

Instrumentenerkennung in Single-Source Audio Streams

Aufgrund der heutigen Rechenmöglichkeiten wurden in den letzten Jahren im Forschungsgebiet Deep Learning, getrieben durch Fortschritte in der Grafikprozessorhardware (GPU), bahnbrechende Ergebnisse in der Bildererkennung und verwandten Bereichen erzielt.

Im Bereich Music Information Retrieval wurden erste Arbeiten veröffentlicht, welche Deep-Learning-Ansätze für die Klassifikation von Musikinstrumenten einsetzen. Klassische Zweischrittverfahren, welche durch die manuelle Wahl von geeigneten Merkmalen und anschließender Klassifikation geprägt sind, gehören heute jedoch nach wie vor zum State-of-the-Art. So genannte Feature-Learning-Verfahren, insbesondere in Form von Convolutional Neural Networks (CNN), vereinen Merkmalsextraktion und Klassifikation in einem geschlossenen System, welches relevante Merkmale direkt aus den Eingabedaten lernt und für die Klassifikation verwendet.

In dieser Bachelorarbeit wird ein System zur automatischen Erkennung von Musikinstrumenten in monophonen Audiosignalen basierend auf einem CNN und Spektrogrammen als Eingabedaten entwickelt. Dieses soll später zur automatischen Konfiguration von Audio-Mischpulten eingesetzt werden, um Toningenieure beim Soundcheck von Musikgruppen zu unterstützen.

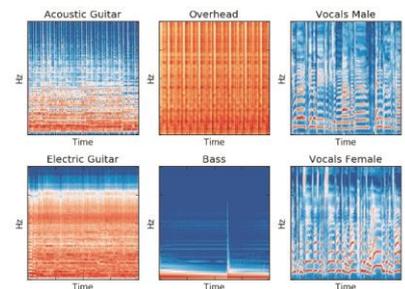
Für das Training des Systems wurde ein eigener Datensatz aus verschiedenen Quellen erstellt. Dieser besteht aus den sechs Klassen akustische und elektrische Gitarre, elektrischer Bass, Schlagzeug sowie männliche und weibliche Stimme.

Mit der gewählten Architektur ist das entwickelte System in der Lage, die Instrumentenklassen mit einer Genauigkeit von über 90% zu unterscheiden. Ein direkter Vergleich mit dem State-of-the-Art ist schwierig, da unterschiedliche Daten verwendet werden. Verwandte Arbeiten erreichen eine vergleichbare oder niedrigere Genauigkeit in der Klassifikation von Orchesterinstrumenten. Die Untersuchungen zeigen, dass CNNs im Zusammenhang mit Spektrogrammen sehr gute Ergebnisse erzielen können und gleichzeitig viel Optimierungspotenzial bieten.

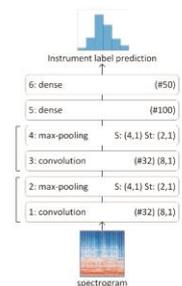


Diplomierende
Yannick Streit
Leotrim Zulfiu

Dozierende
Thilo Stadelmann
Philipp Ackermann



Beispielspektrogramme der zu erkennenden Instrumentenklassen



Verwendete Architektur des Convolutional Neural Network