CONVOLUTION AND POOLING OPERATION MODULE WITH ADAPTIVE STRIDE PROCESSING EFFECT

Anonymous authors
Paper under double-blind review

ABSTRACT

Convolutional neural network is one of the representative models of deep learning, which has a wide range of applications. Convolution and pooling are two key operations in convolutional neural networks. They play an important role in extracting input features and mapping low-level semantic features to high-level semantic features. Stride is an important parameter involved in convolution and pooling operations, which refers to the distance of each slide of the convolution kernel (pooling kernel) during the convolution (pooling) operation. The stride has an impact on the granularity of feature extraction and the selection (filtering) of features, thus affecting the performance of convolutional neural networks. At present, in the training of convolutional neural networks, the content of convolution kernel and pooling kernel can be determined by the optimization algorithm based on gradient descent. However, the stride usually cannot be treated similarly, and can only be selected manually as a hyperparameter. Most of the existing related works choose a fixed stride, for example, the value is 1. In fact, different tasks or inputs may require different stride for better model processing. Therefore, this paper views the role of stride in convolution and pooling operation from the perspective of sampling, and proposes a convolution and pooling operation module with adaptive stride processing effect. The feature of the proposed module is that the feature map finally obtained by convolution or pooling operation is no longer limited to equal interval downsampling (feature extraction) according to a fixed stride, but adaptively extracted according to the changes of input features. We apply the proposed module on many convolutional neural network models, including VGG, Alexnet and MobileNet for image classification, YOLOX-S for object detection, Unet for image segmentation, and so on. Simulation results show that the proposed module can effectively improve the performance of existing models.