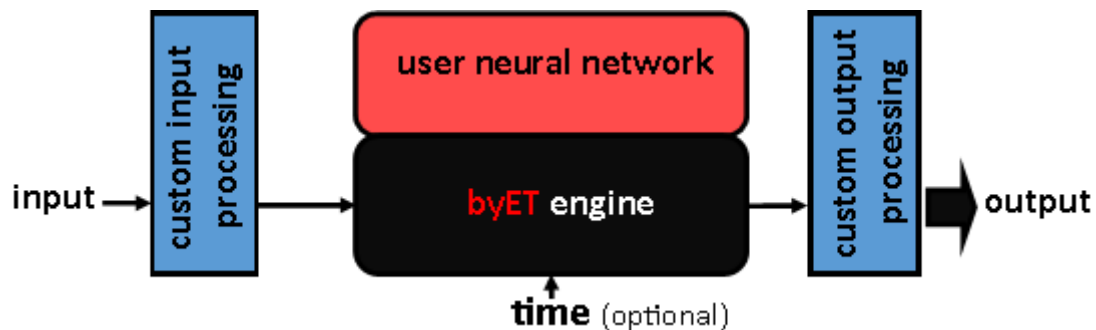


byET-engine

In recent years there has been big advances in Artificial Intelligence (AI) and Machine Learning (ML). Deep neural networks (DNNs) have been shown to successfully solve many difficult problems. DNNs offer the capability of learning any unknown function when there is sufficient data showing the expected outcome for a given input. DNNs can be trained using this data and be able to give correct answers when presented with new, never seen before, data.

These capabilities are also desirable in the embedded field, for example, to develop intelligent sensors. However, the large computation capabilities necessary are a limiting factor. A possible solution is to acquire the data on the embedded device and relay it for processing to a cloud service, but this comes with connectivity costs and reliability questions, increased latency and privacy issues.



byET-engine is a deep neural network engine designed for embedded systems. It is lightweight and fast, making efficient use of the limited resources available. The engine itself uses only a few kilobytes of memory and employs different techniques to reduce the footprint of the deployed DNN. It requires no external libraries and can run directly on bare-metal hardware, an RTOS or a full OS. It can be used to deploy DNNs to deeply embedded systems like ARM Cortex M, Renesas RX, Microchip dsPIC and others.

byET-engine uses DNNs directly avoiding load overhead and duplication of data on the target. The main body of the software is written in portable C, with platform specific customizations in C and assembler to use the available hardware acceleration capabilities. It supports many neural network architectures: FC, CNN, LSTM, GRU. The input and output are processed using other techniques, to reduce the amount of resources necessary to achieve the desired results.

A DNN can be trained on a server with byET-engine, where it will use the CPU or GPU if available, or developed and trained using one of the popular deep learning frameworks and later transformed to the byET-engine format for deployment. byET-engine can be used for inference only or for continued training of a pre-trained DNN when deployed on the embedded device.

