Capstone weekly report3

In the article titled "Sharpness-Aware Minimization for Efficiently Improving Generalization," the authors customized the function for adjusting image sharpness, along with other data augmentation techniques. This inspired me to experiment with filter coefficient adjustments in the kernel. Modifying these filter coefficients can have varying effects on images. For instance, consider the kernel [[1, 2, 3], [1, 6, 1], [3, 2, 1]], which likely combines edge enhancement and weighted averaging. This can result in sharper and more pronounced edges while slightly smoothing other areas of the image. The higher values in the center (6) and surrounding pixels (2, 3, 3, 2) tend to enhance edges and high-contrast regions. The kernel emphasizes differences in pixel values, a characteristic of edge-enhancement kernels. Additionally, this kernel is asymmetric and anisotropic, meaning it has a directional effect, particularly in the vertical and horizontal directions. Over the next two weeks, I plan to conduct further experiments with filter coefficients in the kernel.

In the same article, the authors normalized the kernel to maintain consistent brightness levels in the images. Brightness is a crucial feature for food images since proper lighting can accurately capture color and texture, aiding in food distinction.

To refresh my understanding of convolutional neural networks, activation functions, backpropagation, and related concepts, I read several articles. This knowledge has been instrumental in designing the model architecture for this project, including layer count, customized data augmentation functions, and potential activation functions. Over the next two weeks, I will continue working on different components of the architecture.

Key Takeaways:

Sigmoid and hyperbolic tangent activation functions should be avoided in deep networks due to the vanishing gradient problem.

Rectified linear activation functions effectively mitigate the vanishing gradient problem, leading to faster learning and improved performance.

Rectified linear activation is the default choice for developing multilayer perceptron and convolutional neural networks.

For convolutional operations, input images must have a four-dimensional shape: [batch size, height, width, channels].

related reading:

1. <https://machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks/>
2. <https://en.wikipedia.org/wiki/Activation_function>
3. <https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939>
4. Sharpness-Aware Minimization for Efficiently Improving Generalization <https://arxiv.org/pdf/2010.01412v3.pdf>. SHARPNESS-AWARE MINIMIZATION FOR EFFICIENTLY IMPROVING GENERALIZATION
5. Neural Networks Pt.2: Backpropagation Main Ideas. <https://www.youtube.com/watch?v=IN2XmBhILt4&t=7s>
6. <https://towardsdatascience.com/understanding-backpropagation-algorithm-7bb3aa2f95fd>