

Speech Enhancement based on Deep Neural Networks

Bei der Sprachsignalverarbeitung, insbesondere für Hörbeeinträchtigte, ist es wichtig, die Sprache möglichst von Störgeräuschen (Noise) zu befreien. In dieser Bachelorarbeit wurde das Problem der Störgeräusch-unterdrückung mithilfe von Deep Learning angegangen. Dazu diente eine aktuelle Speech Enhancement-Methode als Vorlage, dessen Architektur so modifiziert wurde, um das bestmögliche Resultat mit den zur Verfügung stehenden Ressourcen und Methoden zu erreichen.

Das Training des Neuronalen Netzes sowie die Aufbereitung des Datensatzes erfolgte in Python auf dem GPU-Cluster der ZHAW. Der DSP-Algorithmus und das trainierte Netz wurden anschliessend in eine C++-Implementation eingebettet. Diese benötigt lediglich 2 % der vorhandenen Rechenleistung auf einem Intel x86 Desktop-Computer. Somit können Audiofiles, wie auch Live-Audio, in Echtzeit verarbeitet und wiedergegeben werden.

Das Ergebnis dieser Bachelorarbeit bietet die Möglichkeit, die Intensität der Geräuschunterdrückung mittels Parameter in Echtzeit zu steuern (siehe Abbildung). Zudem bildet es dank schlankem und flexiblen Design eine gute Grundlage für zukünftige Erweiterungen und demonstriert, dass Machine Learning, insbesondere Deep Learning, auch in Kombination mit klassischen Signalverarbeitungsalgorithmen ein guter Ansatz zur Geräuschunterdrückung in Sprache ist.

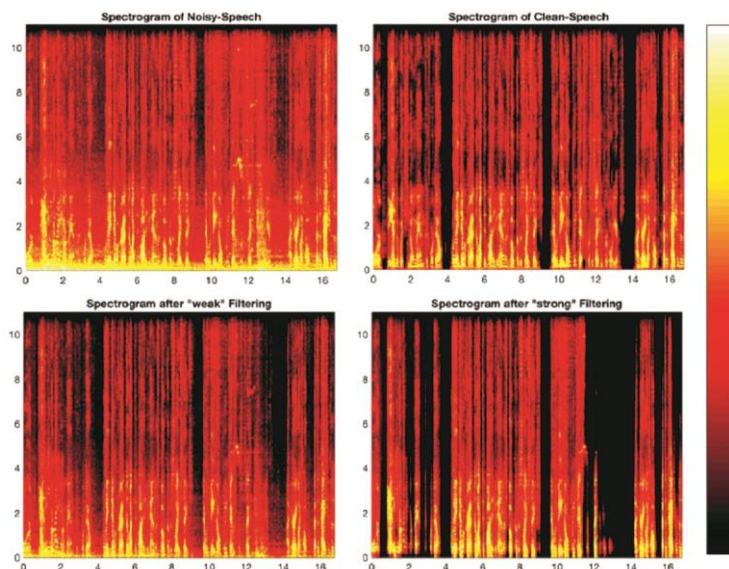


Diplomierende

Nino Blumer
Bastian Frederik Fischer

Dozent

Sigisbert Wyrtsch



Die Abbildung zeigt vier Spektrogramme einer Männerstimme mit starken Hintergrundgeräuschen. Die X-Achse ist die Zeit (in Sekunden) und die Y-Achse zeigt die Frequenz von 0 bis 11.025 kHz. Die Leistung des Spektrogrammes ist so eingefärbt, dass ein starkes Signal gelb und ein schwaches Signal schwarz ist. Oben links: Spektrogramm der verrauschten Männerstimme; Oben rechts: Spektrogramm nur der Männerstimme; Unten links: Spektrogramm mit schwächerer Geräuschunterdrückung; Unten rechts: Spektrogramm mit stärkerer Geräuschunterdrückung