

## 5. Show how to perform ROI Pooling and ROI Align by hand (with your own example)

RoI Pooling 을 구현하기 위해서는 다음의 2개의 input 이 필요하다.

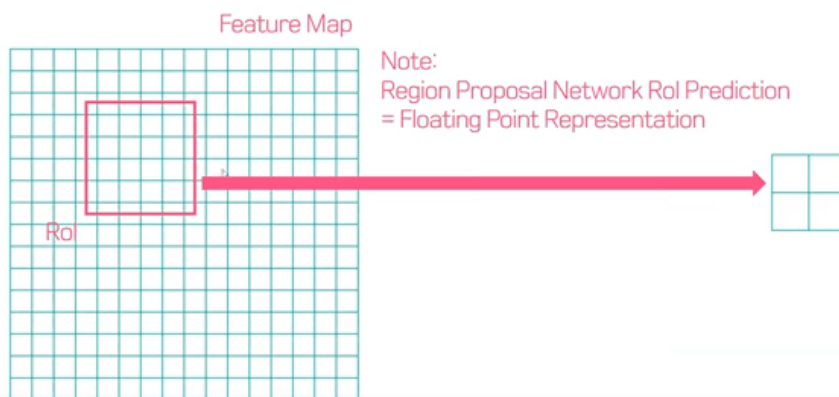
1. deep convolutions 그리고 max pooling layers 를 통해 나온 feature map
2.  $N \times 4$  매트릭스  $\rightarrow$   $N$  은 RoI 개수, 4 는 region 의 위치를 나타내는 coordinate

RoI 로직은 다음과 같다.

1. 각각의 region proposal 을 동일한 크기의 sections 으로 나눈다.
2. 각각의 section 마다 가장 큰 값을 찾는다
  1. 각각의 찾은 maximum 값을 output 으로 만든다.
  - 2.

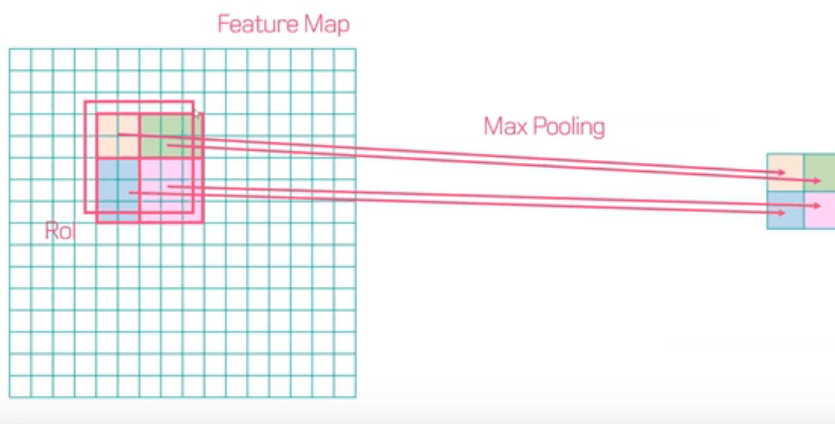
ROI Pooling : Region of Interest Pooling

### RoI Pooling



RPN 에서는 RoI 를 실수값으로 예측하는데 이것을 가지고 RoI Pooling 을 2\*2 로 뽑으려면 반올림해서 칸에 맞추고 Pooling 을 하게 된다.(정수 값의 좌표만을 가지고 있기 때문)

### RoI Pooling

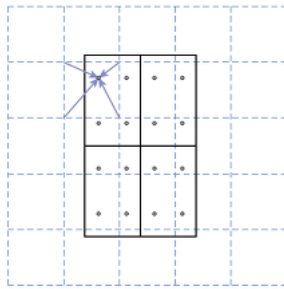


그러면 원본과는 약간 어긋난 상태(aligned 가 아닌 것)

ROI Align : Region of Interest Align

RoI 에서 얻어내고자 하는 정보는 박스 안의 동그라미 점

이미지 데이터는 정수인 좌표의 값만 갖고 있으므로 화살표로 표현된 방법을 통해 동그라미 점의 값을 구하겠다는 것이다.

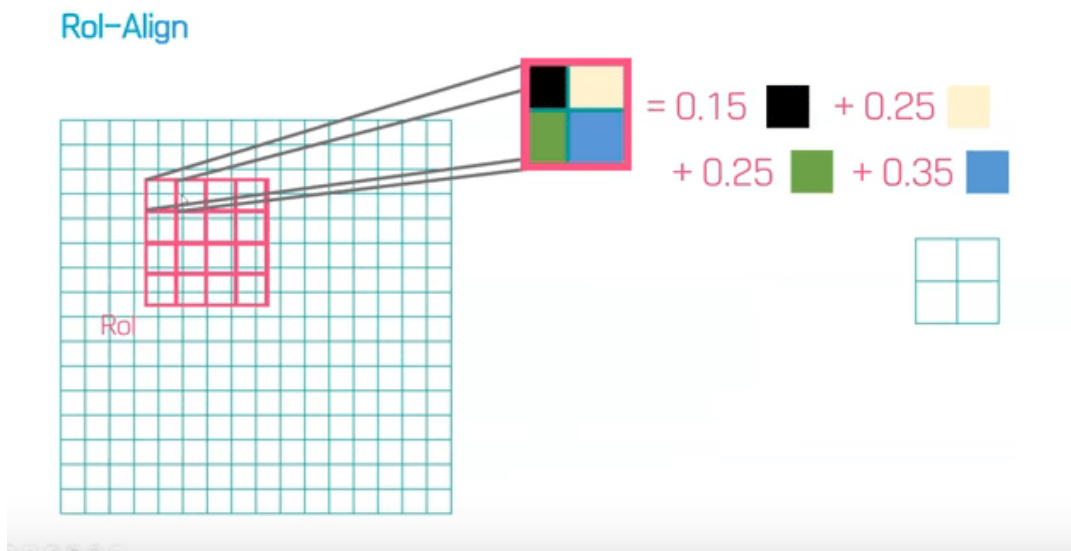


**Figure 3. RoIAlign:** The dashed grid represents a feature map, the solid lines an RoI (with  $2 \times 2$  bins in this example), and the dots the 4 sampling points in each bin. RoIAlign computes the value of each sampling point by bilinear interpolation from the nearby grid points on the feature map. No quantization is performed on any coordinates involved in the RoI, its bins, or the sampling points.

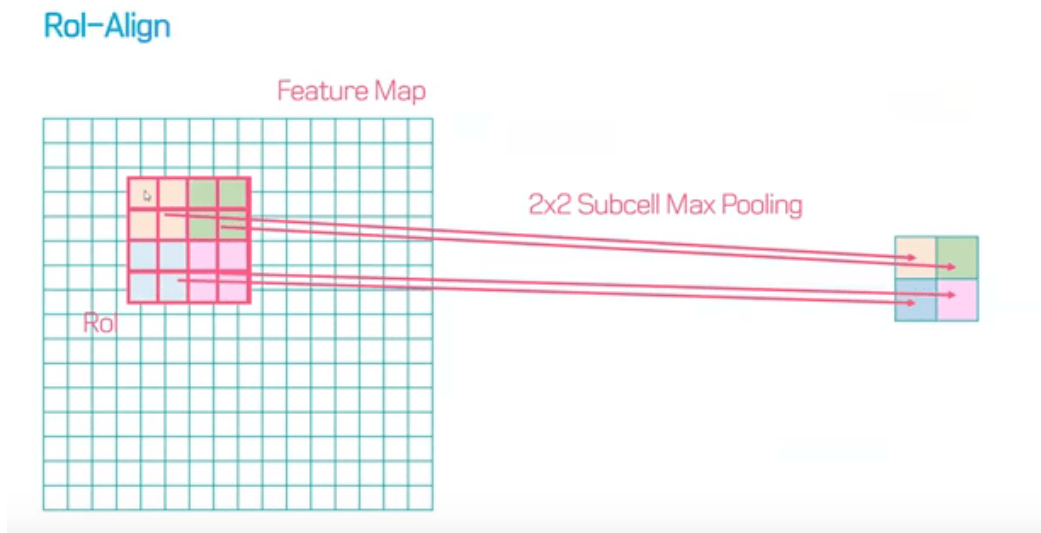
화살표가 의미하는 방법이 bilinear interpolation(이중 선형 보간) 이다.

4등분 한 것을 또 4등분(subset) 한다.

이에 대해 linear interpolation(선형 보간)을 통해 좀 더 정확한 값을 얻으려하는 방식이다.

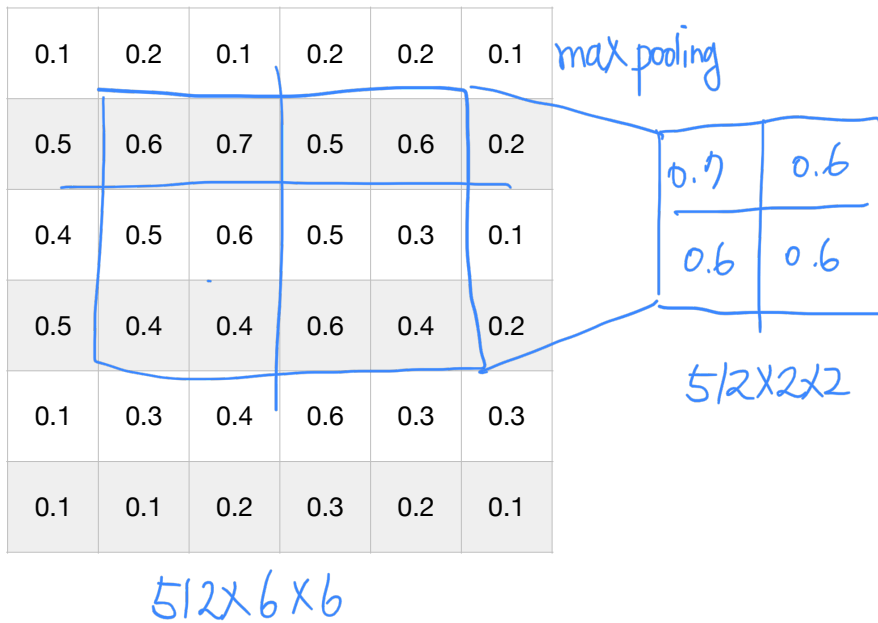


한 subset 을 놓고 보면 검은색은 0.15정도, 파란색 0.35, 나머지 0.25 정도라고 해서 weight sum 을 통해 한 개의 subset 의 값을 정한다.



이렇게 생긴 작은 subset 16개를 4개씩 max pooling 을 해서 2\*2 feature 를 뽑아내는 것이 RoI Align이다.  
 간단한 예를 통해 계산해보면 아래와 같다.

### - RoI Pooling



### - RoI Align

