



对于插值点的选取，作者没有给出统一的规则，在基于mxnet框架实现的ROI align代码中，采用将每个bin长和宽除以3，记为stride_h和stride_w，假设每个bin的起始位置和终点位置为bin_hstart, bin_hend, bin_wstart, bin_wend，则四个插值位置为(bin_hstart + stride_h, bin_wstart + stride_w), (bin_hstart + 2 * stride_h, bin_wstart + stride_w), (bin_hstart + stride_h, bin_wstart + 2 * stride_w), (bin_hstart + 2 * stride_h, bin_wstart + 2 * stride_w)；插值位置选定之后，采用周围最近的四个像素点进行双线性插值，最后再做max pooling输出。