

# Automated Detection of Cardiac Abnormalities in ECG Signals with a Multi-Head Attention-based CNN-RNN Model

**Vanessa Borst**

Department of Computer Science

Chair of Computer Science II (Software Engineering)

**Prof. Dr.-Ing. Samuel Kounev**

First Reviewer

**M. Sc. Marwin Züfle**

First Advisor

**M. Sc. Robert Leppich**

Second Advisor

**Submission**

September 1, 2021

[www.uni-wuerzburg.de](http://www.uni-wuerzburg.de)



# Abstract

Cardiovascular diseases (CVDs) are the leading cause of death worldwide, accounting for an estimated 35.5% of the lives lost in Germany in 2019 and hence, for more than those attributable to all types of cancer together ( $\sim 25\%$ ) [1, 2]. Since the electrocardiogram (ECG) represents an indicative measure for the presence of CVDs and the early detection as well as treatment of affected patients is crucial for the clinical outcome, the need for reliable algorithms for the automated diagnosis of CVDs based on ECG data arises. However, their development is a challenging task, considering the high complexity of ECG signals, which not only exhibit high inter-and intra-patient variabilities but are also prone to various types of acquisition-related noise.

Nevertheless, many approaches for the automated, ECG-based detection of CVDs have recently been proposed. While traditional statistic- and machine learning-based methods usually require a complex pipeline of preprocessing and manual, expert-driven feature extraction beforehand, deep learning has evolved as powerful means of constructing data-driven end-to-end solutions for the task at hand. The variety of different techniques ranges from convolutional neural networks (CNNs) through recurrent neural networks (RNNs) to hybrid architectures combining both. Moreover, the latest advances in deep learning such as skip connections, various types of attention, or Transformer-based architectures are increasingly being used as well. However, many of the existing approaches still do not integrate such advanced concepts at all, miss the combination of several of them, or only exploit their potential to a limited extent. In addition, the majority is not capable of predicting concurrent cardiac abnormalities, albeit they are not uncommon in practice.

In this thesis, a sophisticated architecture for the automated detection of rhythmic and morphological abnormalities in ECG data is developed, where the recognition of concurrent irregularities is explicitly tackled as well. For this purpose, an existing state-of-the-art (SOTA) approach, which among others, comprises both a CNN and an RNN component as well as an attention module, is refined by integrating promising components that proved to be effective in research areas like computer vision and NLP. Based on two groups of experiments, five major modifications are integrated into the baseline. While the first set of studies concerns the extension of the CNN component by skip connections and normalization layers, the second deals with the integration of multi-head attention into the baseline model. Afterwards, both sets of alterations are combined into a single architecture.

The evaluation results clearly show that both groups of modifications are effective and aid in improving the performance of the baseline model even when being used standalone. Moreover, they demonstrate the effectiveness of combining the two sets of alterations into a final model, which further increases its performance. Lastly, the thresholds that are used for predicting different types of cardiac abnormalities are further fine-tuned by means of Bayesian Optimization during a 10-fold cross validation and based on this, the final model is contrasted with existing SOTA methods. The obtained results show that the proposed model performs strikingly well and is able to outperform all except one of the approaches included in the comparison. Especially, it considerably improves the performance of the SOTA model that serves as the foundation for the successive refinements.



# Zusammenfassung

Herz-Kreislauf-Erkrankungen (HKEs) stellen die häufigste Todesursache weltweit dar. Auch in Deutschland waren im Jahr 2019 35.5% aller Todesopfer darauf zurückzuführen, wobei HKEs insbesondere mehr Tode forderten als sämtliche Krebsarten zusammen ( $\sim 25\%$ ) [1, 2]. Vor dem Hintergrund, dass das Elektrokardiogramm (EKG) ein indikatives Maß für die Präsenz von HKEs darstellt und die frühzeitige Erkennung und Behandlung betroffener Patienten einen erheblichen Einfluss auf das klinische Heilungsergebnis hat, wird die Dringlichkeit der Entwicklung zuverlässiger Algorithmen zur automatischen Detektion von HKEs auf Basis von EKG-Daten besonders deutlich. Aufgrund der hohen Komplexität von EKG Signalen, die nicht nur eine hohe Inter- und Intra-Patientenvariabilität aufweisen, sondern auch anfällig gegenüber akquisitionsbedingtem Rauschen sind, stellt die Entwicklung solcher Methoden allerdings eine große Herausforderung dar.

Trotzdem wurden bereits zahlreiche Ansätze zur automatisierten Diagnose von HKEs auf Basis von EKG-Daten entwickelt. Klassische Statistik und Machine Learning Verfahren haben dabei den Nachteil, dass sie üblicherweise vorab eine komplexe Abfolge an Vorverarbeitung sowie von Expertenwissen getriebenen Schritten zur Extraktion von Merkmalen erfordern. Im Gegensatz dazu, ermöglicht modernes Deep Learning die Konstruktion datengetriebener Ende-zu-Ende-Lösungen. Die Bandbreite an verschiedenen Methoden reicht dabei von faltungsbasierten neuronalen Netzen (convolutional neural networks, CNNs) über rekurrente neuronale Netze (RNNs) bis hin zu hybriden Architekturen, die beides miteinander kombinieren. Außerdem finden zunehmend auch neuste Fortschritte aus dem Bereich des Deep Learnings Verwendung, beispielsweise so genannte Skip Connections, verschiedenste Arten von Attention-Mechanismen oder Transformer-Architekturen. Dennoch verwenden viele existierende Ansätze bis heute keines dieser fortschrittlichen Konzepte, lassen sich die Gelegenheit entgehen, mehrere davon zu kombinieren oder nutzen deren Potential nicht vollständig aus. Hinzu kommt, dass der Großteil nicht darauf ausgelegt ist, mit parallel im EKG auftretenden Abnormalitäten umzugehen, obwohl diese im Klinikalltag nicht selten sind.

Im Rahmen dieser Arbeit wird eine fortschrittliche Architektur zur automatischen Detektion von rhythmischen und morphologischen Abnormalitäten in EKG-Signalen entwickelt, wobei auch gleichzeitig auftretende Irregularitäten explizit berücksichtigen werden. Zu diesem Zweck wird eine existierende State-of-the-Art (SOTA) Architektur, welche unter anderem aus einer CNN und einer RNN-Komponente sowie einem Attention-Modul besteht, als Grundgerüst für iterative Verfeinerungen benutzt. Die integrierten Veränderungen basieren dabei auf vielversprechenden Konzepten, die ihre Effektivität bereits in Forschungsgebieten wie der Bild- oder der Sprachverarbeitung bewiesen haben. Insgesamt werden fünf größere Änderungen am Grundgerüst vorgenommen, wobei diese in zwei Gruppen aufgeteilt sind. Der erste Teil beschäftigt sich mit dem Einbauen von Skip Connections und Normalisierungskomponenten in das CNN-Modul, während der zweite auf die Integration sogenannter Multi-Head-Attention in das Grundgerüst abzielt. Anschließend wird beides im Rahmen einer Gesamtarchitektur kombiniert.

Die Evaluationsergebnisse zeigen deutlich, dass beide Gruppen von Modifikationen auch dann effektiv sind und das zugrundeliegende Modell verbessern, wenn sie alleine angewendet werden. Außerdem demonstrieren die noch weiter verbesserten Ergebnisse der Gesamtarchitektur den Mehrwert, der erzielt wird, wenn man beides in einem finalen Modell kombiniert. Basierend auf einer Feinjustierung der Schwellwerte für die Vorhersage verschiedenerer Abnormalitäten im EKG mittels Bayesscher Optimierung und Kreuzvalidierung, wird die finale Gesamtarchitektur außerdem mit existierenden SOTA-Ansätzen verglichen. Die erzielten Ergebnisse zeigen, dass das entwickelte Modell bemerkenswert gut funktioniert und in der Lage ist, alle bis auf eine der im Vergleich berücksichtigten Methoden zu übertreffen. Insbesondere verbessert der entworfene Ansatz die Leistungsfähigkeit der SOTA-Architektur, die als Grundgerüst für die Modifikationen fungiert, erheblich.