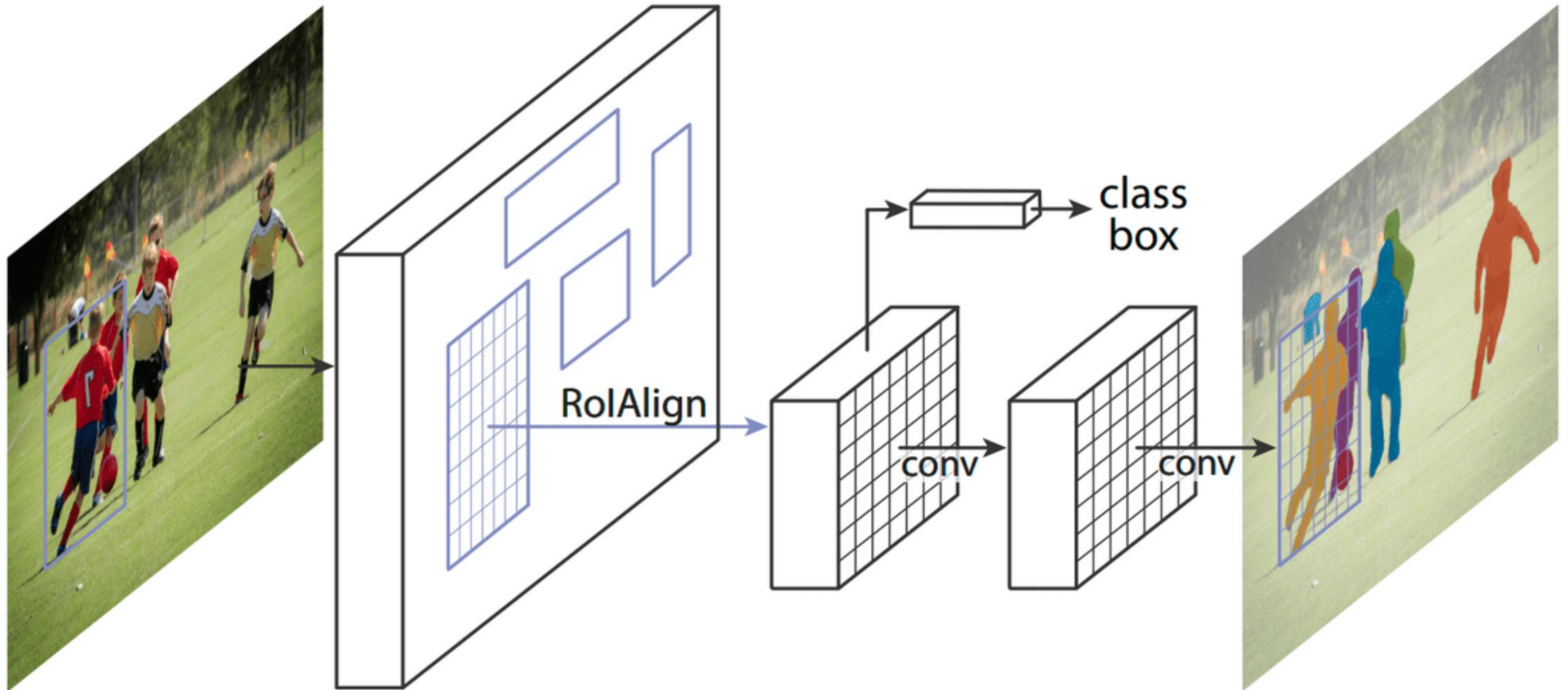


# ROI-Align

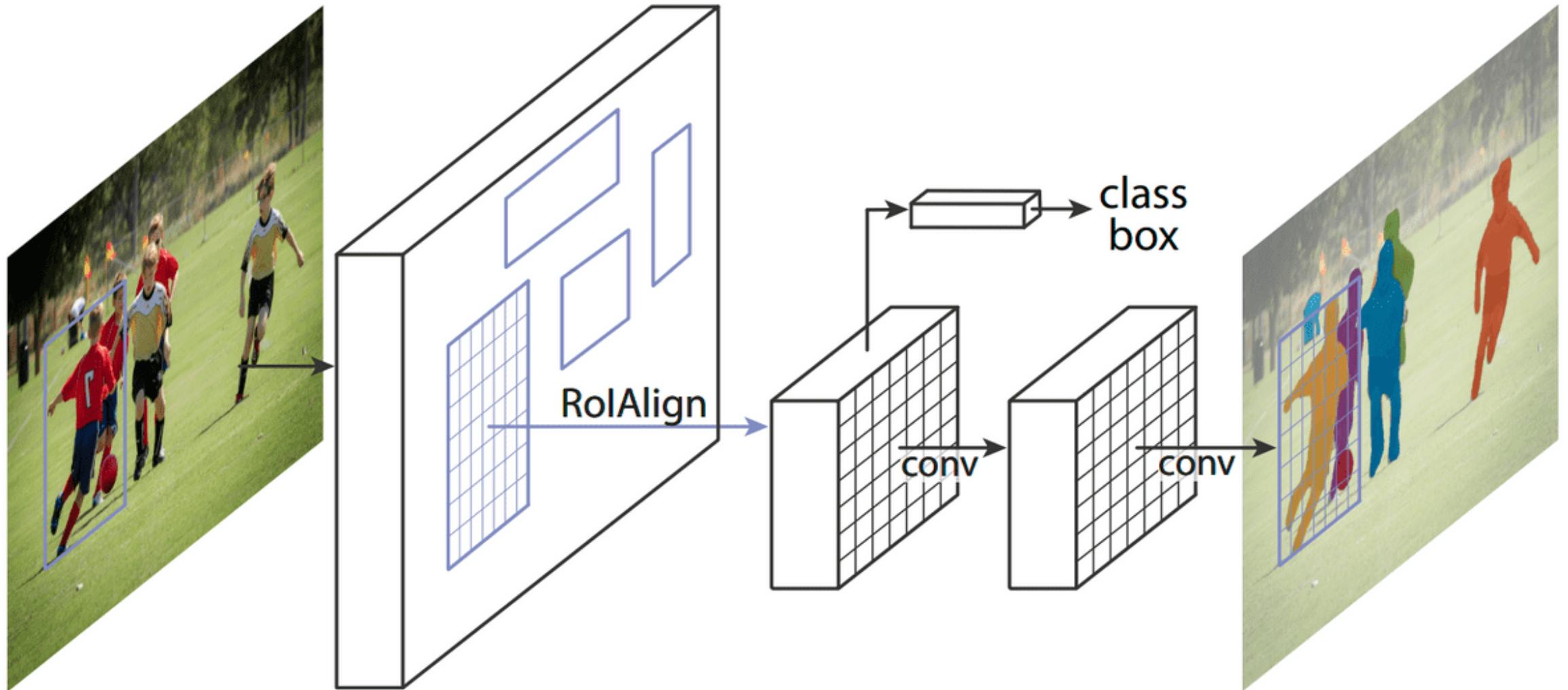
Presented by: Boniface  
Samuel Ijeoma

# Region of Interest (ROI) Align in Object Detection



Introduction

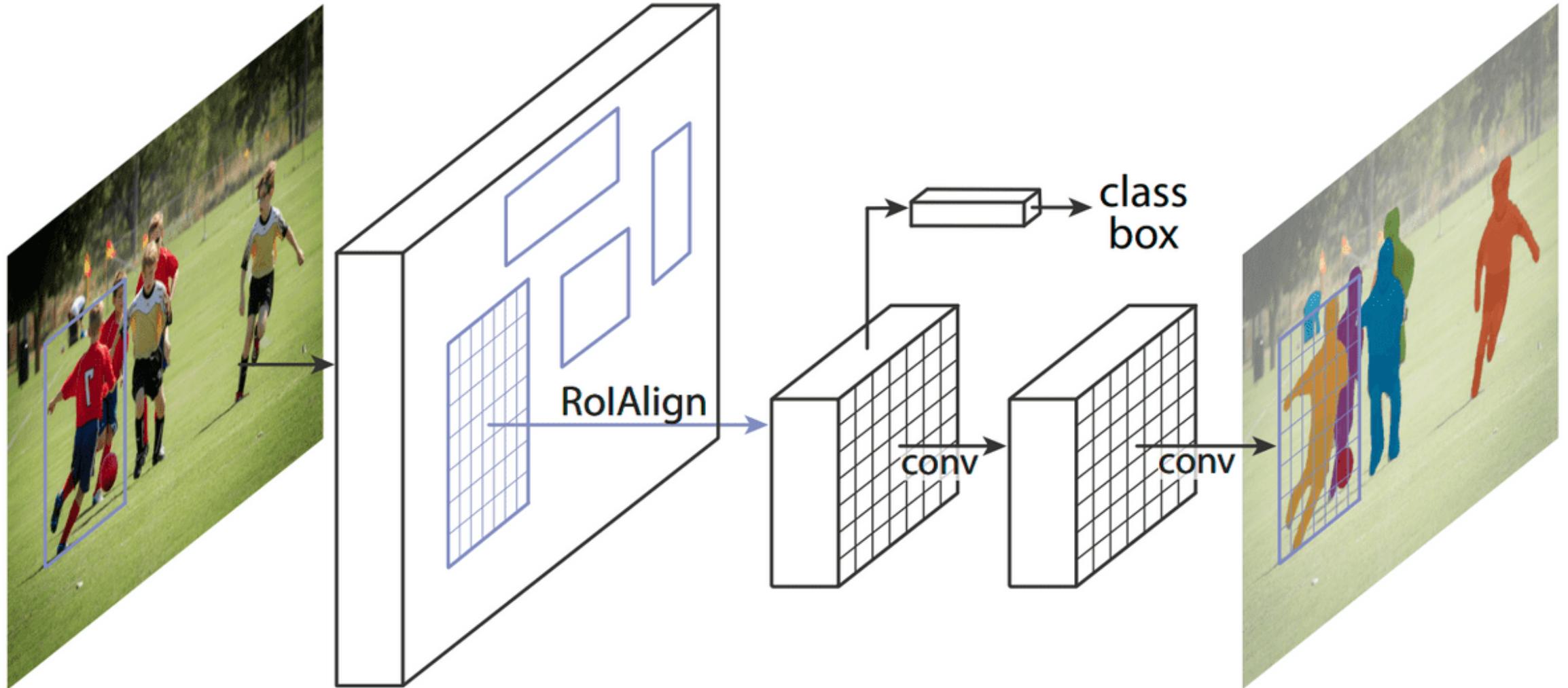
# Region of Interest (ROI) Align in Object Detection



Introduction

What is ROI Align?

# Region of Interest (ROI) Align in Object Detection



Introduction

What is ROI Align?

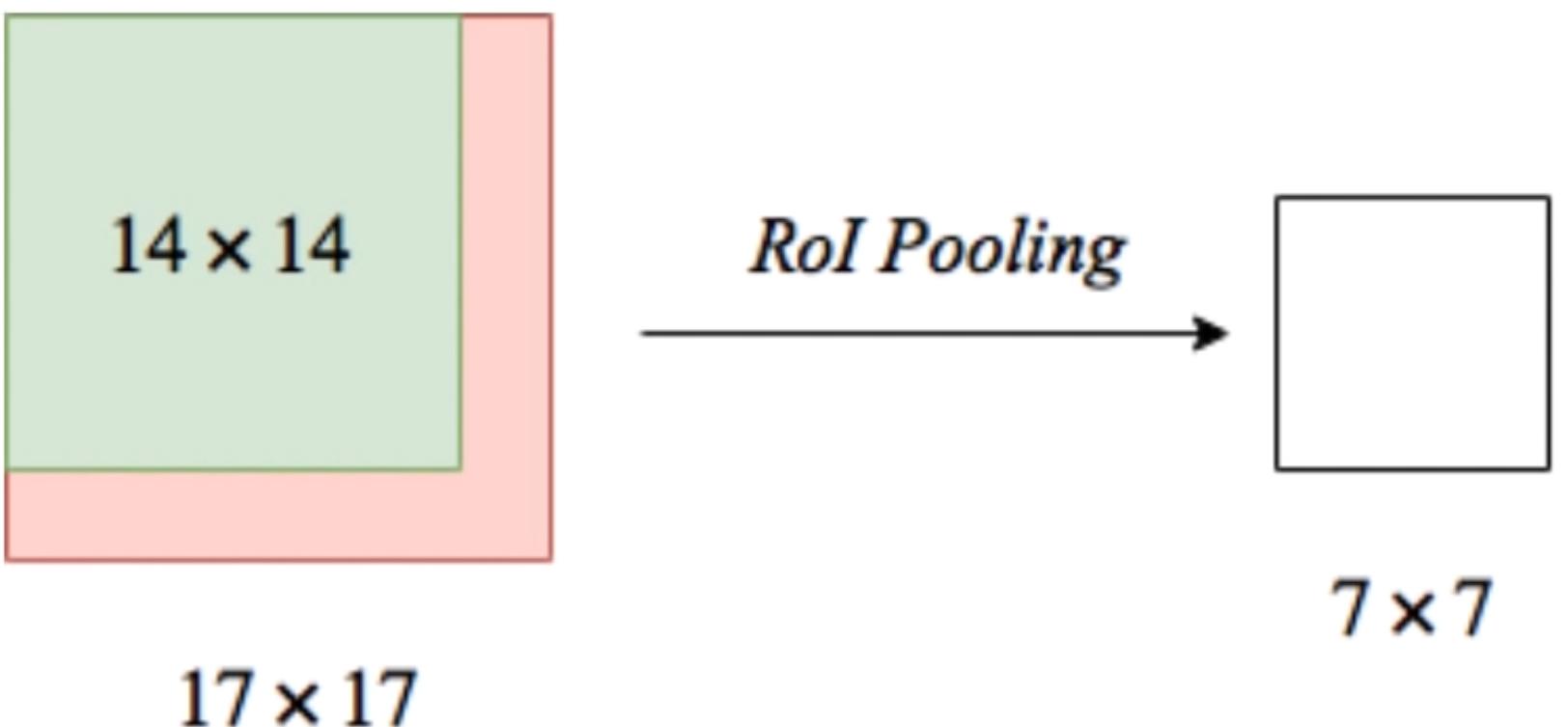
Problem Statement

## Prior Solution

*RoI Pooling*: Stride is quantized.

$$stride = \frac{17}{7} = 2.42$$

$$stride_{RoIPool} = [2.42] = 2$$

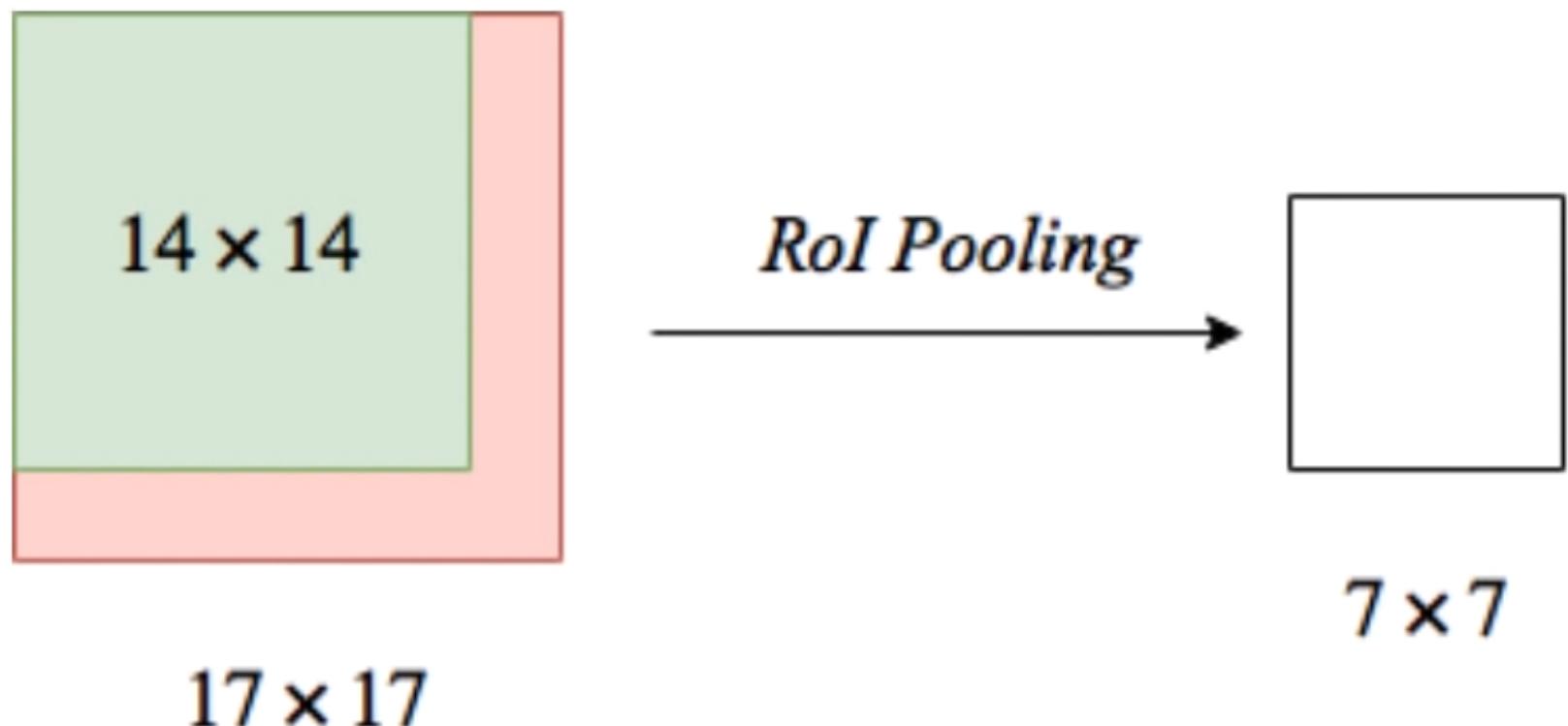


## Prior Solution

*RoI Pooling*: Stride is quantized.

$$stride = \frac{17}{7} = 2.42$$

$$stride_{RoIPool} = [2.42] = 2$$



## Pros

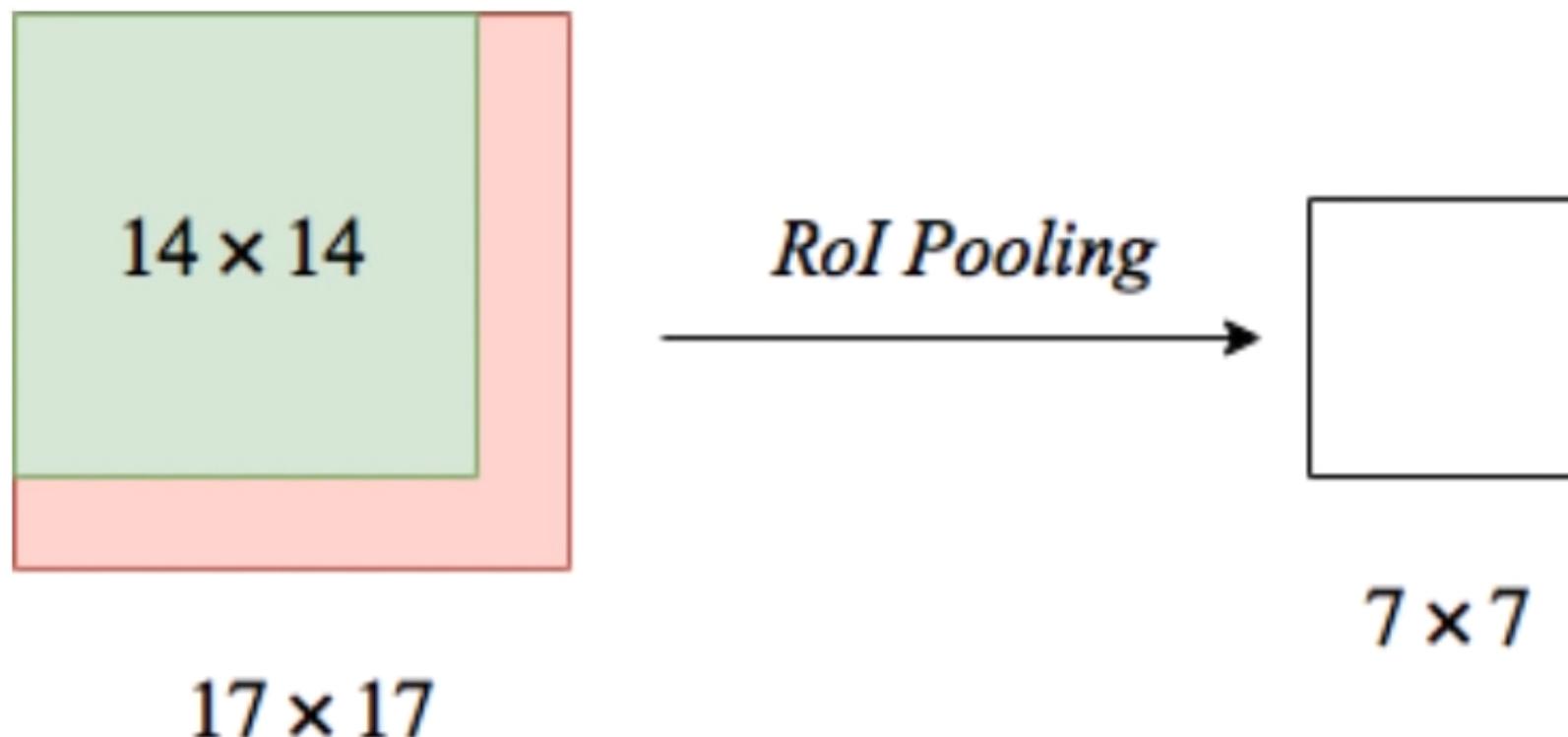
- Speed and Efficiency
- Integration with CNNs
- Simplicity

## Prior Solution

*RoI Pooling*: Stride is quantized.

$$stride = \frac{17}{7} = 2.42$$

$$stride_{RoIPool} = [2.42] = 2$$



## Pros

- Speed and Efficiency
- Integration with CNNs
- Simplicity

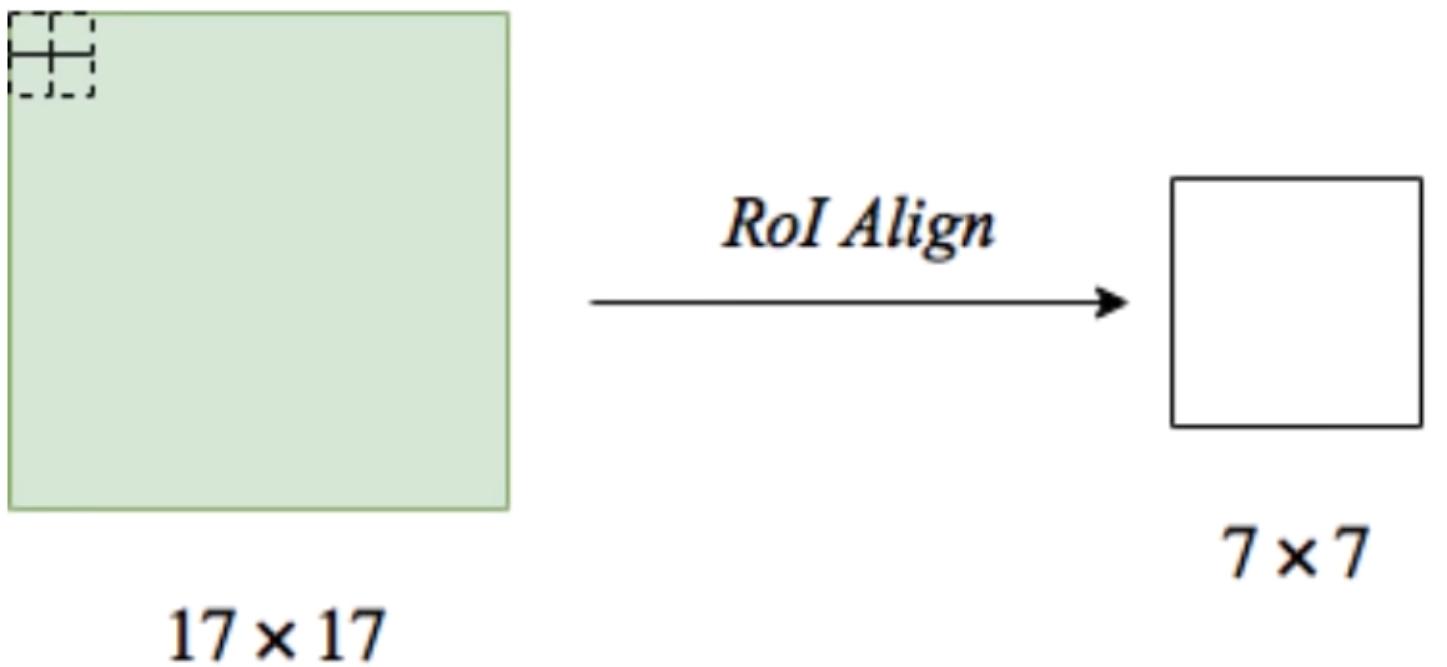
## Cons

- Quantization Error
- Fixed-size Constraint

## ROI Align Solution

*RoI Align*: Stride is *not* quantized.

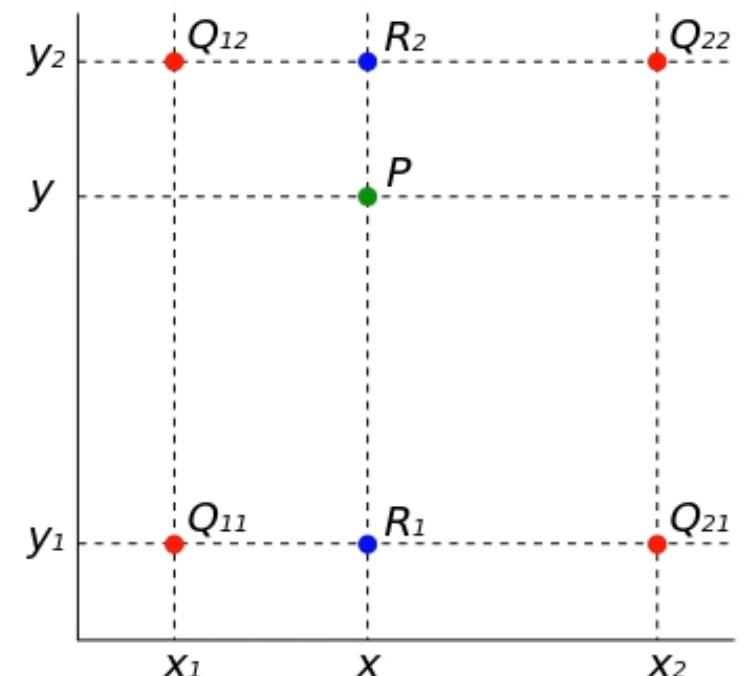
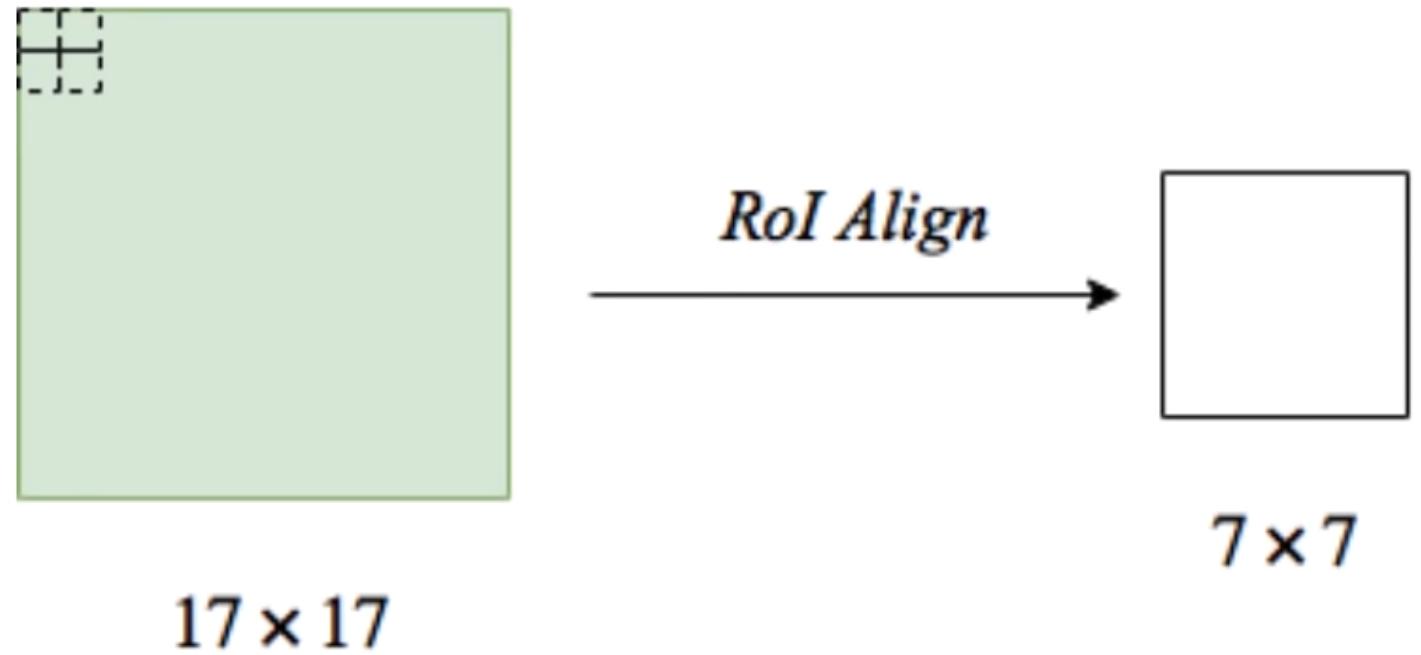
$$stride = \frac{17}{7} = 2.42$$



# ROI Align Solution

*RoI Align:* Stride is *not* quantized.

$$stride = \frac{17}{7} = 2.42$$



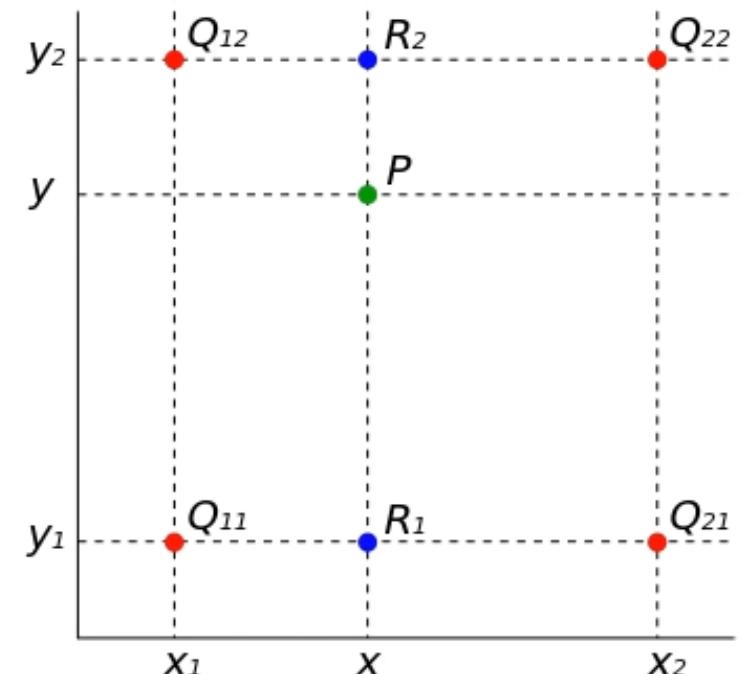
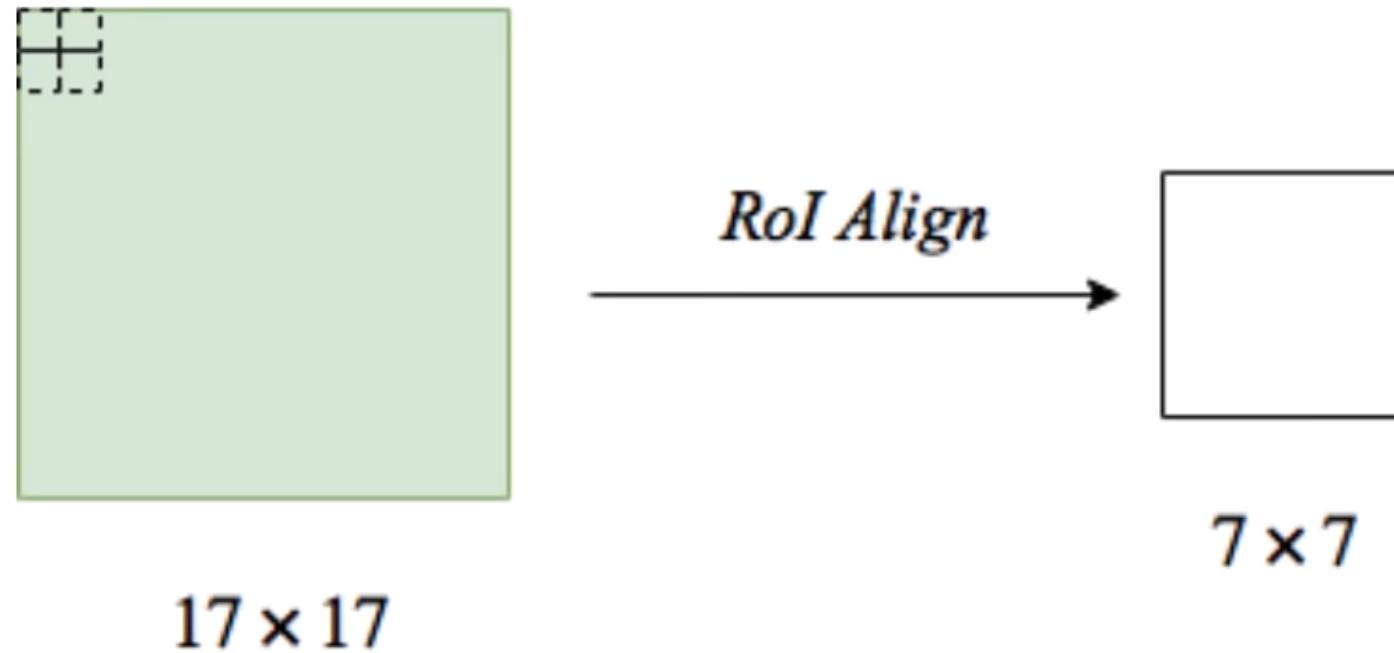
$$V(x, y) = \sum_{i=1}^2 \sum_{j=1}^2 V(x_i, y_j) \times (1 - |x - x_i|) \times (1 - |y - y_j|)$$

where  $V(x_i, y_j)$  is the feature value at point  $(x_i, y_j)$ .

# ROI Align Solution

*RoI Align:* Stride is *not* quantized.

$$stride = \frac{17}{7} = 2.42$$



$$V(x, y) = \sum_{i=1}^2 \sum_{j=1}^2 V(x_i, y_j) \times (1 - |x - x_i|) \times (1 - |y - y_j|)$$

where  $V(x_i, y_j)$  is the feature value at point  $(x_i, y_j)$ .

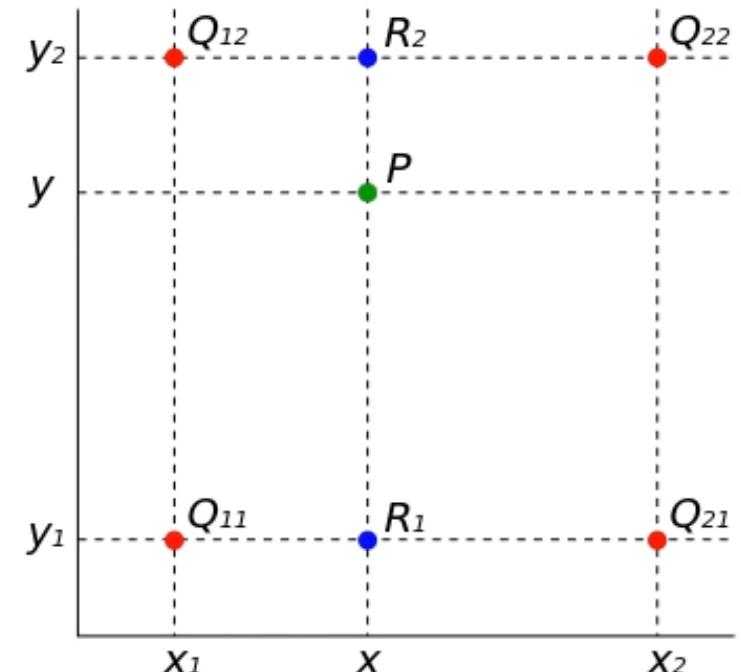
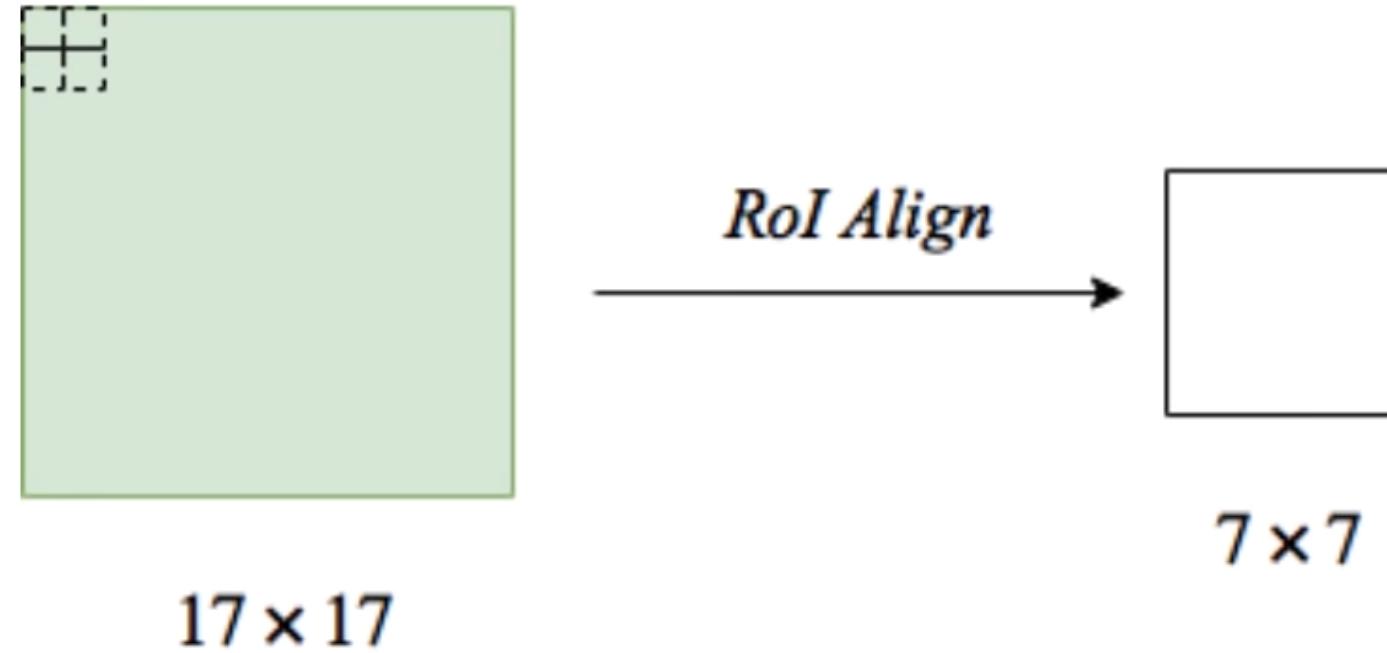
## Pros

- Improved Accuracy
- Flexible Size Handling
- Enhanced Performance

# ROI Align Solution

*RoI Align:* Stride is *not* quantized.

$$stride = \frac{17}{7} = 2.42$$



$$V(x, y) = \sum_{i=1}^2 \sum_{j=1}^2 V(x_i, y_j) \times (1 - |x - x_i|) \times (1 - |y - y_j|)$$

where  $V(x_i, y_j)$  is the feature value at point  $(x_i, y_j)$ .

## Pros

- Improved Accuracy
- Flexible Size Handling
- Enhanced Performance

## Cons

- Computational Complexity
- Implementation Complexity
- Hardware Demands

## Results

		bilinear?/>	agg.	AP	AP <sub>50</sub>	AP <sub>75</sub>
RoIPool			max	26.9	48.8	26.4
RoIWarp		✓	max	27.2	49.2	27.1
		✓	ave	27.1	48.9	27.1
RoIAlign	✓	✓	max	30.2	51.0	31.8
	✓	✓	ave	30.3	51.2	31.5

## Conclusion

The adoption of ROI Align in modern object detection frameworks, such as Mask R-CNN, underscores its effectiveness in achieving more accurate and reliable results.

Thank  
you

