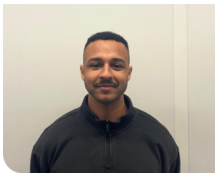
Rauschunterdrückung mit Neuronalen Netzwerken auf Mikrocontrollern

Studenten

Maurin Doswald



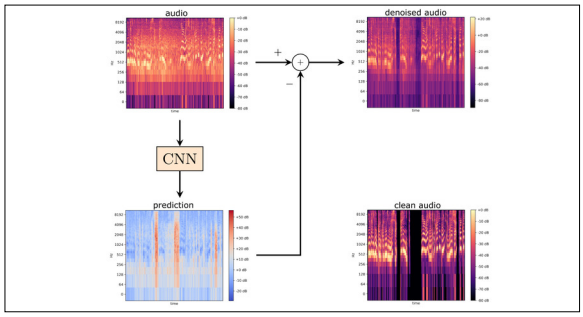
Stephan Oseghale

Einleitung: Neuronale Netzwerke sind heutzutage vor allem für den Einsatz auf leistungsstarken Servern oder spezialisierten Hardwarekomponenten wie GPUs bekannt. Der Bedarf, diese leistungsfähigen Technologien auch auf ressourcenarmen Systemen wie Mikrocontrollern nutzbar zu machen, wächst allerdings stetig, insbesondere in Anwendungen wie der Signalverarbeitung.
Das Ziel dieser Arbeit war es, ein System zu entwickeln, das mithilfe eines Neuronalen Netzwerks Störgeräusche in Audiosignalen in Echtzeit unterdrückt und auf einem ressourcenarmen Mikrocontroller betrieben werden kann.

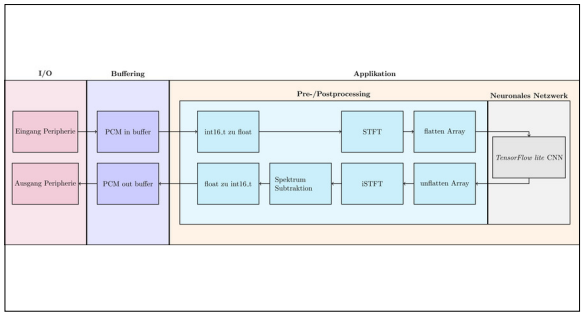
Vorgehen: Zu Beginn wurden verschiedene Ansätze für das Neuronale Netzwerk untersucht. Der aufgrund der Vorgaben vielversprechendste Ansatz wurde weiterverfolgt. Dabei wird das Audiosignal mithilfe der Short-Time Fourier Transform (STFT) in ein Spektrogramm umgewandelt, wodurch eine Darstellung im Zeit- und Frequenzbereich entsteht. Die Netzwerkarchitektur basiert auf einem Convolutional Neural Network (CNN) in U-Net-Struktur. Das Netzwerk wurde so trainiert, dass es die Störgeräusche vorhersagt, anschliessend wird das vorhergesagte Spektrum vom originalen subtrahiert und das resultierende Signal ausgegeben.
Die Herausforderung beim Entwickeln des Systems bestand darin, das funktionierende Netzwerk zu skalieren, sodass es auf dem Mikrocontroller betrieben und die Echtzeitanforderung gewährleistet werden kann. Zunächst wurde eine für den Mikrocontroller optimierte STFT Implementation entwickelt. Das trainierte Neuronale Netzwerk wurde quantisiert, dann in ein Byte-Array exportiert und mittels TensorFlow Lite Micro in das System integriert.
Die Daten werden asynchron über Interrupts und Direct Memory Access (DMA) ein- und ausgegeben, damit die Verarbeitung «quasi parallel» erfolgt. Verschiedene Eingangs- und Ausgangsoptionen wurden im System integriert, darunter die USB Audio Class, ein PDM-MEMS-Mikrofon und ein Line-In/-Out-Anschluss über ein I2S-Board.

Fazit: In dieser Arbeit konnte gezeigt werden, dass es möglich ist, ein CNN zusammen mit Pre-/Postprocessing auf einem STM32H753ZI Mikrocontroller, welcher mit einen Arm Cortex-M7 ausgestattet ist, mit diversen Eingangs- und Ausgangsoptionen in Echtzeit zu betreiben. Das Netzwerk wurde entsprechend seiner Ressourcenbeschränkungen modifiziert und skaliert, benötigt jedoch weitere Anpassungen, Optimierungen und eine Vielzahl an spezifischen Daten, um bessere Ergebnisse zu erreichen.
Das System wurde so konzipiert, dass das Modell bzw. das Netzwerk sehr einfach austauschbar ist, um flexibler auf Optimierungen, neue Anforderungen und Technologien reagieren zu können.

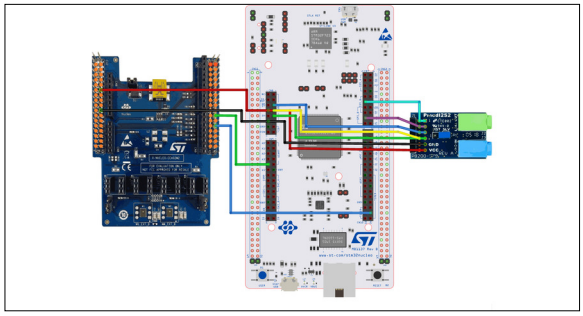
Schätzung des Netzwerks
Eigene Darstellung



Signalfluss
Eigene Darstellung



Eingesetzte Hardware
Eigene Darstellung



Referenten
Prof. Reto Bonderer, Noah Kälin

Themengebiet
Embedded Software Engineering

Projektpartner
ErvoCom AG, Feusisberg, SZ