Dorian PILLARD

Technical exercise

Part 2 : paper summary

I chose to write about this paper, which discusses the replacement of convolution operations with morphological functions in deep neural networks (DNN): https://ieeexplore.ieee.org/abstract/document/9512058.

The goal of this paper was to demonstrate proof of concept that DNNs could be implemented using mathematical morphology (MM), which uses non-linear operations such as erosions and dilations. This way engineers would be able to produce state-of-the-art solutions for some of the more difficult problems in computer vision which cannot be solved using convolutional neural networks (CNN), such as pruning certain objects of certain sizes in images.

After a brief introduction to some of the more common techniques of MM such as dilations and openings, the authors explain how MM has already been used in conjunction with CNNs to preprocess images and extract preliminary features to be fed to the network, features that couldn't be extracted using convolution. Some other works further enmeshed their networks with MM operators, to produce good results compared to the literature, either by outperforming them or by achieving similar results but with fewer parameters to learn. However, the authors of this paper believe they were able to go even further: in essence they were able to produce networks that optimized the structuring elements to perform a broader range of MM operations with a 'flat' structuring element, meaning the network would not be limited to grayscale images.

Replacing the convolution operations with MM operations is not without its drawbacks however: due to the non-linearity of the MM operations, applying the back-propagation algorithm is not possible because calculating the partial derivatives of the operation is not easy. To surmount this problem, the authors offer to combine pointwise and depthwise convolution with depthwise pooling to emulate and optimize MM operations. The depthwise convolution and pooling combined define a single processing unit and allow for the determination of the structuring element and applying erosion or dilation operations. Because of that, the number of potential outcomes gets very high, which is why they use the pointwise convolution to keep only one output after each processing unit.

Based on this initial processing unit, the authors were then able to recreate the more elaborate operations, like opening and closing units, top-hat units, or geodesic reconstruction units. This last unit is the only one that is approximated, due to the iterative nature of the process which is difficult to recreate in networks. As such, only one transformation is performed on the marker image. The main differences between this approach and the traditional CNN are the absence of activation functions in the authors' morphological network, and its increased generalization capacity.

Because they used operations already present in other existing methods, optimizing the weights of their network is easily done by employing traditional deep learning-based approaches like feedforward, backpropagation, and stochastic gradient descent.

In the final part of their article, they experiment with their morphological neuron, after going over the dataset used for their evaluation and their experimental protocol. On the synthetic image dataset, their DeepMorphNet achieved 100% average accuracy, outperforming the other two reference networks. On the image classification datasets, the DeepMorphNet, average accuracy was once again higher than other convolutional networks found in literature. However, the training time and number of parameters was also higher than the CNNs. Finally, the DeepMorphNet outperformed the ConvNet once again on the pixel classification dataset, despite having more learnable parameters and a longer training time.

Such results are proof of concept that morphological operations can replace convolutions in deep neural networks, as they allowed for better results on all the tasks presented in the paper. However, it does raise some new issues such as the high number of parameters, which must be tackled next to consider more complex morphological architectures.

This publication was interesting to me, because I was introduced to image processing through mathematical morphology and completed multiple projects using techniques defined by the field. As such, I was somewhat disappointed to leave it all behind when I started learning about and working with convolutional networks. Seeing as they are widely acknowledged as state-of-the-art in most of today's problems, I feared I would never come back to mathematical morphology. Nevertheless, I always wondered whether CNNs could be coupled with other image processing techniques that I feel have been sidelined in the past years, such as mathematical morphology, SIFT descriptors, or optimization algorithms like those found in the field of metaheuristics for example. Hearing about morphological deep neural networks, which challenge the status quo, renewed my interest for the field of mathematical morphology and inspired me to think outside the box more than previously. It made me understand that we should always try new approaches, because we can always improve our methods and results with the right tools.