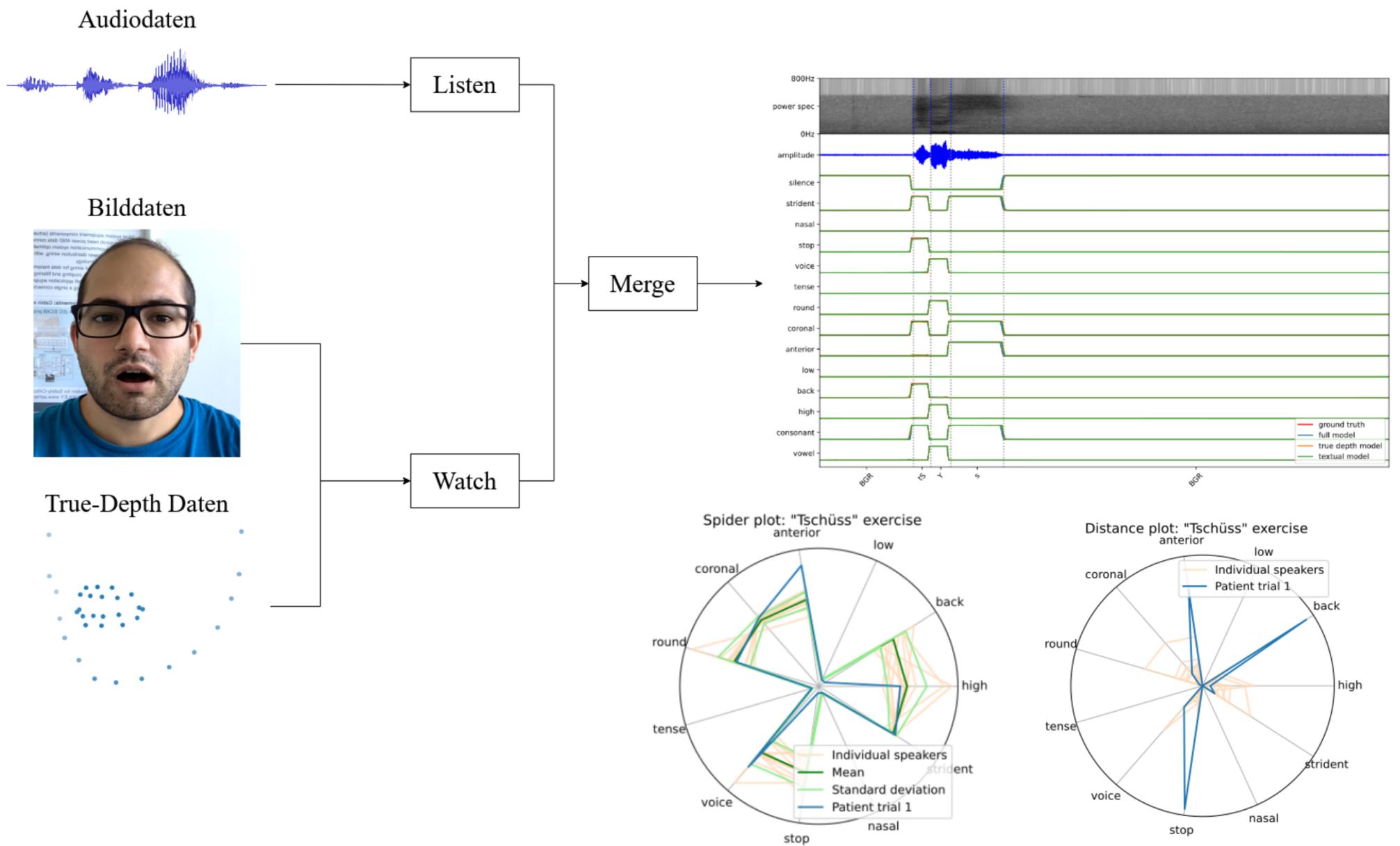


Master-Thesis Engineering, Fachgebiet Information and Communication Technologies

Audiovisuelle Detektion von Aussprachedefiziten



Problemstellung

Mit einer App soll es Schlaganfallpatienten ermöglicht werden, ihre Aussprache selbstständig trainieren zu können. Dafür soll ein neuronales Netzwerk entwickelt werden, welches mit den Audio-, Bild- und True-Depth-Daten die Aussprache einer Aufnahme klassifiziert und dem Benutzer der App ein entsprechendes Feedback gibt.

Lösungskonzept

Als Grundlage für die Lösung dienen die phonological Features. Die Idee dieser Features ist es, die einzelnen Sprachlaute in ihre Komponenten zu zerlegen. Für einen /a/-Laut müssen beispielsweise die Stimmbänder schwingen und die Zunge nach hinten gerollt werden.

Ein neuronales Netzwerk soll dabei so trainiert werden, dass aus den audiovisuellen Eingangsdaten die verschiedenen phonological Features extrahiert werden können.

Dafür wurden die beiden Datenkanäle (Audio und Video) separat untersucht, um möglichst viel Information aus den einzelnen Kanälen zu extrahieren. In einem weiteren Schritt wurden diese Netzwerke zu einem einzelnen Netzwerk fusioniert, um die Erkennungsrate zu verbessern. Es hat sich gezeigt, dass die Erkennungsrate durch den Einbezug beider Datenkanäle signifikant verbessert wird.

Um die Aussprache in einem Video zu klassifizieren wurden die detektierten phonological Features mit Aufnahmen von gesunden Sprechern verglichen.

Durch die Mahalanobis Distanz ist es möglich, allfällige Differenzen zu gesunden Sprechern zu detektieren und klassifizieren. Auch hier hat die Fusion der beiden Datenkanäle das Resultat verbessert.

Eine weitere Erkenntnis war zudem, dass das Netzwerk fähig ist, dieselben Informationen aus den Bild- und True-Depth Daten zu extrahieren. Die Performance des Netzwerkes ist identisch, falls nur True-Depth oder Bildinformationen zur Detektion der phonological Features verwendet werden.

Michael Blättler

Betreuer:
Prof. Dr. Klaus Zahn

Kooperationspartner:
Dr. Amjad Mallisho
Axentless AG
Lehenmattstrasse 248
4052 Basel