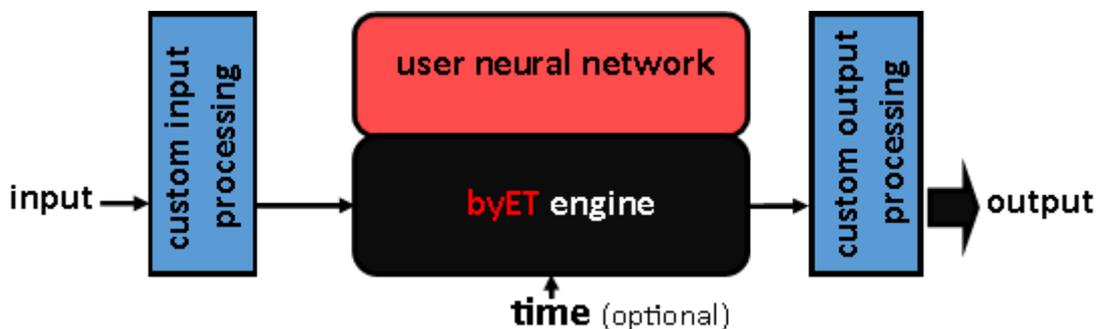


byET-engine

In den vergangenen Jahren sind große Fortschritte im Bereich künstlicher Intelligenz (Artificial Intelligence AI) und maschinellem Lernen (Machine Learning ML) erzielt worden. Mittels Deep Neural Networks (DNNs) ist es gelungen viele komplexe Problemstellungen aufzulösen. Wenn genug Datenmaterial vorhanden ist, welches für ein erwartetes Ergebnis in Bezug auf einen gegebenen Eingangswert steht, ermöglichen diese neuronalen Netze die Fähigkeit bis dahin unbekannte Funktionen zu erlernen. Diese DNNs können derart trainiert werden, dass auch bei neuem, unbekanntem Input korrekte Antworten erzielt werden.

Diese Fähigkeiten sind natürlich auch im embedded Bereich wünschenswert, z.B. um intelligente Sensoren zu entwickeln. Einschränkender Faktor ist aber die hierfür notwendige Rechenleistung. Ein möglicher Ansatz wäre die Eingangsdaten mit dem embedded Gerät nur aufzunehmen und zur Auswertung an einen Cloud-Service weiterzugeben. Dies erhöht aber die Kosten und wirft Fragen bzgl. der Verlässlichkeit der Verbindung, höherer Latenzzeiten und natürlich der Datensicherheit auf.



byEt-engine ist eine DNN-Engine speziell für embedded Systeme. Sie ist schnell, sehr klein im Codeumfang und nutzt äußerst effizient die verfügbaren Ressourcen. Die Engine selbst benötigt nur wenige Kilobytes und verwendet verschiedenste Techniken um die Codegröße des ausgewählten DNN zu reduzieren. Zudem sind keine externen Bibliotheken notwendig. Lauffähig ist byET-engine auf Systemen ohne OS, mit einem RTOS oder einem herkömmlichen Betriebssystem. byET-engine kann auf embedded Systemen wie z.B. ARM Cortex, Renesas RX, Microchip dsPIC und vielen anderen eingesetzt werden.

Um den Overhead und das Duplizieren von Daten auf dem Zielsystem zu vermeiden, werden vorhandene DNNs von unserer byET-engine ohne Umweg verwendet. Der größte Teil der zugrundeliegenden Software ist generisch und in der Hochsprache C geschrieben, zusätzlich gibt plattformspezifische Anteile in C und Assembler, um die vorhandene Hardwarebeschleunigung ausnutzen zu können. byET-engine unterstützt verschiedenste Architekturen neuronaler Netzwerke: FC, CNN, LSTM, GRU. Um die Ressourcen auf das notwendige Maß für das gewünschte Ergebnis zu reduzieren, wird die Ein- und Ausgabe unter Verwendung anderer Techniken verarbeitet.

Das Training eines DNN kann mit der byET-engine sowohl auf einem Server erfolgen, hierbei wird entweder die CPU oder falls vorhanden die GPU verwendet, als auch unter Verwendung populärer deep-learning Frameworks und späterer Umwandlung in das byET-Format. Unsere byET-engine ist geeignet für den Einsatz als reine Inferenz oder zur Laufzeit für das fortgesetzte Training eines vortrainierten DNN.

