note of 《EffConv: Efficient Learning of Kernel Sizes for Convolution Layers of CNNs》

zxp

August 17, 2024

1 content

《EffConv: Efficient Learning of Kernel Sizes for Convolution Layers of CNNs¿》 addresses the kernel sizes of each layer in CNNs, where different sizes of kernels have varying effects and applications on computational power, significantly impacting performance. Initially, the kernel sizes were fixed upon network design, but different scenarios require different kernel sizes. Conventional methods for determining suitable kernel sizes rely on extensive resource searches or heuristic algorithms, which either require many resources or have poor results. The method proposed in the paper involves designing a predictive model that can perform size and kernel predictions. The experimental metrics are speed and accuracy under different parameter scales.

2 Feelings

This paper mainly discusses a method for quickly finding the appropriate sizes of convolutional kernel for every layers during network operation. which I didn't quite understand. It is related to the entire CNN network and involves pruning, but is not very related to the part we've been doing on. This paper is very concise and succinct; the core description is just a page, yet the experimental data section is quite substantial.