The Study of Hardware Resources Usage on Examples of Convolutional Neural Network Implementation

Dmitry Pertsau and Mikhail Tatur

Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus, pertsev@bsuir.by

Keywords: CNN, LeNet-5, hardware implementation

Abstract. Deep learning convolutional neural networks marked another scientific breakthrough in the field of artificial intelligence. Following scientific achievements, technological solutions appeared both software and hardware, which were widely used in solving applied problems. Prominent representatives of software technologies are such projects as TensorFlow, PyTorch and Keras, among hardware projects - the development of GPU, the emergence of the Google Tensor Processing Unit and many other startups. On the whole, their use makes it possible to program quickly and parallelize the computational process on a modern hardware platform, and it is relatively easy to obtain satisfactory complete solutions. However, working at the "top level of technology", the developer is deprived of the opportunity to fully manage key resources, such as performance, hardware costs, etc. Also, software developer often cannot develop original hardware implementations of neural network architectures using FPGA technologies. Software technologies such as NVIDIA CUDA allow you to manage and control the computing process. However, this approach requires a highly qualified developer to work with this technology. In the paper, we will try to demonstrate how it's possible to manage the available resources by using various models of hardware design. In our conducting discussion, we were using the well-known convolutional network LeNet-5. In the paper, we also discuss the challenges and opportunities of introduced models.