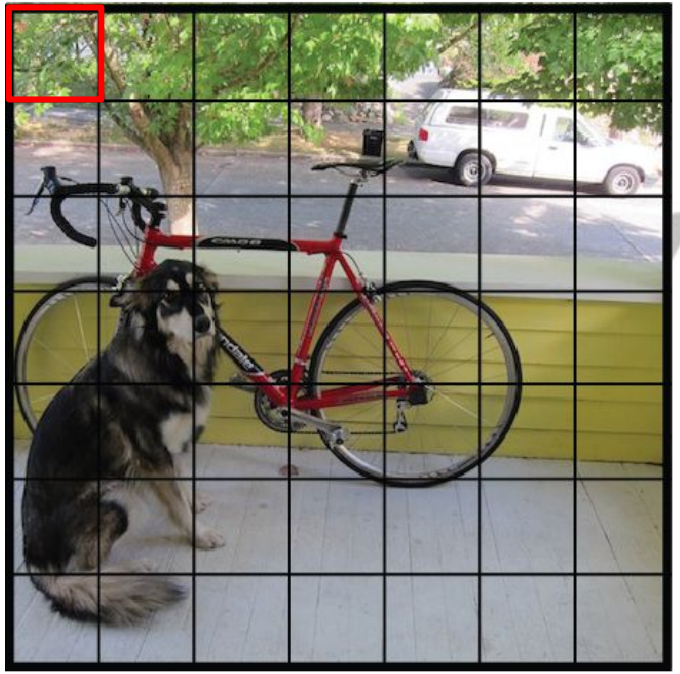


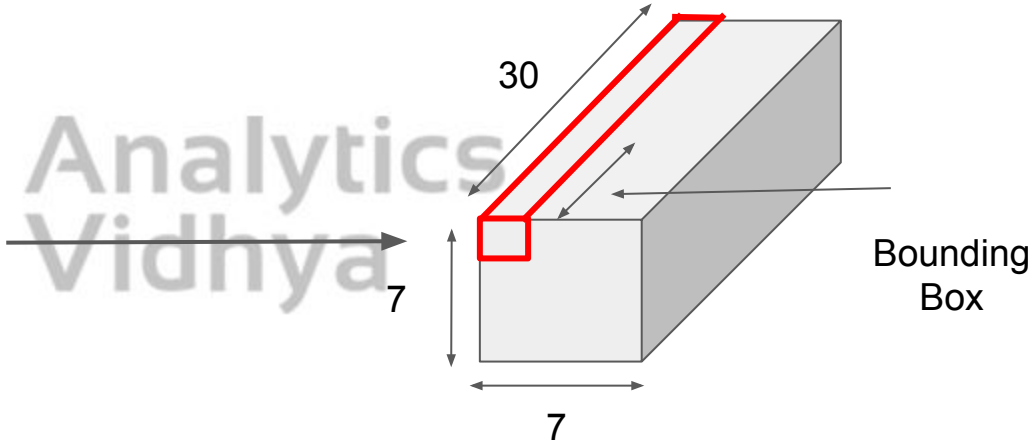
YOLO - Architecture



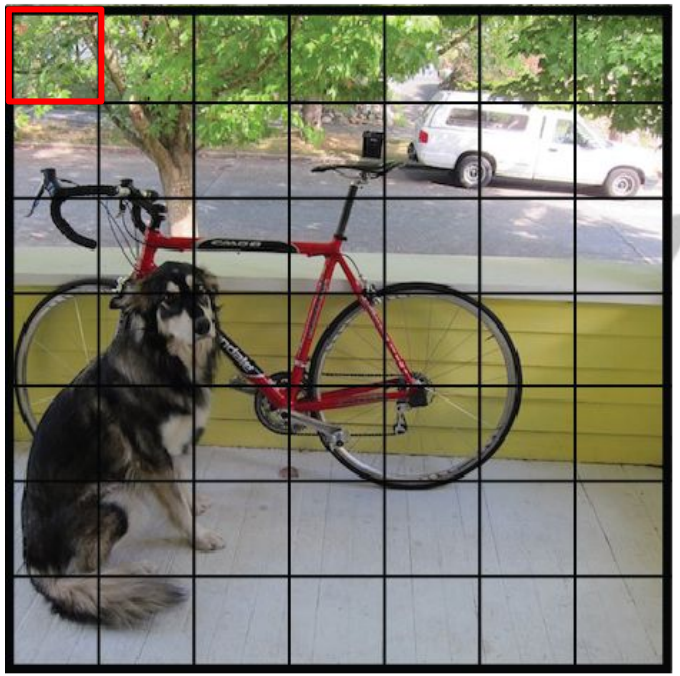


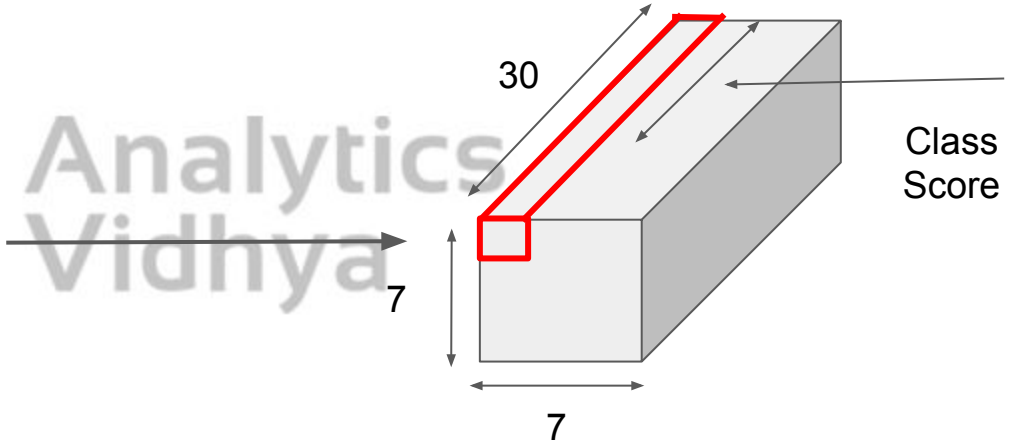
YOLO - Architecture





YOLO - Architecture

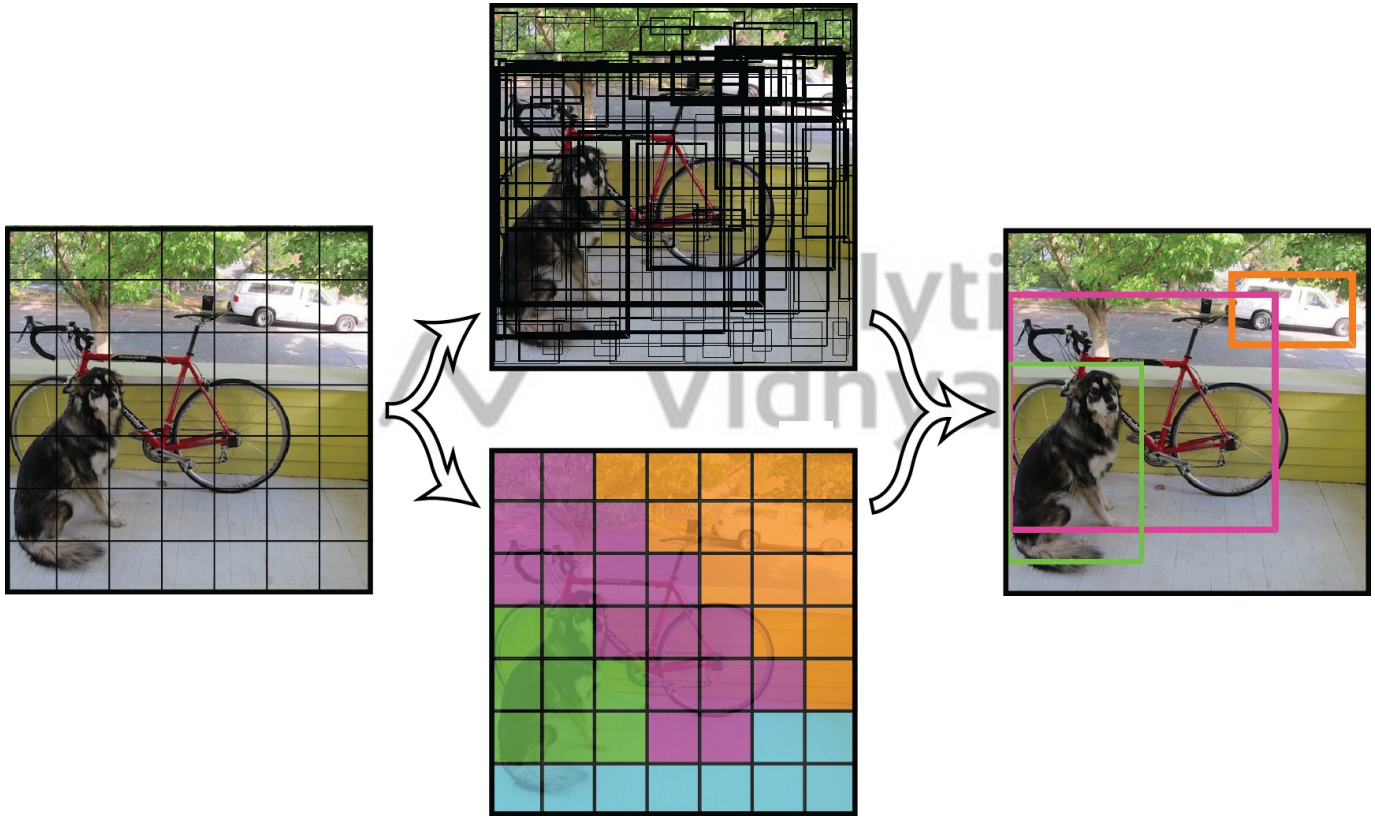




YOLO - Architecture

Source - https://arxiv.org/abs/1612.08242

YOLO - Architecture



Source - https://arxiv.org/abs/1612.08242

Analytics Vidhya

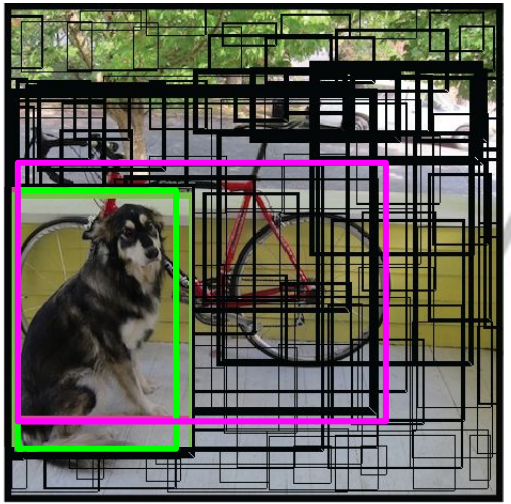
Non-max Suppression

- Select a box with highest score

Non-max Suppression

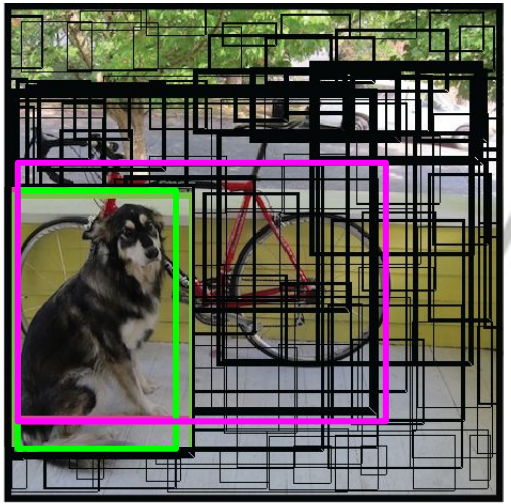
- Select a box with highest score
- Compare IOU with other boxes

Non-max Suppression



- Select a box with highest score
- Compare IOU with other boxes
- Delete all boxes with high overlap with the selected box

Non-max Suppression



- Select a box with highest score
- Compare IOU with other boxes
- Delete all boxes with high overlap with the selected box
- Repeat steps 1 to 3

YOLO - Architecture

Source - https://arxiv.org/abs/1612.08242

Analytics Vidhya

Limitations of YOLO v1 Architecture

- Does not work well for smaller objects

Limitations of YOLO v1 Architecture

- Does not work well for smaller objects
- Unable to detect cluttered objects

Limitations of YOLO v1 Architecture

- Does not work well for smaller objects
- Unable to detect cluttered objects
- Not generalized for unusual aspect ratios

Limitations of YOLO v1 Architecture

- Does not work well for smaller objects
- Unable to detect cluttered objects
- Not generalized for unusual aspect ratios
- Difficulty in detecting object of different scales

Thank You

Analytics Vidhya