A Survey of Pruning Methods for Efficient Person

Re-identification Across Domains

**Abstract**

In recent times we are witnessing tremendous increase in deep learning architectures proposed for vision-based recognition such as person re-identification, in which people are identified with help of distributed shots on several cameras. This paper provides a survey of state-of-the-art pruning techniques that are suitable for compressing deep Siamese networks applied to person re-identification. The computational complexity of CNNs hinders the deployment of Deep Siamese networks on platforms with lesser resource, though they’ve improved accuracy, but cannot be used in applications with real time data constraints, and thus we can compress these without losing accuracy. There are various techniques which could be effective for the compression of the networks which are analysed and compared based on their strategy and pruning criterion, in different design scenarios fine-tuning networks by applying pruning methods for targeted applications. Pruning can drastically reduce the complexities in network according to experimental outcomes from Siamese networks with ResNet feature extractors and keeping track of good accuracy.

Recent years have witnessed a substantial increase in the deep learning architectures proposed for visual recognition tasks like person re-identification, where individuals must be recognized over multiple distributed cameras. Although deep Siamese networks have greatly improved the state-of-the-art accuracy, the computational complexity of the CNNs used for feature extraction remains an issue, hindering their deployment on platforms with limited resources, or in applications with real-time constraints. Thus, there is an obvious advantage to compressing these architectures without significantly decreasing their accuracy. This paper provides a survey of state-of-the-art pruning techniques that are suitable for compressing deep Siamese networks applied to person re-identification. These techniques are analysed according to their pruning criteria and strategy, and according to different design scenarios for exploiting pruning methods to fine-tuning networks for target applications. Experimental results obtained using Siamese networks with ResNet feature extractors, and multiple benchmarks re-identification datasets, indicate that pruning can considerably reduce network complexity while maintaining a high level of accuracy.

In scenarios where pruning is performed with large pre-training or fine-tuning datasets, the number of FLOPS required by the ResNet feature extractor is reduced by half, while maintaining a comparable rank-1 accuracy (within 1% of the original model). Pruning while training a larger CNNs can also provide a significantly better performance than fine-tuning smaller ones. Keywords: Deep Learning, Convolutional Neural Networks, Siamese Networks, Complexity, Pruning, Domain Adaptation, Person Re-identification.