

Zusammenfassung

MOTIVATION: Herz-Kreislaufkrankungen sind die häufigste Todesursache der Welt, eine frühzeitige Erkennung kann dem entgegenwirken. Zur Erforschung solcher kardiovaskulärer Anomalien spielt die Analyse von Elektrokardiogrammen (EKG)s eine entscheidende Rolle. Durch die Anschaffung des Data Warehouse Connects und der Nutzung der Health Data Plattform ist es der Neurologie der Charité möglich, EKGs in hoher Auflösung über lange Zeiträume für die Forschung vorzuhalten. Jedoch stellen die großen Datenmengen und die technischen Artefakte Herausforderungen bei der Analyse dar. Ziele sind daher das Finden einer Methode zur Elimination der technischen Artefakte *Baseline Wander* und *Power-Line-Interferenz* sowie das Entwickeln eines Modells zur Erkennung abnormaler Herzschläge, um die Datenmenge auf die relevanten Stellen zu begrenzen.

METHODEN: Zur Elimination der technischen Artefakte wurden Dekompositionen, Filter und Transformationen evaluiert, die auf diese großen Datenmengen mit moderatem Rechenaufwand zuverlässig arbeiten. Das Modell zur Herzschlagerkennung setzt zur Reduktion des Aufwands auf den neuen Daten, auf die Prinzipien des *Transfer Learnings*. Es wurden Machine-Learning-Modelle auf der MIT-BIH-Arrhythmia-Datenbank trainiert, anschließend auf die Charité-Daten übertragen und anhand einer Teilmenge der vorhandenen Monitoring-Annotationen evaluiert.

ERGEBNISSE: Die Methoden zur Elimination technischer Artefakte wurden auf Basis von vier Metriken verglichen: Signal-to-noise ratio (SNR), Maximum Absolut Error (MAX), Normalised Cross Correlation (NCC) und Mean Squared Error (MSE). Das beste Ergebnis erreichte die Sparse Signal Decomposition (SSD) mit einem SNR von 12,43, einem MAX von 0,15, einer NCC von 0,96 und einem MSE von 0,003. Das nächstbeste Ergebnis erreichte die Diskrete Wavelet-Transformation (DWT) mit einem dmey-Wavelet als Basisfunktion. Bei ihr lag der SNR bei 10,34, der MAX bei 0,31, der NCC bei 0,94 und der MSE bei 0.004. Den geringsten Rechenaufwand zeigte mit Abstand die DWT. Das Modell zur Erkennung abnormaler Herzschläge zeigte eine Accuracy von 99,13 % auf der MIT-BIH-Arrhythmia-Datenbank. Auf den Charité-Daten konnte jedoch nur eine Accuracy von 57,86 % erreicht werden.

FAZIT: Die besten Ergebnisse zur Elimination der technischen Artefakte zeigte die SSD. Die DWT sollte dennoch in Betracht gezogen werden, da der Rechenaufwand deutlich geringer ist. Die Herzschlagerkennung weist auf dem MIT-BIH-Datensatz eine hohe Accuracy auf. Die Übertragung des Modells auf die Charité-Testdaten zeigte noch keine zuverlässigen Ergebnisse.

Abstract

PURPOSE: Cardiovascular diseases (CVDs) are the leading cause of death world wide. Early detection could significantly reduce the mortality caused by CVDs. ECG analysis plays a critical role in the study of cardiovascular abnormalities. By acquiring the Data Warehouse Connect and using the Health Data Platform, Neurology of Charité is able to maintain ECGs at high resolution over long periods of time for research. However, the large data volumes and technical artifacts pose challenges in analysis. Therefore, the goals are to find a method to eliminate the technical artifacts of *Baseline Wander* and *Power Line Interference*, and to develop a model for abnormal heartbeat detection to limit the amount of data to the relevant locations.

METHODS: To eliminate these technical artifacts, decompositions, filters, and transformations were evaluated that work reliably on these large data sets with moderate computational effort. The heartbeat detection model relies on the principles of *Transfer Learning* to minimize the overhead on new data. Machine-learning models were trained on the MIT-BIH Arrhythmia Database, applied to the Charité data and evaluated using a subset of the existing monitoring annotations.

RESULTS: Elimination methods were compared on the basis of four metrics: Signal-to-Noise Ratio (SNR), Maximum Absolute Error (MAX), Normalized Cross Correlation (NCC), and Mean Squared Error (MSE). The best result was achieved by Sparse Signal Decomposition (SSD) with an SNR of 12.43, a MAX of 0.15, an NCC of 0.96, and an MSE of 0.003. The next best result was achieved by the Discrete Wavelet Transform (DWT) with a dmey wavelet as the basis function. The SNR was 10.34, the MAX 0.31, the NCC 0.94, and the MSE 0.004. The model for detection of abnormal heartbeats showed an accuracy of 99.13 % on the MIT-BIH Arrhythmia database, but on the Charité data only an accuracy of 57.86 % could be achieved.

CONCLUSION: The SSD showed the best results for eliminating the technical artifacts. The DWT should still be considered, since the computational effort is significantly lower. The heartbeat detection shows a high accuracy on the MIT-BIH dataset. However, the transfer of the model to the Charité test data did not show reliable results.