1   Median Filtering

Recall that a convolutional kernel in image contexts involves calculating the sum of the elementwise products of the kernel with a window of pixels in the image. The kernel is then "slid" over the image, repeating this process for every pixel in the image.

(a) Consider the following 3x3 kernel:

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

This is often referred to as a "box-blur" kernel. Why does it have a blurring effect?

(b) Now consider a related technique, known as *median filtering*. Here, for a given window of pixels, the center pixel is replaced with the median of the pixels in the window. This process is then repeated across the entire image, just as is done with convolution.

For your curiosity, depicted below are the results of applying 3x3 filters to an image (original left, box-blur center, median right):
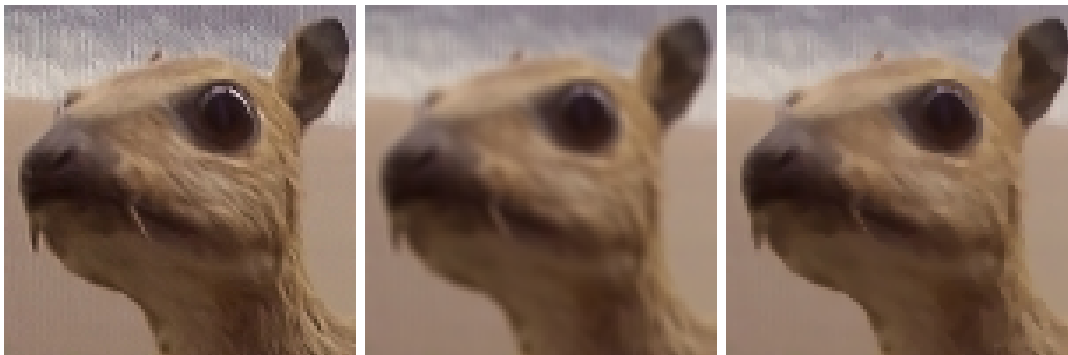
Can you think of a scenario where median filtering would be preferable to box-blur filtering? What about the other way around?

(c) For ConvNets, are median filters similar to convolutional layers? Are they useful? Why or why not?

(d) Let $K \times K$ denote filter size, P padding size, S stride length, and suppose we are operating on an image of size $H \times W \times C$. What is the computational complexity of convolution? How about median filtering? Assume no results are cached across different pixels.

Hint: We can find the median of n numbers in $O(n)$ time.

2   Cross-Entropy Investigation

hello