

## Abstract

Sleep is a cyclic physiological phenomenon, an important aspect of human life activity, which, like sport and diet, is a nutritional element that ensures the growth and development of the organism. Under the influence of various factors such as work and study stress and metabolic disorders, more and more people suffer from various types of sleep disorders. Sleep has become an important research topic in recent years. Sleep stage analysis plays an important role in the early detection and treatment of sleep disorders. However, different age groups show different symptoms of sleep disorders, and different sleep disorders show variability in their different sleep stages. The studies in the literature show that the prevalence of sleep disorders is much higher in children than in adults. Although the classification of sleep stages in adults has been well studied, children show markedly different characteristics of sleep stages. Therefore, there is an urgent need for sleep stage classification in children. With the rapid development of intelligent computing technology, artificial intelligence has found wide application in medical research and health sciences in recent years. In the field of sleep medicine, machine learning-based approaches can efficiently and automatically learn abstracted relevant sleep features from collected sleep data to accurately interpret children's sleep stages accordingly. Compared to traditional sleep data analysis, this saves many manual and time resources for data annotation and helps sleep experts reduce the risk of misdiagnosing sleep disorders based on their prior knowledge. In this context, this thesis presents several deep learning-based approaches for sleep stage classification in children using time series *Polysomnography* (PSG) recordings acquired from clinical sensor devices.

The first part of this thesis (Chapter 2) focuses on two topics: The first is *Annotation Level-based Segmentation* (ALS), which deals with fine-grained annotation of the sleep stage (i.e., the annotation criterion in the dataset is utilized as a default for segment length) to construct non-overlapping trainable segments, where each segment contains only one label to avoid the risk of misclassification due to incorrect selection of a representative label. In comparison, conventional sliding window approaches unnecessarily aggregate such labels into coarse-grained ones. A *DeConvolutional Neural Network* (DCNN) that inversely maps the features of a hidden layer to the input space to predict the sleep stage label for each timestamp is proposed. The second is the necessity of multiple sensor modalities and channels. Different clinical signs, symptoms, or other ancillary examinations may be represented by different PSG recordings, so they should all be analyzed comprehensively. One of the major advantages of the DCNN-based method is that it processes raw PSG recordings and internally extracts features useful for the accurate classification of sleep stages. In this chapter, a comparative study between different

supervised deep learning-based approaches is conducted on the own SDCP dataset on a fair basis. As a result, an evaluation framework is first defined, which is further employed to conduct an extensive experiment on the public benchmark dataset: Sleep-EDFX. The results show that the deep learning-based model can effectively detect sleep stages without hand-crafted features, compared to state-of-the-art studies that performed feature engineering, regardless of whether it is based on children's or adults' sleep data. It also demonstrates the great potential of our method, as it can be widely used for timestamp-based multivariate time series classification in various medical fields.

In the second part of this thesis (Chapter 3), a novel attention-based sleep stage classification system is presented to tackle the following two challenges that have not been fully addressed in the sleep study: The first is the problem of data imbalance, i.e., the highly skewed class distribution with underrepresented minority classes. To this end, strategies for oversampling the data using *Gaussian Noise Data Augmentation* (GNDA) and *Borderline Synthetic Minority Oversampling Technique* (B-SMOTE) are attempted to be applied to PSG recordings to achieve a balance of data sizes for different sleep stages. The second challenge is the difficulty in identifying a minority class of sleep stages (i.e., the sleep transition stage) due to the short sleep duration and the feature similarities in EEG and EMG with other stages. To overcome this problem, a *DeConvolution- and Self-Attention-based Model* (DCSAM) is developed that can invert the feature map of a hidden layer into the input space to extract local features and extract the correlations between all possible pairs of features to distinguish sleep stages. To compare the experimental performance and highlight the improvement, all experiments were further performed on the SDCP and Sleep EDFX datasets. The results imply that the proposed DCSAM based on GNDA has promising potential for performance improvement in various medical fields by considering the issue of data imbalance and correlations between features in time series data.

## Zusammenfassung

Schlaf ist ein zyklisches physiologisches Phänomen und ein wichtiger Aspekt der menschlichen Lebenstätigkeit, der ebenso wie Sport und die Ernährungsweise ein nährendes Element ist, welches das Wachstum und die Entwicklung des Organismus gewährleistet. Durch verschiedene Faktoren wie beispielsweise Stoffwechselstörungen und Arbeitsstress leiden immer mehr Menschen unter verschiedenen Arten von Schlafstörungen. Dementsprechend hat sich der Schlaf in den letzten Jahren zu einem wichtigen Forschungsthema entwickelt. Die Schlafstadienanalyse spielt eine wichtige Rolle bei der Früherkennung und Behandlung von Schlafstörungen. Je nach Altersgruppen zeigen sich unterschiedliche Symptome von Schlafstörungen, und unterschiedliche Schlafstörungen zeigen eine Variabilität in den jeweiligen Schlafphasen. Die Literatur zeigt in mehreren Studien, dass die Prävalenz von Schlafstörungen bei Kindern viel höher ist als bei Erwachsenen. Zwar ist die Klassifizierung der Schlafphasen bei Erwachsenen gut erforscht, jedoch unterscheiden sich die Merkmale der Schlafphasen bei Kinder deutlich gegenüber denen von Erwachsenen. Daher besteht ein dringender Bedarf an einer Klassifizierung der Schlafphasen bei Kindern. Mit der raschen Entwicklung der intelligenten Computertechnologie hat künstliche Intelligenz in den letzten Jahren breite Anwendung in der medizinischen Forschung und den Gesundheitswissenschaften gefunden. Im Bereich der Schlafmedizin können auf maschinellem Lernen basierende Ansätze effizient und automatisch abstrakte, aber relevante Schlafmerkmale aus gesammelten Schlafdaten erkennen, um die Schlafphasen von Kindern mit hoher Genauigkeit zu interpretieren. Im Vergleich zur traditionellen Analyse von Schlafdaten spart dies viele manuelle und zeitliche Ressourcen für die Annotation der Daten und hilft Schlafexperten, das Risiko von Fehldiagnosen von Schlafstörungen aufgrund ihres Vorwissens zu verringern. In diesem Zusammenhang werden in dieser Arbeit mehrere Deep-Learning-basierte Ansätze zur Klassifizierung von Schlafphasen bei Kindern unter Verwendung von Zeitreihen-Polysomnographie (PSG)-Aufzeichnungen vorgestellt, die von klinischen Sensorgeräten erfasst wurden.

Der erste Teil dieser Arbeit (Kapitel 2) konzentriert sich auf zwei Themen: Die erste ist die annotationsbasierte Segmentierung (ALS), die sich mit der feinkörnigen Annotation von Schlafphase befasst (d. h. das Annotationskriterium im Datensatz wird als Default für die Segmentlänge verwendet), um nicht überlappende trainierbare Segmente zu konstruieren, wobei jedes Segment nur eine Bezeichnung enthält, um das Risiko einer Fehlklassifizierung durch die falsche Auswahl einer repräsentativen Bezeichnung zu vermeiden. Im Vergleich dazu aggregieren herkömmliche Schiebefenster-Ansätze solche Bezeichnungen unnötigerweise zu grobkörnigen Bezeichnungen. Es wird ein dekonvolutionäres neuronales Netz (DCNN)

vorgeschlagen, das die Merkmale einer verborgenen Schicht umgekehrt auf den Eingaberaum abbildet, um die Schlafphasenbezeichnung für jeden Zeitstempel vorherzusagen. Das zweite Thema ist die Notwendigkeit, mehrere Sensormodalitäten und -kanäle zu verwenden. Verschiedene klinische Anzeichen, Symptome oder andere Zusatzuntersuchungen können durch unterschiedliche PSG-Aufzeichnungen repräsentiert werden, so dass sie alle umfassend analysiert werden sollen. Einer der Hauptvorteile der DCNN-basierten Methode ist, dass sie rohe PSG-Aufzeichnungen direkt verarbeitet und selbstständig Merkmale extrahiert, die für die genaue Klassifizierung von Schlafphasen nützlich sind. In diesem Kapitel wird eine vergleichende Studie zwischen mehreren überwachten Deep Learning-basierten Methoden anhand dem eigenen SDCP-Datensatz auf einer fairen Basis durchgeführt. Infolgedessen wird zunächst ein Bewertungsrahmen definiert, der dann zur Durchführung eines umfangreichen Experiments mit dem öffentlichen Benchmark-Datensatz verwendet wird: Sleep-EDFX. Die Ergebnisse zeigen, dass das auf Deep Learning basierende Modell Schlafphasen auch ohne manuell erstellte Merkmale effektiv erkennen kann. Diese manuell erstellten Merkmale werden bei anderen state-of-the-art Studien durch entsprechendes Feature-Engineering konstruiert, jedoch unabhängig davon, ob es auf Schlafdaten von Kindern oder Erwachsenen basiert. Es zeigt auch das große Potenzial unserer Methode, da sie für die zeitstempelbasierte multivariate Zeitreihenklassifizierung in verschiedenen medizinischen Bereichen weithin eingesetzt werden kann.

Im zweiten Teil dieser Arbeit (Kapitel 3) wird ein neuartiges, aufmerksamkeitsbasiertes System zur Klassifizierung von Schlafphasen vorgestellt, um die folgenden zwei Herausforderungen zu bewältigen, die in der Schlafstudie noch nicht vollständig gelöst wurden: Die erste ist das Problem der Unausgewogenheit der Daten, d.h. die stark verzerrte Klassenverteilung mit unterrepräsentierten Minderheitsklassen. Zu diesem Zweck werden Strategien für das Oversampling der Daten unter Verwendung der Gaußsches Rauschen Datenerweiterung (GNDA) und der Borderline Synthetic Minority Oversampling Technik (B-SMOTE) auf PSG-Aufzeichnungen angewendet, um ein Gleichgewicht der Datenmenge für verschiedene Schlafphasen zu erreichen. Die zweite Herausforderung ist die Schwierigkeit, eine Minderheitsklasse von Schlafphasen (d. h. die Schlaf-Übergangs-Phase) aufgrund der kurzen Schlafdauer und der Ähnlichkeit der Merkmale in EEG und EMG mit anderen Phasen zu identifizieren. Um dieses Problem zu überwinden, wurde ein auf Dekonvolution und Selbstaufmerksamkeit basierendes Modell (DCSAM) entwickelt, das die Merkmalskarte einer verborgenen Schicht in den Eingaberaum umkehren kann, um lokale Merkmale zu extrahieren und die Korrelationen zwischen allen möglichen Merkmalspaaren zu berechnen und zur Schlafphasenunterscheidung zu nutzen. Alle Experimente werden zusätzlich mit den SDCP- und Sleep-EDFX-Datensätzen wiederholt, um

die experimentellen Ergebnisse zu vergleichen und Verbesserungen hervorheben zu können. Die Ergebnisse zeigen, dass das vorgeschlagene DCSAM auf der Grundlage von GNDA ein vielversprechendes Potenzial zur Leistungsverbesserung in unterschiedlichsten medizinischen Bereichen hat, indem es das Problem des Datenungleichgewichts und der Korrelationen zwischen Merkmalen in Zeitreihendaten berücksichtigt.