

عنوان مقاله:

Improving the performance of deep neural networks by clustering

محل انتشار:

دومین کنفرانس ملی محاسبات نرم (سال: 1396)

تعداد صفحات اصل مقاله: 9

نویسندگان:

Saeedeh Khaleghi - *Masters student of Algorithms and Computation, University of Tehran, School of Engineering Science, Department of Algorithms and Computation, Tehran, Iran*

Mahmood Shabankhah - *Assistant Professor, University of Tehran, School of Engineering Science, Department of Algorithms and Computation, Tehran, Iran*

خلاصه مقاله:

Neural networks applications have been widely employed in many of today's intelligent systems, especially in embedded systems and smart phones. This trend brings with itself many challenges, one of which is the effective handling of large numbers of parameters given the limited size of available memory in such systems. Indeed, networks with high number of adjustable parameters are generally prone to overfitting. A common approach in such cases would be to downsize the network. In this paper, we use clustering to meet this goal. Indeed, we cluster the weights in each layer into a smaller number of groups. In parallel, the network architecture is accordingly modified. The main part of the training procedure is done on this smaller network. Therefore, the training time will be greatly improved. Finally, the initial network's weights will eventually be reconstructed from the weights obtained in the last step. The most important feature of our approach is that although the training is faster but the performance of the network does not suffer

کلمات کلیدی:

Deep neural networks, K-means clustering, learning speed, overfitting

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/696740>

