大连理工大学脑与认识科学概论大作业

Application summary of sleep disorder and deep learning in sleep staging

Department: Faculty of Electronic information and Electrical Engineering

Major: Computer Science-Artificial Intelligence

Names: Liu Kaiyuan Wu Mian Jiang Haiyang

Student Numbers: 201783107 201753004 201788003

Professor: Cong Fengyu

Completion Date: 2019.8.25

大连理工大学

Dalian University of Technology

Application summary of sleep disorder and deep learning in sleep staging

Abstract

Sleep disorder is an epidemic in modern people. It is not only a medical worker's task to study human sleep and provide effective intervention solutions, but also a thorny problem integrating brain science, cognitive science, neuroscience, computer science and other fields. Because manual sleep staging by experts is time-consuming and highly professional, and data acquisition(usually EEG signals) is easily influenced by noise and other bio-signals, the dream of universalizing intelligent clinical care for patients with sleep problems remains tough to realize. Some scientists have studied about automatic sleep stage classification with deep learning. Convolution Neural Networks(CNNs) can extract feature effectively and Long Short-Term Memory(LSTM) has a perfect performance in processing sequence data, which consist the basic modules of their models. Some of the scientists use the raw EEG data as the input of their models which predict the sleep stage. Other scientists convert the raw EEG data into special form matrix such as temporal information matrix before putting the EEG data into their models. We propose three automatic sleep stage annotations based on SleepEEGNet using a single channel EEG signal. Change in the CNN part. (1) Add skip connection to the original CNN part, this can be retained the more feature.(2) Replace the CNN part with a simple fractal network. Through a combination of multiple different depth networks, the shallow layer provides a quicker answer and the deep layer provides a more accurate answer in depth.(3) Added dilated convolution based on fractal networks to monitor the same convolution to a greater time. We evaluated the proposed method on a single EEG channel (ie Fpz-Cz channel) from the Physionet Sleep-EDF datasets published in 2013. The evaluation results show that our method has achieved good results in terms of using fewer parameters or using less training time. The model we developed is ready to be tested with more sleep EEG signals. In the end, we also put forward some conjectures ,which maybe can promote research on this topic. The source code is available at https://github.com/daidaiershidi/Brain-Cognition-Homework

Key Words：Sleep disorder；Automatic sleep staging；Fractal network；Dilated convolution

Catalogue

1 Sleep Disorders

1.1 Background

Roughly one in three American adults doesn't get enough sleep at night. [1] That's a problem, and not just because a lack of sleep leaves many feeling groggy during the day. According to the Centers for Disease Control and Prevention (CDC), poor sleep is associated with elevated risks for a number of diseases—from obesity and diabetes to depression. [2]

A sleep disorder is any recurring issue that disrupts the quality or amount of sleep you get at night. While some sleep disorders are caused by an underlying medical condition, others are not.

By some estimates, nearly 75 percent of adults experience symptoms of a sleep disorder each week, and up to 30 percent of children also grapple with some form of sleep disturbance. [3] While a high percentage of both men and women suffer from sleep disorders, there's some evidence that women—especially those between the ages of 30 and 60—are more likely to suffer from a sleep disorder than men. [4]

For centuries, sleep and dreams have long been topics of immense interest; however, the modern scientific study of sleep began relatively recently. In 1937 an electroencephalograph was used for the first time to observe the electrical activity in the brain during nonrapid eye movement sleep (Loomis et al., 1937) [5]. This opened the field to further advances. Rapid eye movement (REM) was discovered in 1953 by Kleitman and colleagues, and its correlation with dreams was a major step forward in understanding sleep physiology (Aserinsky and Kleitman, 1953) [6]. The culmination of this work came in 1957 when Dement and Kleitman defined the stages of sleep (Dement and Kleitman, 1957) [7]. Since the 1950s a convergence of findings from many fields (e.g., neurology, pulmonology, neuroscience, psychiatry, otolaryngology, anatomy, and physiology) have led to a greater understanding of sleep as a basic universal biological process that affects the functioning of many organ systems (Shepard et al., 2005) [8]. In 1989, a seminal study demonstrated that rats that were subjected to total sleep deprivation developed skin lesions, experienced weight loss in spite of increased food intake, developed bacterial infections, and died within 2 to 3 weeks (Rechtschaffen et al., 1989) [9]. Researchers in sleep and circadian biology continue to work toward a greater understanding of the etiology and pathophysiology of sleep disorders. The field is maturing into an interdisciplinary field in which integration and coordination across the traditional medical specialties, other health care providers (e.g. nurses, dentists), and between basic and clinical science is vita [10].

1.2 Internal Clock

Getting enough sleep is important for people of all ages to stay in good health. People will often cut back on their sleep for work, for family demands, or even to watch a good show on television. But if not getting enough sleep is a regular part of your routine, you may be at an increased risk for obesity, diabetes, high blood pressure, coronary heart disease and stroke, poor mental health, and even early death. Even one night of short sleep can affect you the next day. Not surprisingly, you're more likely to feel sleepy. On top of that, you're more likely to be in a bad mood, be less productive at work, and be involved in a motor vehicle crash.

According to the results from National sleep foundation's sleep time duration recommendations, for healthy individuals with normal sleep, the appropriate sleep duration for newborns is between 14 and 17 hours, infants between 12 and 15 hours, toddlers between 11 and 14 hours, preschoolers between 10 and 13 hours, and school-aged children between 9 and 11 hours. For teenagers, 8 to 10 hours was considered appropriate, 7 to 9 hours for young adults and adults, and 7 to 8 hours of sleep for older adults.

Age and condition	Sleep Needs
Newborns (0-3 months)	14-17 hours
Infants (4-11 momths)	12-15 hours
Toddlers (1-2 years)	11-14 hours
Preschoolers (3-4 years)	10-13 hours
School-age childern (5-12 years)	9-11 hours
Teenagers (13-17 years)	8-10 hours
Adults (18-64 years)	7-9 hours
Older Adults (65 years and over)	7-8 hours

1.3 Common Types of Sleep Disorders

While many people think of insomnia when they hear the term “sleep disorder,” there are several other common conditions that can harm a person’s slumber. In some cases, sleep disorder symptoms may even strike during the daytime—not at night. [4] There are over 90 different types of sleep disorders. The most common include:

Insomnia

Insomnia is an inability to fall asleep, stay asleep, or sleep through the night. It’s the most common sleep disorder, and it comes in two varieties:

Primary insomnia: This is not caused by an underlying medical condition.

Secondary insomnia: This type stems from another medical problem.

Roughly half of all adults report occasional bouts of insomnia, which are stretches of poor sleep lasting anywhere from one night to several weeks. Insomnia becomes “chronic” if it lasts longer than one month. Roughly 10 percent of adults experience chronic insomnia.[11][12][13]

Sleep Apnea

Sleep apnea is a condition in which a person repeatedly stops breathing during the night, which harms the depth and quality of his or her sleep. In some cases, these breathing stoppages can occur hundreds of times each night. There are two types:

Obstructive sleep apnea: This is caused by the partial or total blockage of breathing airways during sleep. It’s the more common type of sleep apnea.

Central sleep apnea: The type of sleep apnea in which the muscles that control a person’s breathing don’t work properly during sleep. This is often a result of "signaling problems" in a person’s brain, or as a result of other conditions, like heart failure and stroke.[3]

Narcolepsy

Narcolepsy is a neurological disorder in which the brain doesn’t properly manage the body’s sleep-wake states. There are two types:

Type 1 narcolepsy: This term is used to describe patients who either have low levels of a brain chemical called hypocretin, which helps regulate the body’s sleep-wake cycles, or those who experience some specific narcolepsy symptoms. (More on those below.)

Type 2 narcolepsy: Patients with this condition do not have low levels of hypocretin, but experience other narcolepsy symptoms.[14]

Restless Leg Syndrome

Restless leg syndrome (RLS) is just what it sounds like: a strong inclination or urge to move your legs while they’re at rest. RLS usually happens once a person lies down in bed or after they’ve been sitting for long periods.

For some, RLS can be so severe that it becomes difficult to fall or stay asleep. This can lead to excessive daytime drowsiness and insomnia.[15]

Circadian Rhythm Disorders

This group of related sleep issues is caused by disruptions to a person's internal circadian rhythm, or sleep clock. This biological clock determines the release of neurochemicals that either initiate sleep or chase it away. For people with a circadian rhythm disorder, there's a mismatch between their internal clocks and their desired sleep-wake schedule, and this leads to problems sleeping.

Circadian rhythm disorders come in a variety of subtypes. These include:

Shift work disorder: This occurs among people who work irregular schedules or night shifts, and are often trying to stay awake or fall asleep at times that don't align with their body's internal clocks.

Delayed sleep phase disorder: This could be thought of as "night owlism." It's most common among teens and young adults, and it's defined as an inability to fall asleep until very late at night—2 or 3 a.m.—and a desire to sleep until midday or later. While this may not seem like a problem for some, it can be a considerable challenge if a person's work or school obligations require them to go to bed and get up at earlier hours.

Advanced sleep phase disorders: This is basically the opposite of delayed sleep phase disorder. This type, which is most common among the elderly, typically involves going to bed at a very early hour—sometime between 6 and 9 p.m.—and rising in the middle of the night.

There are many other types of sleep disorders, but the above are the most common.[16]

1.4 Causes or Risk Factors for Sleep Disorders

While some sleep disorders have a clearly identifiable cause, others emerge from a combination of factors.

Often, sleep disorders are the unwelcome side effect of some other mental or physical medical condition. For example, narcolepsy can stem from a specific kind of autoimmune disorder or, in rare cases, from a brain injury. [14]

For these reasons, the many causes or risk factors for a sleep disorder are too numerous to lay out here, but here are the most common: Iron deficiency anema, Emotional or psychological distress, Inherited genetic factors, Being overweight or obese, Irregular or atypical sleep schedule, Taking certain medications, Stress, Uncomfortable sleeping conditions, Underlying illness, Drinking alcohol or caffeine too close to bedtime, A deviated septum or other structural issue that impacts breathing, Exposures to electronic screens too close to bedtime, et.

1.5 Symptoms of Sleep Disorders

The symptoms of a sleep disorder differ depending on its type and underlying cause.

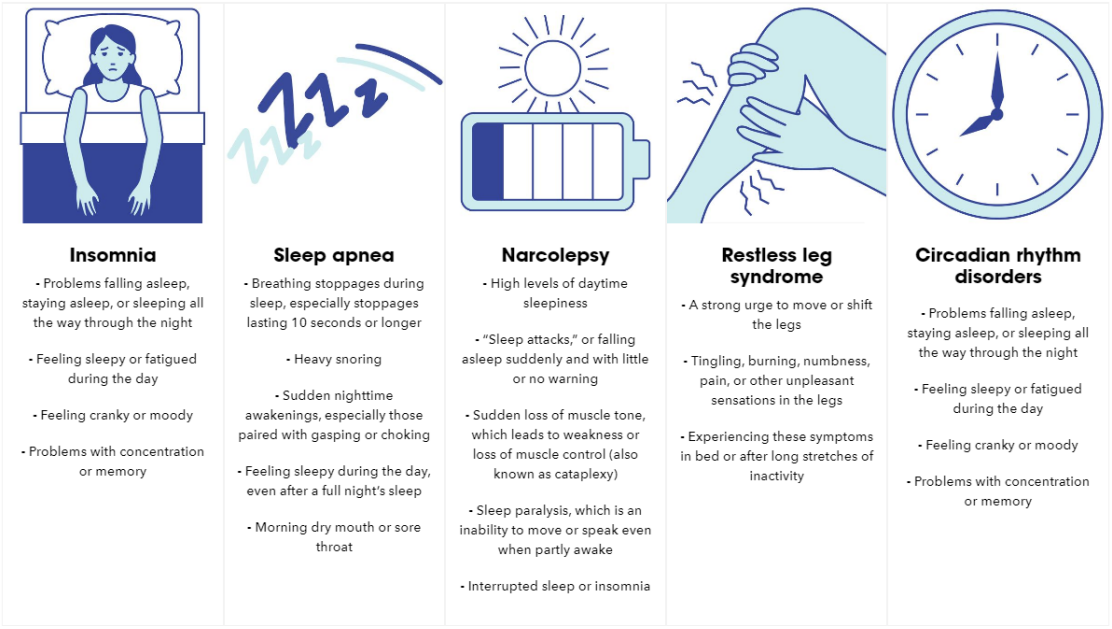

Figure 1 Symptoms of sleep disorders

2 Importance of Getting Enough Sleep

2.1 Overview

Usually sleepers pass through four stages: N1, N2, N3, and REM (rapid eye movement) sleep. These stages progress cyclically from N1 through REM then begin again with stage N1. A complete sleep cycle takes an average of 90 to 110 minutes, with each stage lasting between 5 to 15 minutes. The first sleep cycles each night have relatively short REM sleeps and long periods of deep sleep but later in the night, REM periods lengthen and deep sleep time decreases.

2.2 What Are the Stages of Sleep

There are four stages of sleep: Non-REM (NREM) sleep (Stages N1, N2 & N3) and REM sleep. Periods of wakefulness occur before and intermittently throughout the various sleep stages or as one shifts sleeping position.

Wake is the period when brain wave activity is at its highest and muscle tone is active.

Stage N1 is the lightest stage of NREM sleep. Often defined by the presence of slow eye movements, this drowsy sleep stage can be easily disrupted causing awakenings or arousals. Muscle tone throughout the body relaxes and brain wave activity begins to slow from that of wake. Occasionally people may experience hypnic jerks or abrupt muscle spasms and may even experience sensation of falling while drifting in and out of Stage N1.

Stage N2 is the first actual stage of defined NREM sleep. Awakenings or arousals do not occur as easily as in Stage 1 sleep and the slow-moving eye rolls discontinue. Brain waves continue to slow with specific bursts of rapid activity known as sleep spindlesintermixed with sleep structures known as K complexes. Both sleep spindles and K complexes are thought to serve as protection for the brain from awakening from sleep. Body temperature begins to decrease and heart rate begins to slow.

Stage N3 is known as deep NREM sleep. The most restorative stage of sleep, stage 3 consists of delta waves or slow waves. Awakenings or arousals are rare and often it is difficult to awaken someone in Stage 3 sleep. Parasomnias (sleepwalking, sleep talking or somniloquy and night terrors) occur during the deepest stage of sleep.

REM sleep, also known as rapid eye movement, is most commonly known as the dreaming stage. Eye movements are rapid, moving from side to side and brain waves are more active than in Stages N2 & N3 of sleep. Awakenings and arousals can occur more easily in REM; being woken during a REM period can leave one feeling groggy or overly sleepy.

2.3 What Is a Sleep Cycle

A sleep cycle is the progression through the various stages of NREM sleep to REM sleep before beginning the progression again with NREM sleep. Typically, a person would begin a sleep cycle every 90-120 minutes resulting in four to five cycles per sleep time, or hours spent asleep.

One does not go straight from deep sleep to REM sleep, however. Rather, a sleep cycle progress through the stages of non-REM sleep from light to deep sleep, then reverse back from deep sleep to light sleep, ending with time in REM sleep before starting over in light sleep again.

For a majority of people, a sleep cycle begins with a short period of Stage N1 sleep whereby the body begins to relax and a drowsy state occurs with slow rolling eye movements. Though arousals or awakenings are prevalent, Stage N1 is important as it allows for the body to enter Stage N2; the first quantifiable stage of NREM sleep.

Stage N2 occurs for longer periods than Stage N1. For most, Stage N2 sleep comprises approximately 40-60% of total sleep time.

Moving through the sleep cycle, Stage N3 is most often found next in the progression. This restorative stage does not last as long as Stage N2, lasting between 5-15% of total time asleep for most adults. For children and adolescents Stage N3 is much higher in duration.

REM can occur at at time during the sleep cycle, but on average it begins 90 minutes following sleep onset and is short in duration as it is the first REM period of the night. Following REM, the process resumes starting with periods of Stage N1, N2 & N3 intermixed before returning to REM again for longer periods of time as sleep time continues.

How long is a sleep cycle? The first sleep cycle takes about 90 minutes. After that, they average between 100 to 120 minutes. Typically, an individual will go through four to five sleep cycles a night.

2.4 What Is Deep Sleep

Deep sleep occurs in Stage N3 of NREM sleep. Brain waves during Stage N3 are called delta waves due to the slow speed and large amplitude. Of all of the sleep stages, Stage N3 is the most restorative and the sleep stage least likely to be affected by external stimuli.

Waking a person from deep sleep can be difficult. Following a period of sleep deprivation, a person experiences extensive time in Stage N3 sleep. Parasomnias such as

sleep walking, sleep talking, night terrors and bedwetting can occur. (There is muscle activity, that's how people can talk or kick in their sleep!)

Deep sleep reduces your sleep drive, and provides the most restorative sleep of all the sleep stages. This is why if you take a short nap during the day, you're still able to fall asleep at night. But if you take a nap long enough to fall into deep sleep, you have more difficulty falling asleep at night because you reduced your need for sleep.

During deep sleep, human growth hormone is released and restores your body and muscles from the stresses of the day. Your immune system also restores itself. Much less is known about deep sleep than REM sleep. It may be during this stage that the brain also refreshes itself for new learning the following day.

2.5 When Does REM Sleep Occur

A person's sleep time (approximately 6-8 hours for adults) can be thought of as 2 halves. The first half for a majority of people consists mostly of Stages N2 and N3 with sporadic periods of Stage N1 and short REM periods. As the night progresses, Stage N3 begins to diminish in quantity while Stages N1 and N2 remain with lengthening periods of REM occurring.

A person typically experiences three to five REM periods throughout sleep time with the longest REM period right before awakening for the day. If woken prematurely from completing the REM period a person can experience a period of sleep inertia whereby a heightened sensation of sleepiness can occur for several minutes or even several hours.

In the REM period, breathing becomes more rapid, irregular and shallow, eyes jerk rapidly and limb muscles are temporarily paralyzed. Brain waves during this stage increase to levels experienced when a person is awake. Also, heart rate increases, blood pressure rises, males develop erections and the body loses some of the ability to regulate its temperature.

2.6 In What Stage of Sleep Do Dreams Occur?

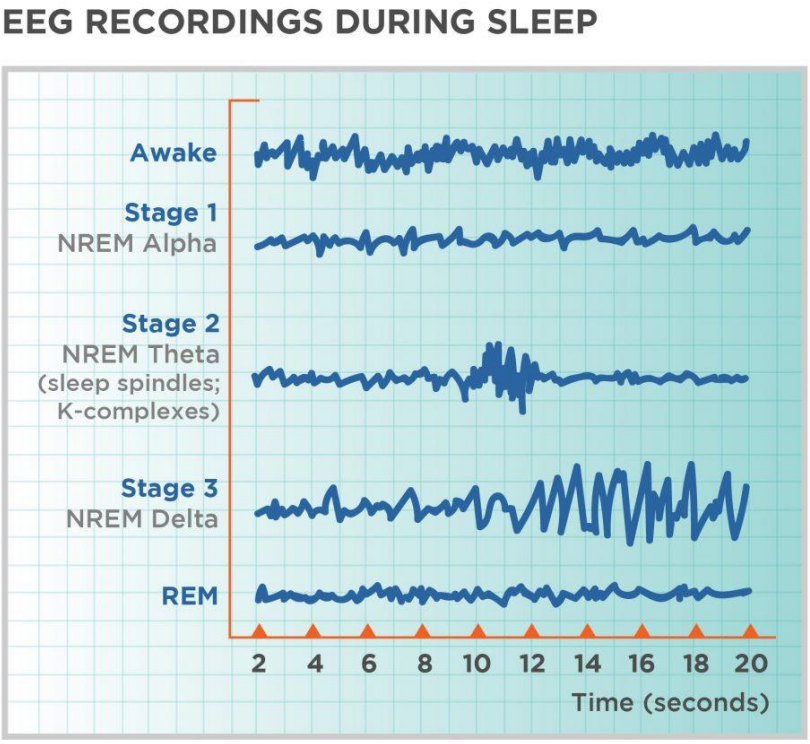

Figure 2 Graphic courtesy of Lumen Learning

As an active sleep state, REM is the time when the most vivid dreams occur. The rapid eye movements that occur can be seen as sharp, rapid movements. Brain waves during REM sleep are considered to be of low amplitude and mixed frequency consistent with higher activity than that seen in Stages N2 and N3.

A person may dream 4 to 6 times each night. A French study found that all people do in fact dream, whether they remember their dreams or not. However, if awoken during REM sleep, a person often can remember their dreams.

Muscle paralysis often accompanies REM sleep. This muscle atonia or muscle paralysis occurs as a protective means to keep one from acting out their dreams. Obstructive Sleep Apnea is often the worst during REM periods due to the lack of muscle tone within the muscles of the airway. Scientists believe this may be to help prevent us from injury while trying to act out our dreams.

During REM respirations are irregular and shallow and irregularities in heart rate and body temperature also occur.

2.7 Brain Waves During REM And Non-REM Sleep

As sleep research is still a relatively young field, scientists did not discover REM sleep until 1953 when new machines were developed to monitor brain activity. Before this discovery it was believed that most brain activity ceased during sleep. Since then, scientists have also disproved the idea that deprivation of REM sleep can lead to insanity and have found that lack of REM sleep can alleviate clinical depression although they do not know why. Recent theories link REM sleep to learning and memory.

2.8 Brain Waves During the Sleep Cycle

STAGE	FREQUENCY (HZ)	AMPLITUDE (MICRO VOLTS)	WAVEFORM TYPE
awake	15-50	<50	
pre-sleep	8-12	50	alpha rhthym
1	4-8	50-100	theta
2	4-15	50-150	splindle waves
3	1-4	100-200	spindle waves and slow waves
REM	15-30	<50	

2.9 How Does Sleep Cycle Change With Age

Sleep changes throughout a person's life. From a newborn, through toddler years, school age, adolescent and adulthood, sleep is changing.

Newborn (0 – approximately 4 months) Do not have distinctive sleep waves. Sleep is categorized as "Active", "Quiet" and "Indeterminate". Active sleep is the equivalent to REM sleep and quiet sleep is equivalent to non-REM sleep. A majority of the time, newborns are in active sleep which allows for frequent arousals or awakenings; this is necessary for regular periods of feeding.

Infants (Approximately 4 months – 1 year) Standard sleep stage distinction is now apparent. Sleep becomes more consolidated and sleeping routines can be developed, sleep is typically 10-13 hours per 24 hour period with 2-3 daytime naps occuring.

Toddlers (1 year – 3 years) With sleeping patterns fully developed, children spend approximately 25% in Stage N3 deep sleep with almost an equal amount of time in REM. Average sleep time is 9.5-10.5 hours per 24 hours period. Typically, naps will reduce to 1 per day most likely occurring early in the afternoon to allow for proper nighttime sleep.

Pre-School (3 – 6 years) Sleep time is similar to that of toddlers, approximately 9-10 hours per 24 hour period. The afternoon nap usually subsides around 3-4 years for a majority of children. Stage N3 sleep still remains high in relation to total sleep time.

School Age (6 years – 12 years) Sleep time remains unchanged; 9-10 hours per 24 hour period and Stage N3 remains approximately 20-25% of total sleep time. Restorative sleep is important for growth and development.

Adolescent (12 years and beyond) Sleep time for adolescents is approximately 9-9.5 hours per 24 hours period. There are physiological changes in circadian rhythm that occur causing sleep onset to be later. This internal shift is the cause for many adolescents to have later lights out and the desire to want to "sleep in" in the morning. As a person ages, the circadian rhythm shifts back and sleep again appears to regulate to approximately 6.5-8 hours of sleep per 24 hour period as adult.

3 Why to Study Sleep Staging

3.1 Challenges

Medical, nursing, and pharmacy students along with individuals in graduate training, residency, and fellowship training programs require greater exposure to the public health burden of sleep loss and disorders; they also must understand the relationship between sleep problems and the proper diagnosis and treatment of a wide range of medical problems throughout an individual's life span. Although the data are limited, they suggest that focused training about sleep can positively influence the performance of medical students (Haponik and Camp, 1994), residents (Owens and Dalzell, 2005), physicians (Haponik et al., 1996; Rosen et al., 2005; Papp and Strohl, 2005), and primary care clinical staff (Zozula et al., 2005). For example, interns who had previous instruction about sleep-related material often asked patients about past or current sleep problems (82 percent of the time), while sleep histories were rarely obtained by interns who did not have any previous instruction (13 percent of the time) (Haponik et al., 1996).

The challenges that lie ahead, outlined below, are many:

- Sustaining educational initiatives begun by the Sleep Academic Award program.
- Monitoring progress and developing new and updated educational materials, such as sleep objective structured clinical examinations.
- Coordinating efforts across institutions.
- Identifying remaining gaps by assessing the impact of sleep education on physician knowledge, skills, and attitudes; clinical practice; and public health.
- Assessing the relative value and effectiveness of sleep curricula that are integrated across other areas versus those that are stand-alone units.
- Developing means of credentialing nurses, psychologists, and other clinicians who will not qualify for American Board of Medical Specialties certification.
- Integrating sleep-related content into continuing education requirements.

To these ends, educational outcomes research grants and partnerships with appropriate medical subspecialty groups for development and dissemi nation of educational programs is essential. Further, many health care-related programs are actively embracing new technologies for teaching (e.g., computer simulations of office practices) that provide an opportunity to ensure that sleep-related materials are incorporated into evolving curricula.

3.2 Electroencephalography (EEG)&Polysomnography (PSG) &hypnogram

Electroencephalography (EEG) is an electrophysiological monitoring method to record electrical activity of the brain. It is typically noninvasive, with the electrodes placed along the scalp, although invasive electrodes are sometimes used, as in electrocorticography. EEG measures voltage fluctuations resulting from ionic current within the neurons of the brain. Clinically, EEG refers to the recording of the brain's spontaneous electrical activity over a period of time, as recorded from multiple electrodes placed on the scalp. Diagnostic applications generally focus either on event-related potentials or on the spectral content of EEG. The former investigates potential fluctuations time locked to an event, such as 'stimulus onset' or 'button press'. The latter analyses the type of neural oscillations (popularly called "brain waves") that can be observed in EEG signals in the frequency domain.

EEG is most often used to diagnose epilepsy, which causes abnormalities in EEG readings. It is also used to diagnose sleep disorders, depth of anesthesia, coma, encephalopathies, and brain death. EEG used to be a first-line method of diagnosis for tumors, stroke and other focal brain disorders, but this use has decreased with the advent of high-resolution anatomical imaging techniques such as magnetic resonance imaging (MRI) and computed tomography (CT). Despite limited spatial resolution, EEG continues to be a valuable tool for research and diagnosis. It is one of the few mobile techniques available and offers millisecond-range temporal resolution which is not possible with CT, PET or MRI.

Derivatives of the EEG technique include evoked potentials (EP), which involves averaging the EEG activity time-locked to the presentation of a stimulus of some sort (visual, somatosensory, or auditory). Event-related potentials (ERPs) refer to averaged EEG responses that are time-locked to more complex processing of stimuli; this technique is used in cognitive science, cognitive psychology, and psychophysiological research.

Polysomnography (PSG), a type of sleep study, is a multi-parametric test used in the study of sleep and as a diagnostic tool in sleep medicine. The test result is called a polysomnogram, also abbreviated PSG. The name is derived from Greek and Latin roots: the Greek πολύς (polus for "many, much", indicating many channels), the Latin somnus ("sleep"), and the Greek γράφειν (graphein, "to write").

Type I polysomnography, a sleep study performed overnight while being continuously monitored by a credentialed technologist, is a comprehensive recording of the biophysiological changes that occur during sleep. It is usually performed at night, when most people sleep, though some labs can accommodate shift workers and people with circadian rhythm sleep disorders and do the test at other times of the day. The PSG monitors many body functions, including brain activity (EEG), eye movements (EOG), muscle activity or skeletal muscle activation (EMG), and heart rhythm (ECG), during sleep. After the identification of

the sleep disorder sleep apnea in the 1970s, the breathing functions, respiratory airflow, and respiratory effort indicators were added along with peripheral pulse oximetry. Polysomnography no longer includes NPT, Nocturnal Penile Tumescence, for monitoring of erectile dysfunction, as it is reported that all male patients will experience erections during phasic REM sleep, regardless of dream content. Limited channel polysomnography, or unattended home sleep tests, should be referenced as Type II – IV channel polysomnography. With some controversy, polysomnography is best performed by technicians and technologists who are specifically licensed and accredited in sleep medicine. However, at times nurses and respiratory therapists are allowed to perform polysomnography despite lack of specific knowledge and training in this area.

In general, many information can be extrapolated from the polysomnography; some can be directly related to the sleep, such as the sleep onset latency (SOL), the REM-sleep onset latency, the number of awakenings during the sleep-period, the total sleep duration, percentages and durations of every sleep stage, and the number of arousals. But there can be also other information, crucial for many diagnostics, that are not directly linked with the sleep, such as movements, respiration, cardiovascular parameters. In any case, through the polysomnographic evaluation, other information can be obtained (such as, for example, body temperature or esophageal pH) according to the patient's or the study's needs.

Video-EEG polysomnography is a technique combining polysomnography and video-recording, which has been described to be even more effective than only polysomnography for the evaluation of some sleep troubles such as parasomnias, because it allows to more easily correlate EEG signals, polysomnography and behaviors.

A hypnogram is a form of polysomnography; it is a graph that represents the stages of sleep as a function of time. It was developed as an easy way to present the recordings of the brain wave activity from an electroencephalogram (EEG) during a period of sleep. It allows the different stages of sleep: rapid eye movement sleep (REM) and non-rapid eye movement sleep (NREM) to be identified during the sleep cycle. NREM sleep can be further classified into NREM stage 1, 2 and 3. The previously considered 4th stage of NREM sleep has been included within stage 3; this stage is also called slow wave sleep (SWS) and is the deepest stage of sleep. Each of the three NREM stages as well as the period of REM sleep and the awake state can be determined and displayed on a hypnogram.

3.3 Related Labs and Top Papers

3.3.1 Geoffrey Hinton& Machine Learning Group

Brief Introduction

Geoffrey Hinton, known as the "father of neural networks" and "father of deep learning", has received a doctorate in artificial intelligence from the university of Edinburgh and is a distinguished professor at the university of Toronto. In 2012, Hinton won Killam Prizes, Canada's top science prize. In 2013, Hinton joined Google and led an AI team. He brought the neural network into the boom of research and application, changed the "deep learning" from the edge topic to the core technology that Google and other Internet giants rely on, and applied the Hinton Back Propagation algorithm to the neural network and deep learning. Recently, he has again attracted wide attention and heated discussion by proposing an capsule concept to overturn back propagation.

Top Paper- Capsule theory

In current neural networks, neurons at each layer do similar things. For example, each neuron in a convolutional layer performs the same convolution operation. Hinton firmly believes that different neurons can focus on different entities or properties, such as different neurons focusing on different categories at the beginning (rather than normalization at the end). Specifically, some neurons focus on location, some on size, and some on direction. This is similar to the human brain, which has separate areas for language and vision, rather than scattered throughout the brain.

To avoid chaos in the network structure, Hinton proposed to package and assemble neurons that focus on the same category or property, like capsules. When the neural network works, the pathways between the capsules form a sparsely activated tree (only partially activated capsules in the entire tree), forming his Capsule theory.

Networks like Capsule fit the intuitive sense of knowing more than one attribute at a time, but it also raises the intuitive question of how different capsules should be trained and how networks can determine the activation relationships between capsules. Hinton's paper focuses on the study of connection weights (routes) between different capsules. [17]

3.3.2 Prof. Yann LeCun& Computational and Biological Learning Lab

Brief Introduction

Yann LeCun, the father of the CNN, New York university professor. Yann, Geoffrey Hinton, and Yoshua Bengio are the "three giants" of deep learning. He is former head of the institute of artificial intelligence of Facebook, the reader of IJCV, PAMI and IEEE Trans. He created ICLR (International Conference on Learning Representations) and chaired meeting with Yoshua Bengio together.

He received his degree in electrical engineering at ESIEE in Paris in 1983 and his doctorate in computer science at Universite P&M Curie in 1987. LeNet5 was developed in

1998. He was awarded IEEE neural network leader award, and in 2019, he was awarded Turing award.

Top Paper- deep neural networks VS glassy systems

The paper analyzes numerically the training dynamics of deep neural networks (DNN) by using methods developed in statistical physics of glassy systems. The two main issues they address are (1) the complexity of the loss landscape and of the dynamics within it, and (2) to what extent DNNs share similarities with glassy systems. Their findings, obtained for different architectures and datasets, suggest that during the training process the dynamics slows down because of an increasingly large number of flat directions. At large times, when the loss is approaching zero, the system diffuses at the bottom of the landscape. Despite some similarities with the dynamics of mean-field glassy systems, in particular, the absence of barrier crossing, they find distinctive dynamical behaviors in the two cases, showing that the statistical properties of the corresponding loss and energy landscapes are different. In contrast, when the network is under-parametrized they observe a typical glassy behavior, thus suggesting the existence of different phases depending on whether the network is under-parametrized or over-parametrized.[18]

3.3.3 Mary A Carskadon& Sleep Research Lab: Sleep for Science

Brief Introduction

Mary A Carskadon is the Director of Chronobiology/Sleep Research at the E.P. Bradley Hospital and Professor of Psychiatry and Human Behavior at Brown Medical School in Providence, RI. Her scientific activities include research examining the interrelation between the circadian timing system and sleep/wake patterns of children and adolescents. Findings of her research have raised public health issues regarding the consequences of insufficient sleep in adolescents as well as concerns about early starting times of schools, prompting many school districts to delay start times. Her new research initiatives assess sleep's impact on mood and epigenetic changes, as well as the influence of sleep and circadian timing on food choices of adolescents.

Top Paper- A deep learning architecture for temporal sleep stage classification

Sleep stage classification constitutes an important preliminary exam in the diagnosis of sleep disorders. It is traditionally performed by a sleep expert who assigns to each 30 s of the signal of a sleep stage, based on the visual inspection of signals such as electroencephalograms (EEGs), electrooculograms (EOGs), electrocardiograms, and electromyograms (EMGs). This paper introduces the first deep learning approach for sleep stage classification that learns end-to-end without computing spectrograms or extracting handcrafted features, that exploits all multivariate and multimodal polysomnography (PSG)

signals (EEG, EMG, and EOG), and that can exploit the temporal context of each 30-s window of data. For each modality, the first layer learns linear spatial filters that exploit the array of sensors to increase the signal-to-noise ratio, and the last layer feeds the learnt representation to a softmax classifier. The model is compared to alternative automatic approaches based on convolutional networks or decisions trees. Results obtained on 61 publicly available PSG records with up to 20 EEG channels demonstrate that their network architecture yields the state-of-the-art performance. Their study reveals a number of insights on the spatiotemporal distribution of the signal of interest: a good tradeoff for optimal classification performance measured with balanced accuracy is to use 6 EEG with 2 EOG (left and right) and 3 EMG chin channels. Also exploiting 1 min of data before and after each data segment offers the strongest improvement when a limited number of channels are available. As sleep experts, their system exploits the multivariate and multimodal nature of PSG signals in order to deliver the state-of-the-art classification performance with a small computational cost.[19]

4 Related Works for Sleep Stage Classification

We have read some papers which introduced some methods for sleep stage classification with EEG data. In these papers, the scientists used the convolutional neural networks as the models they processed the EEG data and made a prediction of sleep stages. Among the scientists, they hold two ideas to process the EEG data. Some scientists used the raw EEG data as the input of the convolutional neural network. While some of them converted the raw EEG data into special form matrix before putting the EEG data into their models.

4.1 Using Raw EEG Data as Input

Some scientists believed that convolutional neural networks can extract the feature efficiently, so they have put the raw EEG data into the models without complex preprocess-ing. And they promoted the accuracy testing the model on the test set data.

4.1.1 DeepSleepNet

DeepSleepNet [20] is one of the models have a good performance, which has been quoted nearly one hundred times since it was published in 2017.

The architecture of DeepSleepNet consists of two main parts as shown in Figure 13. The first part is representation learning, which can be trained to learn filters to extract time-invariant features from each of raw single-channel EEG epochs. The authors employ two CNNs with small and large filter sizes at the first layers to extract time-invariant features from raw single-channel 30-s EEG epochs. The small filter is better to capture temporal information, while the larger filter is better to capture frequency information.

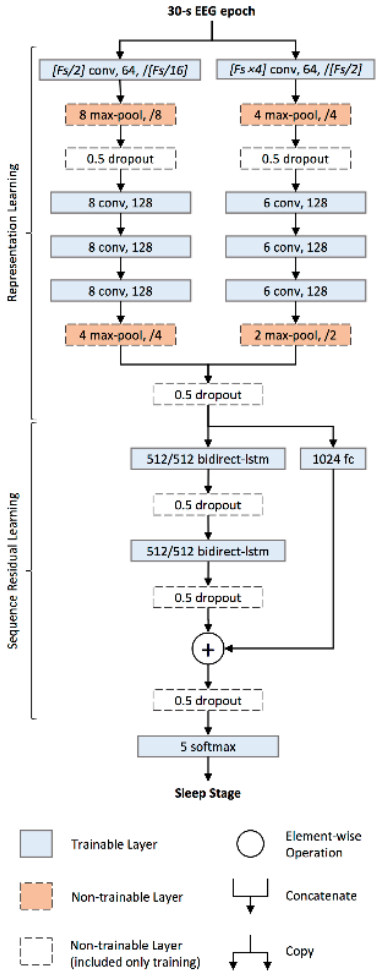

Figure 13 Model of DeepSleepNet

The second part is sequence residual learning, which can be trained to encode the temporal information such as stage transition rules from a sequence of EEG epochs in the extracted features. The authors employ two layers of bidirectional-LSTMs to learn temporal information such as stage transition rules which sleep experts use to determine the next possible sleep stages based on the previous stages. A shortcut connects with fully-connected layer is used to reformulate the computation of this part into a residual function as residual learning framework.

The authors have developed a two-step training algorithm to effectively train their model end-to-end via backpropagation, while preventing the model from suffering class imbalance problem present in a large sleep dataset.

The first step is pre-training. In this step, the two CNNs are extracted from the model and then stacked with a softmax layer, which is not the last layer shown in the model in Figure 13.

These two CNNs stacked with softmax is trained with a class-balance training set which is obtained from duplicating the minority sleep stages in the original training set such that all sleep stage have the same number of samples. At the end of the pre-training, the softmax layer is discarded.

The second step is Fine-Tuning. In this step, the authors trained the whole model with two different learning rate. As the CNNs part has already been pre-trained, the authors used a lower learning rate lr1 for the CNNs part with pre-trained parameters, and a higher learning rate lr2 for the sequence residual learning part, and a softmax layer. Also, the authors use a heuristic gradient clipping technique to prevent the exploding gradients, which is a well-known problem when training RNNs such as LSTMs. This technique rescales the gradients to smaller values using their global norm whenever they exceed a pre-defined threshold. The sequential training set is obtained by arranging the original training set sequentially according to time across all subjects.

Besides, two regularization techniques were used to help prevent overfitting problems. The first technique was dropout that randomly sets the input values to 0 with the specified probability during training. Dropout layers with the probability of 0.5 were used throughout the model as shown in Figure 13 It is important to note that these dropout layers were used for training only, and were removed from the model during testing to provide deterministic outputs.

The second technique was L2 weight decay, which adds a penalty term into a loss function to prevent large values of the parameters in the model. The authors only applied the weight decay on the first layers of the two CNNs. This weight decay helped the model learn smoother filters which resulted in slightly performance gains.

The second step is Fine-Tuning. In this step, the authors trained the whole model with two different learning rate. As the CNNs part has already been pre-trained, the authors used a lower learning rate lr1 for the CNNs part with pre-trained parameters, and a higher learning rate lr2 for the sequence residual learning part, and a softmax layer. Also, the authors use a heuristic gradient clipping technique to prevent the exploding gradients, which is a well-known problem when training RNNs such as LSTMs. This technique rescales the gradients to smaller values using their global norm whenever they exceed a pre-defined threshold. The sequential training set is obtained by arranging the original training set sequentially according to time across all subjects.

Besides, two regularization techniques were used to help prevent overfitting problems. The first technique was dropout that randomly sets the input values to 0 with the specified probability during training. Dropout layers with the probability of 0.5 were used throughout the model as shown in Figure 13 It is important to note that these dropout layers were used for

training only, and were removed from the model during testing to provide deterministic outputs.

The second technique was L2 weight decay, which adds a penalty term into a loss function to prevent large values of the parameters in the model. The authors only applied the weight decay on the first layers of the two CNNs. This weight decay helped the model learn smoother filters which resulted in slightly performance gains.

4.1.2 Architecture Using Multivariate and Multimodal Time Series

In this paper [21], the scientists use multivariate and multimodal time series for sleep stage classification. The input of the model were not only multivariate data but also data sequence.

The length of the data sequence could be adjusted by a parameter k. Whenever k > 0 the neural network has access to the temporal context of the segment of signal to classify, it is the temporal sleep stage classification problem, and when k = 0 the problem boils down to the standard formulation of sleep stage classification.

The deep network architecture proposed to perform sleep stage classification from multivariate time series without temporal context (k = 0) has three key features: linear spatial filtering to estimate so called virtual channels, convolutive layers to capture spectral features and separate pipelines for EEG/EOG and EMG respectively. This network constitutes a general feature extractor the authors denote by $Z{:}R^{C*T} \rightarrow R^{D}$, where D is the size of the estimated feature space. The network can handle various number of input channels and several modalities at the same time. The general architecture is represented in Figure 4

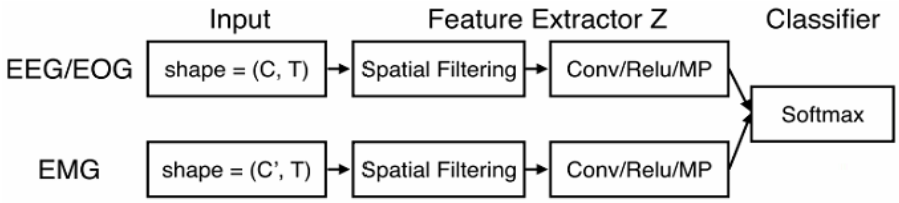

Figure 4 Network general architecture

The resulting outputs are then concatenated to form the feature space of dimension D before being fed into a final layer with 5 neurons and a softmax non-linearity to obtain a probability vector which sums to one. This final layer is referred to as a softmax classifier.

The authors proposed the Time Distributed Multivariate Network to perform temporal sleep stage classification (k > 0). It builds on the Multivariate Network Architecture presented previously and distributes it in time to take into account the temporal context.

The authors proposed the Time Distributed Multivariate Network to perform temporal sleep stage classification (k > 0). It builds on the Multivariate Network Architecture presented previously and distributes it in time to take into account the temporal context.

Formally, let $S_t^k = \{X_{t-k}, ..., X_t, ..., X_{t+k}\} \in \chi_k$ be a sequence of 2k + 1 neighboring samples (k samples in the past and k samples in the future). Distributing in time the features extractor consists in applying Z to each sample in S_t^k and aggregating the 2k + 1 outputs forming a vector of size D(2k +1). Then, the obtained vector is fed into the final softmax classifier. This is summarized in Figure 5

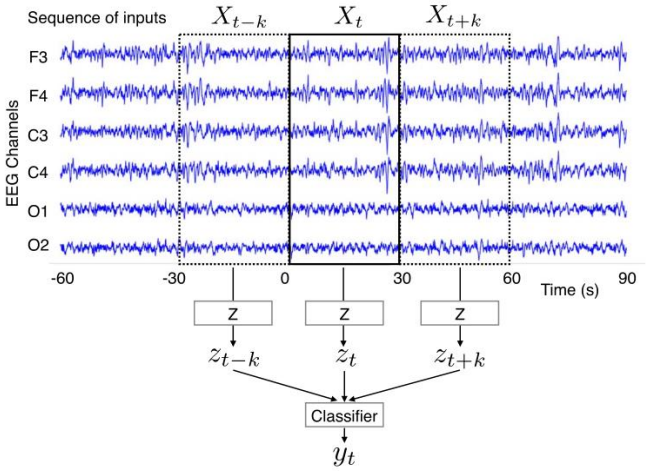

Figure 5 Time distributed architecture to process a sequence of inputs

The training of the time distributed network was done in two steps. First, the author trained the multivariate network, especially its feature extractor part Zt without temporal context (k = 0). The trained model was then used to set the weights of the feature extractor distributed in time. Second, they freezed the weights of the feature extractor distributed in time and they trained the final softmax classifier with aggregated features.

4.1.3 SleepEEGNet

This is one of the newest models proposed in March 2019 [22]. the authors combined some effective methods together, and got good result.

The authors made example by the CNN part of the Deepsleepnet, which had two channels CNN architectures, which are different in the size of the filters and the strides. The CNN consisted of two sections, one with small filters to extract temporal information and another with large filters to extract frequency information. This helped to get benefit from both time

and frequency domain features in the classification task. Figure 6 illustrates the proposed network architecture for automatic sleep stage classification.

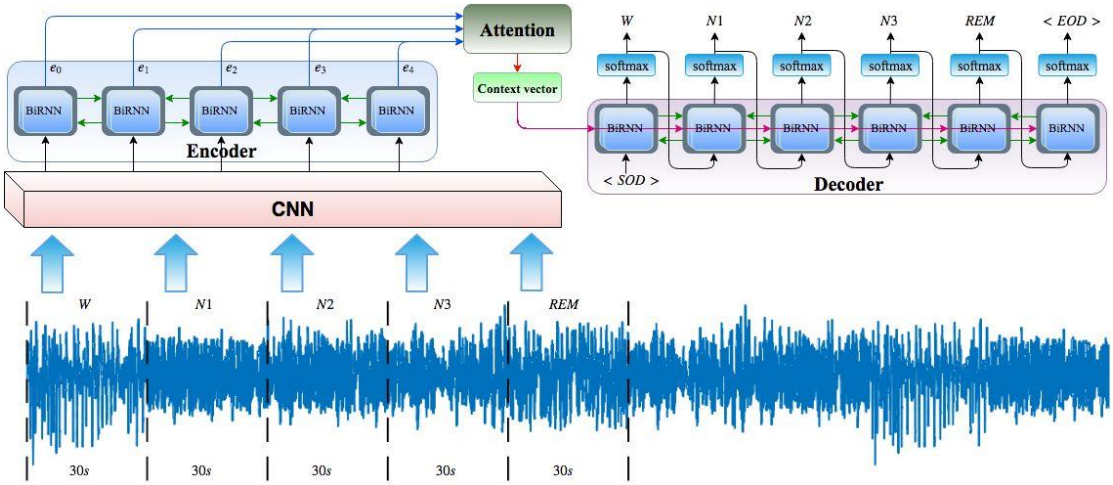

Figure 6 Illustrates the proposed network architecture for automatic sleep stage classification.

The sequence to sequence model is designed based on the encoder-decoder abstract ideas. The encoder encodes the input sequence, while the decoder computes the category of each single channel 30-s EEG of the input sequence. They utilized the bidirectional recurrent neural network (BiRNN) units in the network architecture instead of the standard LSTM (i.e., standard RNN). Standard RNNs are unidirectional, hence they are restricted to use the previous input state. To address this limitation, the BiRNN have been proposed, which can process data in both forward and backward directions. Thus, the current state has access to previous and future input information simultaneously.

The decoder is used to generate the target sequence epoch by epoch. In a basic decoder, at every step of decoding, the decoder gets a new representation of an input element of the sequence generated by the encoder and an element of the target input. The last element of the input sequence is usually the last influence to update the encoder's hidden state. Therefore, the model can be biased to the last element. To address such a problem, the authors have applied an attention mechanism to the model to consider not only the whole encoder representation of the input but also it can learn to put more emphasis on different parts of the encoder outputs in each step of decoding. In a sequence to sequence model without attention approach, the decoder part relies on the hidden vector of the decoder's RNN (or BiRNN), while the sequence to sequence model with the attention is more purposeful. It considers the combination of encoder's representation and decoder hidden vector calling the context vector or the attention vector, (ct).

Similar to other biomedical applications, the sleep stage classification also deals with the problem of class imbalance. To alleviate the effect of this problem on the model, they

calculated new loss functions based on to treat the error of each misclassified sample equally regardless of being a member of the majority or minority class. The authors extended the proposed loss functions, mean false error (MFE) and mean squared false error (MSFE), in for the multi-class classification task. In the most common used loss function, mean squared error (MSE), the loss is calculated by averaging the squared difference between predictions and targets. This way of computing the loss makes the contribution of the majority classes be much more in comparison with the minorities classes in the imbalanced dataset. However, the MFE and MSFE try to consider the errors of all classes equally. By the way, another method is that the dataset is oversampled to nearly reaching a balanced number of sleep stages in each class. The authors have used the synthetic minority over-sampling technique (SMOTE) to generates the synthetic data points by considering the similarities between existing minority samples.

4.2 Convert the EEG Data Aforehand

Some scientists converted the EEG data into special form matrixes so that they got a 2-D data which contained more information. They used CNN models with 2-D convolution layers to extract features.

4.2.1 Time-Frequency Convolutional Neural Network

The authors introduced two novel ways to convert EEG time series to meaningful matrices which CNN can handle with: the dynamical time-frequency spectrum based on Hilbert-Huang transform and the temporal feature matrix, capturing the characteristics of different sleep stages [23].

Since HHT is very suitable for dealing with non-linear and non-stationary EEG signals, the authors will use HHT to extract the time-frequency spectrum of EEG signals. Through HHT, each epoch of sleep stages EEG signals (in 30 s, containing 3000 points of data) is converted into a time-frequency spectrum matrix. The time-frequency matrix characterizes the time-frequency distribution of the instantaneous amplitude of the EEG signal. Then, the time-frequency spectrum matrix will be used as the characteristic input of the Time Frequency Convolution Neural Network, so that an automatic sleep stages scoring model based on time frequency convolution neural network will be further constructed. Figure 7 shows the feature extraction process based on HHT.

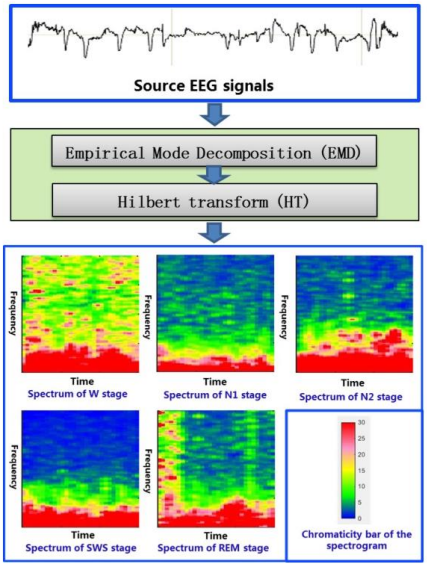

Figure 7 The Dynamical Time-frequency Spectrum Matrix Extraction Process of EEG Signals Based on HHT.

It can be observed that the features extracted based on HHT transform of different sleep stage EEG epoch are distinctly different. Because the method of feature extraction based on HHT transform not only captures the frequency characteristics of the sleep stages EEG epochs, but also captures the changes of frequency domain characteristics as time.

The authors have constructed the temporal information matrix representing the different characteristic waves contained in the EEG signal of each epochs. Based on the temporal information matrix of the EEG signals, an automatic sleep staging model will be further established by Time-Frequency Convolution Neural Network.

In this step, the main objective is to construct the temporal information matrix of each sleep EEG signals epoch (in 30s, containing 3000 points data). In order to keep the local information and long distances change information of each sleep epoch signal of EEG, the authors divide the EEG signal of each epoch into 30 sub-segments in sequence. Each sub-segment contains 100 EEG data points in 1s. The 30 sub-segments are sequentially constructed into a EEG temporal feature information matrix of size (30×100). Figure 8 shows the constructing process of temporal feature information matrix of sleep epoch EEG signals.

The right half of Figure 8 shows the EEG temporal feature information matrix of size (30×100). In this temporal feature information matrix of EEG epoch, each row contains 100 EEG signal data points and each of the two successive points depict the change information of the EEG signal for 0.01s, each column contains 30 EEG signals data points and each of two consecutive points represent the change information of the EEG signals for 1s.

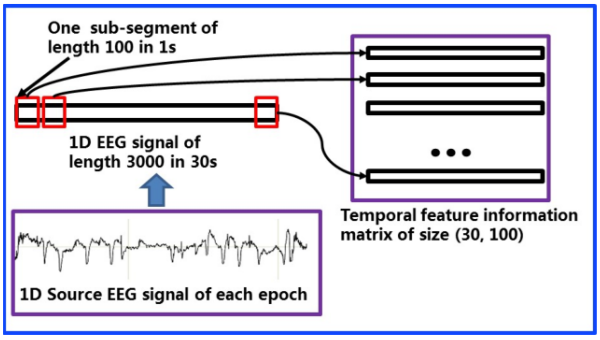

Figure 8 The constructing process of temporal feature information matrix of sleep epoch EEG signals.

Figure 9 shows the model architecture based on convolution neural networks (CNN) and the two types of EEG signal feature matrices that have been constructed earlier. Structurally, the model based on dynamical time-frequency matrix, and the one based on temporary feature matrix are the same, which including the input layer (temporal feature matrix), Cov1, Cov2, Pool1, Con3, FC1, FC2 and output (the probability of belonging to each sleep stage) in order.

The diffierence between the two models is that their filters shape are different.

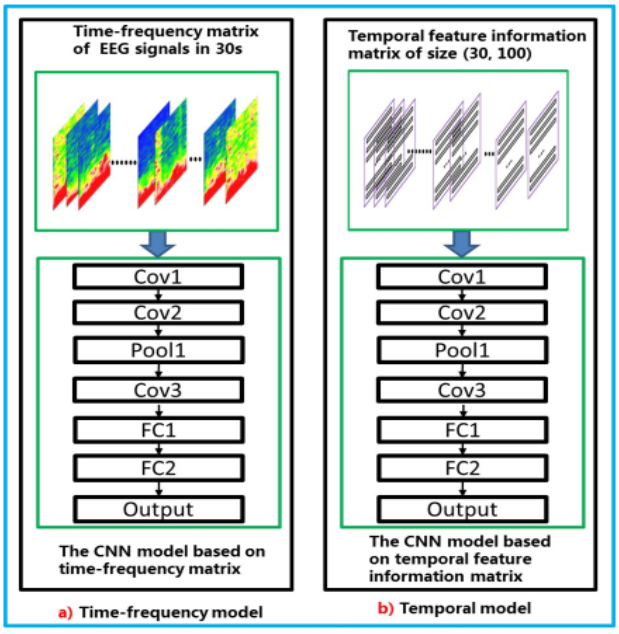

Figure 9 The model architecture.

4.2.2 Using VGGNET to Process EEG Data With Transfer Learning

In this paper [24], the authors proposed a method with transfer learning, they converted the EEG data into time-frequency images, and used pre-trained VGGNET.

EEG data is represented as time-frequency images. A common strategy to reduce the spectrogram's bias is by convolving the raw signal with a window function (called ta-per) before performing spectral estimation. A method called multitaper spectral estimation can be used to reduce both bias and variance by applying multiple taper functions to the raw signal and averaging their results. This technique has recently been proven to be highly successful in analyzing neurophysiological dynamics of sleep using EEG. There exists a set of hyperparameters that highly influence the results of the multitaper spectral estimation: they are the window size ω in seconds, the window step size σ in seconds, the minimum frequency resolution that can be resolved f in Hz, the time-half-bandwidth product, usually defined as $W = \omega f/2$ and the number of tapers used, often set according to the heuristic $L = \lfloor 2W \rfloor - 1$, where $\lfloor x \rfloor$ is a function that rounds x down to the closest integer. After estimating the multitaper spectrogram, they convert its values to a logarithmic scale $x = \log x + 1$ and split it into equally-length bins of sizes, called epochs. Then, the authors convert each epoched spectrogram to an RGB color matrix (i.e., an image) by applying our preferred colourmap (i.e., a lookup table that translates real values in the [0,1] interval to RGB colors).

The CNN of choice to analyze the images created according to the previous block is the VGGNet, due to its simplicity and flexibility. It is composed of 16 weighted layers:

ccm64 ccm128 cccm256 cccm512 cccm512 fcr4096 fcr4096 fcs1000,

where c means a 3 × 3 convolutional filter of stride 1 using a ReLU activation function, m stands for 2 × 2 max-pooling layer with a stride of 2, fcr and fcs correspond to fully-connected layers with ReLU and soft-max activations, respectively; sub-indexed values represent the number of channels in each block.

Transfer learning is employed by using weight values in all convolutional layers that have been previously trained on ILSVRC-2014 data. Fully-connected layers are initialised from scratch using Xavier's initialisation [25] and trained using dropout. The number of final outputs is set according to the task they are tackling.

4.2.3 Joint Classification and Prediction CNN Framework

Usually, in the classification task, the CNN performs one to one classification mapping:

$$F: X_n \rightarrow y_n \in L,$$

where $L = \{1,2,\ldots,Y\}$ denotes the label set of Y sleep stages and Xn is the data need to be classified. In this paper, the authors proposed a joint classification and prediction formulation, the CNN performs the one-to-many mapping:

$$F: X_n \rightarrow (y_{n-\tau}, \ldots, y_n, \ldots, y_{n+\tau}) \in L^{2\tau+1}$$

With their joint classification and prediction formulation, at a certain epoch index n there exists an ensemble of exact 2τ+1 decisions, wherein one classification decision made by itself (i.e. X_n) and 2τ prediction decisions made by its neighbors $(X_{n-\tau},...,X_{n-1},X_{n+1},...,X_{n+\tau})$.These decisions can be aggregated to form the final decision that is generally better that any individual ones.

The authors converted the EEG data into time-frequency image representation. They firstly transformed 30-second signal epoch (i.e. EEG, EOG, or EMG), into a power spectrum using short-time Fourier transform (STFT). The spectrum is then converted to logarithm scale to produce a log-power spectrum image. For frequency smoothing and dimension reduction, the spectrum is filtered by a frequency-domain filter bank.

The authors' recent work presented a simple CNN architecture that was shown efficient for sleep staging. They adapt this architecture by tailoring the last layer and the multi-task softmax layer, to perform joint classification and prediction. The proposed CNN architecture is illustrated in Figure 10. Opposing to typical deep CNN, the proposed CNN consists of only three layers: one over-time convolutional layer, one pooling layer, and one multi-task softmax layer. This simple architecture has three main characteristics.

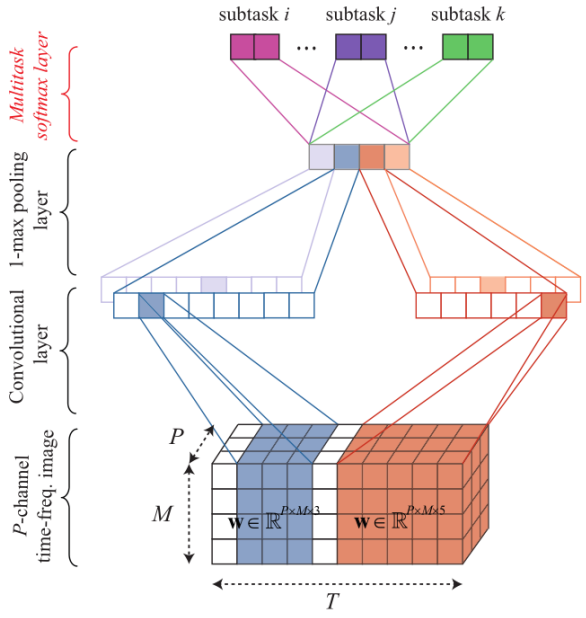

Figure 10 Illustration of the proposed multi-task CNN architecture.

The authors propose a multi-task softmax layer to suit joint classification and prediction. The idea is that the network should be penalized for both misclassification and misprediction on a training example.

5 Our Work

5.1 Methodology

In the following sections, we present a detailed description of our proposed novel model which is improved according to the methods mentioned above and our model is developed to automatically score each sleep stage from a given EEG signal.

Pre-processing

The input to this method is a sequence of 30-s EEG epochs. In order to extract the EEG epochs from a given EEG signal, we follow two simple steps:

(1) Segmenting the continuous raw single-channel EEG to a sequence of 30-s epochs and assigning a label to each epoch (i.e., sleep stage) based on the annotation file.

(2) Normalizing 30-s EEG epochs such that each one has a zero mean and unit variance.

It is worth mentioning that, these pre-processing steps for the sleep stage extraction are very simple and do not involve any form of filtering or noise removal methods.

The architecture

Figure 11 illustrates the proposed network architecture for automatic sleep stage classification. We applied the same bidirectional recurrent neural network and attention decoder architecture provided by [22]. Through this paper, we already know that using CNN helps to benefit from the time domain and frequency domain functions in the classification task.

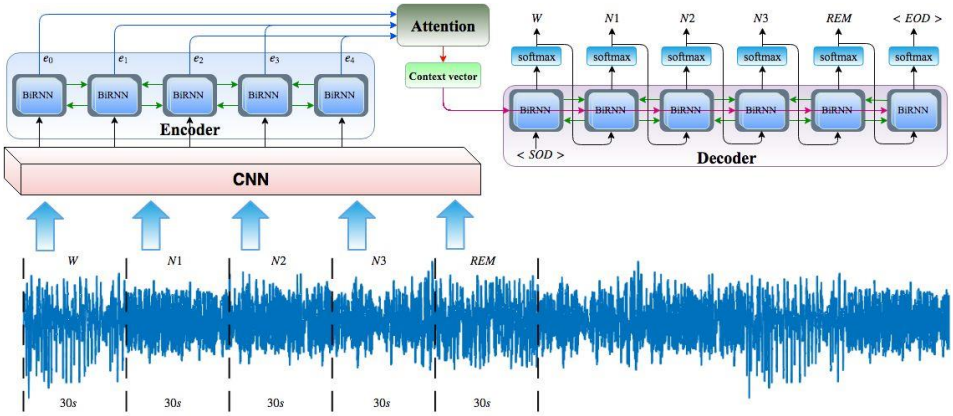

Figure 11 Illustration of sequence to sequence deep learning network architecture used for automated sleep stage scoring.

We hope to deepen the network on this basis, in the hope of getting a higher dimension and more effective features to be sent to BiRNN. Skip connections made the training of very deep networks possible and have become an indispensable component in a variety of neural architectures.

As can be seen in [26], several singularities that hinder the depth of the network have been identified in previous work: (1) overlapping singularities caused by the displacement symmetry of the nodes in a given layer, (2) eliminate singularities corresponding to node cancellation (ie, consistent deactivation), (3) singularities generated by singularities. The linear dependence of the node. These singularities result in the loss of degenerate manifolds in the landscape, which slows down learning. [26] considers that skipping joins eliminates the symmetry of the nodes by eliminating the symmetry of the nodes and eliminating them by reducing the likelihood of node elimination and making the nodes independent of linearity. Based on the above, we first made a slight change in the original CNN structure. Each convolutional layer is passed to a rectified linear unit (ReLU) nonlinearity. The first layer is followed by the largest pooling layer and one missing block. Only one missing block is after the last convolutional layer. After the first largest pooling layer, we added skip connection. At each time step of the training/test model, a sequence of 30 seconds EEG period (the size of the maximum time) is fed to the CNN for feature extraction. Finally, the output of the CNN section is connected, followed by a missing block, so that the encoder network encodes the sequence input. Figure 12 depicts the detailed CNN (skipconnect) structure.

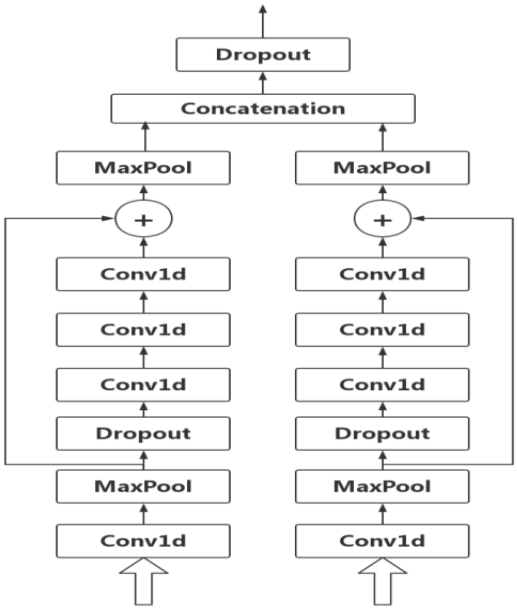

Figure 12 Detailed sketch of the used CNN (skipconnect) model used in the proposed work.

For the second time, we tried a completely different network framework. [27] introduces a neural network macro architecture design strategy based on self-similarity. Repeated application of simple extension rules generates a deep network whose structural layout is a precisely truncated fractal. [27] thinks this network has the ability to transition, during training, from benefits shallow to deep. The second attempt used a convolution to obtain

shallow features, then obtained deep features from FractalNet from shallow to deep, and used them for subsequent models. Figure 13 depicts the detailed CNN (conv) structure.

When using EGG for prediction, the previous signal actually has an effect on the subsequent predictions, so the third attempt is based on the second, replacing the first layer of convolution with a divided convolution. [28] presented a module that uses dilated convolutions to systematically aggregate multiscale contextual information without losing resolution. The architecture is based on the fact that dilated convolutions support exponential expansion of the receptive field without loss of resolution or coverage. Under the same parameters, classified convolution can be associated with longer time features. Figure 13 depicts the detailed CNN (dilatedconv) structure.

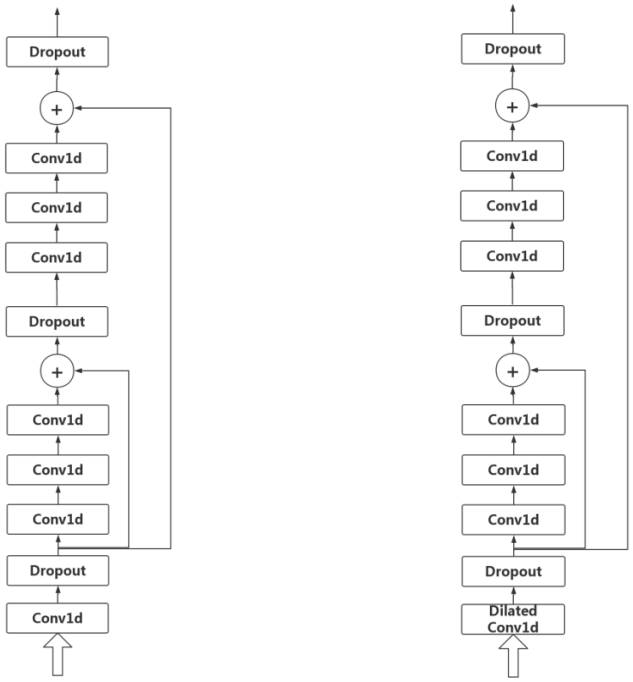

Figure 13 Detailed sketch of the used CNN(conv) model and CNN(dilatedconv) used in the proposed work.

5.2 Experimental Results

Dataset and Data Preparation

In this study, we used the Physionet Sleep-EDF dataset [29] that contributed in 2013 with 61 polysomnograms (PSGs). The Sleep-EDF dataset contains two different studies including (1) study of age effects on sleep in healthy individuals (SC = Sleep Cassette) and (2) study of temazepam effects on sleep (ST = Sleep Telemetry). The dataset includes whole-night polysomnograms (PSGs) sleep recordings at the sampling rate of 100 Hz. Each record contains EEG (from Fpz-Cz and Pz-Oz electrode locations), EOG, chin electromyography (EMG), and event markers. Few records often also contain oro-nasal respiration and rectal

body temperature. The hypnograms (sleep stages; 30-s epochs) were manually labeled by well-trained technicians according to the Rechtschaffen and Kales standard. Each stage was considered to belong to a different class (stage). The classes include W, REM, N1, N2, N3, N4, M (movement time) and '?' (not scored). According to American Academy of Sleep Medicine (AASM) standard, we integrated the stages of N3 and N4 in one class named N3 and excluded M (movement time) and ? (not scored) stages to have five sleep stages. Table I presents the number of sleep stages.

Dataset	W	N1	N2	N3-N4	REM	Total
Sleep-EDF-13	8285	2804	17799	5703	7717	42308

Table I: Details of number of sleep stages.

Experimental Design

The distribution of sleep stages in the Sleep-EDF database is not uniform. Hence, the number of W and N2 stages are much greater than other stages. The machine learning approaches do not perform well with the class imbalance problem. To address this problem, in addition to using the novel loss functions described in Section II-E, the dataset is oversampled to nearly reaching a balanced number of sleep stages in each class. We have used the synthetic minority over-sampling technique (SMOTE) to generates the synthetic data points by considering the similarities between existing minority samples [30].

Our proposed model was evaluated using k-fold cross-validation. We set k to 20 and 10 for version 1 and version 2 of the Sleep-EDF dataset, respectively. In other words, we split the dataset into k folds. Then, for each unique fold, (1) fold is taken as test set and the remaining folds as a training set and (2) trained the model using the training set and evaluated the model using the test set. Finally, all evaluation results were combined.

The network was trained (for each dataset) with a maximum of 400 epochs. RMSProp optimizer was used to minimize the l MFE loss with mini batches of size 20 and a learning rate of $\alpha = 0.0001$. We also applied an additional L2 regularization element with $\beta = 0.001$ to the loss function to mitigate the overfitting. Python programming language and Google Tensorflow deep learning library were utilized to implement our proposed approach.

Evaluation Metrics

We have used different metrics to evaluate the performance of the proposed approach including, overall accuracy, precision, recall (sensitivity), specificity, Cohen's Kappa coefficient (κ) and F1-score. We also computed macroaveraging of F1-score (MF1) which is the sum of per-class F1-scores over the number of classes (i.e., sleep stages).

		Predicted					Per-class Performance (%)			
		W1	N1	N2	N3	REM	Pre	Rec	Spe	F1
Actual	W1	4213	833	553	393	511	77.14	64.78	96.12	70.42
	N1	314	1543	411	83	530	29.72	53.55	89.81	38.23
	N2	531	1853	10107	2280	1673	75.69	61.46	85.41	67.83
	N3	45	27	744	4776	20	63.30	85.10	91.63	72.60
	REM	358	935	1538	12	4410	61.73	60.80	91.30	61.26

Table II: Confusion matrix and per-class performance achieved by the CNN (dilatedconv) using Fpz-Cz EEG channel of the EDF-Sleep-2013 database.

		Predicted					Per-class Performance (%)			
		W1	N1	N2	N3	REM	Pre	Rec	Spe	F1
Actual	W1	7268	317	122	30	176	85.88	91.84	95.97	88.76
	N1	361	1373	196	14	282	48.93	61.68	95.94	54.57
	N2	272	553	12931	955	671	85.64	84.06	90.24	84.84
	N3	67	5	752	4504	7	81.53	84.42	96.83	82.95
	REM	494	558	1097	21	4573	80.10	67.81	96.31	73.45

Table II: Confusion matrix and per-class performance achieved by the CNN (conv) using Fpz-Cz EEG channel of the EDF-Sleep-2013 database.

Method	Dataset	CV	EEG Channel	Overall Performance			Per-class Performance(F1)				
				ACC	MFI	κ	W	N1	N2	N3	REM
SleepEEGNet	Sleep-EDF-13	20-fold CV	Fpz-Cz	84.26	79.66	0.79	89.19	52.19	86.77	85.13	85.02
SleepEEGNet(skipconnect)	Sleep-EDF-13	20-fold CV	Fpz-Cz	84.26	78.70	0.79	89.18	52.18	86.77	85.12	85.02
SleepEEGNet(conv)	Sleep-EDF-13	20-fold CV	Fpz-Cz	81.51	74.87	0.76	88.76	54.57	84.84	82.95	73.45
SleepEEGNet(dilatedconv)	Sleep-EDF-13	20-fold CV	Fpz-Cz	64.72	53.38	0.62	70.42	38.23	67.83	72.60	61.26
Supratak et al.	Sleep-EDF-13	20-fold CV	Fpz-Cz	82.00	76.90	0.76	84.70	46.60	85.90	84.80	82.40
Tsinalis et al.	Sleep-EDF-13	20-fold CV	Fpz-Cz	78.90	73.70	-	71.60	47.00	84.60	84.00	81.40

Table IV: Comparison of performance obtained by our approach with other state-of-the-art algorithms.

5.3 Conclusion

You can see that we used fewer training epochs and relatively fewer parameters (CNN_skipconnect trained 120 rounds, each layer convolved with 128 channels of output, CNN_conv trained 50 rounds, each layer convolved with 128 channels of output, CNN_dilatedconv trained 30 rounds, each layer convolved with 5 channels of output), and we got fairly good results. In fact, due to limitations in computing resources, and my teammate still have other models to run, time is too late. We only used five channels of output in the dilated convolution. Although this greatly reduces the parameters, it also greatly affects the accuracy. Because other networks, including the original network, the output channels are 128 channels. I personally think that if you increase the number of training rounds or increase the amounts of parameters, the model can also achieve a good result.

6 Feelings

We spent more than a month to accomplish this paper. What an unforgettable time!

As a Chinese saying goes: The first step is always the hardest. At the beginning of our work, we were perplexed and did know how to do. The first difficulty was we found the dataset which the professor mentioned in class and it was our first time to process data in file format named ". eth". We even didn't know what data was contained in the files let alone how to process these data. Thanks to our professor, he introduced us two senior schoolmates to help us and we communicated online. The patient senior schoolmates even recorded a video to taught us how to operate the software to check the data. With their help, we learned about the dataset.

Lacking of time was the biggest difficulty we had to overcome. Although we were in summer break, all of us spent a busy time. Wu Mian was preparing for TOEFL and Jiang Haiyang was studying for IELTS. We studied English from morning to evening and did our work until midnight. Liu Kaiyuan was studying driving, he had to get up quite early and did our work dragging a tired body. However, we endeavored to do the best and we communicated online every day. We united together and fight for our common goal which was writing a detailed paper containing much information on the domain we chose. Finally, we completed our excellent paper before the deadline.

Our team had a clear division of labor and every member had a big contribution to our work. Wu Mian looked through lots of websites, read many papers for sleep disorder, sleep stages and EEG data, and organized related top papers well. Jiang Haiyang read some papers on deep learning for sleep stage classification. According to these papers, Liu Kaiyuan proposed an improved method which had good performance. Besides, Jiang Haiyang proposed another model based on the models described in the papers, but it didn't perform well so we didn't mention it in our paper.

We believe that we all try our best to make this assignment to the best of our ability. We all contribute the same in this assignment. The names on the cover page are listed in no particular order.

Reference

[1] https://www.cdc.gov/sleep/index.html
[2] https://www.cdc.gov/sleep/about_sleep/chronic_disease.html
[3] https://my.clevelandclinic.org/-/scassets/files/org/neurological/sleep-disorders/sleep-disorder-treatment-guide-2016.ashx?la=en
[4] https://www.mayoclinic.org/diseases-conditions/sleep-disorders/symptoms-causes/syc-20354018
[5] Loomis AL, Harvey EN, Hobart GA. Cerebral states during sleep as studied by human brain potentials. Journal of Experimental Psychology. 1937; 21:127–144.
[6] Aserinsky E, Kleitman N. Regularly occurring periods of eye motility, and concomitant phenomena, during sleep. Science. 1953;118(3062):273–274. [PubMed]
[7] Dement W, Kleitman N. Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. Electroencephalography and Clinical Neurophysiology Supplement. 1957;9(4):673–690. [PubMed]
[8] Shepard JJW, Buysse DJ, Chesson JAL, Dement WC, Goldberg R, Guilleminault C, Harris CD, Iber C, Mignot E, Mitler MM, Moore KE, Phillips BA, Quan SF, Rosenberg RS, Roth T, Schmidt HS, Silber MS, Walsh JK, White DP. History of the development of sleep medicine in the United States. Journal of Clinical Sleep Medicine. 2005;1(1):61–82. [PMC free article] [PubMed]
[9] Rechtschaffen A, Bergmann BM, Everson CA, Kushida CA, Gilliland MA. Sleep deprivation in the rat: X. Integration and discussion of the findings. Sleep. 1989;12(1):68–87. [PubMed]
[10] Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem.
[11] https://www.sleepfoundation.org/sleep-topics/women-and-sleep
[12] https://my.clevelandclinic.org/health/diseases/12119-insomnia
[13] https://aasm.org/resources/clinicalguidelines/040515.pdf
[14] https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Fact-Sheets/Narcolepsy-Fact-Sheet
[15] https://my.clevelandclinic.org/health/diseases/9497-restless-legs-syndrome
[16] https://my.clevelandclinic.org/health/diseases/12115-circadian-rhythm-disorders
[17] Sara Sabour, et al. Dynamic routing between capsules. NIPS, 2017; 3856-3866.
[18] M Baity-Jesi, et al. Comparing Dynamics: Deep Neural Networks versus Glassy Systems. ICML, 2018; 324-333.
[19] Chambon S, Galtier M N, Arnal P J, et al. A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2018:1-1.

[20] Supratak A, Dong H, Wu C, et al. DeepSleepNet: a Model for Automatic Sleep Stage Scoring based on Raw Single-Channel EEG[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2017:1-1.
[21] Chambon S, Galtier M N, Arnal P J, et al. A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2018:1-1.
[22] Mousavi S, Afghah F, Acharya U R. SleepEEGNet: Automated Sleep Stage Scoring with Sequence to Sequence Deep Learning Approach[J]. 2019.
[23] Wei L, Lin Y, Wang J, et al. Time-Frequency Convolutional Neural Network for Automatic Sleep Stage Classification Based on Single-Channel EEG[C]// 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI). IEEE Computer Society, 2017.
[24] Vilamala A, Madsen K H, Hansen L K. Deep Convolutional Neural Networks for Interpretable Analysis of EEG Sleep Stage Scoring[J]. 2017.
[25] Phan H, Andreotti F, Cooray N, et al. Joint Classification and Prediction CNN Framework for Automatic Sleep Stage Classification[J]. IEEE Transactions on Biomedical Engineering, 2018:1-1.
[26] SKIP CONNECTIONS ELIMINATE SINGULARITIES
[27] FractalNet: Ultra-Deep Neural Networks without Residuals
[28] MULTI-SCALE CONTEXT AGGREGATION BY DILATED CONVOLUTIONS
[29] B. Kemp, A. H. Zwinderman, B. Tuk, H. A. Kamphuisen, and J. J. Oberye, “Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the eeg,” IEEE Transactions on Biomedical Engineering, vol. 47, no. 9, pp. 1185–1194, 2000.
[30] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “Smote: synthetic minority over-sampling technique,” Journal of artificial intelligence research, vol. 16, pp. 321–357, 2002.