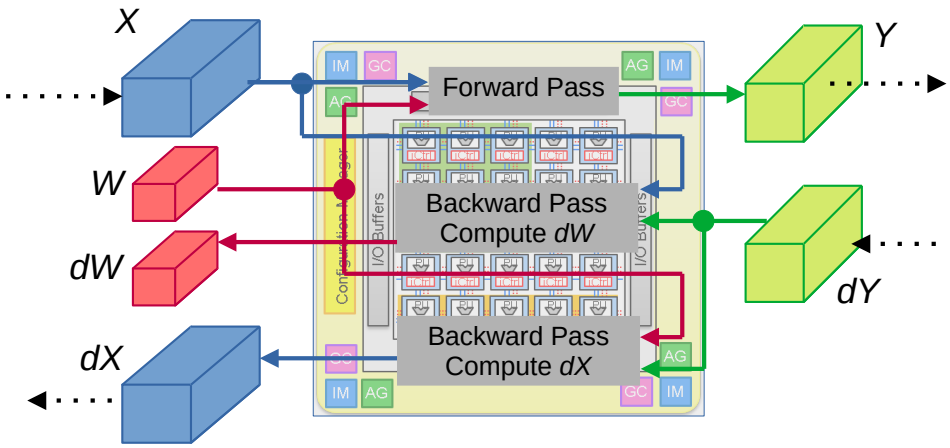
# Training of DNNs on Embedded Processor Arrays

An open application challenge in the Internet-of-Things is on-chip training of Deep Neural Networks (DNNs). For achieving a high energy efficiency, embedded accelerators are required. Tightly Coupled Processor Arrays (TCPAs) are a class of highly customizable coarse-grained reconfigurable architectures used as hardware accelerators perfectly enable fast and low-power acceleration of DNNs.




In the figure, a simplified data flow of a single layer of a DNN is shown. During the forward pass, the input tensor *X* (blue) is multiplied and accumulated with the trainable weight tensor *W* (red) and the resulting output tensor *Y* (green) is produced. During the backward pass (training), the gradients tensor *dY* (green) and the stored activation feature map *X* compute the weight updates *dW* (red).

In this work, a solution for training of DNNs on TCPAs shall be explored and implemented. It is to be investigated how much resources and time the training on an exemplary TCPA architecture requires for exemplary DNN models from the TinyML benchmark suite.

| | |
|---|---|
| Requirements: | Knowledge in C++, Python and Neural Networks |
| Type of thesis: | Theory (20%), concept (40%), implementation (40%) |
| Supervisor: | Christian Heidorn (Christian.Heidorn@fau.de) |


Lehrstuhl für Informatik 12
**Hardware-Software-Co-Design**
Cauerstraße 11
91058 Erlangen

Friedrich-Alexander-Universität
Technische Fakultät