SLEEP DATA ANALYSIS FOR RESPIRATORY EVENTS CLASSIFICATION

BY

SUPASUTA BUSAYAKANON

THAMONWAN KHONGJUI

A PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF BACHELOR OF ENGINEERING IN BIOMEDICAL ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

ACADEMIC YEAR 2020

FACULTY OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

PROJECT CERTIFICATE

Project Title	Sleep Data Analysis for Respiratory Events Classification
Student Name	Miss Supasuta Busayakanon Student ID. 60011273
	Miss Thamonwan Khongjui Student ID. 60011277
Degree	Bachelor of Engineering in Biomedical Engineering

Project Advisor Signed: _____

(Asst. Prof. Dr. Suradej Tretriluxana)

Project Co-Advisor Signed: Ukkrit Jansri

(Ukkrit Jansri, PT, M.Sc)

Committee Signed: _____

(Asst. Prof. Dr. Suradej Tretriluxana)

Committee Signed: _____

(Dr. Wibool Piyawattanametha)

Committee Signed: _____

(Dr. Kasama Srirussamee)

Head of Department Signed: _____

(Asst. Prof. Dr. Sarinporn Visitsattapongse)

Project Title	Sleep Data Analysis for Respiratory Events Classification
Student Name	Miss Supasuta Busayakanon
	Miss Thamonwan Khongjui
Degree	Bachelor of Engineering in Biomedical Engineering
Project Advisor	Asst. Prof. Dr. Suradej Tretriluxana
Project Co-Advisor	Ukkrit Jansri, PT, M.Sc.
Academic Years	2020

ABSTRACT

Obstructive sleep apnea-hypopnea syndrome (OSAHS) is one of the sleep disorders that is likely to be increasing over recent years. Typical symptoms of this disorder are loud snoring and feeling tired during the day after a full night sleep. OSAHS is a serious medical condition that can lead to other health problems such as stroke, hypertension, heart attack etc. Therefore, to be diagnosed and get treated are very important. Normally, patient will undergo whole night polysomnography (PSG) which is a gold standard test for diagnosis sleep disorder. Respiratory events and sleep stages will be scored based on the PSG data by the eyesight of certified sleep doctors and sleep technicians. Our main objective is to use deep learning algorithm to provide the neural network that can automatically score the respiratory events from the PSG data in order to lighten the workload and facilitate both sleep doctor and technician who involve in the scoring process. In this project, 5 PSG data obtained from Excellent Center for Sleep Disorder, King Chulalongkorn Memorial Hospital are pre-processed in two types of image input and trained with pre-trained convolutional neural network models including AlexNet, VGG16, and ResNet50.

ACKNOWLEDGEMENTS

We would like to express our gratitude to our project advisor, Asst. Prof. Suradej Tretriluxana who gives us opportunity to work on Sleep Data Analysis for Respiratory Events Classification project. He always gives us support and recommendation with his patience, kindness and come up with the useful idea when we face some problems.

Beside our project advisor, we would like to thank Mr. Ukkrit Jansri who is our project co-advisor for his generous support and meeting coordination. Without powerful hardware from Mr. Ukkrit, this project would not be possible.

Our deep and sincere gratitude also go to Assoc.Prof.Dr.Naricha Chirakalwasan and Dr. Busarakum Chaitusaney from Excellence Center for Sleep Disorders, King Chulalongkorn Memorial Hospital for providing us important data that are used in this project and give us medical knowledge and advices related to our works.

Lastly, we would like to express our thanks and appreciation to Mr. Boonyawee Sirimaha, a junior from department of Electronics Engineering for helping and guiding us many steps of the way during coding. And he also introduced us different points of view which led to new idea.

Our completion of this project could not have been accomplished without assistance from these people. We are extremely grateful from the bottom of our hearts.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS/ABBREVIATIONS

Symbols/Abbreviations	Terms
OSA	Obstructive Sleep Apnea
CSA	Central Sleep Apnea
MSA	Mix Sleep Apnea
PSG	Polysomnography
AHI	Apnea-Hypopnea Index
AASM	American Academy of Sleep Medicine
AI	Artificial Intelligence
DNN	Deep Neural Network
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
RCNN	Recurrent Convolutional Neural Network

CHAPTER 1
INTRODUCTION

1.1 STATEMENT AND SIGNIFICANCE OF THE PROBLEMS

Whole Night Polysomnography (PSG) or sleep lab test is the gold standard test for diagnosis sleep disorders and sleep evaluation. This test is a recording of physical function change during sleeping. Parameters that PSG can measure are electroencephalogram (EEG), electrooculogram (EOG), chin and leg electromyogram (EMG), electrocardiogram (ECG), respiratory effort and pulse oximetry [1]. These whole night polysomnograms will be observed and analyzed by the expert known as sleep doctor or sleep technician. The process of the sleep data analyzation is called sleep scoring which is the process that the experts will carefully look through the polysomnogram to identify the sleep stages, respiratory events, limb movements, arousals, cardiac events and evaluate Apnea-Hypopnea index (AHI) to determine the severity of the patient from those signals.

Obstructive sleep apnea (OSA) is one of the sleep disorders that is likely to be increasing over coming years. According to the Estimation of the global prevalence and burden of obstructive apnea research from Lancet Respiratory Medicine [2], they estimated the prevalence based on American Academy of Sleep Medicine (AASM) 2012 diagnostic criteria and AHI threshold value and found that around 936 million adult people aged between 30 to 69 including men and women have mild to severe symptom (AHI $\geq$ 5 events per hour) of obstructive sleep apnea and around 425 million men and women adults aged between 30 to 69 have moderate to severe symptom (AHI $\geq$ 15 events per hour) of obstructive sleep apnea globally [2]. Top ten countries with the highest estimated number of obstructive sleep apnea are shown in figure 1.1 below.

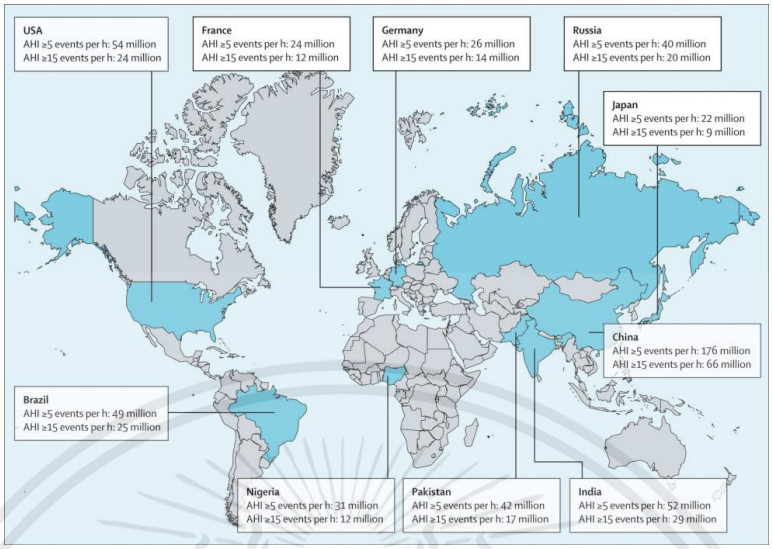


Figure 1.1 Top ten countries with the highest estimated number of OSA in 2019 [2].

In Thailand, the estimate prevalence of obstructive sleep apnea from this research is shown in table 1.1.

	Population aged 30 -69 years	AHI ≥ 5 events per hour	AHI ≥ 15 events per hour
Thailand	37,728,597	13,743,556 (36.4%)	5,531,503 (14.7%)

Table 1.1 Number of patients with OSA in Thailand 2019 [2].

According to the information mentioned above, if we assume that all of them underwent the polysomnography, it means that the sleep doctors and sleep technicians would have hard times using their eyesight to score all of the sleep data (8 hours of polysomnogram take about one hour of expertise's time to complete the score report). Therefore, to have the system that can automatically scoring the sleep data is necessary for reducing the workload of both doctor and technician and also has a benefit in eliminating the disagreement of the scoring between each person which may lead to differences and uncertain in diagnose result in some cases.

In recent years, Deep neural networks have achieved impressive performance and widely used in several applications as its ability in making decision and having human-like intelligence [3]. Consequently, this method is the most chosen to be use in medical field for development in sleep scoring system.

1.2 GOAL AND OBJECTIVE

1. To analyze the signals from polysomnography.

2. Using Artificial Intelligence (AI) to provide deep neural network that can automatically score the respiratory event of patient data.

3. To lighten the workload and facilitate both doctor and technician who involves in this field.

4. To eliminate the disagreement of the scoring between each person

5. To study and practice in MATLAB programming.

1.3 HYPOTHESIS TO BE TESTED

By using artificial intelligence network, the accuracy of model should be greater than 80%.

1.4 SCOPE OR LIMITATION OF THE STUDY

1. Only 5 patients' data are used in this study.

2. Research and develop the system or network for scoring respiratory events by using deep learning method as well as finding the suitable input type for it.

1.5 EXPECTED BENEFIT GAIN

1. The system or network can be use in real clinical diagnosis.

2. Develop MATLAB programming skill.

1.6 THESIS STRUCTURE

This thesis is divided in to 5 parts, including:

Chapter 1 Introduction consists of statement and significance of the problems, goal and objective, hypothesis, scope of the study and expected benefit gain.

Chapter 2 Review of theories and principles consists of theory that involve in this research including pathology of sleep disorder conditions, polysomnography, AHI information, AASM manual, information of artificial intelligence, transfer learning training and literature review.

Chapter 3 Research methodology states our method for pre-processing the data into two type of image input, Training and Testing process.

Chapter 4 Experiment results demonstration.

Chapter 5 Discussion and conclusion.

CHAPTER 2
REVIEW OF THEORIES AND PRINCIPLES

2.1 Pathology

Sleep Apnea is a sleep disorder in which breathing stops at least 10 seconds and starts repeatedly several times during the night. As the body is not getting enough oxygen, the brain will trigger patient to wake up and start breathing again. When the body is frequently triggered throughout the night, a strain will be occurred as the result and lead to snoring loudly, feeling tired even after a full night's sleep, hardly concentrate and lack of energy which are typical symptoms of sleep apnea. Moreover, the cumulative effects from these symptoms will develop into other serious health problems [4].

Sleep Apnea can be divided into 3 types which are obstructive sleep apnea (OSA), central sleep apnea (CSA) and complex or mixed sleep apnea [4].

Obstructive sleep apnea (OSA) occurs when there is a complete occlusion of the upper airway [5]. So, the air cannot flow to the lungs because of the obstruction. It can be occurred from many factors. For instance, accumulated of excess fluid in the laryngeal area, temporary block of falling back against the back of the throat of the tongue or soft palate due to too much relaxing of the muscles, obesity which leads to deposition of the fat in the upper airway and decreasing lung volume, inflammation of the tonsils that causes swollen and enlarged of tonsils [5]. During OSA, a breathing effort of the patient can be observed while an airflow does not exist.

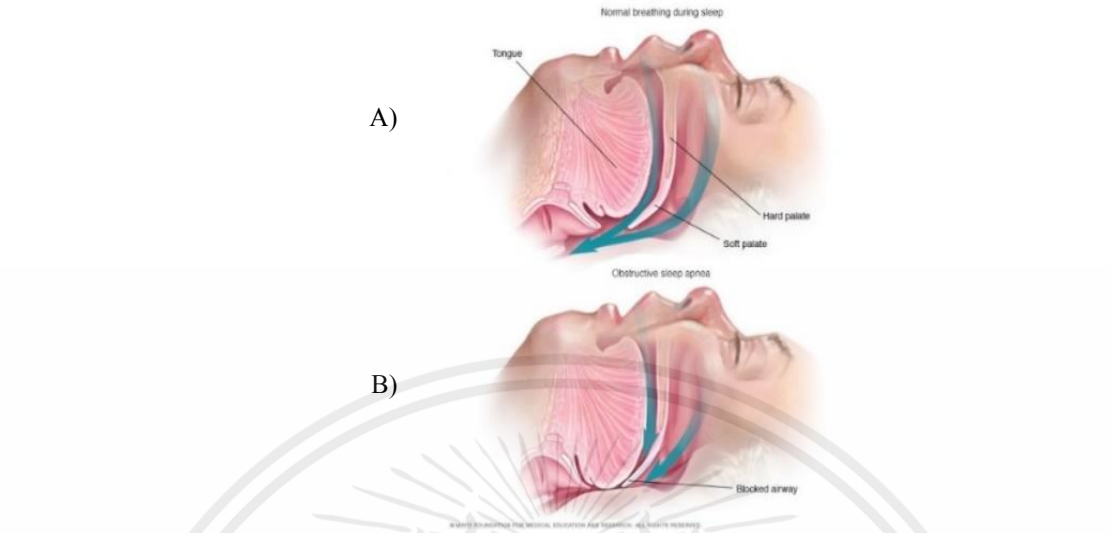


Figure 2.1 Normal and Abnormal Breathing: A) Normal Breathing B) Obstructive sleep apnea causing from the blockage of soft palate [4].

Central sleep apnea caused by dysfunction of the brain. It is the condition that the brain stops sending signal to the muscle that control breathing for the limited period of time leading to stop or abnormal breathing [5]. This type of sleep apnea is less common than obstructive sleep apnea and regularly occurs in the people who are already have certain health problems, for instance, problems that involve with brainstem (one of the main functions of brainstem is control respiratory system), atrial fibrillation, heart failure and stroke. In this period, both respiratory effort and airflow are absent [5].

Mixed or complex sleep apnea is a combination of both OSA and CSA [5]. Some patients are being treated for OSA with CPAP machines develop symptoms of CSA upon therapy. There are many effects occur to the body as consequences of sleep apnea. First, it causes low blood oxygen levels, high blood pressure or hypertension and heart problems (e.g., Atrial fibrillation) due to the repeated of stop breathing. Moreover, it also causes memory loss, depression, fatigue, breathing troubles and so forth [5].

Another type of sleep disorder is hypopnea. Hypopnea is the narrowing or partially block of the airways which leading to slow or shallow breathing. Majority of patients with sleep apnea condition also experience an episode of hypopnea before the apnea event is occurred [5].

From above information, we can notice that both sleep apnea and hypopnea are serious medical condition thus, it is necessary for the patients to get diagnose and treat as soon as possible.

2.2 Polysomnography

As mentioned earlier, polysomnography is an important tool for sleep disorder diagnosis as well as monitoring the effective of treatment. In detail, participants with suspected in sleep disorder condition need to attach with many electrodes over the entire night for measuring all the electrical activities and signals from their body.

Electroencephalogram (EEG) is a brain activity signal that can be obtain from the surface electrode placing around the scalp according to the 10-20 electrode placement system. This electrical signal uses for detecting the wakefulness and sleep [6]. Electrooculogram and Electromyogram are used to identify the REM (rapid eye movement) stage of sleep since it consists of both rapid eye movement and reduction in skeleton muscle tone. Electrodes of EOG are placed at left and right canthus, one slightly above than another one [6]. For EMG, two electrodes are placed near the chin to record movement of mentalis/sub-mentalis muscle which relate to muscle atonia in REM stage. And addition EMG electrode (such as Leg EMG) can be used for specific needs. For respiratory signal, it is an assessment of many parameter including 1.) oxygenation, 2.) ventilation, 3.) respiratory center drive and 4.) respiratory related sound [6].

1.) Oxygenation is a measurement of arterial oxygen saturation by pulse oximetry at the fingertip or ear lobe. Pulse oximetry also use to continuously measure the arterial blood pressure. The increasing in blood pressure is related to breathing disorder [6].
2.) Ventilation or air flow from mouth and nose can be measured from the thermistor which will detect the temperature different when the flow is presence and absent and infer the presence of breathing. Nasal catheter sensor is also used to measure pressure of airflow of the patient which suspected relate to OSA [6].
3.) Respiratory center drive needs to be assessed from ribcage-abdominal motion which obtain by measurement of respiratory inductive plethysmography or

other external sensor in order to differentiate sleep disturbance whether it is occurred from OSA or CSA. (there is no respiratory effort in CSA since the respiratory center is not activate the ribcage and abdominal muscle to move) Two sensor belt will be placed at thoracic and abdominal to detect the movement [6].

4.) Sound recordings use to observe the snoring sound to confirm sleep apnea condition in the patient. This signal can be record via the microphone or vibration sensor taped to a certain skin location such as at the level of the larynx, forehead etc [6].

Electrocardiogram is a recorded via three lead system to measure heart rate and rhythm. This measured signal will be combined with other measurement the assess the normal and abnormal physiological events during sleep [6].

All the electrode placements are indicated in figure below.

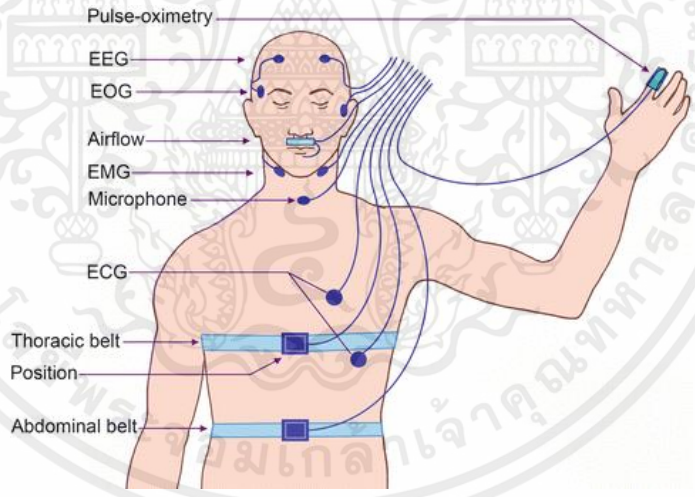


Figure 2.2 Electrode placement in polysomnography [7].

Other parameters such as continuous video monitoring using camera is also perform for assessment of parasomnias and nocturnal seizure disorder [6].

Then, all these recorded signals will be demonstrated to the doctor or sleep technician via specific program e.x. ProFusion PSG 4 (shown in figure 2.3) which will preprocess the signal by filtering and combine some signal channel as a different mode

reference such as 6 channels of EEG including F3, F4, C3, C4, O1, and O2, each need to be referenced with the contralateral mastoid which are M1 or M2. For example, two frontal EEG can be indicated by F3-M2 and F4-M1 [6].

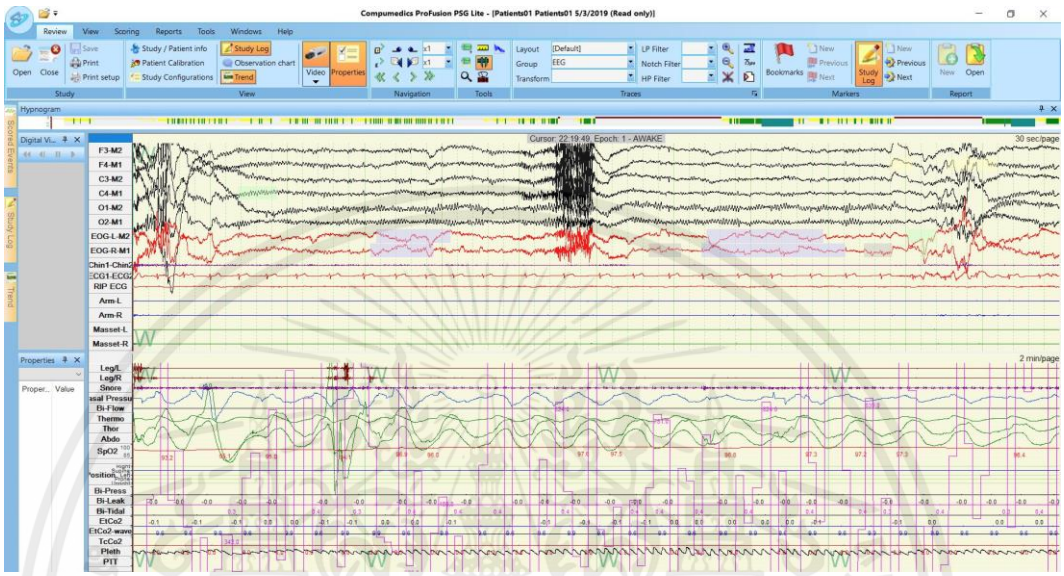


Figure 2.3 Polysomnogram in ProFusion PSG 4 program: the upper part is 30 second or epoch window for observe the sleep stage and the lower part is two minutes window for observe respiratory event.

Then doctor or sleep technician will visualize screening through all the signals and perform scoring according to American Academy of Sleep Medicine (AASM) standards manual [8] which will explain in the next topic. The respiratory event consists of 6 types which are hypopnea, Obstructive Apnea, Central Apnea, Mix Apnea, SpO2 Desat (low level of oxygen in blood) and Non (not scored or non-respiratory class).

Beside of basic polysomnography, split night polysomnography is sometimes used for both diagnosis and treatment purpose. It is performed by undergo basic polysomnography in the first half of the night, followed by a Continuous Positive Airway Pressure (CPAP) titration study in the second half. When OSA is occurred, CPAP device will determine the appropriate positive pressure to recover the condition of unbreathing [9]. It has been reported that CPAP is the most effective treatment available for OSA [10].

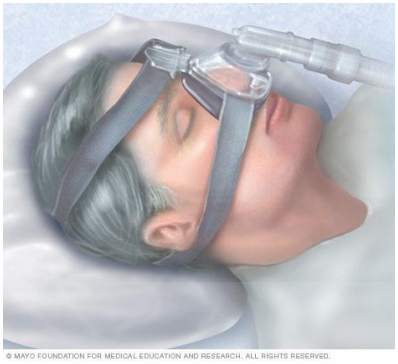


Figure 2.4 Continuous Positive Airway Pressure (CPAP) [10].

The severity of sleep apnea can be calculated from the Apnea-Hypopnea Index or AHI which is a sum of apnea and hypopnea events throughout the night per the total sleep time in hour. According to AHI, the severity can be divided in to 4 stages [11] as shown in the table below.

AHI	Severity
<5	Normal
5 - 15	Mild
15 - 30	Moderate
>30	Severe

Table 2.1 Severity of sleep apnea according to AHI calculation [11].

2.3 The AASM Manual for the Scoring of Sleep and Associated Events (Version2.4)

This manual is published to be the guideline for all users in performing the procedure of polysomnography (PSG) or sleep test and also for interpretation and evaluation of signals that obtained from PSG. This manual is like the universally standard that can create uniformity and reliability in the process of diagnostic and appropriate sleep disorders treatment for overall of sleep centers in this world [8].

For the identification of apnea and hypopnea, 6 main signals are required.

1. Airflow signal which is measured by the airflow sensor or nasal pressure transducer.

2. The temperature of the airflow can be measured by oronasal thermal sensor such as thermistors and thermocouples (the temperature is related to the presence of airflow).
3. The signals from abdominal and thoracic are recorded by thoracoabdominal belts which is indicated respiratory effort of the patients.
4. The saturation of blood oxygen is observed by using pulse oximetry.
5. The snoring will be monitored by using acoustic sensor such as microphone.

The criteria for the scoring of apneas and hypopneas and for measuring the duration of the event are clearly stated in this scoring manual. It states that the period from which the first breath that is noticeably decreased to the first breath that approaches the baseline breathing amplitude is considered as one event duration [8].

When there is an event with a reduction of peak signal more than or equal to ninety percent of the pre-event baseline from the signals that are measured by an oronasal thermal sensor with more than or equal to ten seconds of the duration will be scored as apnea [8]. The apnea will be further scored as obstructive if the inspiratory effort is occurred continuously throughout the whole period that the airflow is absent, as central if there is no inspiratory effort from the entire apnea duration and as mixed if there is no inspiratory effort in the first part of the apnea event and followed by the renewal of inspiratory effort in the second part of the apnea event [8].

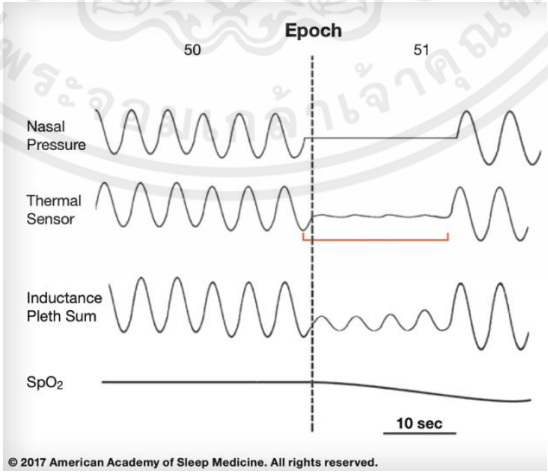


Figure 2.5 A respiratory event that should be scored as an apnea: The red bracket indicated the full duration of the apnea event [8].

When there is an event with a reduction of peak signal more than or equal to thirty percent of the pre-event baseline of signals that are measured by using an oronasal thermal sensor with more than or equal to ten seconds of the duration together with more than or equal to three percent of oxygen desaturation from the baseline of the prior event will be scored as hypopneas [8]. The hypopneas will be scored as obstructive if there is snoring and/or if there is increased inspiratory flattening of the nasal pressure and/or if there is related thoracoabdominal paradox occurred during the hypopnea event but not during the breathing of the prior event, and as central if there are no snoring, increased inspiratory flattening of the nasal pressure and related thoracoabdominal paradox occurred during the hypopnea event but not during the breathing of the prior event [8].

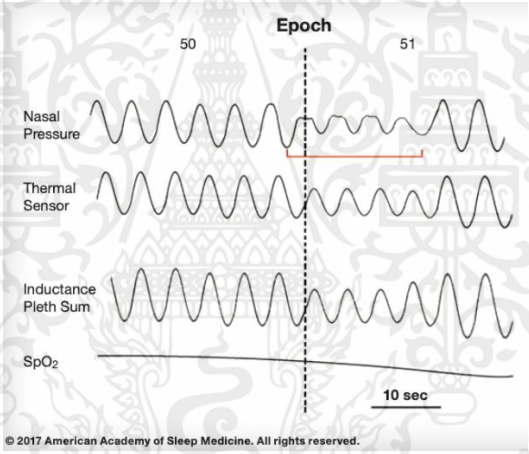


Figure 2. 6 A respiratory event that should be scored as hypopnea: The red bracket indicated the full duration of the hypopnea event [8].

2.4 Artificial intelligence

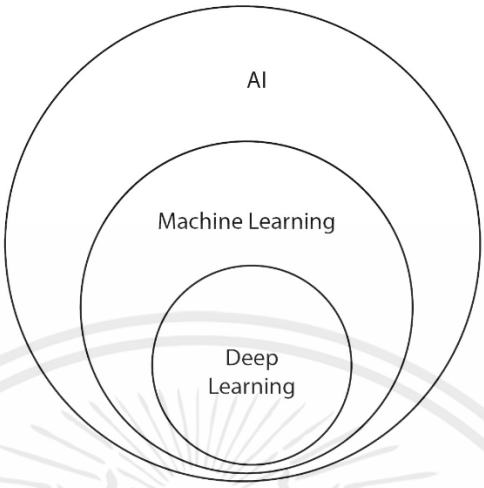


Figure 2.7 Relation of Artificial intelligence, machine learning and Deep learning [12].

Artificial Intelligence (AI) is a technique that enables machine to mimic human intelligence and behavior. It is programmed to have potentials like human for instance, logical thinking, understanding, problem solving and planning. The objective of AI is to maximize the chance for successfully achieving the specific goals [13]. AI plays a crucial role in healthcare. It can cope with very complex and large data and also has a wide range of applications starting from organizing the surgical schedules to improving diagnostic procedure [13].

For Machine learning (ML), it is subset of AI that provides systems the ability to automatically learn and improve from experience. Machine learning focuses on the development of computer programs that can access data and use it to learn for themselves [13, 14]. It achieves AI ability through algorithm trained with the input data [14]. Thus, the better algorithm, the more accurate the decisions and predictions will become as it processes with more data.

Deep learning (DL) is a type of machine learning which inspired by the structure of the real human brain [14]. Unlike ML, DL can learn to automatically extract the feature and perform classification and regression tasks directly from the input that pass through them. While ML need to have a manually extractor to select the relevant

features in the data before doing a classification [15]. As a result, DL is the preferred method in most cases since it allows people with non-expertise to process on it. However, DL required large amount of data in order to get more precise result [14, 15].

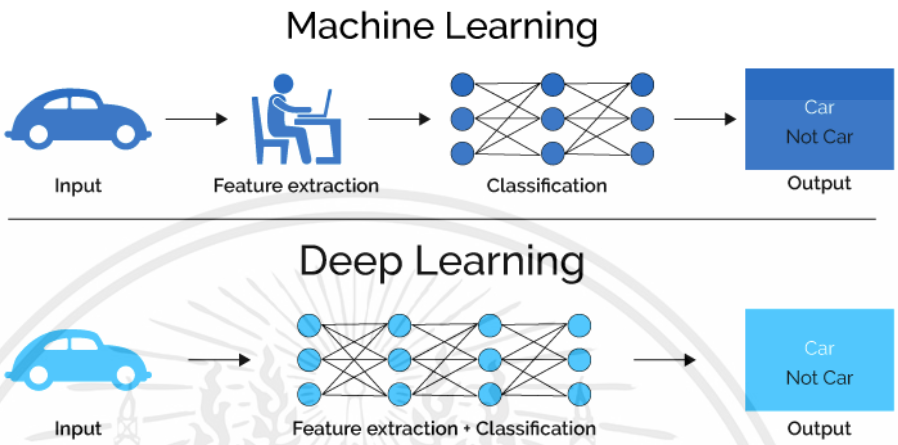


Figure 2.8 Difference between machine learning and Deep learning [16].

Deep learning is accomplished using a "neural network" architecture. Deep Neural Networks (DNN) are made up from layer of neuron which is the core processing unit of the network as structure of human brain and the term "deep" refers to the number of layers in the network [15]. The more layers, the deeper the network. Three main layers of DNN are input layer, hidden layer and output layer shown in figure 2.8. Most of the computational operation is done in hidden layer [15].

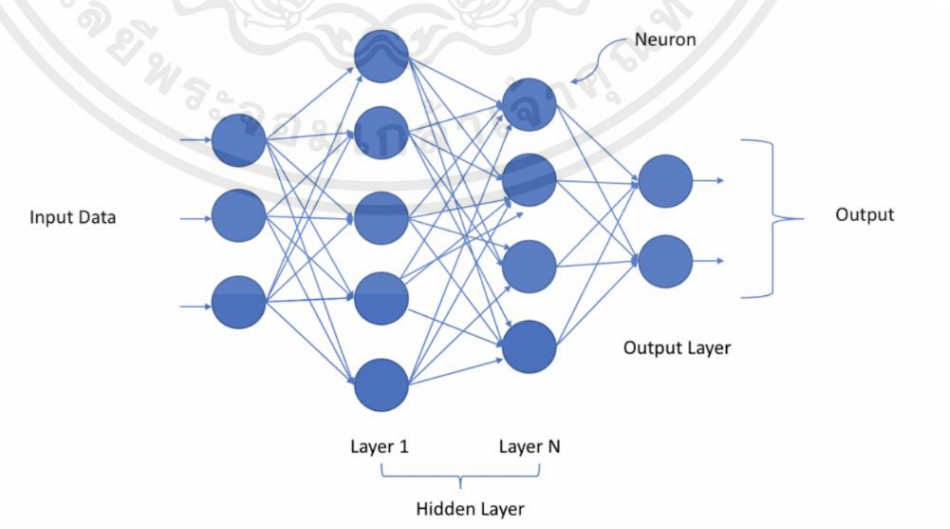


Figure 2.9 Architecture of deep neural network [17].

Each layer connects through the channels which these channels are assigned with a numerical value called weight. The inputs are multiplied to corresponding weight and their sum is sent to the neurons in hidden layer. These neurons are associated with bias value which is then added to the previous input sum. This value is passed through a threshold function called activation function (AF) [18]. The result of AF determines whether the neuron will get activate or not then an activated neuron will transmit data to the next layer. This manner of transmission is called forward propagation. The network also has a back propagation to adjust the weight in each channel and it will be repeated until weight are assigned such that the network can correctly predict in most of the cases [18].

There are several types of deep neural network. All of them using the same main layer as mentioned above but there are some specific layers in each type for improving the feature extraction task.

Two of the most popular deep learning networks are:

Convolutional neural network (CNN): Three of the most common layers in the first part of CNN are stack of convolution, activation or ReLU (non-linear activation function), and pooling. These layers are repeated several times to detect different features in the input data. For the second part, fully connected (FC) layers are used to perform the classification [15].

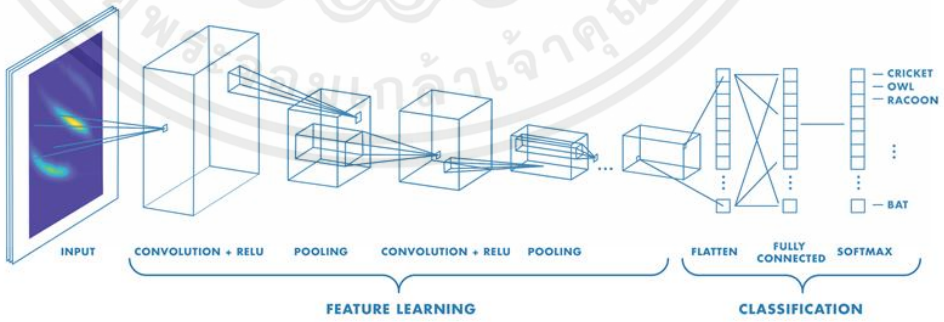


Figure 2.10 CNN architecture [15].

In convolution layer, the kernels or filters is a sliding window which contains a set of random weight vectors to convolved (computed dot product) and shift along with the

input data. Stride number is a parameter that indicate the size of shifting on the data after one convolute is done. The depth of the filters can be determined by the depth of the input [19, 20]. For example, RGB or color image will have the filter of depth 3 applied to it. More number of filters means more feature types can be extract and more computer resource. The output from convolution is sometimes referred to feature map and these feature maps are stacked to create the output volume which corresponding to the amount of filter [19, 20]. Mostly CNN is applicable to image or 2D information [15], although It can be use with 1D time series data as well.

Activation layer of CNN need to be a non-linear type function only, otherwise it will not be any learning of the network. Therefore, Rectifier Linear Unit or ReLU is the most commonly uses for activation function. Next pooling layer. This layer involves in down sampling of the features to reduce the computational cost by using sliding window to take either average value from the window (average pooling) or the maximum value (max pooling) [19, 20]. The max pooling is typically applied to control the overfitting and reduce the weight parameters. After several repeated of these three layers are finished, the output will be flatted to 1D vector for fully connected layer which generally followed by SoftMax classifier for processing the classification task [19, 20].

Various CNN network such as AlexNet, GoogLeNet, VGGNet, and ResNet have been utilized for many applications in medical field including medical image recognition and bio-signal classification.

In this study, we use AlexNet, VGG16 and ResNet50 as main networks for respiratory classification and here are the briefly explanation about these three networks:

• AlexNet: AlexNet was the first deep convolutional neural network proposed by Alex Krizhevsky and it has made this type of model become widespread and popular [21]. It is quite similar to LeNet model but AlexNet has a deeper network structure [22]. AlexNet network consists of 5 convolutional layers followed by 3 fully connected layer. In traditional neural network, the activation function is performed by Tanh or Sigmoid

function which can be saturate and causing the network to suffer with the Vanishing Gradient or VG problem (the gradient become smaller as the network become deeper resulting in failing to update the weight at the earlier layer which makes training process become more difficult or stop) [23, 24]. AlexNet uses ReLu instead of Sigmoid to solve this VG problem and also reduce the overfitting problem (network become more familiar with particular training set than other set) by applying the drop out layer after every fully connected layer. The max pooling is placed after the convolution and ReLu layer as a pooling layer [23, 24].

• VGGNet: VGG network is the improved version of AlexNet proposed by K. Simonyan and A. Zisserman from the university of Oxford [24]. It was born to reduce the number of parameters in convolutional layer and improve the training time [23]. The kernel size of 11x11, 5x5 and 3x3 with 4 and 2 strides are used to performs a convectional process in AlexNet. While in VGG, the kernel size is fixed at 3x3 with 1 stride only [22, 23, 24]. As the size of the kernel is reduced, the number of parameters is decrease and this resulting in faster learning, more robust to over-fitting which mean that the network can perform more effective to providing the better result when compare with AlexNet [23]. There are multiple variants of VGGNet including VGG16 (use in this study), VGG19 and etc. The difference among them is the total number of the layers in the network (not include maxpool and softmax layer) [23].

• ResNet: ResNet or Residual Network is developed by Kaiming He et al. It is purposed to increase the accuracy by increase the depth of the network while avoiding the vanishing gradient problem by using batch normalization and shortcut connection between every two convolutional layers as shown in the figure [23, 24]

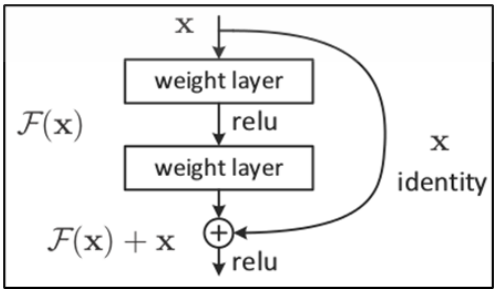


Figure 2.11 Residual Block in ResNet model [25].

This residual connection directly adds the value at the beginning of the block (x) to the output or the end of the block (F(x)+x). As the beginning value does not go through the activation, the derivative respect to the input becomes higher as a result, the VG problem is solved [23, 25, 26]. Batch normalization also reduce the problem. It is performed after each convolution and before activation to maintain the derivative [26]. The existing of these solution allowed CNNs to have deeper and deeper layer to perform the classification. However, the deeper the network, the more likely to causing the overfitting. Thus, it is importance to take this problem in to account to make sure that the training and classification result is efficient when using the deep network.

ResNet has the same kernel size as VGG and the network is ended with a global average pooling layer and only one layer of FC (1000-way fully connected layer) with a softmax classifier [25].

Recurrent Neural Network (RNN): RNN is another type of deep neural network. The name “Recurrent” indicates the operation that repeated over time on the input. This network is suited with the sequence data in which the order or the entire flow of the input is necessary to be considered [15]. Unlike a CNN, RNN can remember the state of the network between predictions like a real brain since the output of the next layer will be fed as input along with the new forward information in the current layer (RNN architecture is shown in figure 2.11). Therefore, the neurons can have an ability to stores and merge all information for processing the network [15, 27]

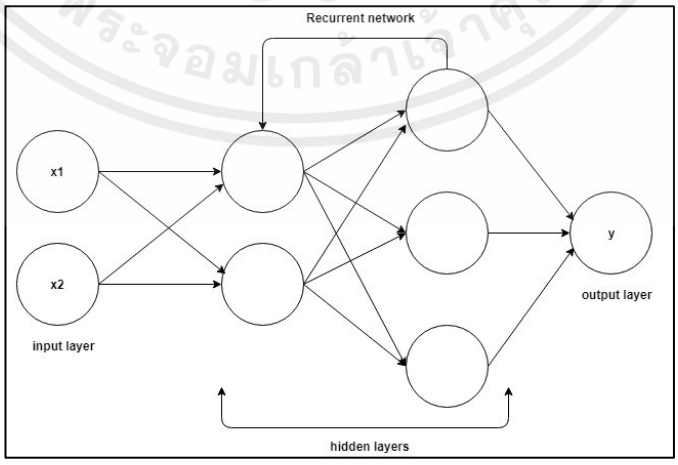


Figure 2.12 RNN architecture [28].

RNN are frequently used in speech recognition, handwriting recognition, and machine translation. It also has had a great success in healthcare application as well.

Two common types of RNN are Long short-term memory (LSTM) and Gated Recurrent Unit (GRU).

These two approaches are emerged to solve vanishing gradient problem of traditional RNN which occurred when the sequence data is long enough. During back propagation of the long sequence data, the gradients that are the values use to update the weight becomes small. As a result, it does not contribute the network to learn and cause RNN to forget what it seen [27, 29]. Thus, short-term memory solution was created to allow network to have potential in learning of long-term dependencies between time steps for long sequence data. In order to do that, both LSTM and GRU have an internal mechanism called gate which will make a decision on whether which data will be keep or throw away. The amount of gate is a different between LSTM and GRU [27, 29].

The Combination of CNN and RNN is called Recurrent Convolutional Neural Network (RCNN). This is a hybrid model that allow the network to have both feature extraction and preserve long-term relationship present in the input data at the same time [30].

2.5 Transfer learning

Transfer Learning or TL is a process of modifying and retaining the existing pre-trained network to train with a new data to let them can recognize more classes and adopt it for use in specific classification [31]. This training method is an efficient solution for many problems since it requires less data and computer time when compare with training from scratch [22, 32]. Components that needed for transfer learning are 1. Pre-trained network layers for modifying, 2. Training data and 3. Algorithm option which is a variable containing the options that control the behavior of the training algorithm such as amount of image in one training step or batch size, the maximum

number of iterations to take or max iteration, how many times in training or max epoch, learning rate and etc. [31].

2.6 LITERATURE REVIEW

2.6.1 SLEEPNET : Automated Sleep Staging System via Deep Learning (Biswal et al., 2017)

In this research, they developed a network called SLEEPNET for sleep staging by using various features of six channel EEG data (i.e. F3, F4, C3, C4, O1 and O2 each referenced with by M1 or M2) as an input with many different algorithms including basic machine learning methods such as logistic regression, tree boosting, multilayer perceptron and deep learning method which including Convolutional Neural Network (CNN), Long short term memory (LSTM) of Recurrent Neural Network (RNN) and Recurrent-Convolutional Neural Network (RCNN) which construct by connect final layer of CNN with 2 layers of RNN [30]. All of these approaches have been compared about the accuracy result and chosen to be the best model for use in clinical application. Over 10,000 patients data from the Massachusetts General Hospital (MGH) Sleep laboratory were obtained in this development. The system architecture of SLEEPNET consists of training module and deployment module as shown in figure below [30].

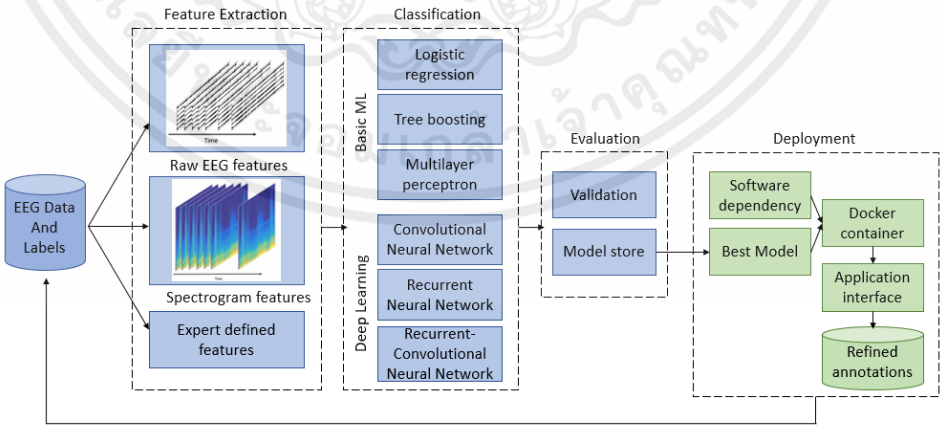


Figure 2.13 System architecture of SLEEPNET: Blue color corresponds to training model and green color corresponds to deployment module [30].

In training model, the input type of EEG in each classification algorithm is different for identify best feature and algorithm configuration for use. Different types of inputs including raw data EEG and spectrogram. Spectrogram is a represent of data in frequency domain by using Fast Fourier Transform (FFT). Raw EEG data input is segmented to be 30 seconds epoch and Spectrogram was segmented to be 2 seconds each by using window with length of 2 seconds and shift 1 second at a time, so 30 seconds epoch will be segmented in to 29 sub epochs [30]. Then these two forms of inputs will be labeled about the sleep stages and pass through deep neural network for feature extraction and classification. For machine learning, feature extraction of both time domain and frequency domain input was determined by sleep expert. Input data is split in to 9K patients for train and validation set and 300 patients for test set. After all the process in training model, the best performance network will be store for deployment. The result of classification is displayed with human prediction and input data on web interface [30].

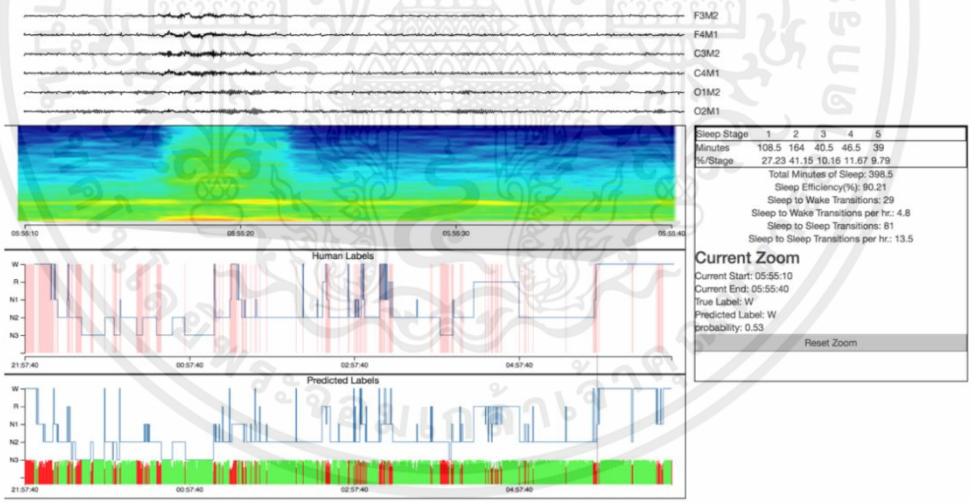


Figure 2.14 SLEEPNET's web interface: The first two panel are raw EEG data with corresponding to spectrogram. Next two panel are staging from clinical expert and SLEEPNET network respectively while right hand side panel is showing summary statistics [30].

From this research, RNN with expert-defined features has the best performance in sleep stages classification. The accuracy is 85.76%. RNN also has an outstanding

performance with accuracy greater than 79% in all type of input. This can confirm the power of deep learning over the traditional machine learning for classify EEG data [30].

	Expert-defined Features	Spectrogram	Waveform of raw data
Model	Accuracy		
LR	68.54%	66.54%	67.43%
TB	75.67%	71.61%	72.36%
MLP	72.23%	70.23%	69.56%
CNN	-	77.83%	77.31%
RNN	85.76%	79.21%	79.46%
RCNN	81.67%	81.47%	79.81%

Table 2.2 Accuracy of each network with different types of input [30].

2.6.2 Expert-level sleep scoring with deep neural networks (Biswal et al., 2018)

The objectives of this research team are to create the model that can automatically score sleep stages, sleep disorders breathing, and limb movements which are three main categories of sleep scoring. The data that they used in this research were derived from the Massachusetts General Hospital (MGH) sleep laboratory which are 10,000 in-lab polysomnograms and the Sleep Heart Health Study (SHHS) which are 5804 at-home polysomnograms [33].

For the methodology, they used the RCNN system which is the combination of convolutional neural network (CNN) and recurrent neural network (RNN). The CNN will be used to extract features from data due to local connections and parameter sharing through convolution operations. And all features will be passed down to the RNN to classify those three primary categories due to the abilities of the network that are suitable for sequential data and time series. They used the raw data and the spectrogram representation of data as input of model [33].

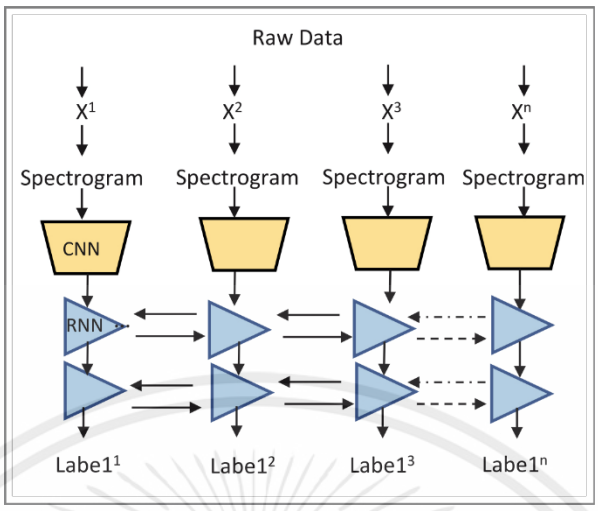


Figure 2.15 Architecture of the RCNN model [33].

For sleep stage scoring, EEG signals are scored to be five stages including wake (W), rapid eye movement (REM), non-REM stage 1 (N1), non-REM stage 2 (N2), and non-REM stage 3 (N3) in 30-second epoch without overlapping of signals [33].

For the detection of respiratory events, they used four signals which are thoracic movement, abdominal movement, oxygen saturation, and airflow. The obstructive apnea, central apnea, mixed apnea and hypopnea are combined together to be one class called apnea event class. And the rest data that occurred outside the respiratory events will be considered as the non-apnea event class. In conclusion, they classified the respiratory events in terms of presence or absence of respiratory event. To imitate the calculation of apnea-hypopnea index, events will be detected every one second consecutively without overlapping. Events that occurred in consecutive time windows are combined into a single apnea event [33].

For the detection of limb movement, they classified EMG signals as presence or absence of the movement. Movement will be detected every one second consecutively without overlapping. Movement that occurred in consecutive time windows are combined into a single apnea event [33].

The ratio between train sets and test sets are divided in to 90 percentage to 10 percentage. It means that 9000 polysomnograms of the MGH are used to train and 1000 polysomnograms are used for testing. For the data from the SHHS, 5224 of polysomnograms are used as train sets and 580 are used as test sets [33].

After trained all the train sets and tested with test sets, the RCNN model can produce the accuracies for sleep staging, sleep apnea, and limb movements of 87.6%, 88.2% and 84.7% respectively. They concluded that their deep learning model is accurate and scalable and able to widen access to crucial medical care [33].

2.6.3 Real-time apnea-hypopnea event detection during sleep by convolutional neural networks (Choi et al.,2018)

Choi and team proposed to find an alternative method for manual minute-by-minute scoring by developed a new approach for real-time apnea-hypopnea event detection using single-channel nasal pressure signal with convolutional neural network (CNN). Polysomnographic recording of 179 patients was used as a dataset, 129 PSG recordings were obtained from internal dataset of Central for Sleep and Chronobiology, Seoul National University hospital and another 50 data came from Multi-Ethinic Study of Atherosclerosis (MESA) dataset. They only included subjects whose ages were ≥ 20 years with absence of mental disorder, unstable vital sign, heart failure and sleep disorders other than sleep apnea and hypopnea syndrome (SAHS). From total amount of data, 50 recording were split for training, 25 for validation and the remaining 104 recording for testing [34].

Development process of this research can be classified in three phases

(shown in figure 2.16) :

1) Preprocessing signal and fed to CNNs [34].

2) Each segment is classified from fully connected layer after feature extraction of previous layers (convolutional and max pooling layer) [34].

3) validate classification events are used for AHI computation [34].

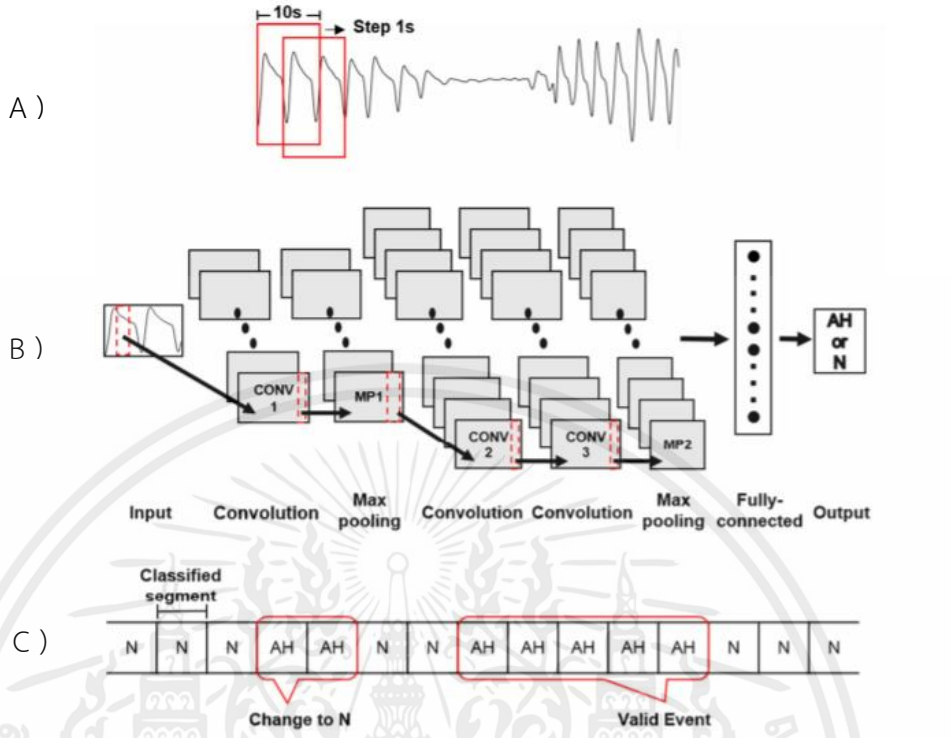


Figure 2.16 Overview of development process : A) Data segmentation of preprocessed signal, B) Feed input in to CNN and demonstrate of CNN architecture and C) Valid event detector [34].

In preprocessing process, all the signals were down sampled at 16 Hz to reduce the training time. Then operated with 0.01 Hz high pass filter and 3 Hz low pass filter to reduce baseline drift and noise. The adaptive normalization also used to save the small amplitude part which occur from the same sleeping posture for a long time. The preprocessed signals were then segmented from overlapping window. The window length is 10 seconds best performance from adjusting between 5 to 10 seconds) and shifted by 1 seconds at a time as it passes through the signal. Each 10 seconds segment were labeled into two classes which are AH (apnea-hypopnea) and N (non). The segment will be categorized to be AH only if at least 80% (best event detection from adjusting between 50 to 100%) of the segment occurred in apnea or hypopnea event (based on manual score) and others will be labeled as N. Class segments with more segments were randomly subsampled to avoid overfitting that may cause from imbalance data [34].

In this study, time series data were fed into 1D CNN which consists of three convolution layers, two max pooling layer and two fully connected layer (shown in figure 2.16B). All the hyperparameter that involved in training were determined from their experiment [34].

After testing overall 3008004 test segments are evaluated the validity as shown in figure 2.16C). If at least five continuous segments are classified as AH, it considered to be valid otherwise, it will be marked as N. These valid results are then use for AHI calculation to determine severity of sleep apnea condition. The accuracy obtained from confusion matrix (table 2.3) is about 96.6% [34].

		Reference	
	Class	AH	N
Predicted	AH	270372	40518
	N	63160	2633954

Table 2.3 Confusion matrix across all test data segments [34].

This research is another study confirmed the great performance of deep neural network for sleep scoring. They provide more convenient way to conduct scoring system by using only nasal pressure signal and also conclude that the using of overlapping approach is more precisely than non-overlapped one [34].

2.6.4 Sleep Staging from Electrocardiography and Respiration with Deep Learning (Westover et al., 2020)

The propose of this research is to find the alternatives to sleep stage scoring. Normally, the sleep staging is performed by using electroencephalography to acquire the electroencephalogram (EEG) and use them as data for scoring the sleep stage [35].

The alternatives that they chose to study in this research are electrocardiogram (ECG) and respiratory signals. Cardiorespiratory signals can be obtained in numerous ways from contact recording or non-contact radar-type application. Moreover, ECG is recorded in a number of medical settings. On the contrary, EEG is uncommon medical setting among ill populations [35].

The data that they used in this research were derived from the Sleep Laboratory at Massachusetts General Hospital (MGH) during 2009 to 2016 including 3 types of sleep tests which are diagnostic, full-night CPAP and split-night CPAP. They also used 1,000 recording data from Sleep Heart Health Study (SHHS) for the external validation dataset [35].

For the preprocessing of data, their deep neural network models will determine the sleep stage to 30-second time window by using information expanding 120 seconds on both sides of 30-second time window, making 270-second time window with 30 second time steps [35].

They created five networks based on five types of input signals: 1) ECG; 2) chest respiratory effort; 3) abdominal respiratory effort; 4) ECG+chest respiratory effort; 5) ECG+abdominal respiratory effort. Each network composed of convolutional neural network (CNN) which responsible for learning features combined with long-short term memory (LSTM) recurrent neural network (RNN). The structure of the CNN networks for a single type of input signal and the two types of input signal are different but the structure of LSTM networks is same for any types of input signal [35].

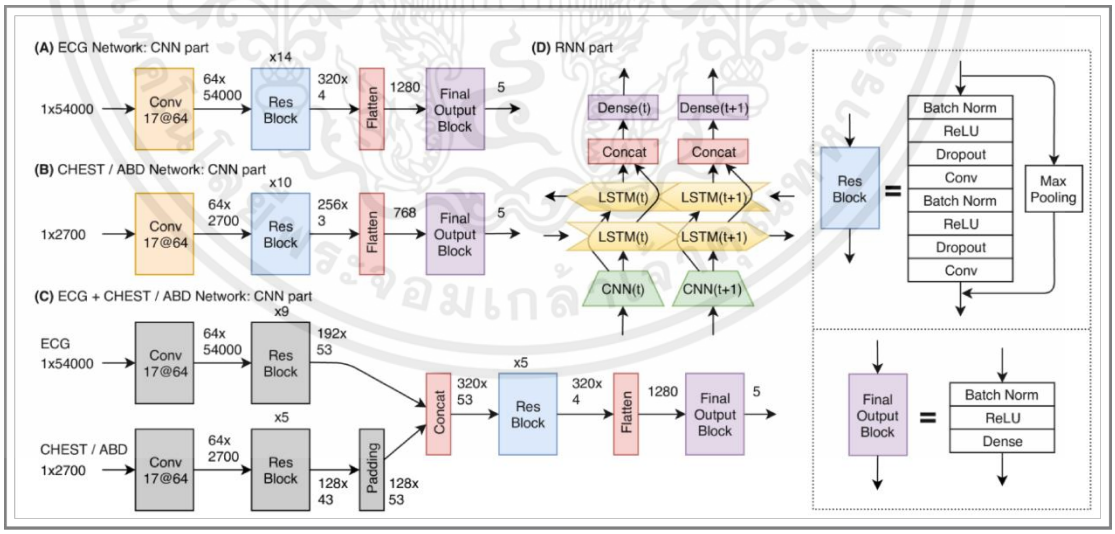


Figure 2.17 Architecture of the CNN and LSTM model [35].

The networks were trained by using a 32 mini-batch size, 10 of maximum number of epochs and learning rate of 0.001. They also reduced learning rate by 10%

of the loss in the validation set does not decrease for 3 consecutive epochs and they will stop training when the validation loss does not decrease for 6 consecutive epochs [35].

After training and evaluating the networks, the results showed that using both ECG and abdominal respiratory effort as input produced the best accuracy of prediction results: For MGH dataset, the network is correct in 78.9% of wake, 54.8% of N1, 68.4% of N2, 58.8% of N3 and 90.6% of REM. And for external validation dataset or SHHS dataset, the network is correct in 62.6% of wake, 59.1% of N1, 66.4% of N2, 43.4% of N3 and 88.2% of REM [35].

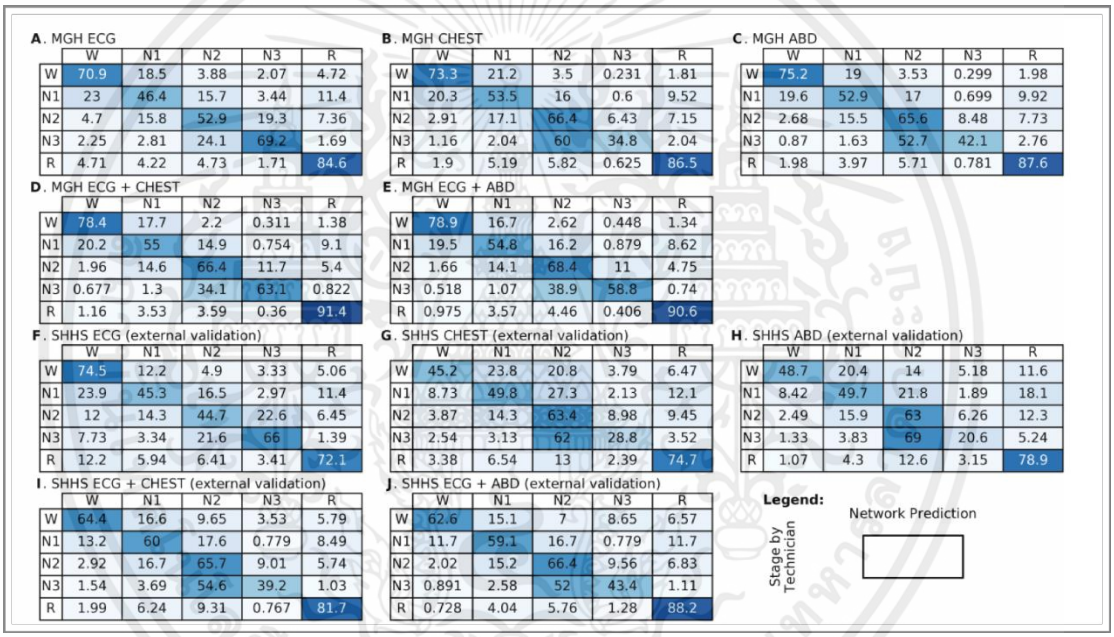


Figure 2.18 Confusion matrixes [35].

They believed that there is a possibility to stage sleep by using the electrocardiogram (ECG) and respiratory signals with deep neural networks. And this creates new possibilities in sleep research and for the applications when electroencephalography is not available [35].

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Introduction

The idea of this project is to research and develop a model or network to have a decision ability as clinical expert in sleep scoring as well as finding the suitable input type for it. In this study, we converted polysomnogram of the patient into 2D image. First, we start with time domain image input because we think the networks should learn the same representation of data as what's doctor seen in order to allow it to have the same ability as an expert in scoring. And later, we prepare input data in spectrogram or in frequency domain image and compare the result among two types.

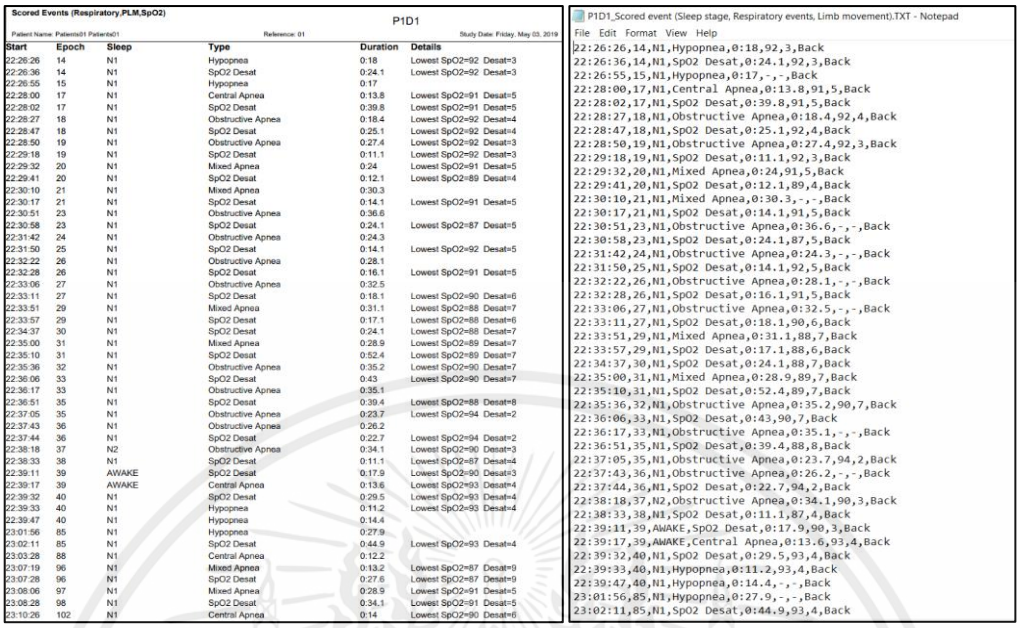
We use the transfer learning in training with three pre-trained networks including AlexNet, VGG16 and ResNet50. The networks are modified some layer to classify the respiratory events.

After the training is done, the testing data (exclude from training data) is classified by the network and the performance of network is then determined and compared by the accuracy.

3.2 Subject and Data Obtained

We have polysomnographic recording of five patients obtained from Excellence Center for Sleep Disorders, King Chulalongkorn Memorial Hospital. All the signals were recorded with the standard system and operate under the control of sleep expert. Patient number 1,2,3 and 5 are underwent with split night study. Patient number 4 underwent no split night study.

Both annotation (or scored file) of respiratory events (in ASCII and PDF file format) and amplitude of PSG signal (in ASCII and EDF file format) are exported from PSG by ProFusion PSG 4 software (Compumedics, Abbotsford, Australia) for use in this study.



Scored Events (Respiratory,PLM,SpO2)

P1D1

Patient Name: Patients01 Patients01 Reference: 01 Study Date: Friday, May 03, 2019

Start	Epoch	Sleep	Type	Duration	Details
22:26:26	14	N1	Hypopnea	0:18	Lowest SpO2=92 Desat=3
22:26:36	14	N1	SpO2 Desat	0:24.1	Lowest SpO2=92 Desat=3
22:26:55	15	N1	Hypopnea	0:17	
22:28:00	17	N1	Central Apnea	0:13.8	Lowest SpO2=91 Desat=5
22:28:02	17	N1	SpO2 Desat	0:39.8	Lowest SpO2=91 Desat=5
22:28:27	18	N1	Obstructive Apnea	0:18.4	Lowest SpO2=92 Desat=4
22:28:47	18	N1	SpO2 Desat	0:25.1	Lowest SpO2=92 Desat=4
22:28:50	19	N1	Obstructive Apnea	0:27.4	Lowest SpO2=92 Desat=3
22:29:18	19	N1	SpO2 Desat	0:11.1	Lowest SpO2=92 Desat=3
22:29:32	20	N1	Mixed Apnea	0:24	Lowest SpO2=91 Desat=5
22:29:41	20	N1	SpO2 Desat	0:12.1	Lowest SpO2=89 Desat=4
22:30:10	21	N1	Mixed Apnea	0:30.3	
22:30:17	21	N1	SpO2 Desat	0:14.1	Lowest SpO2=91 Desat=5
22:30:51	23	N1	Obstructive Apnea	0:36.6	
22:30:58	23	N1	SpO2 Desat	0:24.1	Lowest SpO2=87 Desat=5
22:31:42	24	N1	Obstructive Apnea	0:24.3	
22:31:50	25	N1	SpO2 Desat	0:14.1	Lowest SpO2=92 Desat=5
22:32:22	26	N1	Obstructive Apnea	0:28.1	
22:32:28	26	N1	SpO2 Desat	0:16.1	Lowest SpO2=91 Desat=5
22:33:06	27	N1	Obstructive Apnea	0:32.5	
22:33:11	27	N1	SpO2 Desat	0:18.1	Lowest SpO2=90 Desat=6
22:33:51	29	N1	Mixed Apnea	0:31.1	Lowest SpO2=88 Desat=7
22:33:57	29	N1	SpO2 Desat	0:17.1	Lowest SpO2=88 Desat=6
22:34:37	30	N1	SpO2 Desat	0:24.1	Lowest SpO2=88 Desat=7
22:35:00	31	N1	Mixed Apnea	0:28.9	Lowest SpO2=89 Desat=7
22:35:10	31	N1	SpO2 Desat	0:52.4	Lowest SpO2=89 Desat=7
22:35:36	32	N1	Obstructive Apnea	0:35.2	Lowest SpO2=90 Desat=7
22:36:06	33	N1	SpO2 Desat	0:43	Lowest SpO2=90 Desat=7
22:36:17	33	N1	Obstructive Apnea	0:35.1	
22:36:51	35	N1	SpO2 Desat	0:39.4	Lowest SpO2=88 Desat=8
22:37:05	35	N1	Obstructive Apnea	0:23.7	Lowest SpO2=94 Desat=2
22:37:43	36	N1	Obstructive Apnea	0:26.2	
22:37:44	36	N1	SpO2 Desat	0:22.7	Lowest SpO2=94 Desat=2
22:38:18	37	N2	Obstructive Apnea	0:34.1	Lowest SpO2=90 Desat=3
22:38:33	38	N1	SpO2 Desat	0:11.1	Lowest SpO2=87 Desat=4
22:39:11	39	AWAKE	SpO2 Desat	0:17.9	Lowest SpO2=90 Desat=3
22:39:17	39	AWAKE	Central Apnea	0:13.6	Lowest SpO2=93 Desat=4
22:39:32	40	N1	SpO2 Desat	0:29.5	Lowest SpO2=93 Desat=4
22:39:33	40	N1	Hypopnea	0:11.2	Lowest SpO2=93 Desat=4
22:39:47	40	N1	Hypopnea	0:14.4	
23:01:56	85	N1	Hypopnea	0:27.9	
23:02:11	85	N1	SpO2 Desat	0:44.9	Lowest SpO2=93 Desat=4
23:03:28	88	N1	Central Apnea	0:12.2	
23:07:19	96	N1	Mixed Apnea	0:13.2	Lowest SpO2=87 Desat=9
23:07:28	96	N1	SpO2 Desat	0:27.6	Lowest SpO2=87 Desat=9
23:08:06	97	N1	Mixed Apnea	0:28.9	Lowest SpO2=91 Desat=5
23:08:28	98	N1	SpO2 Desat	0:34.1	Lowest SpO2=91 Desat=5
23:10:26	102	N1	Central Apnea	0:14	Lowest SpO2=90 Desat=6

P1D1_Scored event (Sleep stage, Respiratory events, Limb movement).TXT - Notepad

File Edit Format View Help

```
22:26:26,14,N1,Hypopnea,0:18,92,3,Back
22:26:36,14,N1,SpO2 Desat,0:24.1,92,3,Back
22:26:55,15,N1,Hypopnea,0:17,-,-,Back
22:28:00,17,N1,Central Apnea,0:13.8,91,5,Back
22:28:02,17,N1,SpO2 Desat,0:39.8,91,5,Back
22:28:27,18,N1,Obstructive Apnea,0:18.4,92,4,Back
22:28:47,18,N1,SpO2 Desat,0:25.1,92,4,Back
22:28:50,19,N1,Obstructive Apnea,0:27.4,92,3,Back
22:29:18,19,N1,SpO2 Desat,0:11.1,92,3,Back
22:29:32,20,N1,Mixed Apnea,0:24,91,5,Back
22:29:41,20,N1,SpO2 Desat,0:12.1,89,4,Back
22:30:10,21,N1,Mixed Apnea,0:30.3,-,-,Back
22:30:17,21,N1,SpO2 Desat,0:14.1,91,5,Back
22:30:51,23,N1,Obstructive Apnea,0:36.6,-,-,Back
22:30:58,23,N1,SpO2 Desat,0:24.1,87,5,Back
22:31:42,24,N1,Obstructive Apnea,0:24.3,-,-,Back
22:31:50,25,N1,SpO2 Desat,0:14.1,92,5,Back
22:32:22,26,N1,Obstructive Apnea,0:28.1,-,-,Back
22:32:28,26,N1,SpO2 Desat,0:16.1,91,5,Back
22:33:06,27,N1,Obstructive Apnea,0:32.5,-,-,Back
22:33:11,27,N1,SpO2 Desat,0:18.1,90,6,Back
22:33:51,29,N1,Mixed Apnea,0:31.1,88,7,Back
22:33:57,29,N1,SpO2 Desat,0:17.1,88,6,Back
22:34:37,30,N1,SpO2 Desat,0:24.1,88,7,Back
22:35:00,31,N1,Mixed Apnea,0:28.9,89,7,Back
22:35:10,31,N1,SpO2 Desat,0:52.4,89,7,Back
22:35:36,32,N1,Obstructive Apnea,0:35.2,90,7,Back
22:36:06,33,N1,SpO2 Desat,0:43,90,7,Back
22:36:17,33,N1,Obstructive Apnea,0:35.1,-,-,Back
22:36:51,35,N1,SpO2 Desat,0:39.4,88,8,Back
22:37:05,35,N1,Obstructive Apnea,0:23.7,94,2,Back
22:37:43,36,N1,Obstructive Apnea,0:26.2,-,-,Back
22:37:44,36,N1,SpO2 Desat,0:22.7,94,2,Back
22:38:18,37,N2,Obstructive Apnea,0:34.1,90,3,Back
22:38:33,38,N1,SpO2 Desat,0:11.1,87,4,Back
22:39:11,39,AWAKE,SpO2 Desat,0:17.9,90,3,Back
22:39:17,39,AWAKE,Central Apnea,0:13.6,93,4,Back
22:39:32,40,N1,SpO2 Desat,0:29.5,93,4,Back
22:39:33,40,N1,Hypopnea,0:11.2,93,4,Back
22:39:47,40,N1,Hypopnea,0:14.4,-,-,Back
23:01:56,85,N1,Hypopnea,0:27.9,-,-,Back
23:02:11,85,N1,SpO2 Desat,0:44.9,93,4,Back
```

A) B)

Figure 3.1 Example of Scored Events: A) Annotation of respiratory event in PDF file format and B) in ASCII file format.

Scored event is consists of six headers including start time of the event, epoch of the event, type of respiratory event (Hypopnea, Obstructive Apnea, Central Apnea, Mix apnea and SpO2 Desaturation), Duration of the event and other details.

Amplitude of all signals can be exported in to two types as ACSII and EDF file format. ASCII file is raw data while the EDF is modified data. At first, we use the ASCII file along with the EDF file to perform our project but later we use only EDF file instead since we found that EDF file amplitude representation is more similar to the real output or what's doctor see on the screen (see figure 3.2)

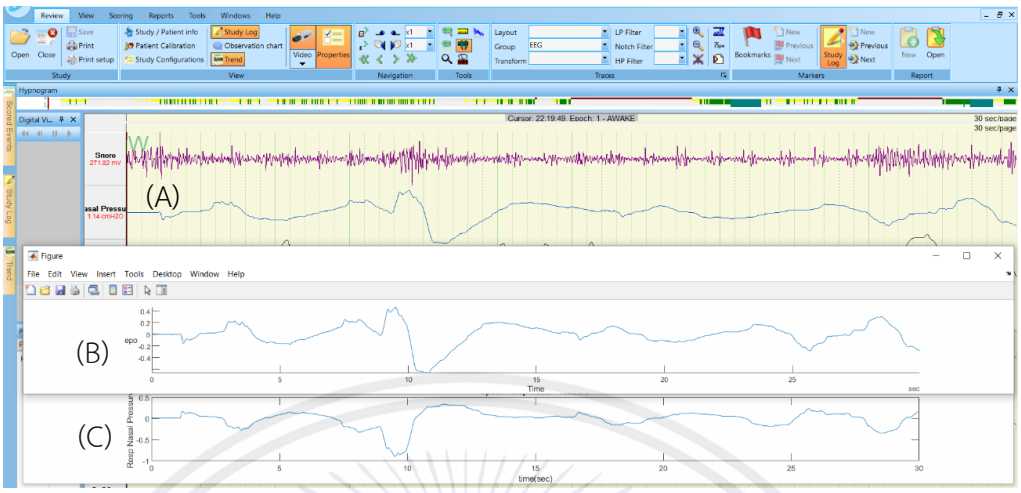


Figure 3.2 Comparison of real representation (A) with EDF (B) and ASCII (C) plot from MATLAB program.

3.3 Methodology

Step to achieve this project can be demonstrate in diagram below.

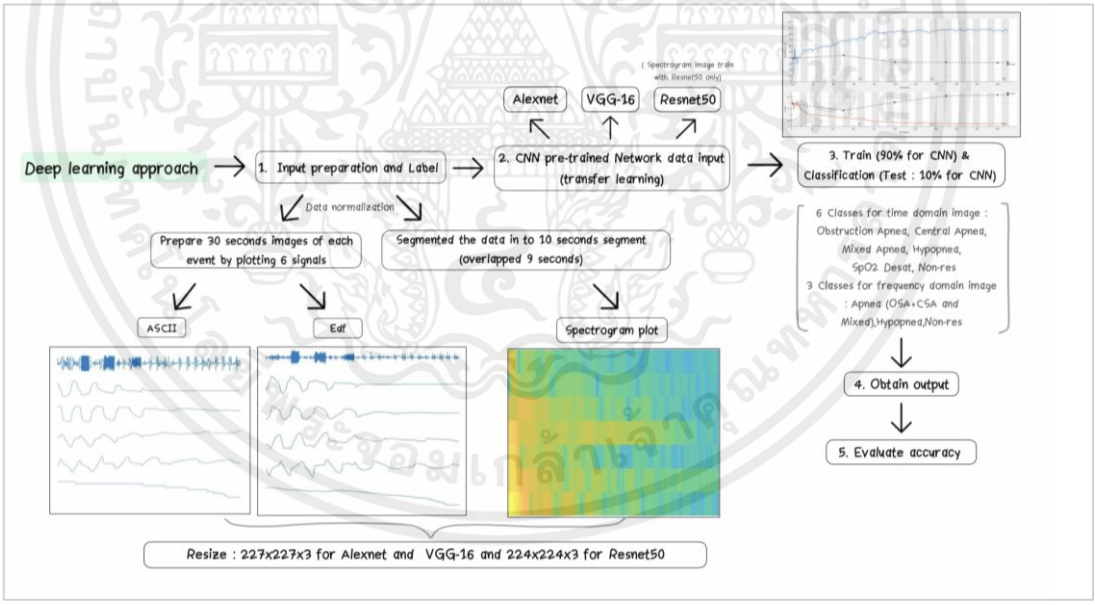


Figure 3.3 Methodology of project.

3.3.1 Input preparation and Labeling process

Before feeding the data of five patients into the deep neural network, the preprocessing need to be done first. In this project, the signal is converted into image by 2 approaches: plotting (time domain image) and spectrogram (frequency domain image). All of the operations are done in MATLAB environment.

3.3.1.1 Time domain image preparation

Images of time domain comprise of 30 seconds segment or epoch of six signals including Snore (256 Hz), Nasal Pressure (32 Hz), Thermistor (32 Hz), Thoracic (32 Hz), Abdominal (32 Hz) and Oxygen saturation or SpO2 (16 Hz). Each image is labeled according to the annotation file and store in its class folder.

First, we considered six classes of respiratory event including Obstructive apnea, Central apnea, Mixed apnea, Hypopnea, Oxygen Desaturation (SpO2 Desat) and Non-respiratory (normal). Later, we considered to reduce amount of class in to three by combining three types of sleep apnea (OSA, CSA and mixed apnea) into a single apnea class and cut SpO2 Desat out. Therefore, the remaining three classes are Apnea, Hypopnea and Non-respiratory.

For six classes, the amount of image in each class is unequal or unbalance and the equalization (balancing the image in each class to be the same amount) is not possible since there are only few images in some classes such as central apnea and mixed apnea. As a result, the equalizing will be useless because it will remain only small amount of data in training. Therefore, instead of equalizing, we decided to reduce images from the class with the largest amount (non-respiratory to be the same as the amount of the rest classes combined. However, this un-equalize problem can be solved as the number of class is reduced and this is the main reason why we modify the number of class into three. When all apnea classes are combined, the amount of image in each class is now enough to equalize and the total of image also remain in sufficient amount for training.

Both ASCII and EDF file format are used in preparing of time domain image (see figure 3.3) and their procedure can be illustrated in the flow chart of figure 3.4

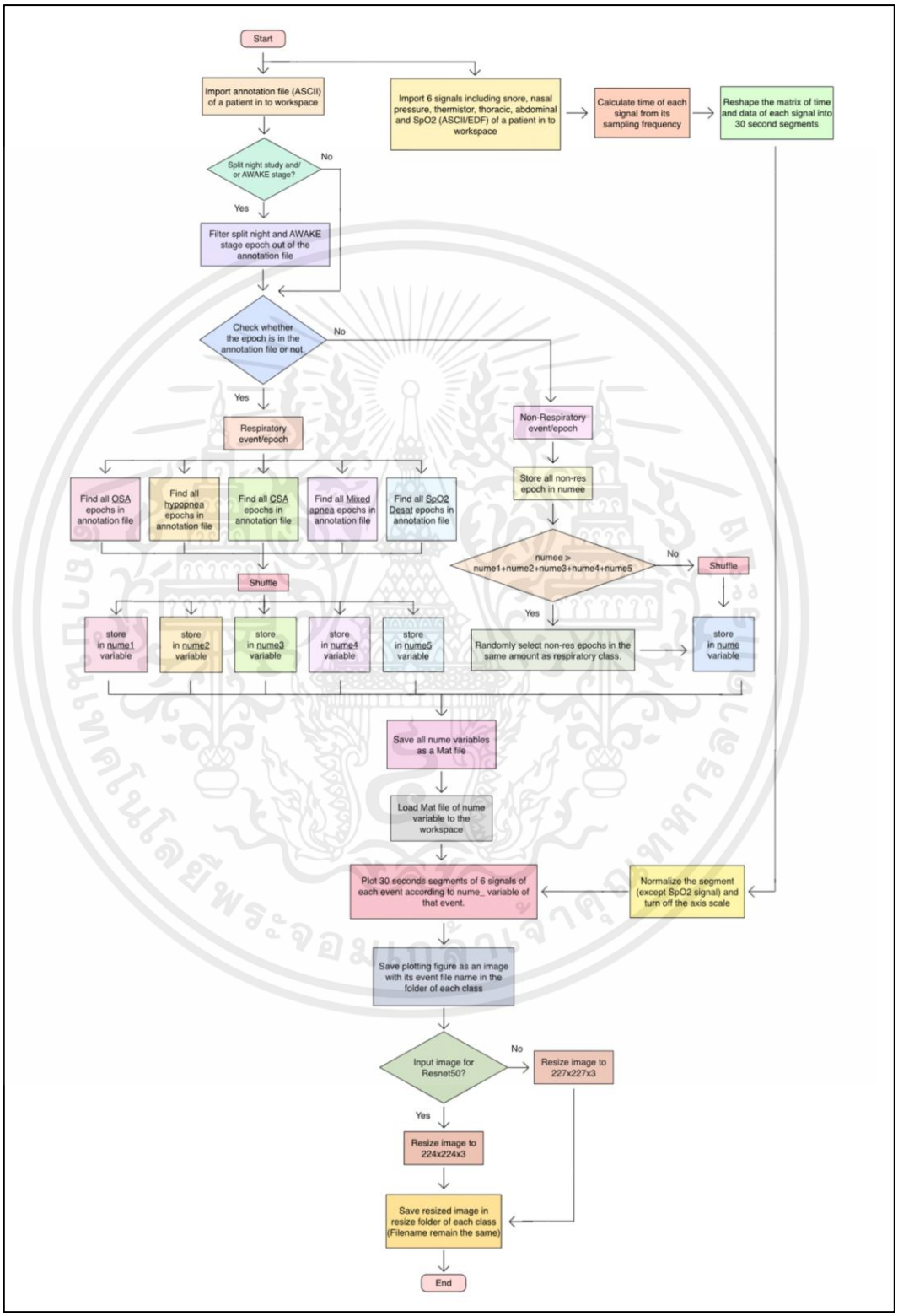


Figure 3.4 Flow chart for preparing time domain image (six classes)

All the process start from the importing of annotation file and six signals data of a patient (one patient at a time) either in ASCII and EDF file format into the workspace. Then the epoch number in 'AWAKE' stage and/or in CPAP study of the second half of the night (Nasal pressure and Thermistor signal are constant) are excluded from the annotation file either by coding or excel filteration. After rejected unused epochs, finding epoch number of each interest event from the annotation file (except for non respiratory event that we use epoch that are excluded from the annotation file), assort them and store in matrix variable named 'nume_' for later use. Epoch numbers of Non-respiratory, OSA, hypopnea, CSA, Mixed apnea and SpO_2 Desat are stored in nume, nume1, nume2, nume3, nume4 and nume5 respectively. Note that this process is done once a time to ensure that image order of different file types are the same when input to the network. For the signal data that have been imported previously, we calculated time of all the signal by using sampling frequency of each signal and segmented it into 30 seconds length. The amplitude of the signal also segmented to be 30 seconds as well. Thus, the total amount of 30 seconds segment of both time and amplitude matrix will be equal to the total epoch of that patient. After knowing epoch of each class (nume variable) and segmented data in proper length, the images of each class can be separately prepared by plotting 30 seconds segment of six signal against its time on the figure according to the nume variable of each class. We also normalize each signal segmented to make sure all data is plotted in the same scale (except for SpO_2 signal because most of the amplitude are repeated and will remain in zero after normalize) and all the axis is turned off before plotting because unequal value of time axis may interfere the training process and effect the accuracy result.

The plotted image is then saved in the folder of each class demonstrate in figure below.

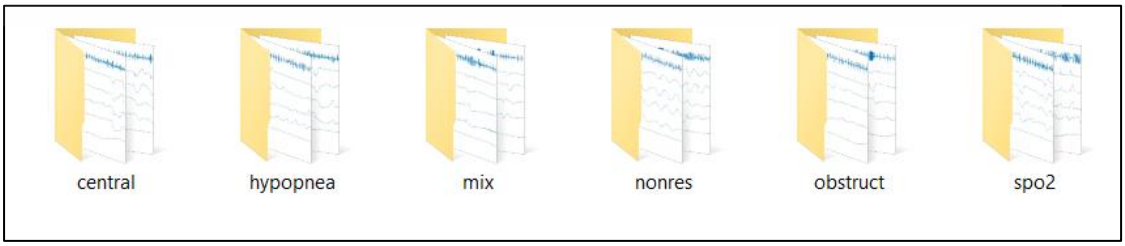


Figure 3.5 Image in each classes folder

And after the saving process is done, all images are resized to meet the requirement of the network. For alexnet , the require input size is 227x227x3 while VGG and Resnet is 224x224x3.

An example of image in each class before resize are shown in figure 3.6-3.17

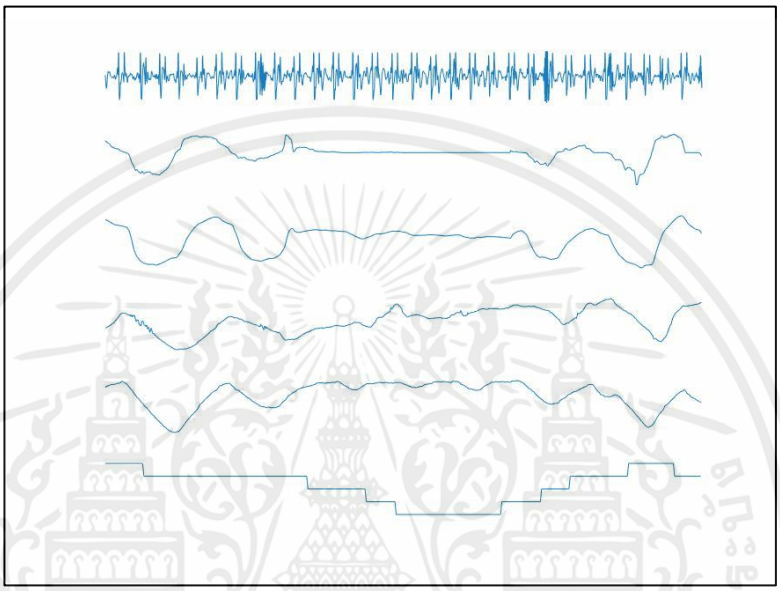


Figure 3.6 Obstructive apnea image from ASCII file format

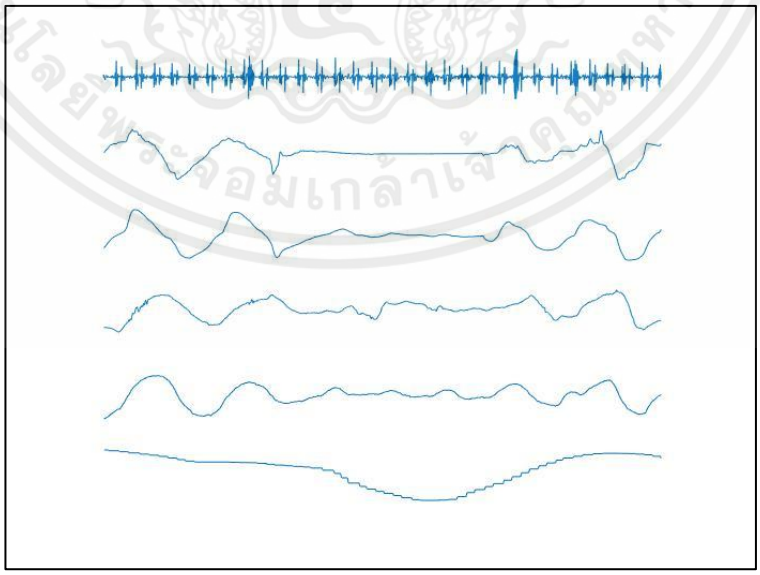


Figure 3.7 Obstructive apnea image from EDF file format

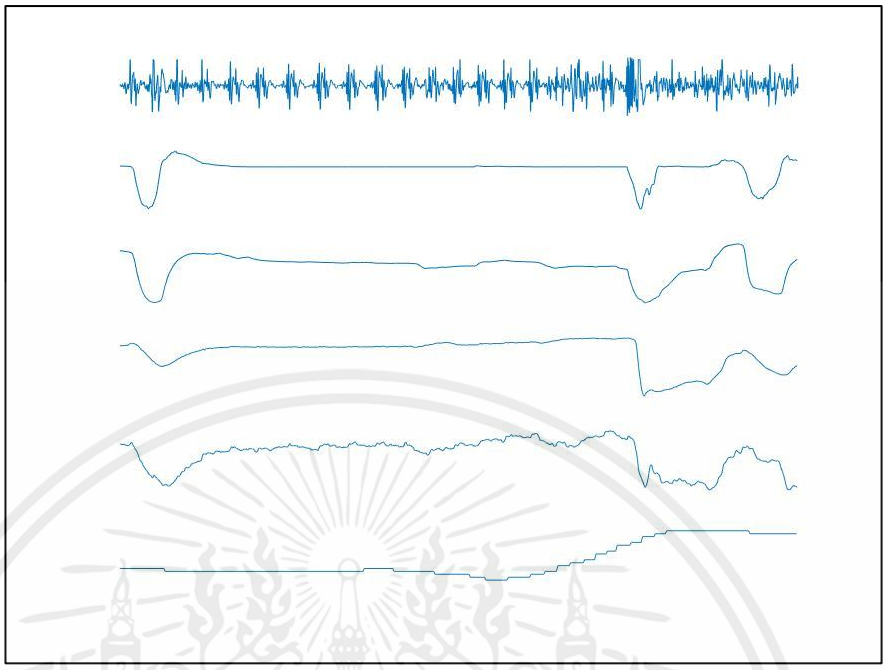


Figure 3.8 Central apnea image from ASCII file format

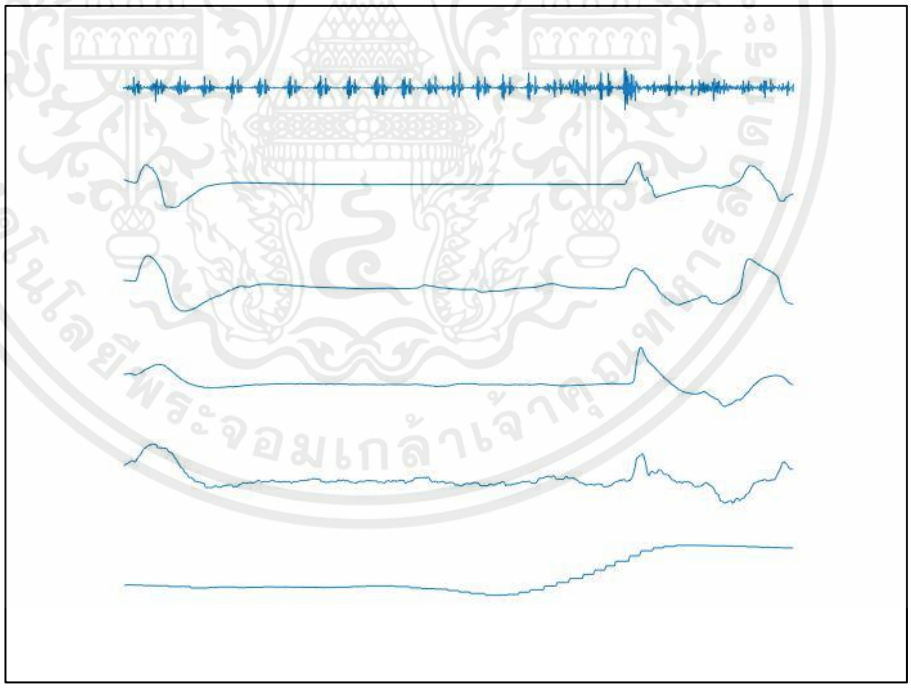


Figure 3.9 Central apnea image from EDF file format

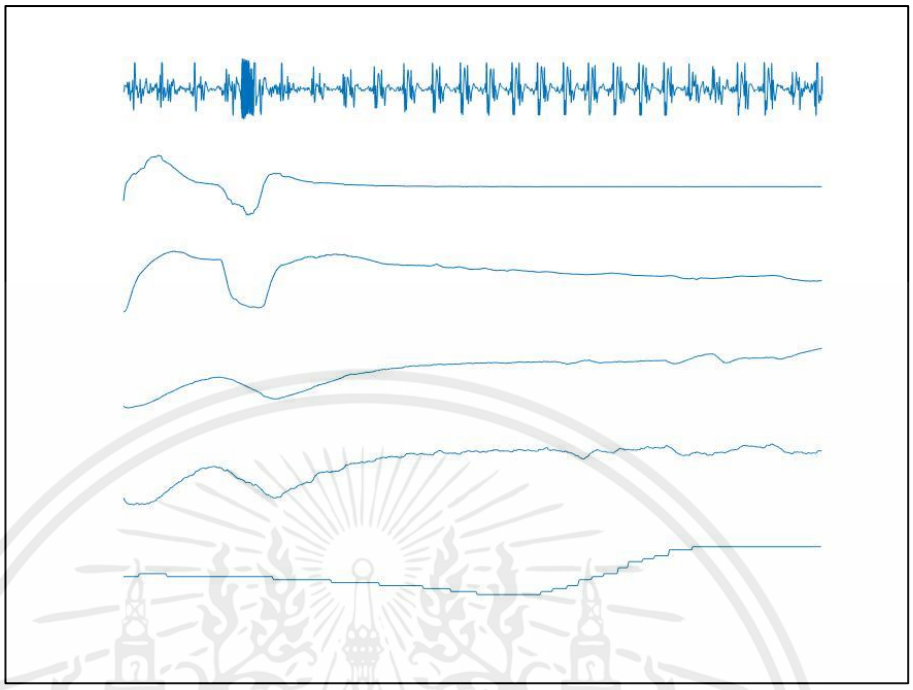


Figure 3.10 Mixed apnea image from ASCII file format

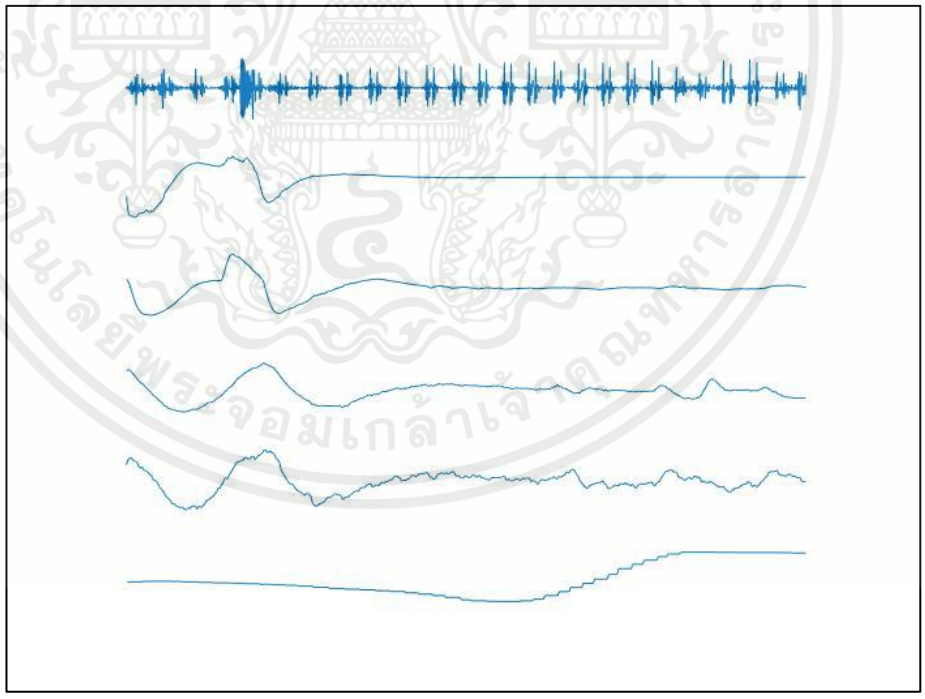


Figure 3.11 Mixed apnea image from EDF file format

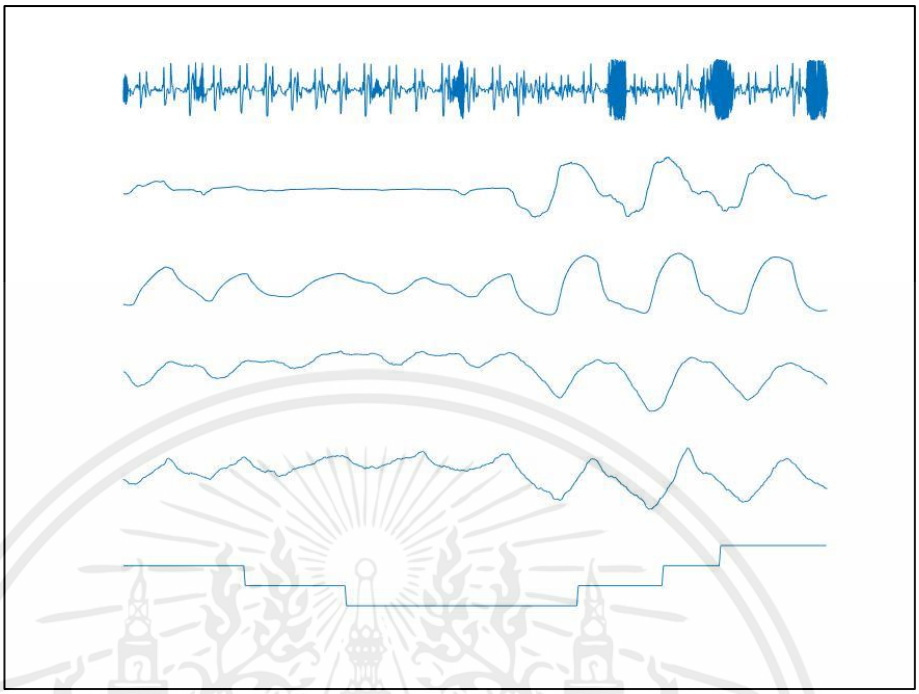


Figure 3.12 Hypopnea image from ASCII file format

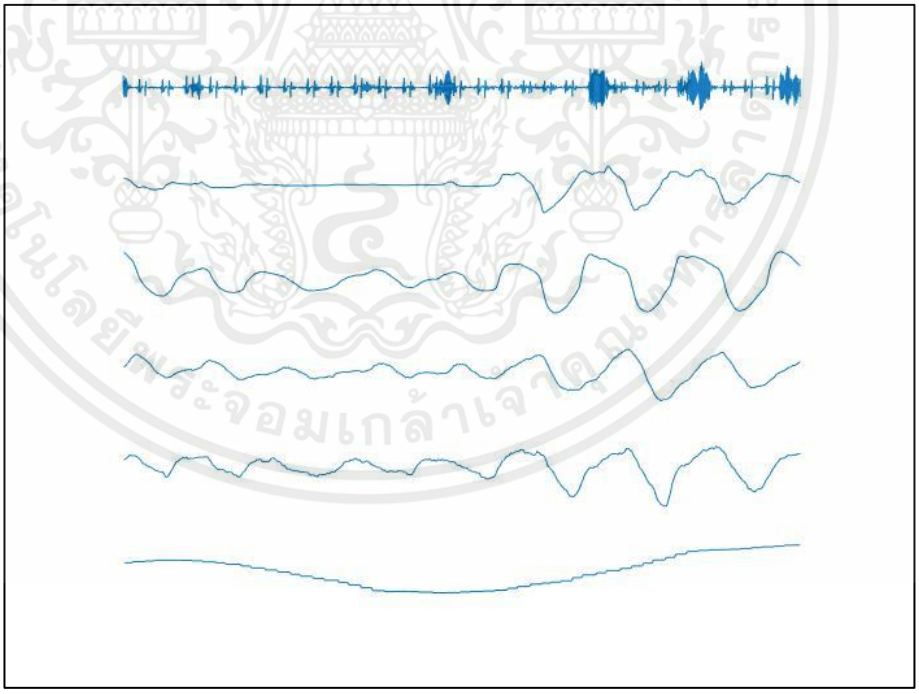


Figure 3.13 Hypopnea image from EDF file format

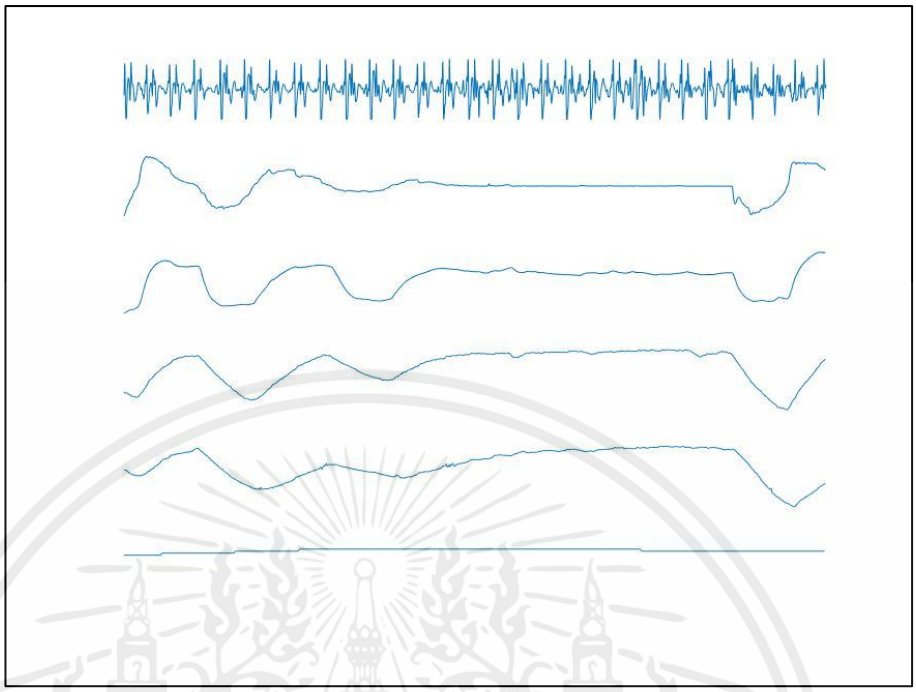


Figure 3.14 SpO2 Desaturation image from ASCII file format

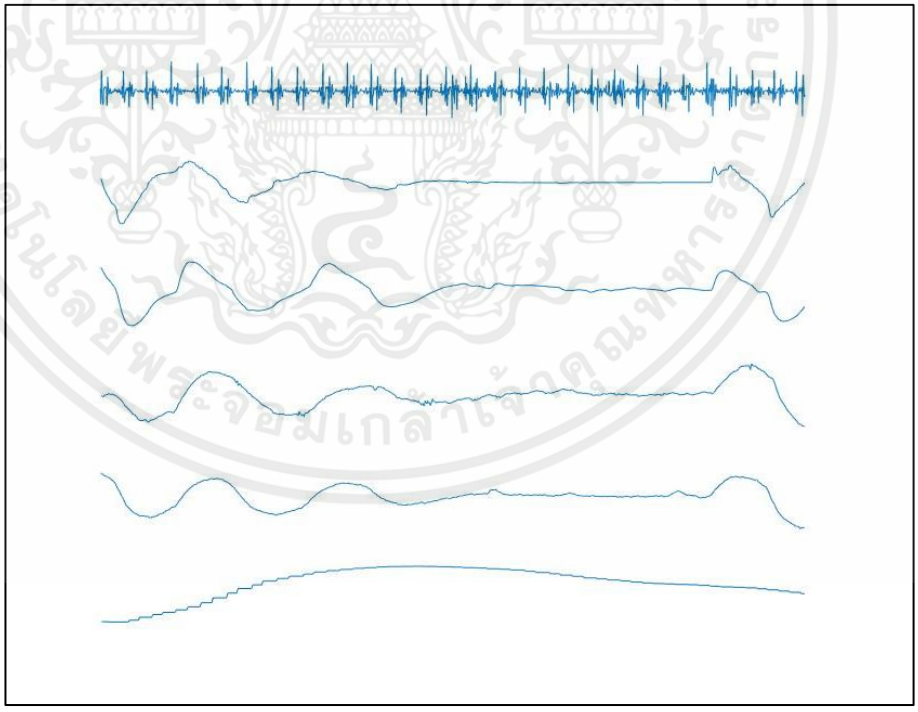


Figure 3.15 SpO2 Desaturation image from EDF file format

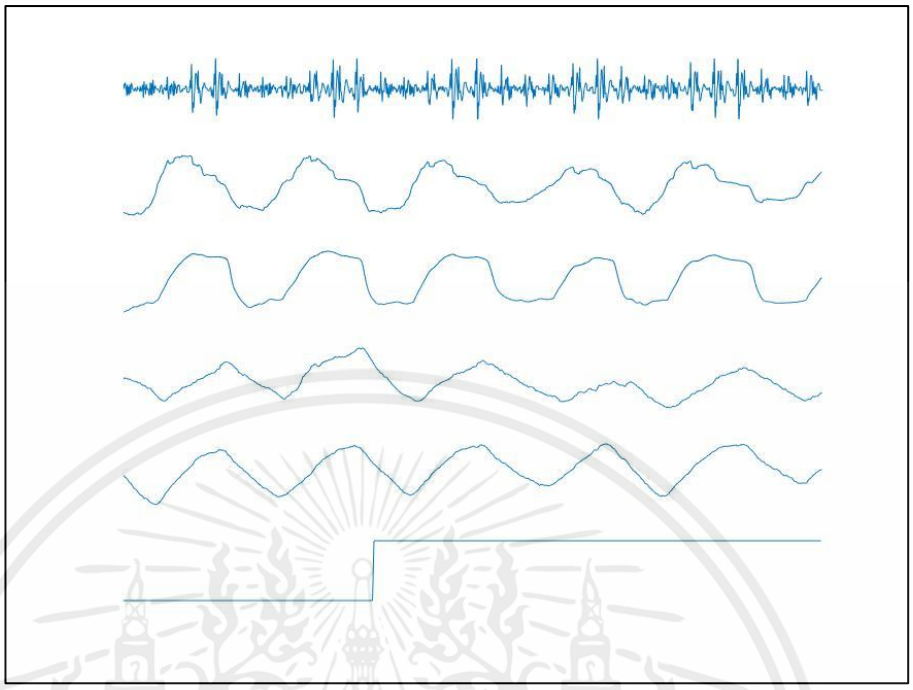


Figure 3.16 Non-respiratory image from ASCII file format

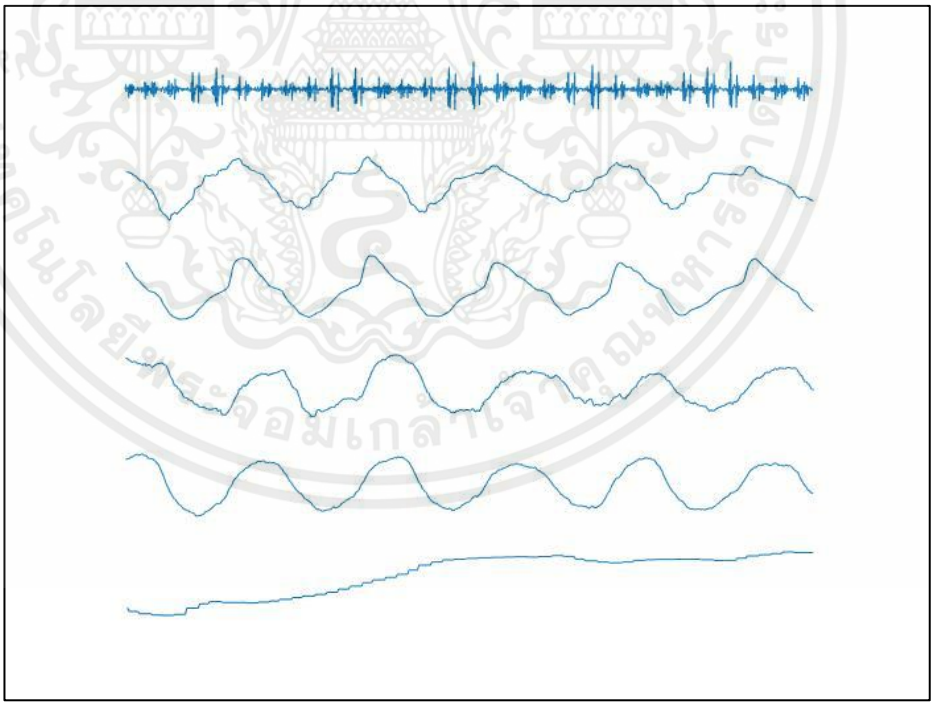
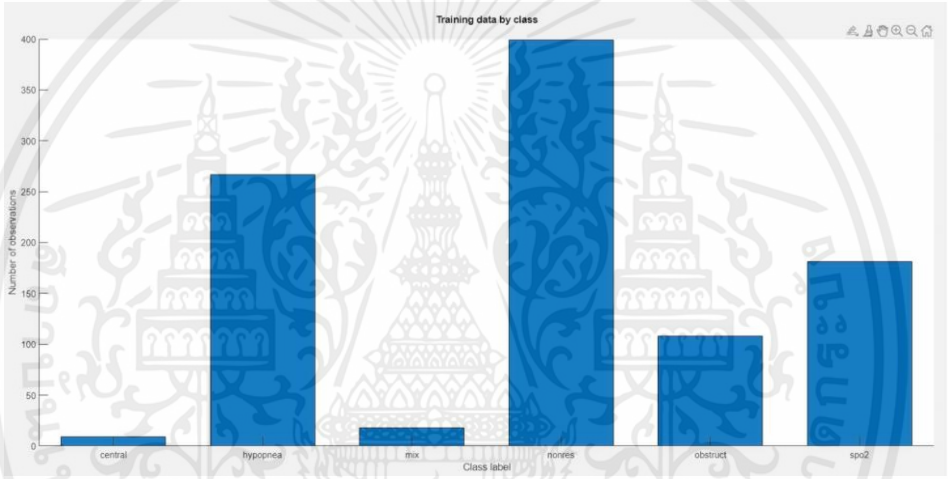
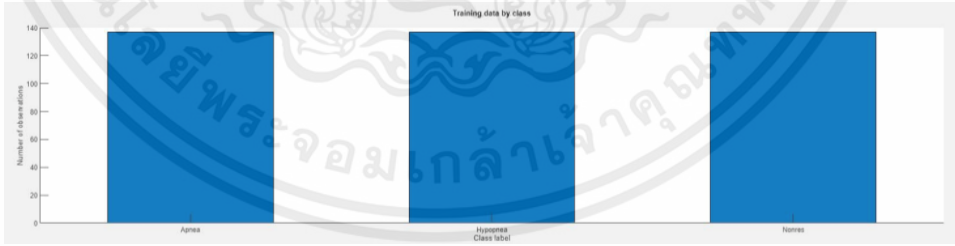


Figure 3.17 Non-respiratory image from EDF file format

For 3 classes image preparation, the image in three folders (OSA, CSA and Mixed apnea) are combined as one folder and the equalization were then applied based on the class with minimum data by coding in MATLAB (note that the order of images for training with three networks are the same after equalizing). The class with the minimum data is apnea which contain 217 images in total. Therefore, the overall dataset for 3 classes of time domain image are 217 samples of apnea, 217 samples of hypopnea and 217 samples of non-respiratory while the samples in 6 classes remain unbalance.



A)

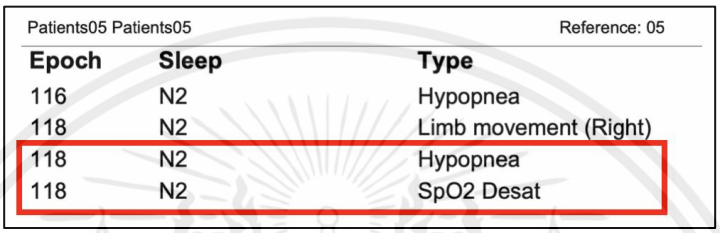


B)

Figure 3.18 Amount of training data for A) 6 classes (unbalance) and B) 3 classes (balance).

Although the time domain images are similar to what's doctor seen on screen but we faced with many problems which resulting in low accuracy (see result in chapter 4). The first problem is the duration of the segmented data is too long. To clarify, some

epochs or 30 seconds duration can be scored to have more than one event. Hence, it results in repeating of the same image in different classes which can cause the network to be confused and eventually make a wrong prediction. Another problem is a low color content of image. Since most area of this type of image are in white color, it is difficult for the network to classify the respiratory events from the convolute of the filter with only blue line and white background.


Patients05 Patients05 Reference: 05

Epoch	Sleep	Type
116	N2	Hypopnea
118	N2	Limb movement (Right)
118	N2	Hypopnea
118	N2	SpO2 Desat

Figure 3.19 Problem of using 30 seconds duration (several events in one epoch).

To solve all these problems, we come up with another type of image input called frequency domain image or spectrogram plot. A detail and preparing process of this type of image will be explained in the next topic.

3.3.1.2 Frequency domain image (Spectrogram) preparation.

According to the highest accuracy that is belong to the combination of ResNet50 with EDF file format (see chapter 4), the frequency domain images are plotted from frequency domain value, generated by Short Time Fourier Transform or STFT via MATLAB command, of individual signal (from six) of EDF file format and trained with ResNet50 only. And unlike time domain image, frequency domain images are the segment of 10 seconds preparing by using 10 seconds width window that will shift 1 second at a time (overlapped for 9 seconds) once it passes through the sequence data. The length of 10 seconds was chosen according to the AASM standard where they indicated that the event (both apnea and hypopnea) will be valid if it undergoes equal or beyond 10 seconds [8] and we labeled each 10 seconds segment using event of scored or annotation file.

The normalization of frequency domain image is difference from the time domain image. Instead of normalized each segment, the whole length of the signal of each patient

was underwent with Z-Score normalization before the segmentation process. This normalization method was chosen because it can provide the standard scale of all patients and reduced the significant of highest and lowest value that may cause by the movement of the patient and from the noise of the sensor.

Flowchart of frequency domain image preparation process are indicated below.

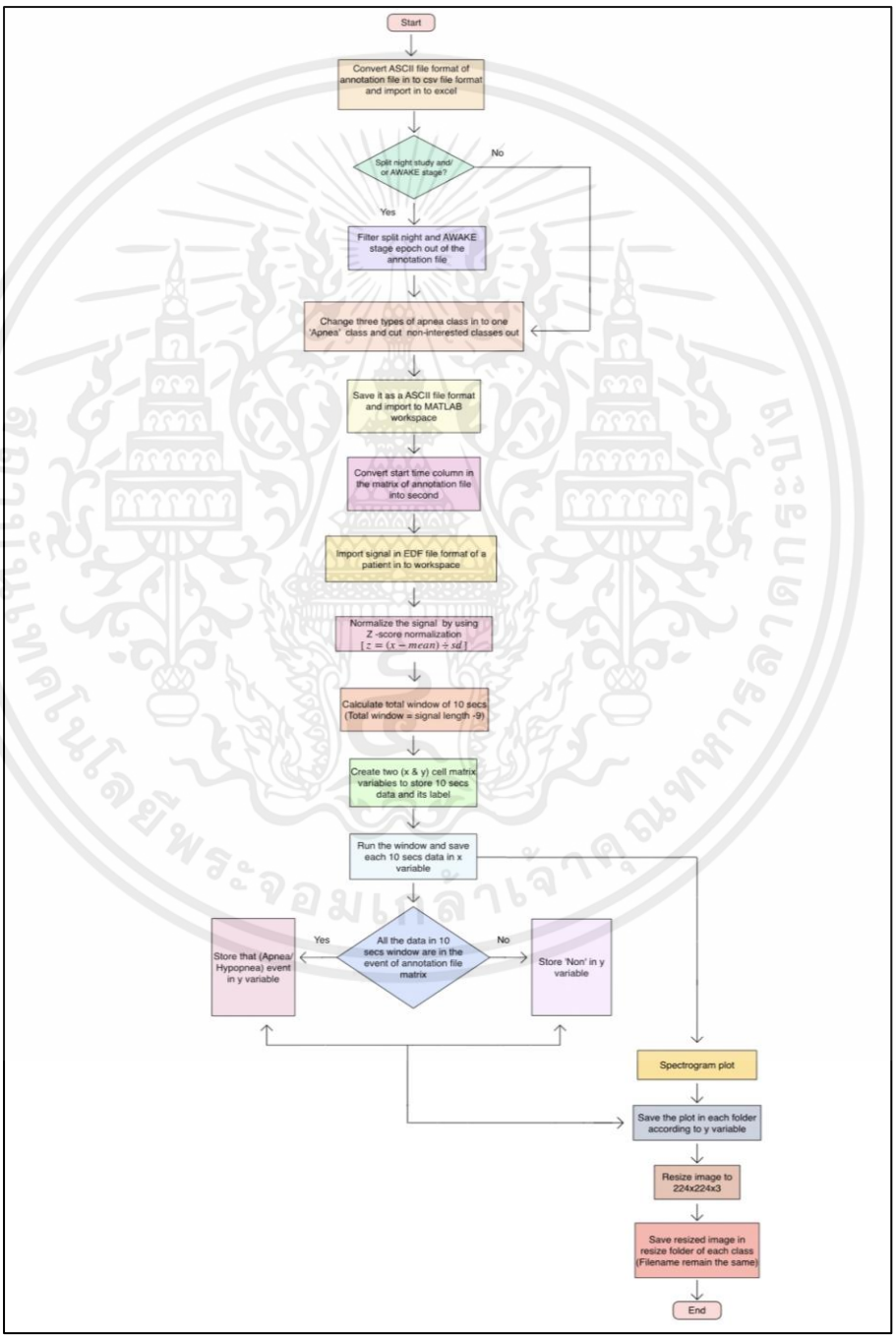


Figure 3.20 Flow chart for preparing frequency domain image.

First, the unwanted epochs (split night and awake stage) and non-interested classes were removed from the patient's annotation file by using excel. Then all three classes of apnea were renamed in to 'Apnea' class before imported into MATLAB-workspace. The processes are now separated into two parts, the preparation of segment data for spectrogram plotting and the labeling part.

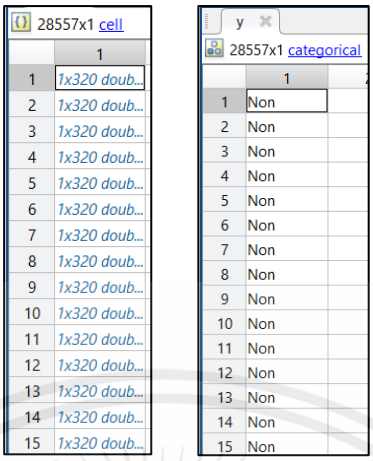
Total amount of 10 seconds sliding window or total amount of image were calculated based on the total length of the signal after excluded the split night study and 90% overlapping. Each window with its time stamp ran on the imported signal to obtain the data and stored them in X variable.

For labeling part, we created score table which consists of three headers: 1.) Start time in second unit (calculated from the difference between start time of recording and start time of the event). 2.) Type of respiratory event including Apnea, Hypopnea and Non-respiratory and 3.) Duration of the event.

	Start	Event	Duration
1	397	Hypopnea	18.0000
2	426	Hypopnea	17.0000
3	491	Apnea	13.8000
4	518	Apnea	18.4000
5	541	Apnea	27.4000
6	583	Apnea	24.0000
7	621	Apnea	30.3000
8	662	Apnea	36.6000
9	713	Apnea	24.3000

Table 3.1 Scored table for labeling.

This table is used to compare the position of window related to the event time. If start position of the window is equal to start time of the event and less than the time of the next event and the end of the window is less than the sum of start time and its duration, that segment will be labeled to be the event of the start time. Meaning that 100% of the data in the segment need to be occurred in the range of the event only otherwise, it will label as Non-respiratory class. All the label of each segment were stored in Y variable for later use.



A) B)

Figure 3.21 Scored table for labeling A) X variable stores the data of 10 seconds segment and B) Y variable stores label of each segment.

X variable was used to plot the spectrogram image while the Y variable was used to indicate the folder for saving those images. After saved, the images were resized to have a size of 224x224x3 for input to ResNet50 model.

The equalization for this type of image was taken place in the same approach as in 3 classes classification of time domain image but since we segmented the data in different way, the remaining of image after equalized are now greater than the time domain image. The overall dataset are 5457 samples of apnea, 5457 samples of hypopnea and 5457 samples of non-respiratory.

An example of spectrogram of each signal from patient number 1 before resize are shown in figure 3.10-3.15 Note that the order of an image from left to right are Apnea, Hypopnea and Non-respiratory class respectively.

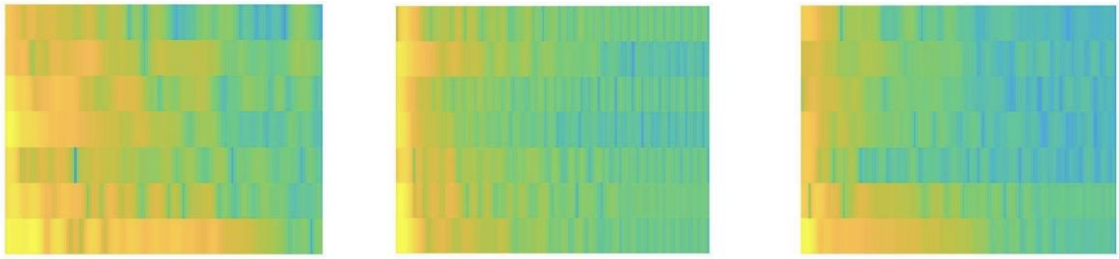
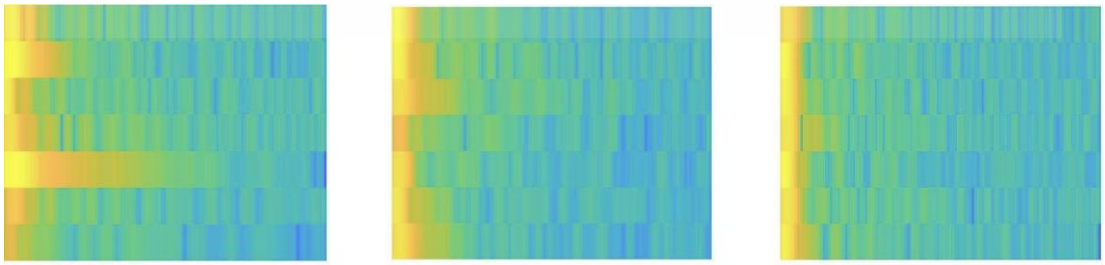


Figure 3.22 Spectrogram from Nasal Pressure signal.

Figure 3.23 Spectrogram from Thermistor signal.

Figure 3.24 Spectrogram from SpO2 signal.

Figure 3.25 Spectrogram from Thoracic signal.Figure 3.26 Spectrogram from



Abdominal signal.

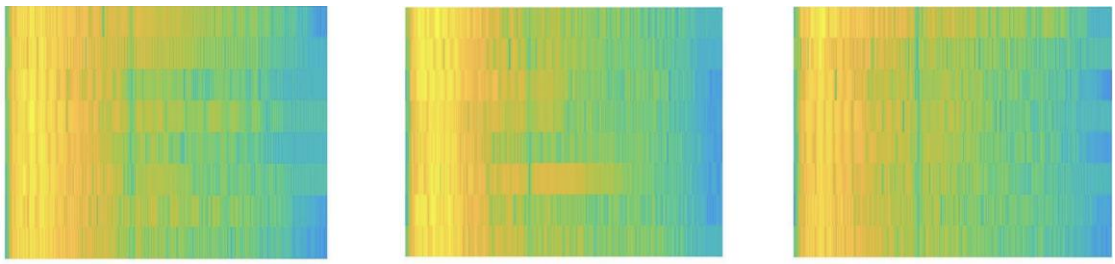


Figure 3.27 Spectrogram from Snore signal.

From the sample images above, the color content of some signals can give us a hint about what's the class of that image belong. Usually the image with more yellow area will belong to Apnea class while the image with more green and blue will be in Non-respiratory class (The color area of Hypopnea class is in between of Apnea and Non-respiratory class). But this is not all the cases. We just use this observation criteria only when we want to classify some of images with our naked eye.

Images from SpO_2 signal are difference from others and difficult to observe because its amplitudes are range in 90 to 100 and some are repeated when the oxygen in patient's blood is constant therefore, the color in each image is quite the same.

3.3.2 Training (Transfer Learning) and Testing

In this study, we trained pre-trained network via transfer learning approachs by modifying the last fully connected layer from 1,000 neurons to 6 or 3 neurons (depends on total amount of class) and retaining the remaining layer to learning with our two types of dataset images for respiratory events classification. We started with time domain image and followed with frequency domain image when training. (All the code and the step by step operation of training and testing process are explained in Appendices B).

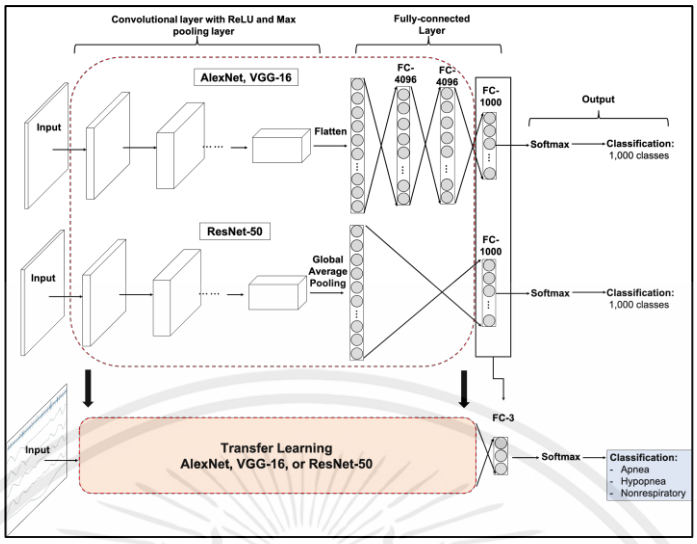
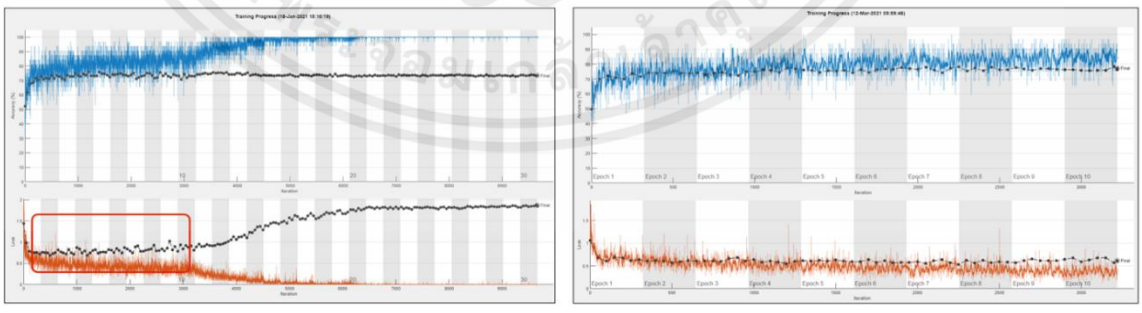


Figure 3.28 Configuration of transfer learning.

All the dataset was divided into training/validation and test dataset with the ratio of 9:1. The training data is splited into 70 percent for training and 30 percent for validation.

When training, the training set of either time domain image or frequency domain image were fed as an input to the pre-trained network models.Time domain images were trained with all three networks with 30 maximum epochs only while frequency domain images were trained with ResNet50 only but learning with both 30 and 10 maximum epochs. The latter maximum epoch was chosen according to the starting epoch of overfitting which can be observed in the 30 maximum epochs training progress.



A) B)

Figure 3.29 Training progress of Nasal Pressure signal with A) 30 and B) 10 maximum epochs.

To indicate the starting point of overfitting, we can observe at the mismatching point or separation point between training loss in red line and validation loss in black dash line of the training progress. From figure 3.29A), the overfitted was occurred at about epoch number 10 therefore, 10 maximum epochs were chosen for training to ensure that the accuracy result will not be overestimated as well as the network will not overfitted with this dataset (training progress with 10 maximum epochs is shown in figure 3.29B)).

In addition of maximum training epochs, we use adaptive moment estimation (Adam) as an optimizer and use cross-entropy loss function to train these models with learning rate of 0.001. However, a mini-batch size of 32 was used instead of 128 due to the GPU memory limitation.

After training, the testing dataset was predicted or classified by the trained network and the network performance was evaluated from the accuracy which can be derived from the confusion matrix according to the accuracy formula:

$$\text{Accuracy} = (TP+TN)/(TP+TN+FP+FN)$$

Where, TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

In this study, Both Training and Testing process were conducted in MATLAB 2020b software on Intel Core i7 with 16 GB of RAM and GeForce RTX 2070 of GPU.

CHAPTER 4
EXPERIMENTAL RESULT

4.1 Introduction

This project is conducted by training preprocessed data with pre-trained models including AlexNet, VGG16 and ResNet50. In this chapter, we present results in forms of training progress and confusion matrix which describe the classification performance of our models.

4.2 Result

Training Progress result shows how well the network learned with the input dataset. The training accuracy and training loss in blue and red line should proceed in different direction in proper training while both validation accuracy and validation loss in black dash line need to be contiguous with its training line as long as the network is not overfit. For Confusion Matrix, the rows of it correspond to the true class and the columns correspond to the predicted class. All correct predictions or true positive are located in the diagonal of the matrix. A table at the right side of confusion matrix is a row summary that displays the percentages of correctly and incorrectly classified observations for each true class. A table at the bottom of the confusion matrix is the column summary for displays the percentages of correctly and incorrectly classified observations for each predicted class.

4.2.1 Time domain inputs (six-class classification)

This section includes the results from ASCII and EDF format input that are trained with AlexNet, VGG16 and ResNet50 to perform 6-class classification.

4.2.1.1 ASCII format

1. AlexNet model

When trained on the clinical dataset, the AlexNet reproduces PSG diagnostic scoring for sleep apnea with accuracy of 52.90%.

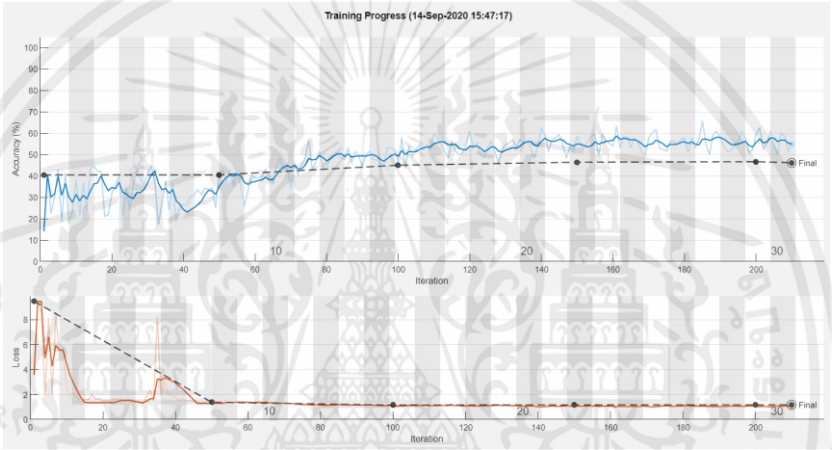


Figure 4.1 Training progress of 6-class classification of AlexNet model with ASCII format.

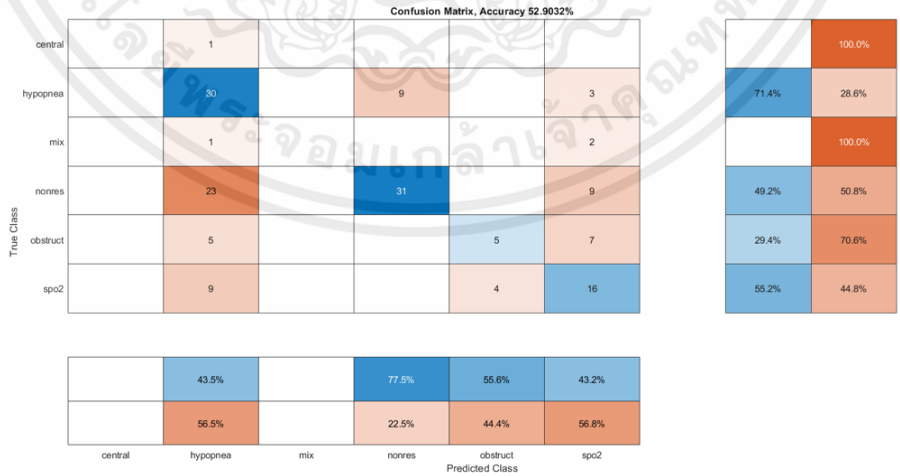


Figure 4.2 Confusion Matrix from six-class classification of the AlexNet model with ASCII file inputs.

2. VGG16 model

When trained on the clinical dataset, the VGG16 reproduces PSG diagnostic scoring for sleep apnea with accuracy of 49.68%.

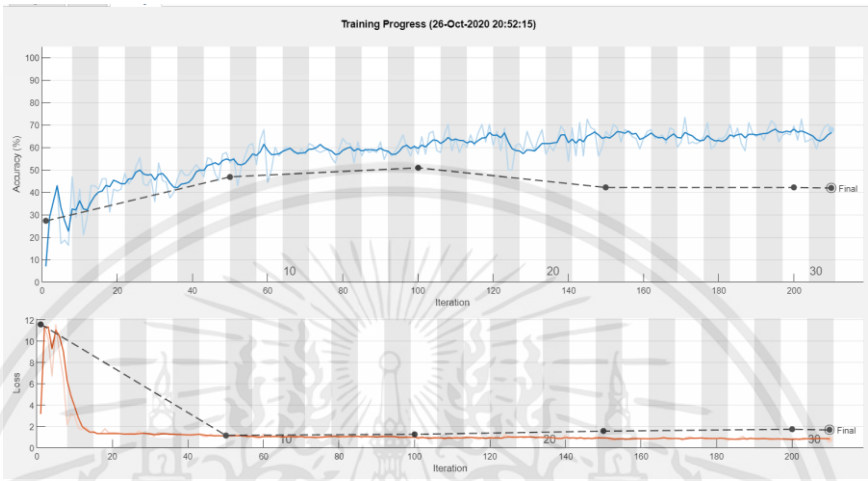


Figure 4.3 Training progress of 6-class classification of VGG16 model with ASCII format.

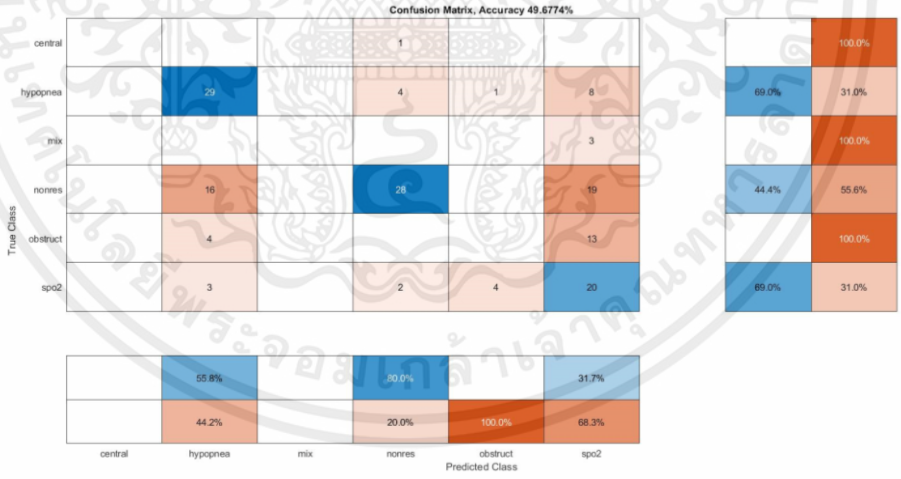


Figure 4.4 Confusion Matrix from six-class classification of the VGG16 model with ASCII file inputs.

3. ResNet50 model

When trained on the clinical dataset, the ResNet50 reproduces PSG diagnostic scoring for sleep apnea with accuracy of 61.94%.

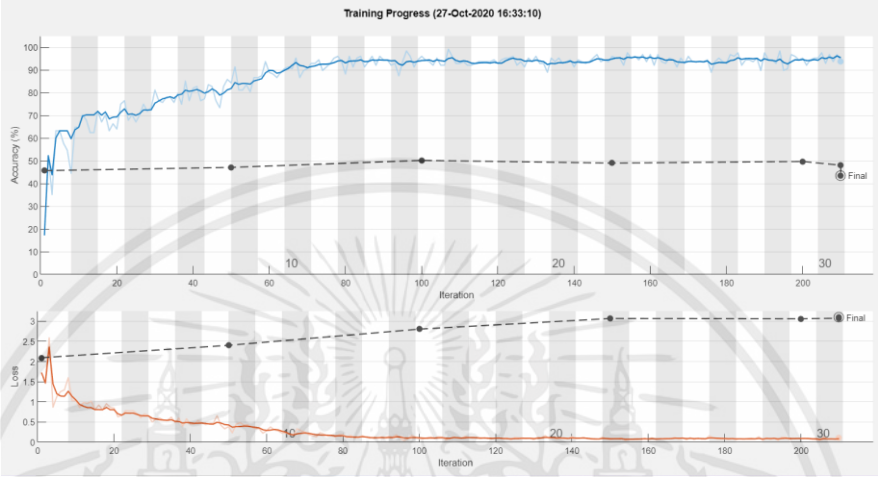


Figure 4.5 Training progress of 6-class classification of ResNet50 model with ASCII format.

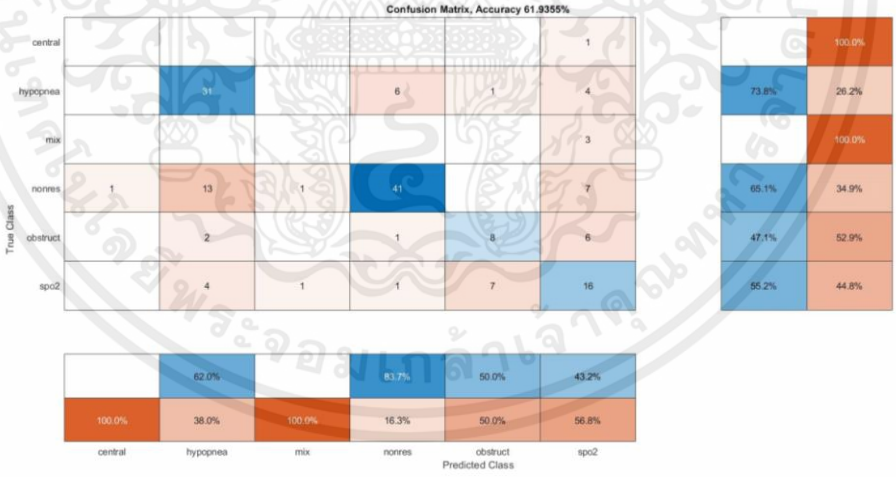


Figure 4.6 Confusion Matrix from six-class classification of the ResNet50 model with ASCII file inputs.

4.2.1.2 EDF format

1. AlexNet model

When trained on the clinical dataset, the AlexNet reproduces PSG diagnostic scoring for sleep apnea with accuracy of 40.76%.

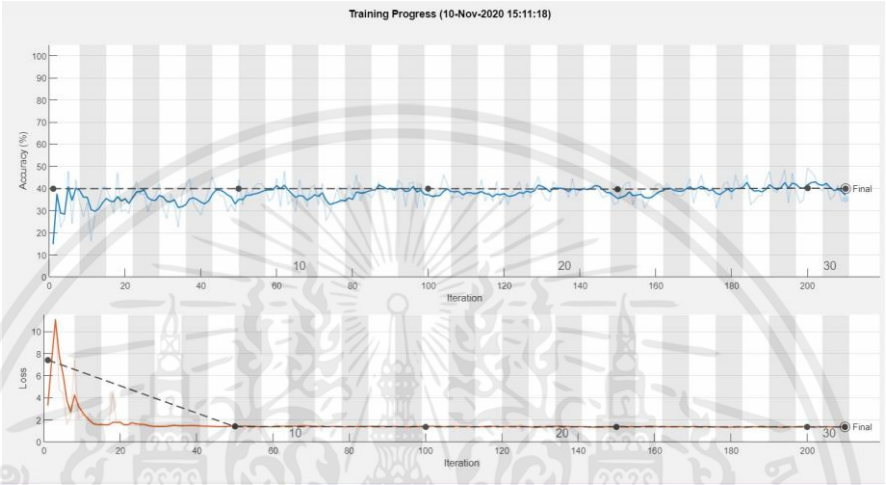


Figure 4.7 Training progress of 6-class classification of AlexNet model with EDF format.

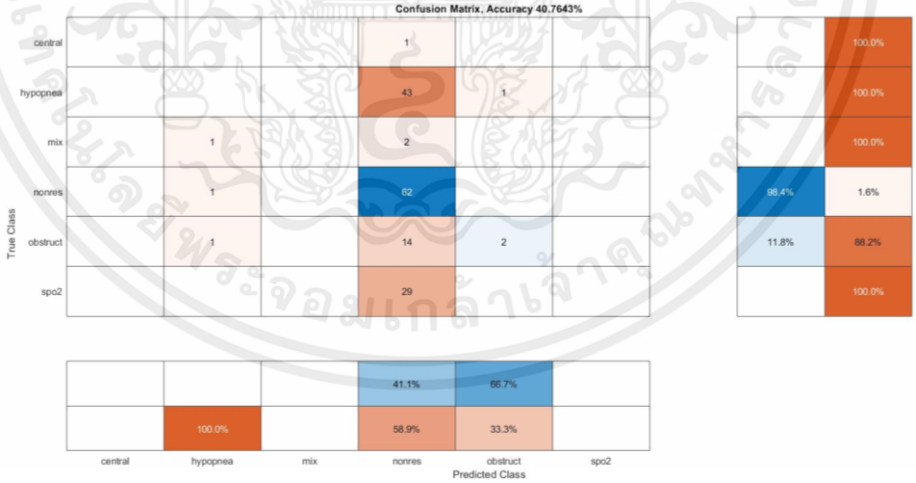


Figure 4.8 Confusion Matrix from six-class classification of the AlexNet model with EDF file inputs.

2. VGG16 model

When trained on the clinical dataset, the VGG16 reproduces PSG diagnostic scoring for sleep apnea with accuracy of 46.50%.

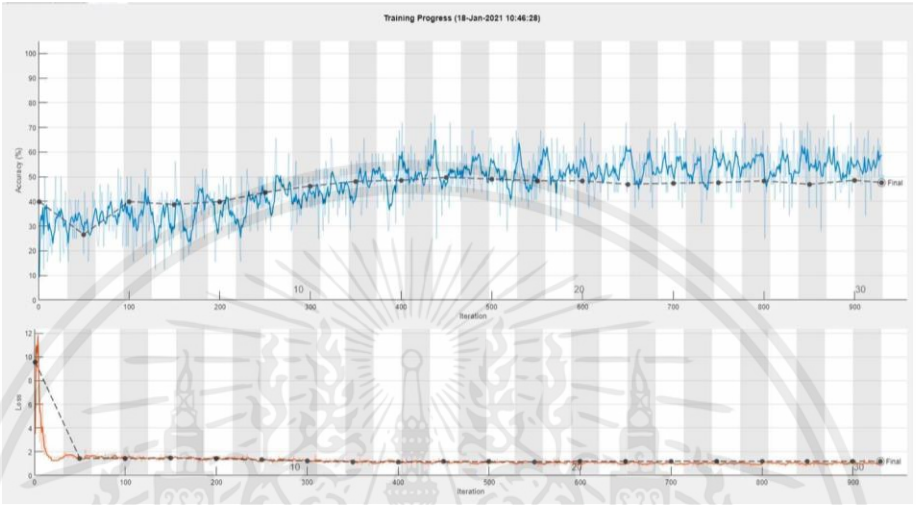


Figure 4.9 Training progress of 6-class classification of VGG16 model with EDF format.

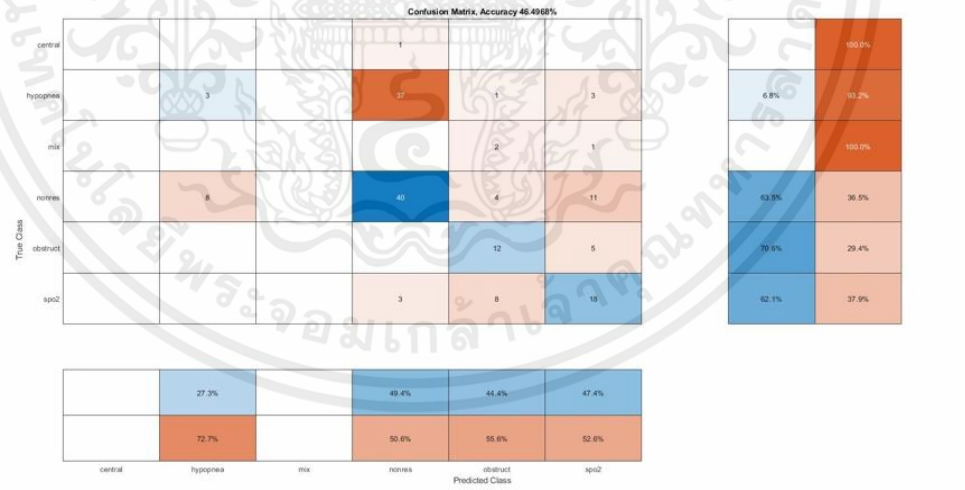


Figure 4.10 Confusion Matrix from six-class classification of the VGG16 model with EDF file inputs.

3. ResNet50 model

When trained on the clinical dataset, the ResNet50 reproduces PSG diagnostic scoring for sleep apnea with accuracy of 45.22%.

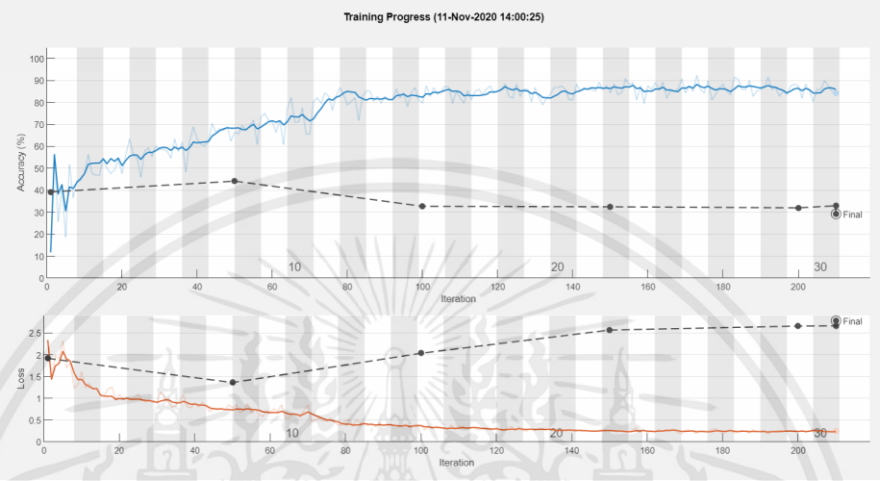


Figure 4.11 Training progress of 6-class classification of ResNet50 model with EDF format.

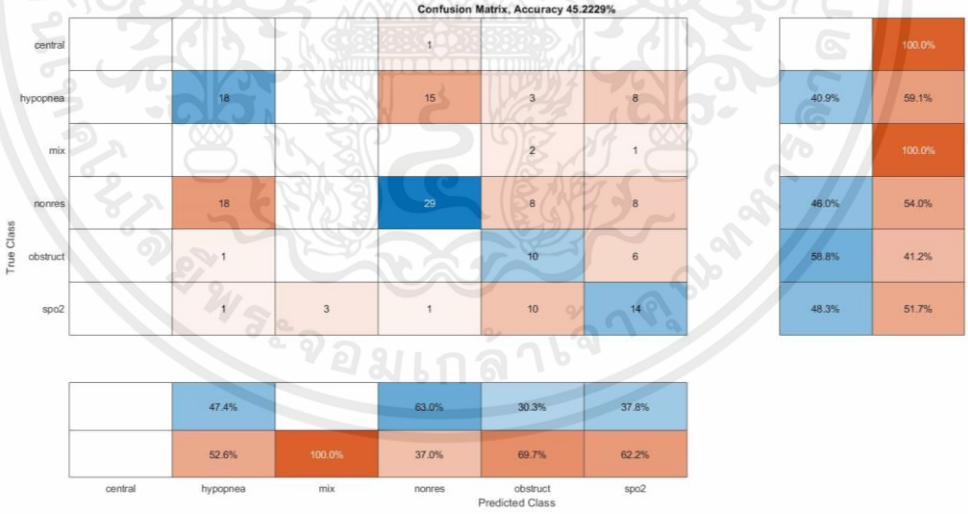


Figure 4.12 Confusion Matrix from six-class classification of the ResNet50 model with EDF file inputs.

4.2.2 Time domain input (three-class classification)

This section includes the results from ASCII and EDF format input that are trained with AlexNet, VGG16 and ResNet50 to perform 3-class classification.

4.2.2.1 ASCII format

1. AlexNet model

When trained on the clinical dataset, the AlexNet reproduces PSG diagnostic scoring for sleep apnea with accuracy of 33.33%.

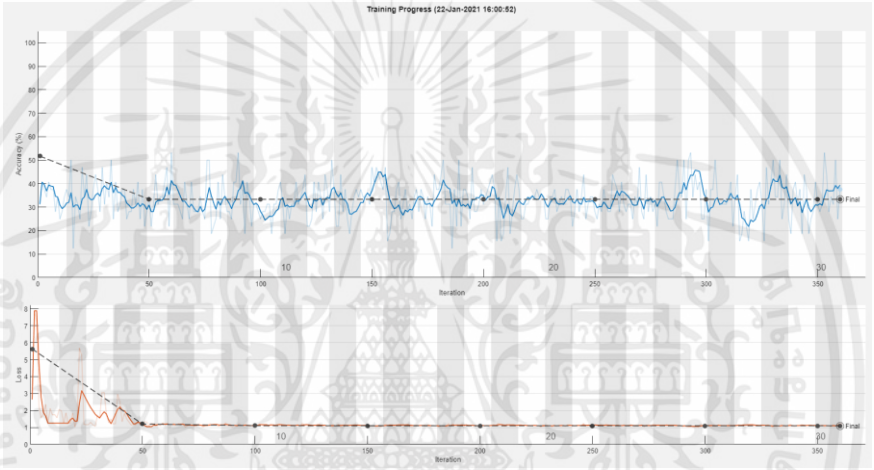


Figure 4.13 Training progress of 3-class classification of AlexNet model with ASCII format.

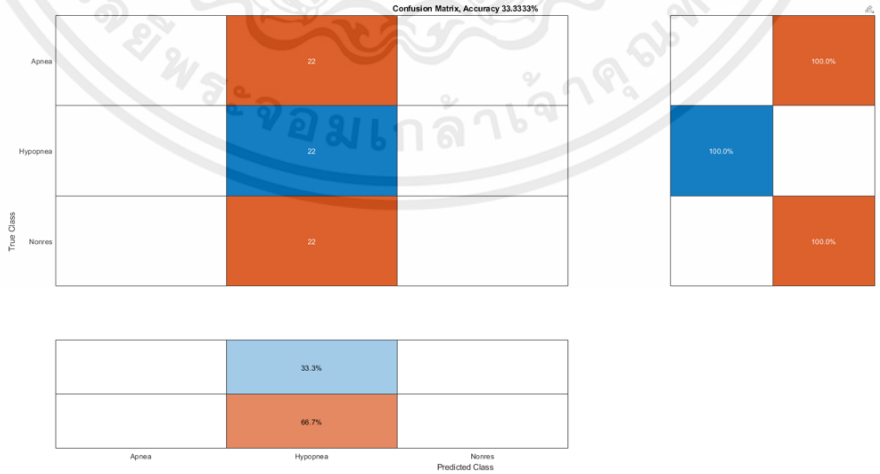


Figure 4.14 Confusion Matrix from three-class classification of the AlexNet model with ASCII file inputs.

2. VGG16 model

When trained on the clinical dataset, the VGG16 reproduces PSG diagnostic scoring for sleep apnea with accuracy of 54.55%.

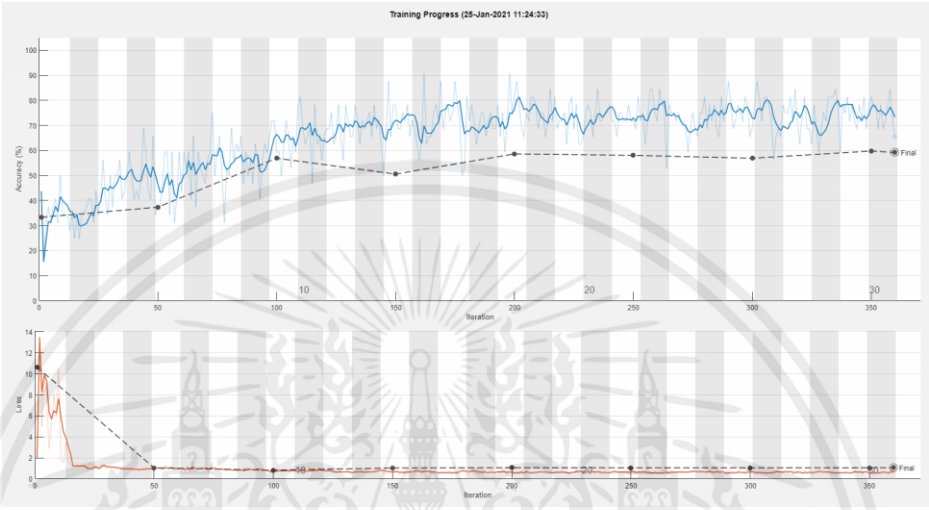


Figure 4.15 Training progress of 3-class classification of VGG16 model with ASCII format.

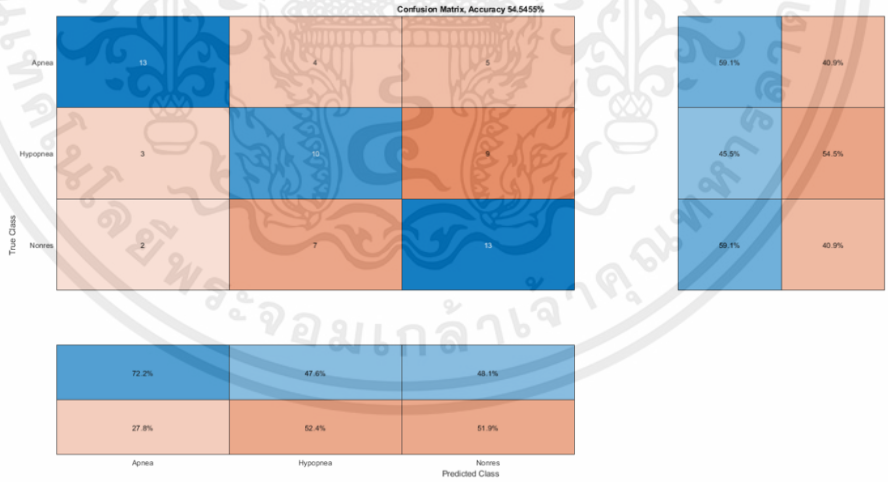


Figure 4.16 Confusion Matrix from three-class classification of the VGG16 model with ASCII file inputs.

3. ResNet50 model

When trained on the clinical dataset, the ResNet50 reproduces PSG diagnostic scoring for sleep apnea with accuracy of 72.73%.

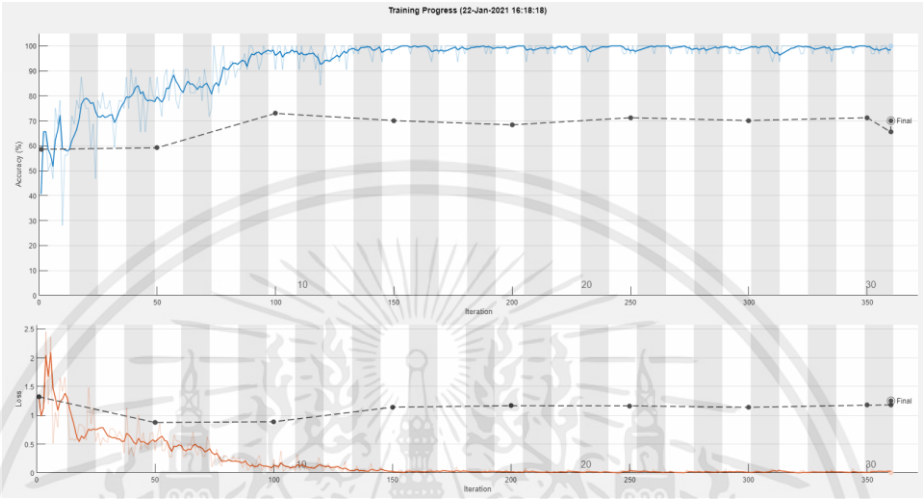


Figure 4.17 Training progress of 3-class classification of ResNet50 model with ASCII format.

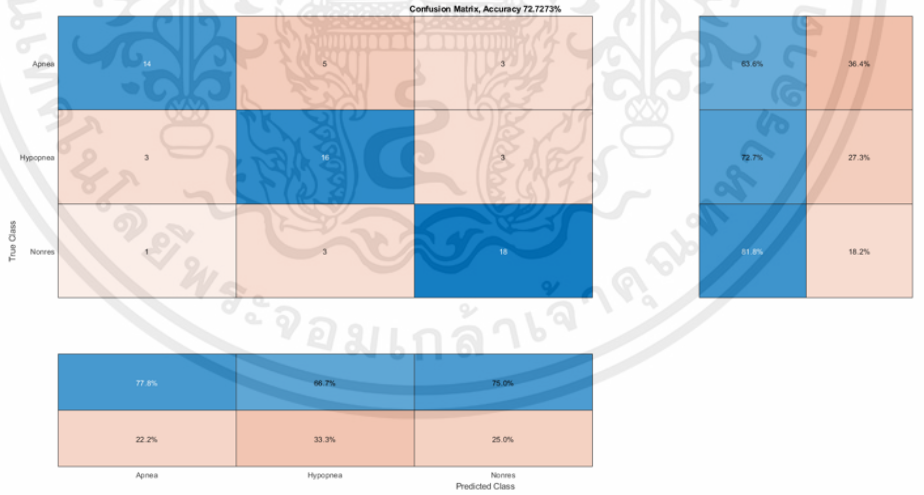


Figure 4.18 Confusion Matrix from three-class classification of the ResNet-50 model with ASCII file inputs.

4.2.2.2 EDF format

1. AlexNet model

When trained on the clinical dataset, the AlexNet reproduces PSG diagnostic scoring for sleep apnea with accuracy of 62.12%.

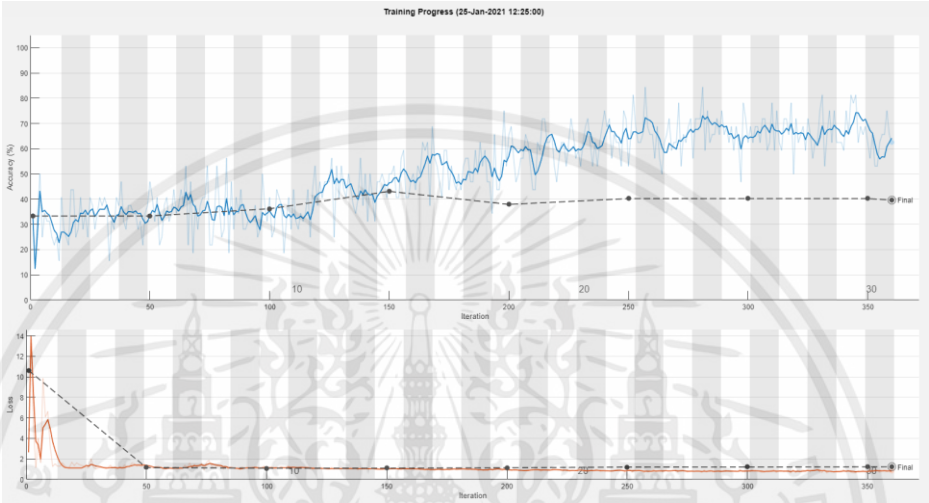


Figure 4.19 Training progress of 3-class classification of AlexNet model with EDF format.

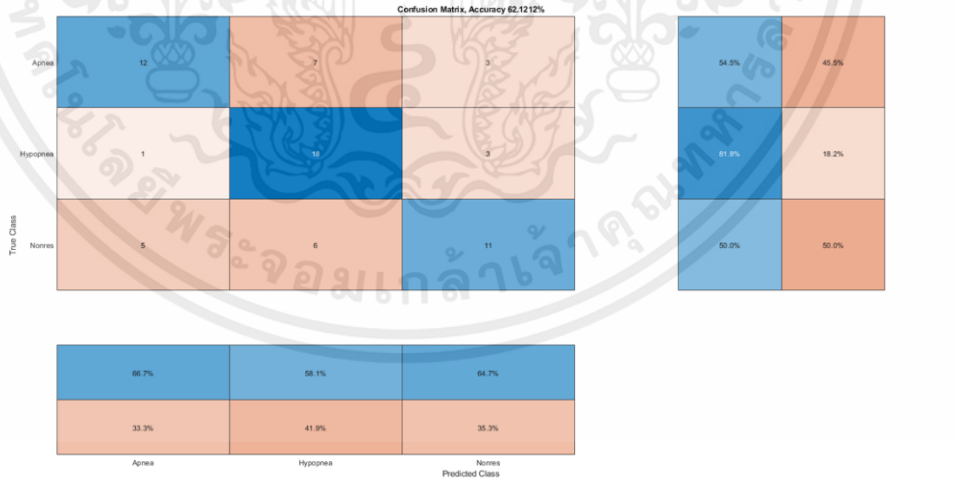


Figure 4.20 Confusion Matrix from three-class classification of the AlexNet model with EDF file inputs.

2. VGG16 model

When trained on the clinical dataset, the VGG16 reproduces PSG diagnostic scoring for sleep apnea with accuracy of 68.18%.

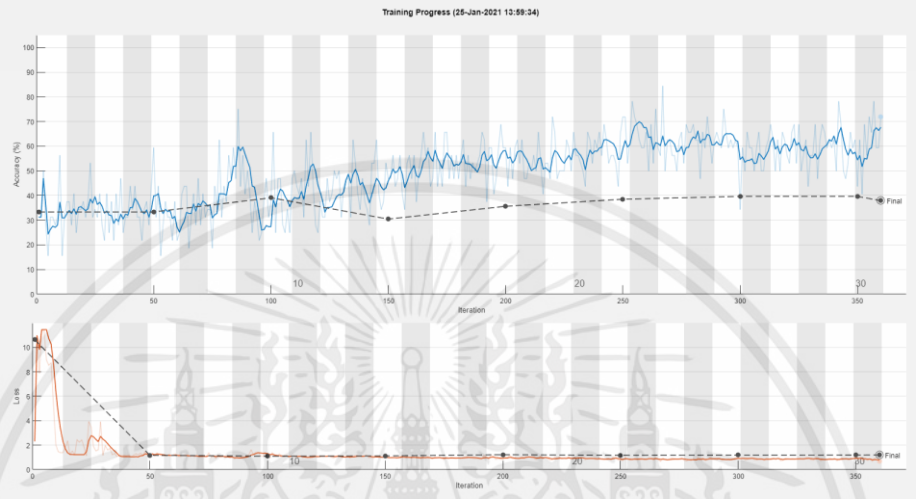


Figure 4.21 Training progress of 3-class classification of VGG16 model with EDF format.

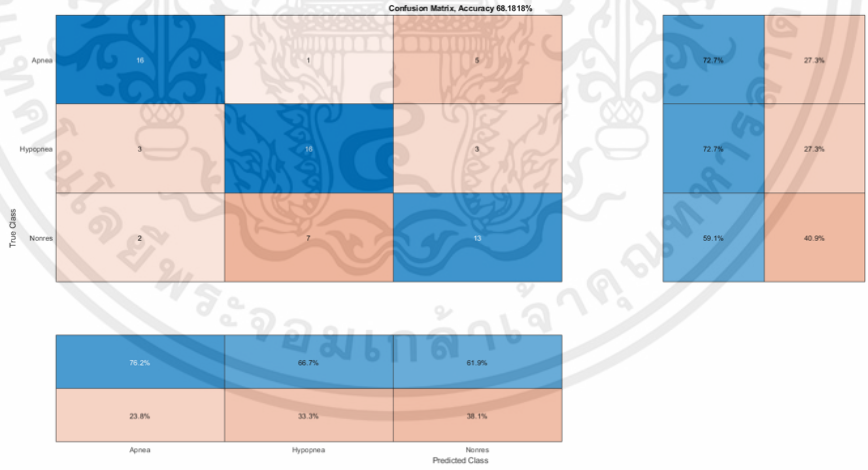


Figure 4.22 Confusion Matrix from three-class classification of the VGG-16 model with EDF file inputs.

3. ResNet50 model

When trained on the clinical dataset, the ResNet50 reproduces PSG diagnostic scoring for sleep apnea with accuracy of 83.33%.

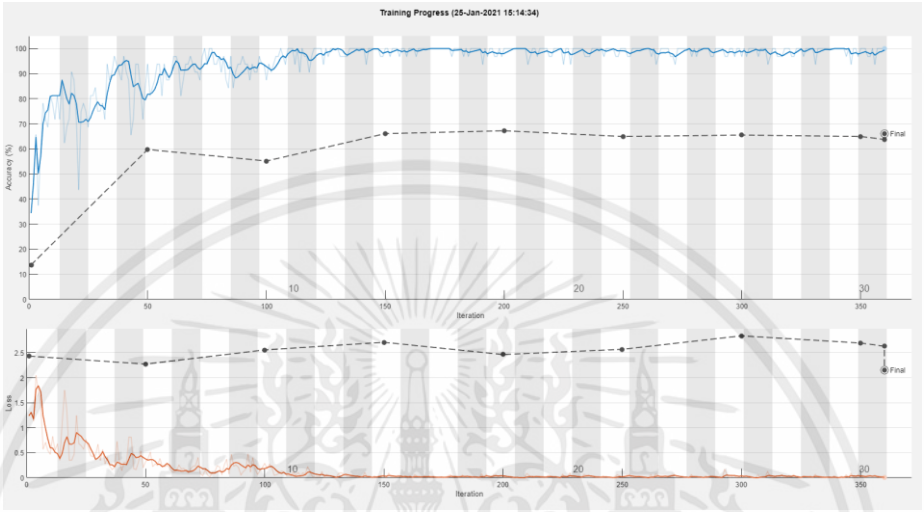


Figure 4.23 Training progress of 3-class classification of ResNet50 model with EDF format.

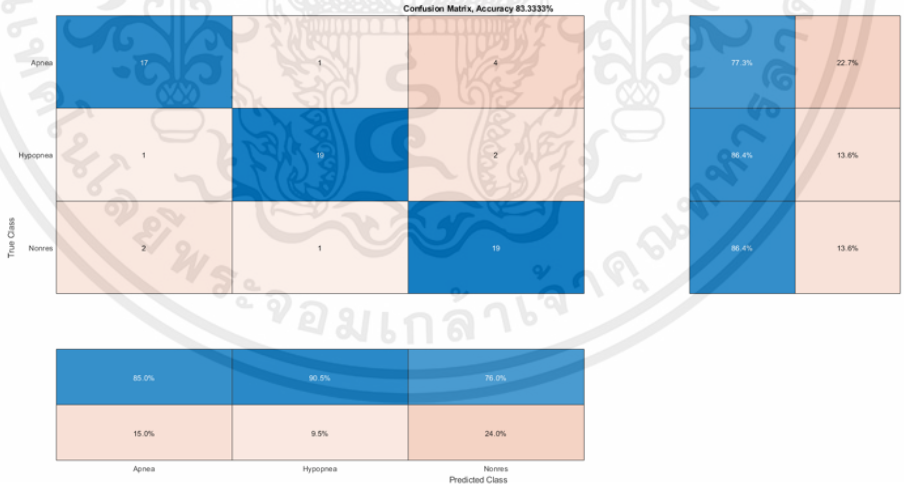


Figure 4.24 Confusion Matrix from three-class classification of the ResNet-50 model with EDF file inputs.

4.2.3 Frequency domain input (Balanced and Normalized data)

This section includes the results from spectrogram of 6 signals that are trained with ResNet50 by using normalized data with 10 maximum epochs to perform balanced data classification.

4.2.3.1 Nasal Pressure signal

When trained on the clinical dataset with Nasal Pressure signal, ResNet50 reproduces PSG diagnostic scoring for sleep apnea with accuracy of 80.16%

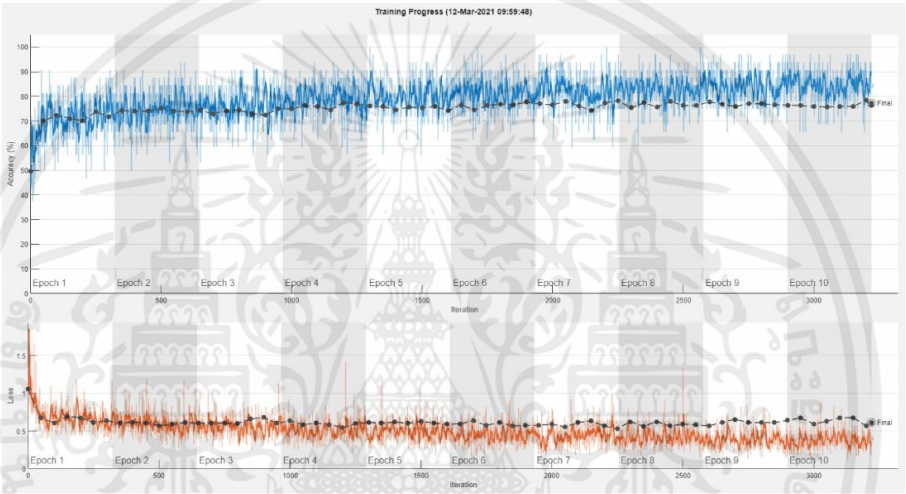


Figure 4.25 Training progress from spectrogram of nasal pressure signal

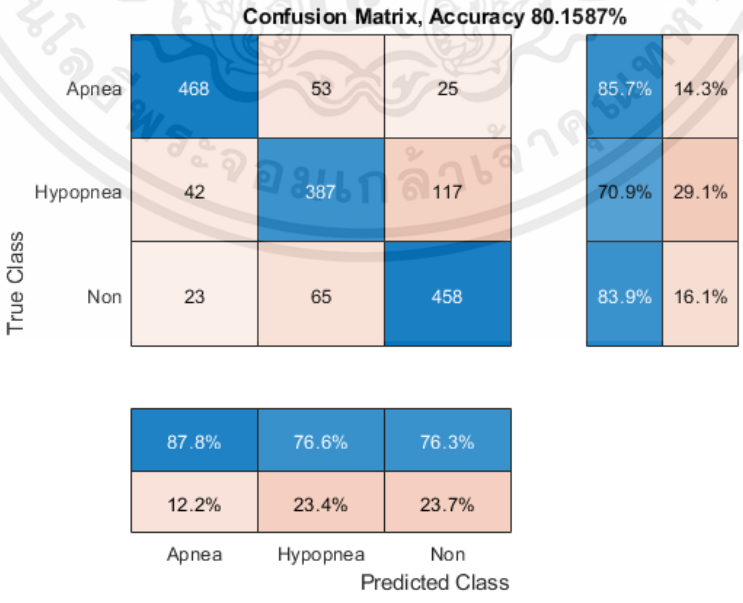


Figure 4.26 Confusion Matrix from classification with ResNet50 model using spectrogram of Nasal Pressure signal.

4.2.3.2 Snore signal

When trained on the clinical dataset with snore signal, ResNet50 reproduces PSG diagnostic scoring for sleep apnea with accuracy of 78.14%

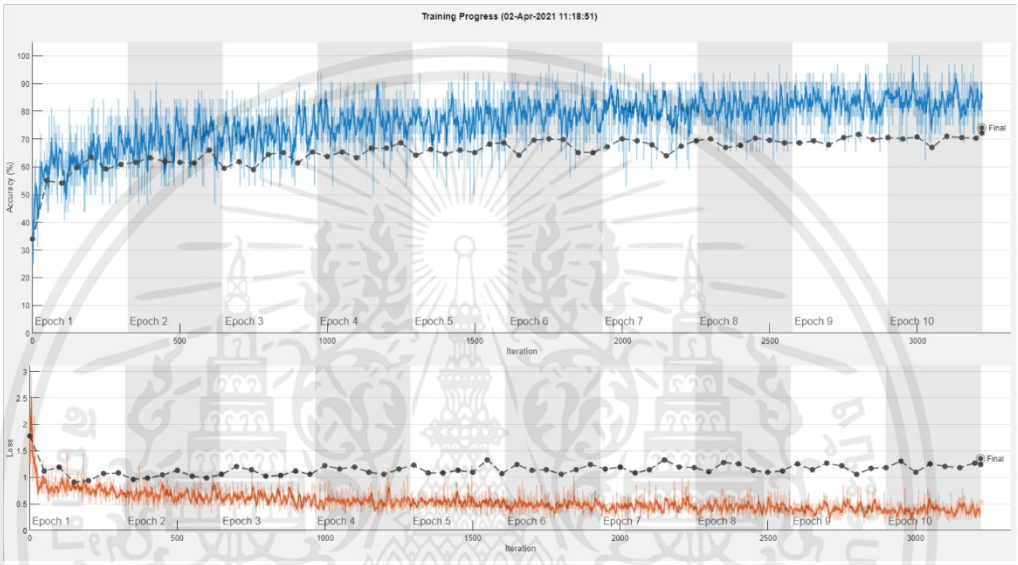


Figure 4.27 Training progress from spectrogram of snore signal.

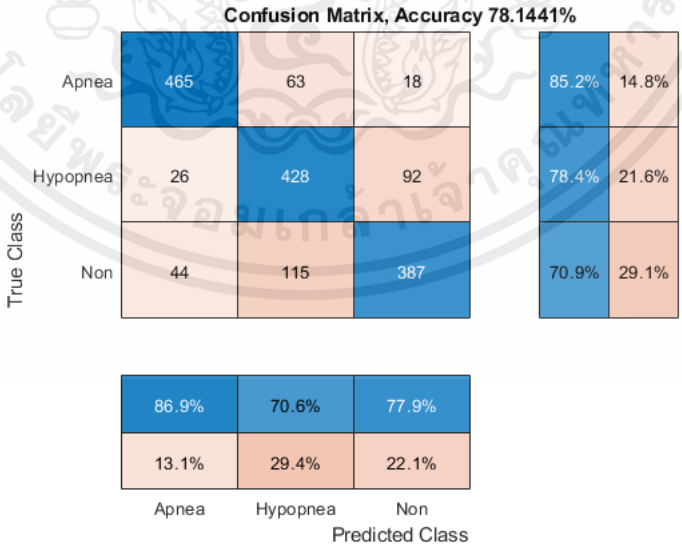


Figure 4.28 Confusion Matrix from classification with ResNet50 model using spectrogram of snore signal.

4.2.3.3 Thoracic signal

When trained on the clinical dataset with thoracic signal, ResNet50 reproduces PSG diagnostic scoring for sleep apnea with accuracy of 75.03%

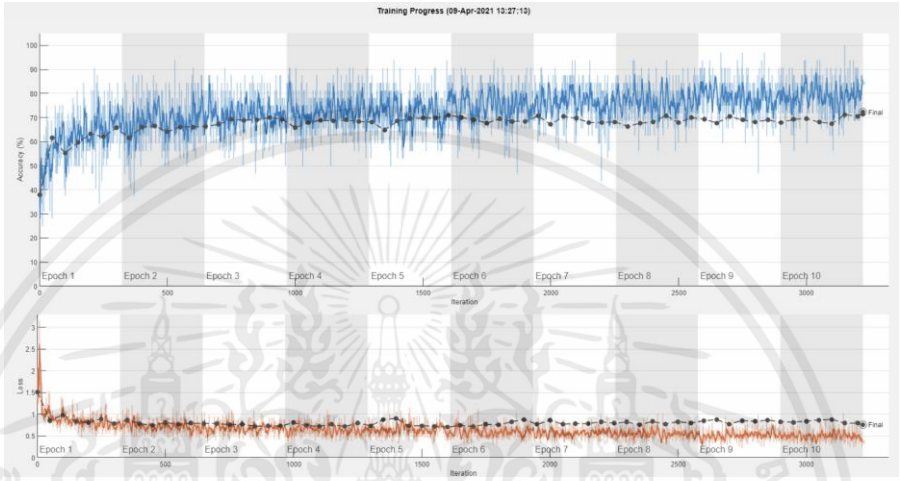


Figure 4.29 Training progress from spectrogram of thoracic signal.

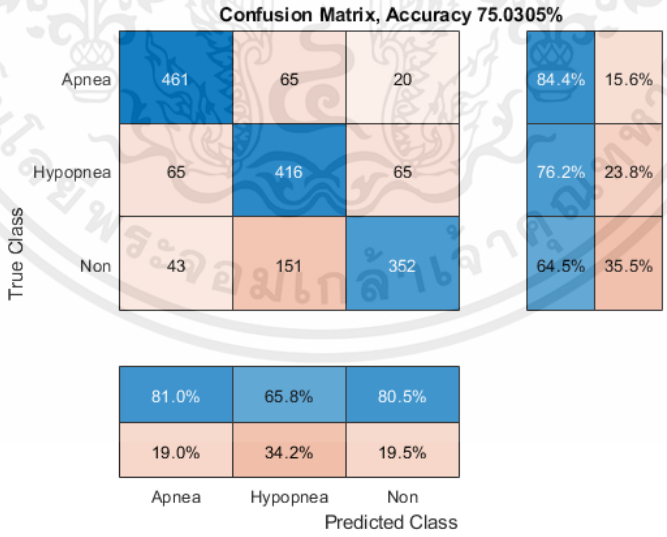


Figure 4.30 Confusion Matrix from classification with ResNet50 model using spectrogram of thoracic signal.

4.2.3.4 Thermistor signal

When trained on the clinical dataset with thermistor signal, ResNet50 reproduces PSG diagnostic scoring for sleep apnea with accuracy of 75.27%

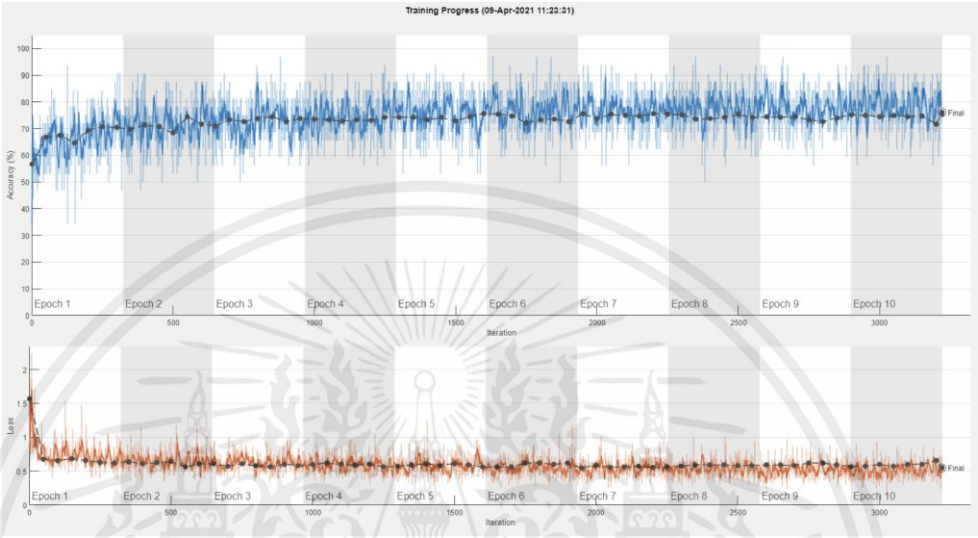


Figure 4.31 Training progress from spectrogram of thermistor signal.

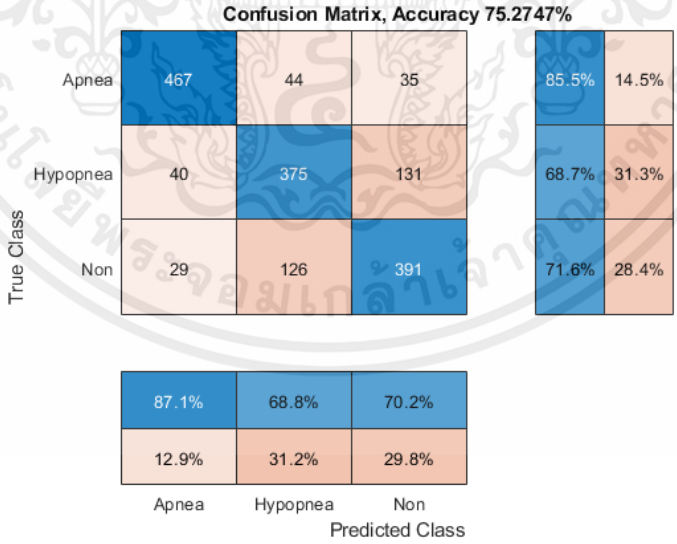


Figure 4.32 Confusion Matrix from classification with ResNet50 model using spectrogram of thermistor signal.

4.2.3.5 Abdominal signal

When trained on the clinical dataset with abdominal signal, ResNet50 reproduces PSG diagnostic scoring for sleep apnea with accuracy of 69.72%

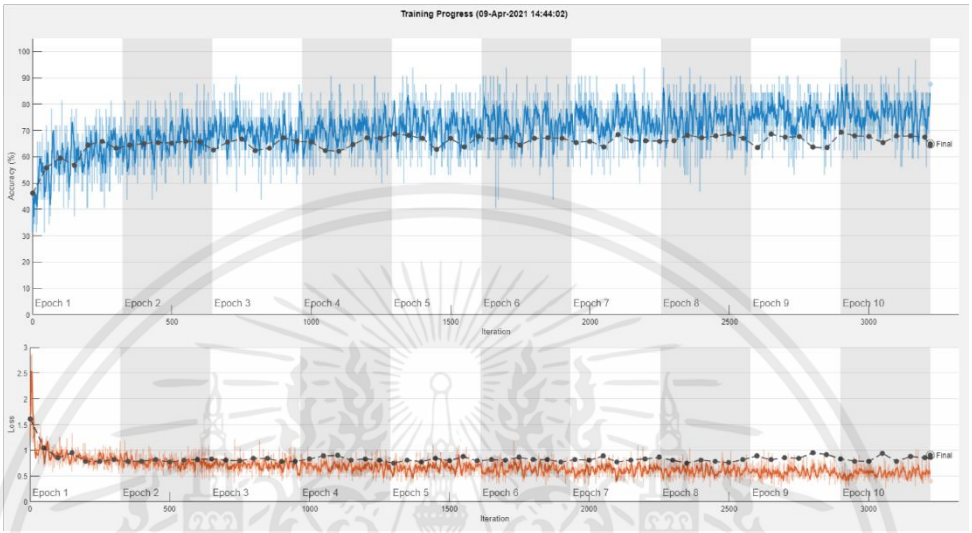


Figure 4.33 Training progress from spectrogram of abdominal signal.

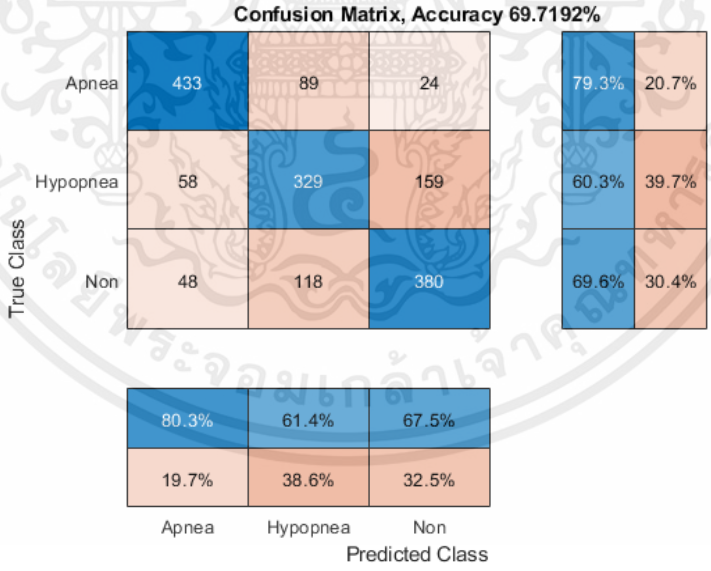


Figure 4.34 Confusion Matrix from classification with ResNet50 model using spectrogram of abdominal signal.

4.2.3.6 SpO2 signal

When trained on the clinical dataset with SpO2 signal, ResNet50 reproduces PSG diagnostic scoring for sleep apnea with accuracy of 63.13%

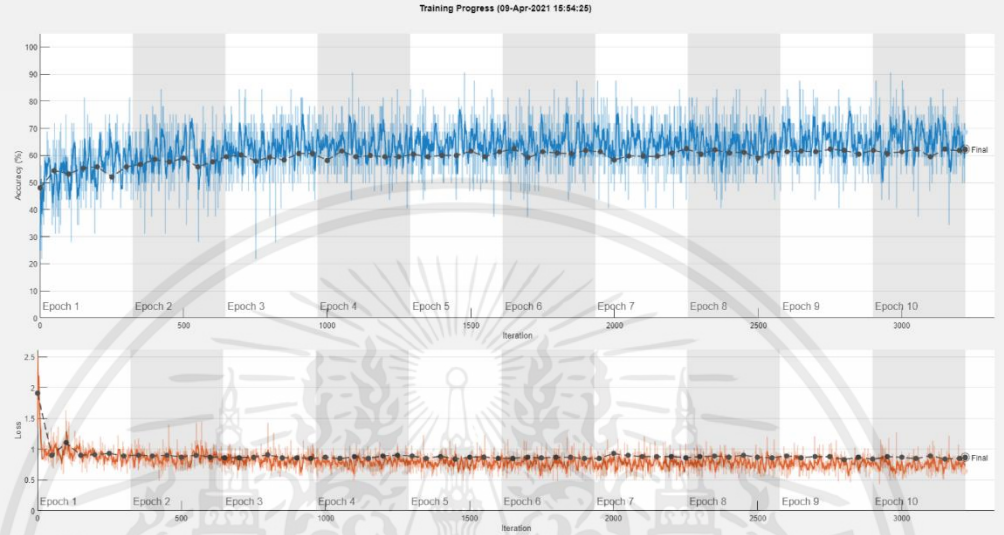


Figure 4.35 Training progress from spectrogram of SpO2 signal

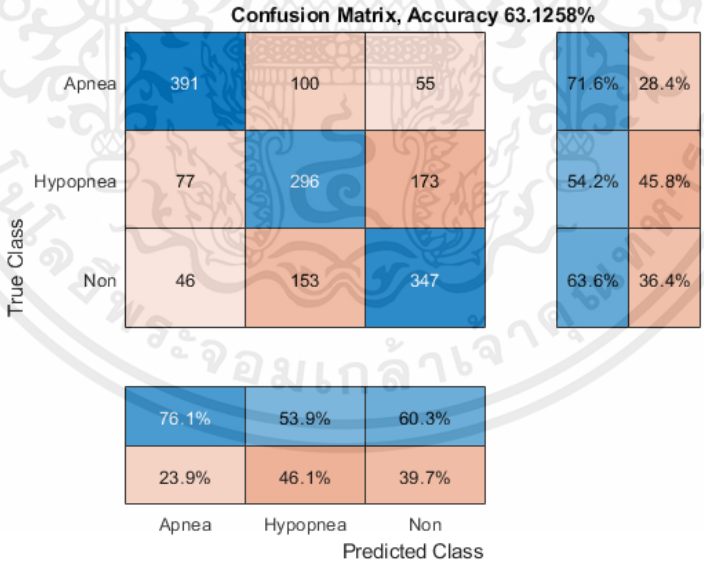


Figure 4.36 Confusion Matrix from classification with ResNet50 model using spectrogram of SpO2 signal.

CHAPTER 5
DISCUSSION AND CONCLUSION

5.1 Introduction

In this chapter, we summarize our project and the results of each type of classification into tables for convenience in observation. Next, we conclude and discuss the key points of our work. Lastly, we state future work and how we see our model can support doctors and technicians.

5.2 Summary

Sleep apnea is a serious condition that tends to be increasing in global population year by year. It can cause various life-threatening diseases to patients if they get improperly diagnose and left untreated. The gold standard that is used to diagnosis sleep disorders is polysomnography which is the record of physiological data or polysomnogram. The records will be scored by sleep doctors to identify the severity of each patients and provide proper treatment.

Our study aims to facilitate sleep experts and reduce the conflict in scoring between person by researching and developing the deep neural network for respiratory events scoring and calculate the AHI to predict the severity. The three pre-trained convolution neural network (CNN); AlexNet, VGG16 and ResNet50 were adopted to train with two types of 2D images.

Time domain images were constructed first as they are same with the one that presented to the scorers. It is a plotting of several normalized signals from 5 patients, including Snore, Nasal Pressure, Thermistor, Thoracic, Abdominal and SpO2, of both ASCII and EDF file format. Each image is a window of 30 seconds with non-overlap. Images of time domain were classified into 6 classes and 3 classes respectively for training with 30 maximum epochs by three networks.

Table 5.1 Results from 6-class classification using ASCII and EDF file input with AlexNet, VGG16 and ResNet50

	AlexNet	VGG-16	ResNet-50
ASCII	52.90%	49.68%	61.93%
EDF	40.76%	46.50%	45.22%

Table 5.2 Results from 3-class classification using ASCII and EDF file input with AlexNet, VGG16 and ResNet50.

	AlexNet	VGG-16	ResNet-50
ASCII	33.33%	54.55%	72.73%
EDF	62.12%	68.18%	83.33%

Frequency domain images or spectrograms were prepared later by using only one signal of EDF file format to plot. The 10 seconds window is sliding on each normalized signal by moving 1 second at a time (90% overlap) to generate the frequency domain image. The type of image underwent balanced data classification with ResNet50 by using 30 maximum epochs first and later altered to 10 maximum epochs to avoid overfitting problem.

Table 5.3 Results from classification using spectrograms of all signals with ResNet50.

Nasal Pressure	Thermistor	SpO2	Thoracic	Abdominal	Snore
80.16%	75.27%	63.13%	75.03%	69.72%	78.14%

5.3 Discussion and Conclusion

The accuracy results of 6 classes classification are unreliable for chosen both type of file format and network and still not satisfy because most true-positive classification belong to only one class with biggest amount of data which is Non-respiratory class while the precision of the remaining classes are very low especially in

the classes with small amount of data such as Central and Mixed apnea. So, we decided to reduce the amount of class to be three instead of six and considered to equalize amount of sample in each class to be the same to solve the bias from imbalanced data.

The results of three-class classification from time domain images are clearer. The combination of ResNet50 model with the plotting of EDF file format showed the highest accuracy at 83.33% with each class precision is higher than 70%. We believe that this combination can perform better than others because firstly, the EDF file format images are more similar to what's doctor saw on screen when compare with the plot from ASCII file. Secondly, ResNet50 has the residual function, which is not found in other models and lastly, the data in each class were equally trained to the network.

Although this type of image gives quite satisfy accuracy, but we still faced with several problems. First, the duration of image of this type is too long and it is the cause of repeated image in different classes because it is possible that several events are occurred within 30 seconds, which finally resulted in difficulty for the network to classify. Next, low content of color may affect the convolution process for feature extraction and another problem is the overfitting of ResNet50 with this set of data. As a result of deeper layer of ResNet50 and the low remaining amount of data in each class after equalized, it can cause the network to overfit and give the accuracies that are exaggeration.

Frequency domain images or Spectrograms were prepared and trained according to the best combination from the previous result. It also relied on the equalization of data and trained with 10 maximum epochs to ensure that we will eliminate all the preceding problem and get the best and reliable accuracy. Furthermore, the adjusting of this image type into 10 seconds can help in shorten of duration and increase the amount of image after equalized whereas the color of the spectrogram can increase the content of an image.

Our result reveals that the spectrogram of Nasal Pressure is the best candidate for use. It provides the accuracy of 80.16% while the worst accuracy belongs to spectrogram of SpO_2 signal with 63.13% accuracy.

In our point of view, normalization is very important since it can minimize bias from the fluctuation of signals from noise of 5 patients and determine the standard scale

for plotting. The nasal pressure is the main respiratory signal when observe the sleep apnea event therefore, it is obvious that this signal will give the best accuracy among others. On the other hand, the amplitude of SpO_2 signal is relatively the same in every period (range between 90 to 100). Thus, the colors of the different events are quite similar and difficult to distinguish.

Besides of signal type, we have found that the sampling frequency of the signal also affect the classification accuracy. As a result of the highest sampling frequency in snore signal (256 Hz) when compared to others (16 and 32 Hz), we noticed about much larger in content of an image which can be difficult for the network to predict and related to the overfitted in training progress. Therefore, an appropriate sampling frequency, which in study is equal to 32 Hz, should be optimized and considered before prepared and input to the network.

This preliminary study, only 5 Thai PSG data sets were analyzed. Our results suggest that frequency domain image or spectrogram of normalized Nasal Pressure signal from EDF file format training with ResNet50 is better used than other signals in the same type of image and also superior to the time domain image type.

5.4 Future work

For the future work, 100 more of PSG data will be recruited to ensure that the data in each class will be sufficient for use without reducing number of the class for equalization as well as allowing us to obtain better model that will be able to predict more precisely and accurately so, it is possible for AHI calculation in order to determine the severity of patients. In addition of increasing amount of data, all six signals can be simultaneously converted into spectrogram and plot in the same frame to create an image. Furthermore, the window size and the covering percentage approach for event validation of the segment can also be adjusted from 10 seconds and 100% respectively to compare with this study. The down-sampling of snore signal will be considered to compare in training progress with other signal candidates. And lastly, training with other interesting networks such as GoogLeNet and other type of neural network such as Recurrent neural network (RNN) can be performed to observe the improvement of classification accuracy.

REFERENCES

[1] I. JESSICA VENSEL RUNDO AND RALPH DOWNEY, "Polysomnography," *Handbook of Clinical Neurology,,* vol. 16, p. 1, 2019.

[2] N. T. A. P. R. E. R. H. M. S. M. I. M. J. M. C. M. N. S. R. P. T. P. Adam V Benjafield, "Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis," *Lancet Respir Med ,* vol. 7, pp. 687-698, 2019.

[3] F. M. A. G. R.-G. a. F. M.-D. Sheikh Shanawaz Mostafa, "A Systematic Review of Detecting Sleep Apnea Using Deep Learning," *Sensors,* vol. 19, pp. 1-22, 2019.

[4] M. C. Staff, "Sleep apnea," Mayo Clinic (Mayo Foundation for Medical Education and Research), 28 July 2020. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/sleep-apnea/symptoms-causes/syc-20377631. [Accessed 22 October 2020].

[5] M. F. B. M. F. C.-R. M. J. A. D. P. Shahrokh Javaheri, "Sleep Apnea : Types, Mechanisms, and Clinical Cardiovascular Consequences," *Journal of the American College of Cardiology,* vol. 69, no. 21 February 2017, pp. 841-843, 2017.

[6] K. E. Bloch, "Polysomnography: a systematic review," *Technology and Health Care,* pp. 285-292, 1997.

[7] D. W. S. ,. S.R. Pandi-Perumal, "Polysomnography: An Overview," *Primary Care Sleep Medicine,* p. 33, 2014.

[8] M. (. Richard B. Berry, M. R. R. Rita Brooks, M. Charlene E. Gamaldo, M. Susan M. Harding, M. Robin M. Lloyd, M. Stuart F. Quan, D. Matthew M. Troester and M. Bradley V. Vaughn, The AASM Manual for the Scoring of

Sleep and Associated Events, Darien: American Academy of Sleep Medicine, 2017.

[9] G. B. M. Shrinivas Kambali MDa, "Pitfalls of Split-Night Polysomnography," *THE SOUTHWEST RESPIRATORY AND CRITICAL CARE CHRONICLES (in the clinic),* vol. 1, p. 36, 2013.

[10] M. C. Staff, "Obstructive sleep apnea," Mayo Foundation for Medical Education and Research (Mayo Clinic), 5 June 2019. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/obstructive-sleep-apnea/diagnosis-treatment/drc-20352095. [Accessed 25 October 2020].

[11] Sleepadmin, "What does “AHI” represent?," American Sleep Apnea Association, 31 August 2017. [Online]. Available: https://www.sleephealth.org/ufaqs/what-is-ahi-represent/#:~:text=For%20an%20Apnea%2DHypopnea%20Index,%2C%20on%20average%2C%20each%20hour.. [Accessed 25 October 2020].

[12] W. Meints, The relationship between AI, machine learning, and deep learning, Birmingham, United Kingdom: Packt, 2019.

[13] M. Dhande, "What is the difference between AI, Machince learning and deep learning?," Geospatial world, 07 03 2020. [Online]. Available: https://www.geospatialworld.net/blogs/difference-between-ai%EF%BB%BF-machine-learning-and-deep-learning/. [Accessed 31 October 2020].

[14] M. Gaurav, Director, *What's deep learning?.* [Film]. San Francisco, California, United States: Simplilearn, 2019.

[15] MathWorks, "Deep Learning for Signal Processing with MATLAB," *White paper,* pp. 2-3, 2019.

[16] D. Patel, "Insights Of The Machine Learning And The Deep Learning," Thinkwik, 2 July 2018. [Online]. Available: https://blog.thinkwik.com/insights-of-the-machine-learning-and-the-deep-learning/. [Accessed 31 October 2020].

[17] J. J. Moolayil, "A Layman's Guide to Deep Neural Networks," Medium, 24 July 2019. [Online]. Available: https://towardsdatascience.com/a-laymans-guide-to-deep-neural-networks-ddcea24847fb. [Accessed 31 October 2020].

[18] M. Gaurav, Director, *What's Neural Network?*. [Film]. San Francisco, California, United States: Simplilearn, 2019.

[19] F. G. S. P. C. R. a. S. A. A. M. Sokolovsky, "Deep learning for automated feature discovery and classification of sleep stages," *IEEE/ACM Trans. Comput. Biol. Bioinform,* vol. 17, pp. 1835-1845, 2020.

[20] S. Z. a. A. A. N. Jmour, "Convolutional neural networks for image classification," *IC_ASET,* pp. 397-402, 2018.

[21] G. L. Team, "AlexNet: The First CNN to win Image Net," Great Learning, 24 June 2020. [Online]. Available: https://www.mygreatlearning.com/blog/alexnet-the-first-cnn-to-win-image-net/. [Accessed 1 November 2020].

[22] B. A. S. T. Arslan Tuncera, "A deep learning-based decision support system for diagnosis of OSAS using PTT signals," *Medical Hypotheses,* vol. 127, pp. 19-21, 2019.

[23] A. Anwar, "Difference between Alexnet, VGGNet, ResNet and Inception.," Medium, 7 June 2019. [Online]. Available: https://towardsdatascience.com/the-w3h-of-alexnet-vggnet-resnet-and-inception-7baaaecccc96. [Accessed 1 November 2020].

[24] Koustubh, "ResNet, AlexNet, VGGNet, Inception: Understanding various architectures of Convolutional Networks," CV-tricks, 3 August 2018. [Online].

Available: https://cv-tricks.com/cnn/understand-resnet-alexnet-vgg-inception/. [Accessed 1 November 2020].

[25] P. Nepal, "Deep Residual Learning for Image Recognition (ResNet paper explained)," Medium, 1 January 2021. [Online]. Available: https://medium.com/analytics-vidhya/deep-residual-learning-for-image-recognition-resnet-paper-explained-26b29e0fa73e. [Accessed 28 March 2021].

[26] C.-F. Wang, "The Vanishing Gradient Problem," Medium, 8 January 2019. [Online]. Available: https://towardsdatascience.com/the-vanishing-gradient-problem-69bf08b15484. [Accessed 28 March 2021].

[27] A. Mittal, "Understanding RNN and LSTM," Medium, 12 October 2019. [Online]. Available: https://aditi-mittal.medium.com/understanding-rnn-and-lstm-f7cdf6dfc14e. [Accessed 1 November 2020].

[28] Debarko, "RNN or Recurrent Neural Network for Noobs," HackerNoon, 1 March 2018. [Online]. Available: https://hackernoon.com/rnn-or-recurrent-neural-network-for-noobs-a9afbb00e860. [Accessed 1 November 2020].

[29] M. Phi, "Illustrated Guide to LSTM's and GRU's: A step by step explanation," Medium, 25 September 2018. [Online]. Available: https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21. [Accessed 1 November 2020].

[30] J. K. H. S. B. G. M. B. W. M. T. B. a. J. S. S. Biswal, "SLEEPNET: automated sleep staging system via deep learning," *arXiv:1707.08262v1 [cs.LG] ,* pp. 1-17, 2017.

[31] E. Byrne, Director, *Deep Learning Onramp.* [Film]. Natick, Massachusetts, United States: MathWorks, 2019.

[32] A. S. a. V. B. Fatih Demir, "Convolutional neural networks based efficient approach for classification of lung diseases," *Health Information Science and Systems,* p. 4, 2020.

[33] H. S. B. G. M. B. W. J. S. a. M. T. B. Siddharth Biswal, "Expert-level sleep scoring with deep neural networks," *American Medical Informatics Association,* vol. 25, pp. 1643-1650, 2018.

[34] H. Y. H. S. K. H. B. K. H. B. K. S. M. O. Y. J. L. K. S. P. Sang Ho Choia, "Real-time apnea-hypopnea event detection during sleep by convolutional neural network," *Computers in Biology and Medicine,* vol. 100, pp. 123-131, 2018.

[35] W. G. E. P. M. J. L. S. A. Q. B. G. R. A. T. O. A. R. J. T. M. B. W. Haoqi Sun, "Sleep staging from electrocardiography and respiration with deep learning," *Sleep Research Society ,* vol. 43, no. 7, pp. 1-28, 2019.

APPENDICES

APPENDIX A

PRE-PROCESSING CODE

Sample code of patient 1 for preparing time domain image from ASCII file format (six classes):

Prepare respiratory image:

Import annotation file.

```
%Set up the Import Options and import the data (Can manually import by click 'import data' in home tap.
opts = delimitedTextImportOptions("NumVariables", 8);

% Specify range and delimiter
opts.DataLines = [1, Inf];
opts.Delimiter = ",";

% Specify column names and types
opts.VariableNames = ["Start", "Epoch", "Sleep", "Type", "Duration", "Detail1", "Detail2", "Back"];
opts.VariableTypes = ["datetime", "double", "categorical", "categorical", "string", "double", "double", "categorical"];

% Specify file level properties
opts.ExtraColumnsRule = "ignore";
opts.EmptyLineRule = "read";

% Specify variable properties
opts = setvaropts(opts, "Duration", "WhitespaceRule", "preserve");
opts = setvaropts(opts, ["Sleep", "Type", "Duration", "Back"], "EmptyFieldRule", "auto");
opts = setvaropts(opts, "Start", "InputFormat", "HH:mm:ss");
```

```
% Import the data
P1D1Scoredevent = readtable("D:\Sleep Patient 01_Doctor-20200827T050521Z-001\Sleep Patient 01_Doctor\P1D1_Scored event (Sleep stage, Respiratory events, Limb movement).TXT", opts)
```

Prepare epoch of all classes (except non respiratory event).

```
P1D1Scoredevent(P1D1Scoredevent.Sleep == 'AWAKE',:)=[]; % exclude awake
P1D1Scoredevent(P1D1Scoredevent.Epoch > 381,:)=[]; % exclude CPAP titration study
nume=P1D1Scoredevent.Epoch(find(P1D1Scoredevent.Type =='Obstructive Apnea'));
nume1=nume(randperm(size(nume,1)));
nume=P1D1Scoredevent.Epoch(find(P1D1Scoredevent.Type =='Hypopnea'));
nume2=nume(randperm(size(nume,1)));
nume=P1D1Scoredevent.Epoch(find(P1D1Scoredevent.Type =='Central Apnea'));
nume3=nume(randperm(size(nume,1)));
nume=P1D1Scoredevent.Epoch(find(P1D1Scoredevent.Type =='Mixed Apnea'));
nume4=nume(randperm(size(nume,1)));
nume=P1D1Scoredevent.Epoch(find(P1D1Scoredevent.Type =='SpO2 Desat'));
nume5=nume(randperm(size(nume,1)));
%save as nume 1 to 5 in shuffleepochp1.mat
```

Plot image of obstructive apnea class (changing the nume variable to change the class).

```
load shuffleepochp1.mat %import nume variable
% import all six signals
importdata("ASCii\P1D1_Abdo\P1_D1_Abdo_1-952.txt");
Abdo=ans.data(:,1);
importdata("ASCii\P1D1_Thor\P1_D1_Thor_1-952.txt");
Thor=ans.data(:,1);
```

```
importdata("ASCii\P1D1_Nasal Pressure\P1_D1_Nasal Pressure_1-952.txt");
NasalP=ans.data(:,1);
importdata("ASCii\P1D1_SpO2\P1_D1_SpO2_1-952.txt");
SpO2=ans.data(:,1);
importdata("ASCii\P1D1_Snore\P1_D1_Snore_1-952.txt")
Snore=ans.data(:,1);
importdata("ASCii\P1D1_Thermo\P1_D1_Thermo_1-952.txt")
Thermo=ans.data(:,1);

%Calculate time for Abdo Thor Nasal
N=913920;
Fs = 32;
dt = 1/Fs;
tsec=dt*(0:N-1)';
%rearrange time to have length of 30 seconds.
time=(reshape(tsec,960,952))';

%Calculate time for SpO2
N1=456960;
Fs1 = 16;
dt1 = 1/Fs1;
tsec1=dt1*(0:N1-1)';
time1=(reshape(tsec1,480,952))';

%Calculate time for Snore
N2=7311360;
Fs2 = 256;
dt2 = 1/Fs2;
tsec2=dt2*(0:N2-1)';
time2=(reshape(tsec2,7680,952))';
```

```
%rearrange each signal to have length of 30 seconds.
EPOCHAbdo = (reshape(Abdo,960,952))';
EPOCHThor = (reshape(Thor,960,952))';
EPOCHNasalP = (reshape(NasalP,960,952))';
EPOCHSpO2=(reshape(SpO2,480,952))';
EPOCHSnore=(reshape(Snore,7680,952))';
EPOCHThermo=(reshape(Thermo,960,952))';

%plot image
for i=1:137 %total image in obstructive apnea class

   figure(i)

   x1=time(nume1(i),:)
   x2=time1(nume1(i),:)
   x3=time2(nume1(i),:)

   ax1=subplot(6,1,1)
   ESnore=EPOCHSnore(nume1(i),:)
  %Normalize segmented data
   s=min(ESnore)
   t=max(ESnore)
   ESnore=(ESnore-s)/(t-s)
   plot(x3,ESnore)
   axis off

   ax2=subplot(6,1,2)
   ENasalP=EPOCHNasalP(nume1(i),:)
  %Normalize segmented data
   p=min(ENasalP)
   d=max(ENasalP)
```

```
ENasalP= (ENasalP-p)/(d-p)
plot(x1,ENasalP)
axis off

ax3=subplot(6,1,3)
EThermo=EPOCHThermo(nume1(i),:)
%Normalize segmented data
o=min(EThermo)
l=max(EThermo)
EThermo= (EThermo-o)/(l-o)
plot(x1,EThermo)
axis off

ax4=subplot(6,1,4)
EThor=EPOCHThor(nume1(i),:)
%Normalize segmented data
c=min(EThor)
v=max(EThor)
EThor= (EThor-c)/(v-c)
plot(x1,EThor)
axis off

ax5=subplot(6,1,5)
EAbdo=EPOCHAbdo(nume1(i),:)
%Normalize segmented data
m=min(EAbdo)
y=max(EAbdo)
EAbdo= (EAbdo-m)/(y-m)
plot(x1,EAbdo)
axis off
```

```
    ax6=subplot(6,1,6)
    ESpO2=EPOCHSpO2(nume1(i),:)
    plot(x2,ESpO2)
    axis off


   % save an image
    filename=sprintf("Obstructp1%d.jpg",i)
    saveas(figure(i),filename)

end
% respiratory images are done.
```

Prepare Non-respiratory image:

Import annotation file.

```
opts = delimitedTextImportOptions("NumVariables", 8);

% Specify range and delimiter
opts.DataLines = [1, Inf];
opts.Delimiter = ",";

% Specify column names and types
opts.VariableNames = ["Start", "Epoch", "Sleep", "Type", "Duration", "Detail1",
"Detail2", "Back"];
opts.VariableTypes = ["datetime", "double", "categorical", "categorical", "string",
"double", "double", "categorical"];

% Specify file level properties
opts.ExtraColumnsRule = "ignore";
opts.EmptyLineRule = "read";
```

```
% Specify variable properties
opts = setvaropts(opts, "Duration", "WhitespaceRule", "preserve");
opts = setvaropts(opts, ["Sleep", "Type", "Duration", "Back"], "EmptyFieldRule",
"auto");
opts = setvaropts(opts, "Start", "InputFormat", "HH:mm:ss");

% Import the data
P1D1Scoredevent = readtable("D:\Sleep Patient 01_Doctor-20200827T050521Z-
001\Sleep Patient 01_Doctor\P1D1_Scored event (Sleep stage, Respiratory events,
Limb movement).TXT", opts)
```

Prepare epoch of Non-respiratory class (exclude from respiratory file):

```
A1=P1D1Scoredevent.Epoch
A2=[1:952]';

for j=1:925
   A2(A1(j),1)=0;
end

X1=find(A2)

for a=1:287
    if X1(a) > 381 % split night
        X1(a)= 0
     end
end

 numee=X1(X1~=0) % all amount of non-res event
 nume=numee(randperm(size(numee,1))); %random non-respiratory epoch (use all
amount of epochs) and save nume variable in shuffleepochnonresp1.mat
```

If sum of five classes epochs is lower than non-res epochs, the amount of non-res class will be randomly select in the equal amount of five combination. If not, all amount of non-res epochs will be use.

Plot image of Non-respiratory class.

```
load shuffleepochnonres.mat
importdata("ASCii\P1D1_Abdo\P1_D1_Abdo_1-952.txt");
Abdo=ans.data(:,1);
importdata("ASCii\P1D1_Thor\P1_D1_Thor_1-952.txt");
Thor=ans.data(:,1);
importdata("ASCii\P1D1_Nasal Pressure\P1_D1_Nasal Pressure_1-952.txt");
NasalP=ans.data(:,1);
importdata("ASCii\P1D1_SpO2\P1_D1_SpO2_1-952.txt");
SpO2=ans.data(:,1);
importdata("ASCii\P1D1_Snore\P1_D1_Snore_1-952.txt")
Snore=ans.data(:,1);
importdata("ASCii\P1D1_Thermo\P1_D1_Thermo_1-952.txt")
Thermo=ans.data(:,1);

%Calculate time for Abdo Thor Nasal
N=913920;
Fs = 32;
dt = 1/Fs;
tsec=dt*(0:N-1)';
time=(reshape(tsec,960,952))';

%Calculate time for SpO2
N1=456960;
Fs1 = 16;
dt1 = 1/Fs1;
```

```
tsec1=dt1*(0:N1-1)';
time1=(reshape(tsec1,480,952))';

%Calculate time for Snore
N2=7311360;
Fs2 = 256;
dt2 = 1/Fs2;
tsec2=dt2*(0:N2-1)';
time2=(reshape(tsec2,7680,952))';

%rearrange each signal to have length of 30 seconds
EPOCHAbdo = (reshape(Abdo,960,952))';
EPOCHThor = (reshape(Thor,960,952))';
EPOCHNasalP = (reshape(NasalP,960,952))';
EPOCHSpO2=(reshape(SpO2,480,952))';
EPOCHSnore=(reshape(Snore,7680,952))';
EPOCHThermo=(reshape(Thermo,960,952))';


% plot image
for i=1:109 %total image in non-res class

   figure(i)

   x1=time(nume(i),:)
   x2=time1(nume(i),:)
   x3=time2(nume(i),:)

    ax1=subplot(6,1,1)
    ESnore=EPOCHSnore(nume(i),:)
   %Normalize segmented data
```

```
 s=min(ESnore)
 t=max(ESnore)
 ESnore=(ESnore-s)/(t-s)
 plot(x3,ESnore)
 axis off

 ax2=subplot(6,1,2)
 ENasalP=EPOCHNasalP(nume(i),:)
%Normalize segmented data
 p=min(ENasalP)
 d=max(ENasalP)
 ENasalP= (ENasalP-p)/(d-p)
 plot(x1,ENasalP)
 axis off

 ax3=subplot(6,1,3)
 EThermo=EPOCHThermo(nume(i),:)
%Normalize segmented data
 o=min(EThermo)
 l=max(EThermo)
 EThermo= (EThermo-o)/(l-o)
 plot(x1,EThermo)
 axis off

 ax4=subplot(6,1,4)
 EThor=EPOCHThor(nume(i),:)
%Normalize segmented data
c=min(EThor)
v=max(EThor)
EThor= (EThor-c)/(v-c)
plot(x1,EThor)
```

```
    axis off

    ax5=subplot(6,1,5)
    EAbdo=EPOCHAbdo(nume(i),:)
   %Normalize segmented data
    m=min(EAbdo)
    y=max(EAbdo)
    EAbdo= (EAbdo-m)/(y-m)
    plot(x1,EAbdo)
    axis off

   ax6=subplot(6,1,6)
   ESpO2=EPOCHSpO2(nume(i),:)
   plot(x2,ESpO2)
   axis off

    filename=sprintf("Nonresp1%d.jpg",i)
    saveas(figure(i),filename)

end
% non-respiratory images are done.
```

Sample code of patient 1 for preparing time domain image from EDF file format (six classes):

Prepare respiratory image:

Plot image of obstructive apnea class (changing the nume variable to change the class).

```
 load D:\'Sleep Patient 01_Doctor-20200827T050521Z-001'\'Sleep Patient 01_Doctor'\shuffleepochp1.mat %use same mat file from ASCII image version.
```

```
data=edfread("D:\seniorproject\Raw Unclipped EDF\P1\P1RawData.edf");

for i = 1:137

 f = figure;
 figure(f);

 ax1=subplot(6,1,1);
 temp1 = EDFepoch(data.Snore,nume1(i,1),256);
 axis off;

 ax2=subplot(6,1,2);
 temp2 = EDFepoch(data.NasalPressure,nume1(i,1),32);
 axis off;

 ax3=subplot(6,1,3);
 temp3 = EDFepoch(data.Thermo,nume1(i,1),32);
 axis off;

 ax4=subplot(6,1,4);
 temp4 = EDFepoch(data.Thor,nume1(i,1),32);
 axis off;

 ax5=subplot(6,1,5);
 temp5 = EDFepoch(data.Abdo,nume1(i,1),32);
 axis off;

 ax6=subplot(6,1,6);
 temp6 = EDFepoch(data.SpO2,nume1(i,1),16);
 axis off;
```

```
  filename=sprintf("Obstruct%d.jpg",i);
  imwrite(frame2im(getframe(f)),strcat("D:\Sleep Patient 01_Doctor-
20200827T050521Z-001\Sleep Patient 01_Doctor\Classedf\obstruct",filename));
  clf
 end

 function epo1 = EDFepoch(data,epochNum,fs)
    epo1=ones(fs*30,1);
    c=1;
    for i=((epochNum*30)-29):(epochNum*30)
       epo1((c*fs)-(fs-1):c*fs,1) = data{i,1};
       c=c+1;
    end
    plot(epo1)
 end
```

Prepare Non-respiratory image:

Plot image of Non-respiratory class.

```
  load D:\'Sleep Patient 01_Doctor-20200827T050521Z-001'\'Sleep Patient
01_Doctor'\shuffleepochnonresp1.mat
 %use same mat file from ASCII image version
 data=edfread("D:\seniorproject\Raw Unclipped EDF\P1\P1RawData.edf");

 for i = 1:109


  f = figure;
  figure(f);
```

```
ax1=subplot(6,1,1);
temp1 = EDFepoch(data.Snore,nume1(i,1),256);
axis off;

ax2=subplot(6,1,2);
temp2 = EDFepoch(data.NasalPressure,nume1(i,1),32);
axis off;

ax3=subplot(6,1,3);
temp3 = EDFepoch(data.Thermo,nume1(i,1),32);
axis off;

ax4=subplot(6,1,4);
temp4 = EDFepoch(data.Thor,nume1(i,1),32);
axis off;

ax5=subplot(6,1,5);
temp5 = EDFepoch(data.Abdo,nume1(i,1),32);
axis off;

ax6=subplot(6,1,6);
temp6 = EDFepoch(data.SpO2,nume1(i,1),16);
axis off;

filename=sprintf("Nonres%d.jpg",i);
imwrite(frame2im(getframe(f)),strcat("D:\Sleep Patient 01_Doctor-
20200827T050521Z-001\Sleep Patient 01_Doctor\Classedf\nonres",filename));
clf
end

function epo1 = EDFepoch(data,epochNum,fs)
```

```
    epo1=ones(fs*30,1);
    c=1;
    for i=((epochNum*30)-29):(epochNum*30)
        epo1((c*fs)-(fs-1):c*fs,1) = data{i,1};
        c=c+1;
    end
    plot(epo1)
end
```

Sample code of patient 1 for preparing frequency domain image of Nasal Pressure signal (from EDF file format):

Import annotation file that was modified by excel into workspace.

```
opts = delimitedTextImportOptions("NumVariables", 3);

% Specify range and delimiter
opts.DataLines = [2, Inf];
opts.Delimiter = ",";

% Specify column names and types
opts.VariableNames = ["Start", "Event", "Duration"];
opts.VariableTypes = ["datetime", "categorical", "datetime"];

% Specify file level properties
opts.ExtraColumnsRule = "ignore";
opts.EmptyLineRule = "read";

% Specify variable properties
opts = setvaropts(opts, "Event", "EmptyFieldRule", "auto");
opts = setvaropts(opts, "Start","InputFormat", "HH:mm:ss");
```

```
opts = setvaropts(opts, "Duration", "InputFormat","mm:ss.S");

% Import the data
Testscore1 = readtable("D:\Sleep Patient 01_Doctor-20200827T050521Z-001\Sleep Patient 01_Doctor\Testscore1.txt", opts)
```

Build new scored table.

```
stime = hours(22) + minutes(19) + seconds(49); %start time
time = []; %header
event = categorical([]); %header
duration = []; %header
for t = 1:182
    sec = second(Testscore1.Start(t) - stime);
    min = minute(Testscore1.Start(t) - stime);
    hr = hour(Testscore1.Start(t) - stime);
    interval = hr*3600 + min*60 + sec;
    dur = second(Testscore1.Duration(t)) + minute(Testscore1.Duration(t))*60;
    time(end+1) = interval;
    event(end+1) = Testscore1.Event(t);
    duration(end+1) = dur;
end
Nasal = table(time',event',duration');
Nasal.Properties.VariableNames = ["Start","Event","Duration"] %determine header
```

Label and store data of 10-second segment.

```
data = edfread("D:\seniorproject\Raw Unclipped EDF\P1\P1RawData.edf");
Nasal=data.NasalPressure
mnasal=cell2mat(Nasal)
```

```
nnasal=normalize(mnasal) %z-scorenormalize
cnasal=(reshape(nnasal,32,28566))';
for i=1:28566
   nasal{i,1} = double(cnasal(i,:))';
end

totalwindow = (381*30)-9; %calculate total window (total amount of image)
x = cell(totalwindow,1);
y = categorical([ones(totalwindow,1)]);
for i = 1:totalwindow
   x(i,1) = {window(NasalP, i)'};
   for j = 1:length(Nasal.Start)-1
     if ((i-1 >= Nasal.Start(j)) && (i-1 < Nasal.Start(j+1)) && (i+8 <=
(Nasal.Start(j)+Nasal.Duration(j)))) %condition for labeling
        y(i,1) = Nasal.Event(j);
        break
     end
   end
   if y(i,1) == categorical(cellstr('1'))
      y(i,1) = categorical(cellstr('Non'));
   end
end

%function  for segmentation
function result = window(d, N) %N=number of window

   temp = [ones(320, 1)];
   c = 1;
   for i = N:N+9
      index = 1;
      for j = ((c*32)-31):(c*32)
```

```
            temp(j,1) = d{i,1}(index,1);
            index = index + 1;
        end
        c = c + 1;
    end
    result = temp;
end
%save X & Y variable in P1Labelednasalnorm.mat
```

Spectrogram plot.

```
Load P1Labelednasalnorm.mat
f = figure;
for i = 1:13731 %total amount of image
spectrogram(x{i,1})
colorbar("off")
axis("off")
filename=sprintf("specp1%d.jpg",i);
imwrite(frame2im(getframe(f)),strcat("D:\Tae\specspo2normp5\",folder,"\",filenam
e));
clf
end
```

Sample code for resize an image.

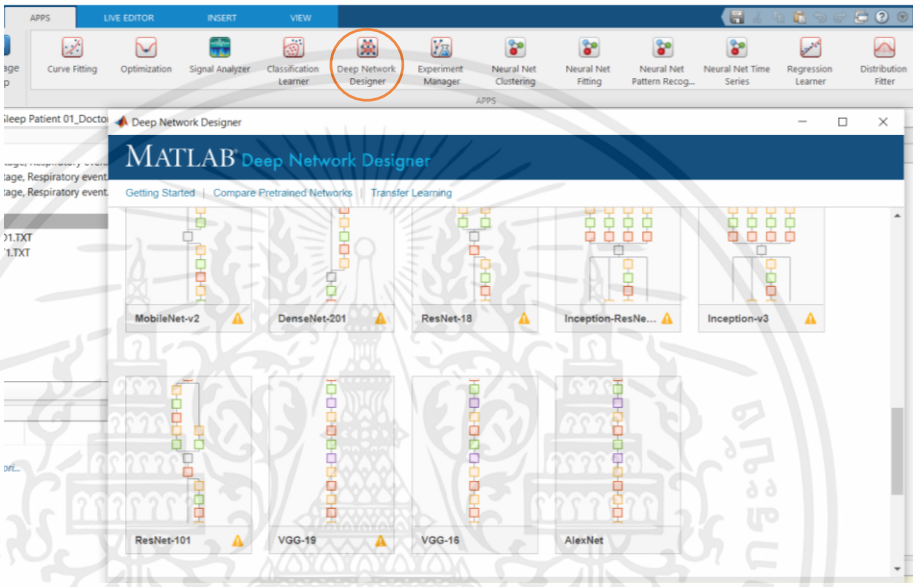
```
srcFiles=dir('Tae\specspo2normp5\Apnea\specp1*.jpg'); %directory to folder that
we want to resize.
for i=1:length(srcFiles)
    filename=strcat('Tae\specspo2normp5\Apnea\',srcFiles(i).name);
    im=imread(filename);
    k=imresize(im,[224,224]); % modify image size to be 224x224x3
```

```
    newfilename=strcat('resizespec\specspo2normp5\Apnea\',srcFiles(i).name);
    imwrite(k,newfilename,'jpg'); %save to new folder with same name
end
```

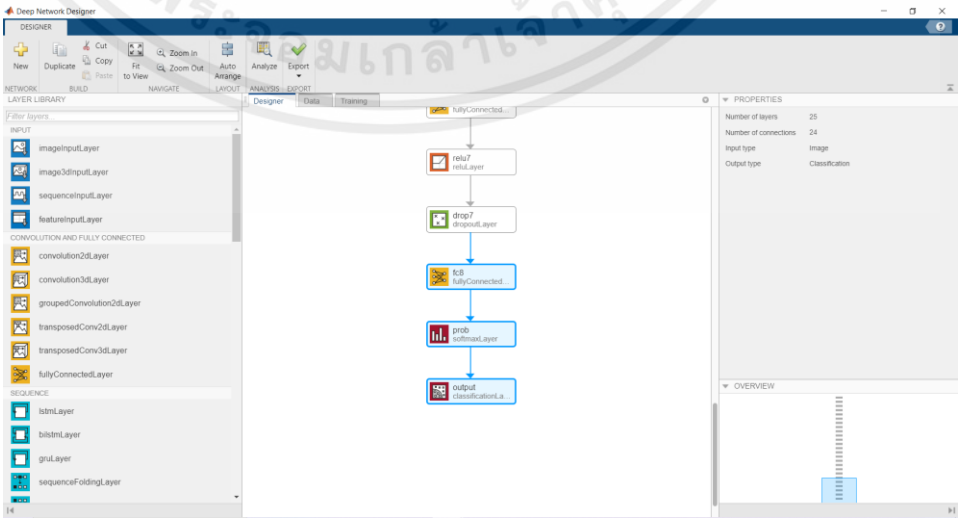
APPENDIX B
TRAINING AND TESTING

Step by step operation for training:

Click on deep network designer option and select the pre-trained network.



Modify last three layers of the network for matching with amount of our class. For example, in AlexNet network, the output size of fully connected layer needs to be changed into 3 or 6 instead of 1000. Click analyze after modified to check whether the network is correct or not.



Then import the data into workspace by coding.

Sample code for importing 6 classes time domain image (unequalized):

```
allImages=imageDatastore("Trainedfresnet\","IncludeSubfolders",true,'LabelSource'
,'foldernames');
[trainingImages,testImages]=splitEachLabel(allImages,0.9); %training 90%
%Preview image of each class
sampleImages = splitEachLabel(trainingImages,1,'randomized');
montage(sampleImages,'Size',[1 Inf])
title("Image for training (" +join(string(sampleImages.Labels'),"-")+ ")")
```

Sample code for importing 3 classes time domain image (equalized):

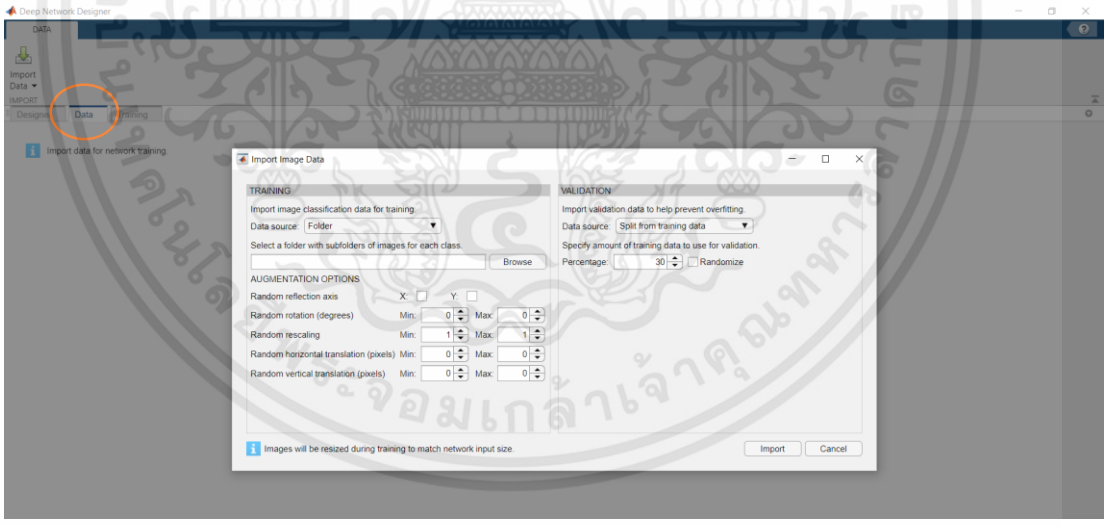
```
apnea=imageDatastore("C:\227newnormEDFimage3classes\Apnea\","LabelSource",
"foldernames");
hypopnea=imageDatastore("C:\227newnormEDFimage3classes\Hypopnea\","LabelS
ource","foldernames");
non=imageDatastore("C:\227newnormEDFimage3classes\Non\","LabelSource","fold
ernames");
 nhyp=imageDatastore(hypopnea.Files(1:217),"LabelSource","foldernames");
 nnon=imageDatastore(non.Files(1:217),"LabelSource","foldernames");
 imds=imageDatastore(cat(1,apnea.Files,nhyp.Files,nnon.Files));
 imds.Labels=cat(1,apnea.Labels,nhyp.Labels,nnon.Labels);
[trainingImages,testImages]=splitEachLabel(imds,0.9);
sampleImages = splitEachLabel(trainingImages,1,'randomized');
montage(sampleImages,'Size',[1 Inf])
title("Image for training (" +join(string(sampleImages.Labels'),"-")+ ")")
```

Sample code for importing 3 classes frequency domain image (equalized):

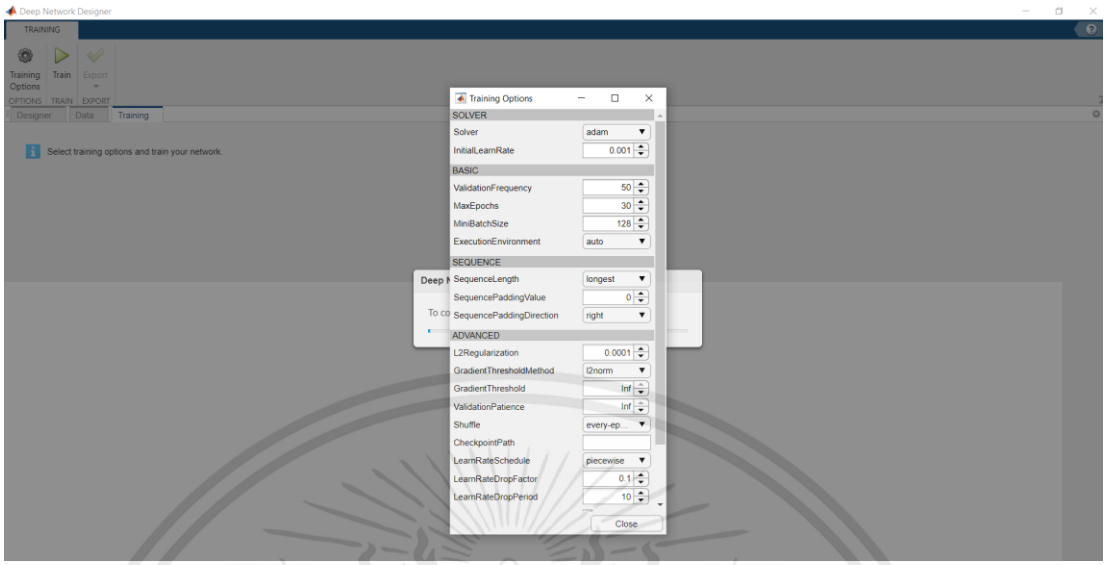
```
apnea=imageDatastore("C:\specnasalnorm\Apnea\","LabelSource","foldernames");
hypopnea=imageDatastore("C:\specnasalnorm\Hypopnea\","LabelSource","folderna
mes");
```

```
non=imageDatastore("C:\specnasalnorm\Non\","LabelSource","foldernames");
hypopneaa=shuffle(hypopnea);
nonn=shuffle(non);
nhyp=imageDatastore(hypopneaa.Files(1:5457),"LabelSource","foldernames");
nnon=imageDatastore(nonn.Files(1:5457),"LabelSource","foldernames");
imds=imageDatastore(cat(1,apnea.Files,nhyp.Files,nnon.Files));
imds.Labels=cat(1,apnea.Labels,nhyp.Labels,nnon.Labels);
[trainingImages,testImages]=splitEachLabel(imds,0.9,"randomized");
sampleImages = splitEachLabel(trainingImages,1,'randomized');
montage(sampleImages,'Size',[1 Inf])
title("Image for training (" +join(string(sampleImages.Labels'),"-")+ ")")
```

After ran the code, click ‘Data’ option. Then selected the data source from workspace and browse as training image set and click ‘import’ (validation ratio is selected to be 30%)



Click ‘Training’ header and go to ‘Training Options’ to adjust the option. Lastly, click ‘Train ▷’ button to training the network.



Sample code for classification or testing:

```
predictedLabels=classify(trainedNetwork_1,testImages)
accuracy=sum(predictedLabels == testImages.Labels) / numel(predictedLabels)
cm = confusionchart(testImages.Labels, predictedLabels, ...
    "Title","Confusion Matrix, Accuracy " + accuracy*100 + "%", ...
    "RowSummary", "row-normalized", "ColumnSummary","column-normalized");
```

APPENDIX C

PROCEEDING RESEARCH PAPER IN ICEAST2021

Automatic Sleep Data Scoring by Artificial Intelligence: A Pilot Study in Thai Population

Ukkrit Jansri
Department of Biomedical Engineering, School of Engineering
King Mongkut's Institute of Technology Ladkrabang
Faculty of Medicine
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
ukkrit.ja@kmitl.ac.th

Naricha Chirakalwasan
Division of Pulmonary and Critical Care Medicine, Department of Medicine, Faculty of Medicine Chulalongkorn University
Excellence Center for Sleep Disorders, King Chulalongkorn Memorial Hospital, Thai Red Cross Society
Bangkok, Thailand
narichac@hotmail.com

Busarakum Chaitusaney
Excellence Center for Sleep Disorders, King Chulalongkorn Memorial Hospital, Thai Red Cross Society
Bangkok, Thailand
busarakum.c@chula.ac.th

Supasuta Busayakanon
Department of Biomedical Engineering, School of Engineering
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
60011273@kmitl.ac.th

Thamonwan Khongjui
Department of Biomedical Engineering, School of Engineering
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
60011277@kmitl.ac.th

Suradej Tretriluxana
Department of Electronics Engineering, School of Engineering
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
suradej.tr@kmitl.ac.th
ORCID ID: 0000-0002-4359-5759

***Abstract*— Sleep apnea, a sleep-disordered breathing (SDB), is defined as repeatedly intermittent cessation of breathing during sleep. It causes various life-threatening diseases. The American Academy of Sleep Medicine (AASM) releases the manual for sleep data scoring. Patients with SDB are prescribed to be monitored at the sleep clinic where several physiological data are recorded, called polysomnogram (PSG). The massive PSG data must be scored by the well-trained expert before being diagnosed by the physician. Our research is to use the Artificial Intelligence (AI) in sleep data scoring, particularly in respiratory events detection. Three ready-made Convolution Neural Networks (CNN); AlexNet, ResNet-50, and VGG-16, with transfer learning were applied to classify 5 overnight PSG data from Chulalongkorn hospital. Our preliminary results showed that all networks provide higher classification result in European Data Format (EDF) than in the text (ASCII) formats (71% vs 54%). The ResNet-50 model structure performed better than the other two networks on both data formats. As expected, the visualized (EDF) data is better than the unconditioned (ASCII) data. Our future development is modifying learning model to increase the scoring performance from more recruited PSG data.**

Keywords—Deep learning, convolutional neural networks, sleep apnea classification, data format, transfer learning

I. INTRODUCTION

Sleep, which is the state of quiescent behavior, usually consumes about 8 hours a night in a normal daily cycle to conserve energy, consolidate memory formation, and reallocate energy from brain and metabolism [1]. Moreover, sleep is also an active process that continuously occurs from shallow to deep sleep and overturn [2]. However, the disruption during sleep by sleep disorders, such as sleep-disordered breathing (SDB), insomnia, and narcolepsy. [3], can substantially increase the risk for stress, mood disorders, hypertensions, cardiovascular disease, obesity, cognitive disorders, cancer, and mortality [4]. Sleep apnea, which is the one of SDB, is defined as a completely cessation of breathing for at least 10 seconds and then resuming to breathe again. Sleep apnea is classified into obstructive sleep apnea (OSA), which is associated with a persistence of inspiratory effort and is the most common type, central sleep apnea (CSA), which is associated with an absence of inspiratory effort, and mixed apnea, which is the combination of CSA and, followed by, OSA [5]. These types share the clinical symptoms as excessive daytime sleepiness (EDS), snoring, choking, awakening, and fatigue [6]. For the global prevalence, sleep apnea, especially in OSA, impacts on almost 1 billion of the global population, who has obesity or higher body mass index (BMI), male gender, older age, snoring, daytime sleepiness, and large neck circumference as the risk factors [7]. According to the global population, the prevalence studies in Thai population, who is in central [8] and southern [9] Thailand, also has been revealed the risk factors that attribute to sleep apnea. Consequently, the long-term effect of undiagnosed and untreated sleep apnea leads to various diseases including mental illness, metabolic disease, stroke, cancer, neurodegenerative disease, and cardiovascular disease [10]. Therefore, the precise and rapid diagnosis, which is associated with sleep study, is critically to plan and treat sleep apnea and to decrease these consequences.

Polysomnography (PSG), which is a gold standard tool for exploration, investigation, and diagnosis of sleep disorders in the sleep test at the laboratory and clinic, comprises of multiple signals including electroencephalogram (EEG), electro-oculogram (EOG), electromyogram (EMG), electrocardiogram (ECG), airflow, respiratory effort, and oxygen saturation [11]. These signals are recorded simultaneously and used to score the sleep stages, arousal events, respiratory events and cardiac events following the criteria of American Academy of Sleep Medicine (AASM), which recommends to use 30-second length of window or an epoch to score the sleep stage and events [12]. For scoring the respiratory events as the sleep apnea, the signals of airflow,

978-0-7381-4484-9/21/$31.00 ©2021 IEEE

respiratory effort, and oxygen saturation are considered, and its severity is calculated by Apnea-Hypopnea Index (AHI) [13]. However, most of the sleep scoring requires the highly experienced certificated technicians and specialized physicians to perform manually visual scoring throughout the period of sleep for making a decision to interpret and diagnose the sleep staging and respiratory events. This process consumes the time and has intensive labors.

To address this problem, many studies have concentrated on automatic sleep staging classification based on deep learning (DL) with developing the model, for instance, from raw single channel, such as DeepSleepNet [14] and SleepEEGNet [15] as 1-Dimensional (1D) data, to several channels of EEG, such as SLEEPNET [16] as 1D and 2-Dimensional (2D) data. Futhermore, a few studies have focused the scoring on respiratory events that deploying the convolutional neural networks (CNN) to learn the feature from 1D or 2D data and classify presence of sleep apnea and its severity [17], [18]. Accordingly, 1D data represents signal (sequence) of time series and 2D data represents spectrogram [19]. However, there is no use CNN to learn feature of the 2D image of waveforms from multiple channels, as human visually scoring the events from the monitor, including snore (for sound), nasal pressure (for airflow), thermistor (for airflow), thoracic (for respiratory effort), abdomen (for respiratory effort), and oxygen saturation channel. Thus, in this study, our propose is to deploy CNN to detect respiratory events in supervised approach. Following the limitation of training data, we adopt the feature extraction of transfer learning (TL) from the pre-trained network models, which include AlexNet [20], VGG-16 [21], and ResNet-50 [22], to learn the characteristics of these signals to make a classification of the respiratory events that are apnea, hypopnea, and nonrespiratory events. Moreover, we also propose to compare the performance among these pre-trained network models with American Standard Code for Information Interchange (ASCII) and European Data Format (EDF) data formats.

The following sections in this study comprise of dataset in Section II, methods and fundamental function of CNN in Section III, results in Section IV, discussion and conclusion in Section V.

II. Dataset

We collected five PSG recording data from Excellence Center of Sleep Disorders in King Chulalongkorn Memorial Hospital (Bangkok, Thailand). The raw data consists of snore (256 Hz), nasal pressure (32 Hz), thermistor (32 Hz), thoracic (32 Hz), abdominal (32 Hz), and oxygen saturation (16 Hz) signals that were exported from PSG by ProFusion PSG 4 software (Compumedics, Abbotsford, Australia). These signals are recommended by AASM and specialized physicians to use in classifying types of sleep apnea. For ASCII, which is the standard code set of representing characters and has been assigned a unique binary string, each single signal is exported and is read on a third-party application, such as Excel. For EDF, this is used in non-Compumedics PSG readers and is generally used for storage and exchange of digitized signal data [23]. However, we can import both ASCII and EDF with Matlab 2020b (MathWorks, Natick, MA) for creating 2D images, in which each image comprises of 6 signals, and categorized into each class following the annotation as a true label. We consider to combine the types of sleep apnea (OSA, CSA, and mixed apnea) into a single apnea class and we also use hypopnea and non-apnea classes in this study. Moreover, the size of images must be resized according to the size of input for feeding to the pre-trained network models. Fig. 1 shows the example of apnea class in ASCII and EDF. For this study, because of the unbalance data in each class, thus we consider to use a minimum data to balance all classes. Overall, the dataset contains 217 samples of apnea, 217 samples of hypopnea, and 217 samples of nonrespiratory.

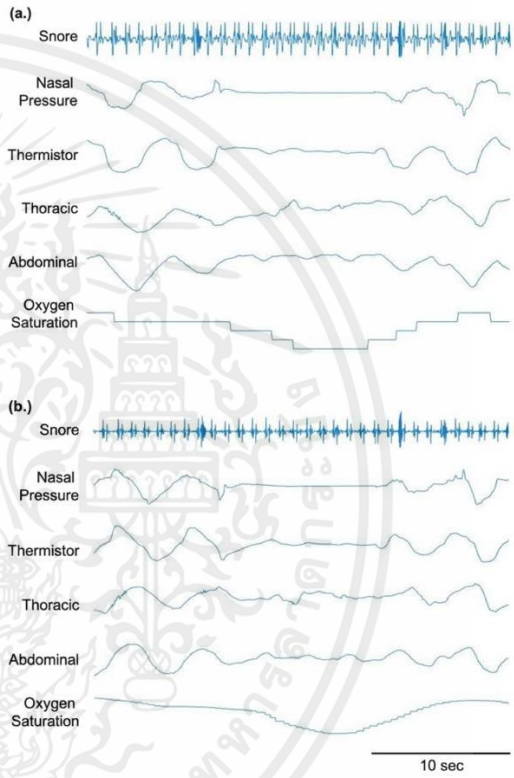


Fig. 1. Illustration of 6 signals image in OSA event at the same 30-s epoch from ASCII (a) and EDF (b).

III. Methods

In this study, we show the overall pipeline in Fig. 2.

A. CNN architecture

Convolutional Neural Networks (CNN), as a computer vision (CV), use an algorithm to process data which is in the form of multiple arrays and to solve the problem of image recognition. The popular CNN including AlexNet, VGG-16, and ResNet-50, has the common major parts that comprise of consecutively stacks of convolutional layers and pooling layers as the first part, which performs feature learning, and

fully connected layers as the second part, which performs classification [19], [24], [25].

- **Convolutional layers**: the image input size [W x H x D], where W is the width, H is the height, and D is the depth that represents the 3 color channels (Red, Green, and Blue), is locally patched to extract the feature with the filter, which contains a set of random weight vectors, of small size kernel. The kernel filter slides across the entire input image to compute the dot product, which is the linear transformation, to produce the output, which is the so-called a feature map. The feature map is followed by rectified linear unit (ReLU) to produce the non-linear transformation. Generally, the single convolutional layer consists of many feature maps [26], [27].
- **Pooling layers**: the pooling layers apply the down-sampling operation to reduce the magnitude of the feature maps along the W and H of spatial dimensions. For optimization, max pooling is typically to apply to control the overfitting and reduce the wight parameters [26], [27].
- **Fully connected layers**: the feature outputs, which are flattened into a column vector, of the preceding layer are connected to the fully connected layers and followed by SoftMax classifier to classify the image [26], [27].

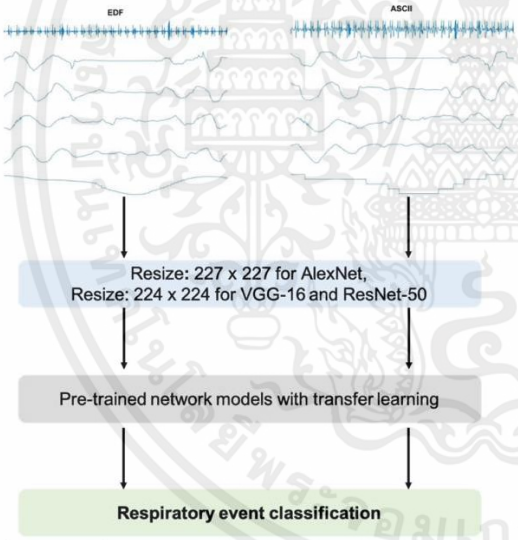


Fig. 2. Illustration of overall pipline of respiratory event classification

B. Transfer Learning

The transfer learning (TL) is the transferring prior knowledge from pre-trained network that is trained on the enormous datasets including ImageNet to perform the recognition tasks. We use the transfer learning from AlexNet, VGG-16, and ResNet-50, which is shown in Fig. 3, with replacing the last fully connected layer from 1,000 neurons to 3 neurons and retraining on our dataset to classify the respiratory events. Moreover, the weights of the pre-trained networks also are fine-tuned with continuous of backpropagation [26] - [28].

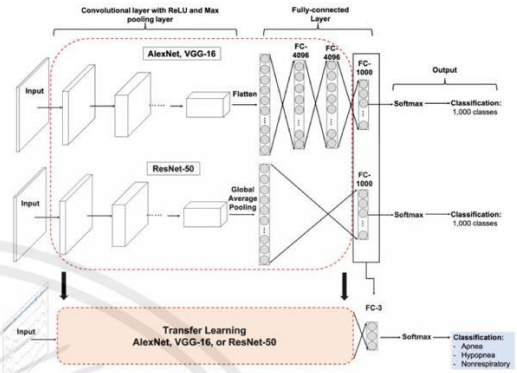


Fig. 3. Configuration of the transfer learning from AlexNet, VGG-16, and ResNet-50 with replacing the last fully connected (FC) layer from 1,000 neurons to 3 neurons to classify the dataset for respiratory events.

C. Training and Evaluation

We conduct the training and evaluation on Matlab 2020b. We divide the dataset into training/validation and test data with ratio 9:1 of dataset. The training data is splited into 70 percent for training and 30 percent for validation.

For training, we feed the training data to the pre-trained network models with appropriate size of [227 x 227] for AlexNet and [224 x 224] for VGG-16 and ResNet-50. We use adaptive moment estimation (Adam) for an optimizer and use cross-entropy loss function to train these models with learning rate of 0.001 and epoch of 30. However, we use a mini-batch size of 32 due to the GPU memory limitation.

For evaluation, the test data is used to evaluate the performance of each model by using accuracy value, which is the key measurement of this study, derived from confusion matrix. Moreover, we also use confusion matrix from the pre-trained network model with the highest accuracy to evaluate sensitivity, specificity, and precision. The formulas for calculating accuracy, sensitivity, specificity, and precision are shown:

$$Accuracy = (TP+TN)/(TP+TN+FP+FN) \quad (1)$$

$$Sensitivity = TP/(TP + FN) \quad (2)$$

$$Specificity = TN/(TN + FP) \quad (3)$$

$$Precision = TP/(TP + FP) \quad (4)$$

Where, TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

In this study, the pre-trained network models are trained on Intel Core i7 with 16 GB of RAM and GeForce RTX 2070 of GPU.

IV. Results

Our study utilized the pre-trained network models with transfer learning on data format. The key measurement is the accuracy as shown in Table 1.

TABLE I. COMPARISON ON ACCURACY IN DIFFERENT PRE-TRAINED NETWORK MODELS WITH 2D IMAGE FROM ASCII AND EDF

Model	Data Format	
	ASCII	EDF
AlexNet	33.33%	62.12%
VGG-16	54.55%	68.18%
ResNet-50	72.73%	**83.33%**

The best accuracy is given in bold.

This result showed that AlexNet, VGG-16, and ResNet-50 have the accuracy of 33.33%, 54.55%, and 72.73% in ASCII, respectively, whereas the AlexNet, VGG-16, and ResNet-50 have the accuracy of 62.12%, 68.18%, and 83.33% in EDF. Moreover, this result indicated that ResNet-50 has a highest accuracy among these models. Overall, all models showed higher accuracy of 71% in EDF than accuracy of 54% in ASCII.

Fig.4 shows the performance of ResNet-50 in EDF in classifying 3 classes of the respiratory. Other performances are shown in Table 2. ResNet-50 also showed sensitivity, specificity, and precision of 0.77, 0.63, and 0.85 in apnea, 0.86, 0.96, and 0.91 in hypopnea, and 0.86, 0.86, and 0.76 in non-respiratory, respectively.

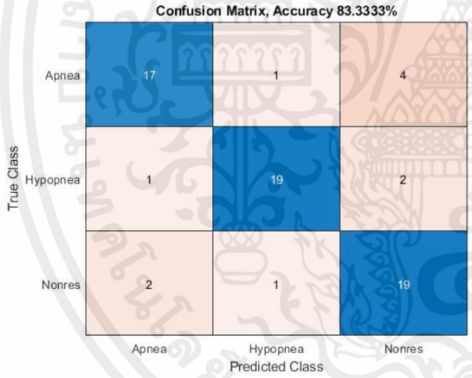


Fig. 4. Confusion matrix of ResNet-50 performing classification in the EDF. Nonres = nonrespiratory.

TABLE II. THE OTHER PERFORMANCE OF THE RESNET-50 MODEL FOR RESPIRATORY EVENT CLASSIFICATION

	Sensitivity	Specificity	Precision
Apnea	0.77	0.63	0.85
Hypopnea	0.86	**0.96**	0.91
Nonrespiratory	0.86	0.86	0.76
Average	0.83	0.82	0.84

The best result is given in bold.

V. Discussion and Conclusion

This present study demonstrates on the automatic sleep scoring especially in respiratory events by using the pre-trained network models with transfer learning. We consider to use 2D images because we need to mimic the task of human, which is manually visual scoring on sleep from the monitor. ResNet-50 shows highest accuracy and shows a sensitivity of 0.83, specificity of 0.82, and precision of 0.84. ResNet-50 achieves the best performance in our study would be its residual network that can be trained on much deeper layers and might reduce the degradation of gradient while backpropagation to update the weight parameters [22], whereas the AlexNet and VGG-16 have not the residual networks. However, the ResNet-50 model shows the quite performance due to the limitation of amount of data. Additionally, this study is a preliminary study on Thai population that we collect only five PSG recording data, whereas many studies use the open-source data or their local data [17]. Moreover, for data format, not surprisingly, EDF has shown the best format and is globally utilized for exchange in data of digital signals [29]. For the future work, we need to recruit more PSG data to attempt to improve the accuracy and need to compare 2D image data and spectrogram data.

In conclusion, ResNet-50 has shown the best the pre-trained network model with transfer learning to classify the respiratory events in EDF file and would be suitable to use for the starting point in development of the model for scoring the sleep in Thai population.

Acknowledgment

The first author would like to thank the Excellence Center of Sleep Disorders in King Chulalongkorn Memorial Hospital (Bangkok, Thailand) for their guidance in this study.

References

[1] R. C. Anafi, M. S. Kayser, and D. M. Raizen, "Exploring phylogeny to find the function of sleep," Nat. Rev. Neurosci., vol. 20, no. 2, pp. 109–116, Febuary 2019.

[2] C. P. Lee, and J. Abishegenaden, "A brief history of the biology of sleep," Ann. Acad. Med. Singap., vol. 45, no. 9, pp. 413–415, September 2016.

[3] M. W. Mahowald and C. H. Schenck, "Insights from studying human sleep disorders," Nature, vol. 437, no. 7063, pp. 1279–1285, October 2005.

[4] G. Medic, M. Wille, and M. E. H. Hemels, "Short- and long-term health consequences of sleep disruption," Nat. Sci. Sleep, vol. 9, pp. 151–161, May 2017.

[5] J. A. Dempsey, S. C. Veasey, B. J. Morgan, and C. P. O'Donnell, "Pathophysiology of sleep apnea," Physiol. Rev., vol. 90, pp. 47–112, January 2010.

[6] C. Guilleminault, A. Tilkian, and W. C. Dement, "The sleep apnea syndromes," Annu. Rev. Med, vol. 27, pp. 465–484, 1976.

[7] A. V. Benjafield, N. T. Ayas, P. R. Eastwood, R. Heinzer, M. S. M. Ip, M. J. Morrell, C. M. Nunez, S. R. Patel, T. Penzel, J. L. Pepin, P. E. Peppard, S. Sinha, S. Tufik, K. Valentine, and A. Malhotra, "Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis," Lancet Respir. Med., vol. 7, no. 8, pp. 687–698, August 2019.

[8] C. Neruntarat and S. Chantapant, "Prevalence of sleep apnea in HRH Princess Maha Chakri Srinthorn Medical Center, Thailand," Sleep Breath, vol. 15, no. 4, pp. 641–648, December 2011.

[9] K. Tawaranurak, P. Leelasawatsuk, and V. Chaiyarukjirakun, "Prevalence, risk factors and clinical manifestation of patients suspected as having obstructive sleep apnea in Songklanagarind Hospital Sleep Center," J. Health Sci. Med. Res., vol. 37, pp. 305–312, 2019.

[10] L. S. Gasper, A. R. Alvaro, J. Moita, and C. Cavadas, "Obstructive sleep apnea and hallmarks of aging," Trends Mol. Med., vol. 23, no. 8, pp. 675–692, August 2017.

[11] M. Hirshkowitz, "Polysomnography: understanding this technology's past might guide future developments," IEEE Pulse, vol. 5, pp. 26–28, September 2014.

[12] B. Jafari and V. Mohsenin, "Polysomnography," Clin. Chest Med., vol. 31, no. 2, pp. 287–297, June 2010.

[13] R. B. Berry, R. Brooks, C. E. Gamaldo, S. M. Harding, R. M. Lloyd, S. F. Quan, M. M. Troester, and B. V. Vaughn; for the American Academy of Sleep Medicine, "The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specification," version 2.4, Darien, IL: American Academy of Sleep Medicine, 2017.

[14] A. Supratak, H. Dong, C. Wu, and Y. Guo, "DeepSleepNet: a model for automatic sleep stage scoring based on raw single-channel EEG," IEEE Trans. Neural. Syst. Rehabil. Eng., vol.25, no.11, pp. 1998–2008, November 2017.

[15] S. Mousavi, F. Afghah, and U. R. Acharya, "SleepEEGNet: automated sleep stage scoring with sequence to sequence deep learinng approach," PLoS One, vol. 14, no.5, pp. 1–15, May 2019.

[16] S. Biswal, J. Kulas, H. Sun, B. Goparaju, M. B. Westover, M. T. Bianchi, and J. Sun, "SLEEPNET: automated sleep staging system via deep learning," arXiv: 1707.08262v1[cs.LG], pp. 1–17 July 2017.

[17] S. Biswal, H. Sun, B. Goparaju, M. B. Westover, J. Sun, and M. T. Bianchi, "Expert-level sleep scoring with deep neural networks," J. Am. Med. Inform. Assoc., vol.25, no.12, pp. 1643–1650, December 2018.

[18] D. Dey, S. Chaudhuri, and S, Munshi, "Obstructive sleep apnoea detection using convolutional neural network based deep learning framework," Biomed. Eng. Lett., vol. 8, no.1, pp. 95–100, December 2017.

[19] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol.521, pp. 436–444, May 2015.

[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," Adv. Neural. Inf. Process. Syst., pp. 1097–1105, 2012.

[21] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv: 1409.1556v6 [cs.CV], pp. 1–14 April 2014.

[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," arXiv: 1512.03385v1 [cs.CV], pp. 1–12, December 2015.

[23] B. Kemp, A. Varri, A. C. Rosa, K. D. Nielsen, and J. Gade, "A simple format for exchange of digitized polygraphic recordings," Electroencephalogr. Clin. Neurophysiol., vol. 82, no. 5, pp. 391–393, May 1992.

[24] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, and J. Dean, "A guide to deep learning in healthcare," Nat. Med., vol. 25, no. 1, pp. 24–29, January 2019.

[25] A. Elhassouny and F. Smarandache, "Trends in deep convolutional neural networks architectures: a review," 2019 ICCSRE, Agadir, Moroco, pp. 1–8, 2019.

[26] M. Sokolovsky, F. Guerrero, S. Paisarnsrisomsuk, C. Ruiz, and S. A. Alvarez, "Deep learning for automated feature discovery and classification of sleep stages," IEEE/ACM Trans. Comput. Biol. Bioinform., vol.17, no. 6, pp. 1835–1845, November – December 2020.

[27] N. Jmour, S. Zayen, and A. Abdelkrim, "Convolutional neural networks for image classification," IC_ASET, Hammamet, pp. 397–402, 2018.

[28] F. Demir, A. Sengur, and V. Bajaj, "Convolutional neural networks based efficient approach for classification of lung diseases," Health Inf. Sci. Syst., vol. 8, no.1, pp. 1–8, December 2020.

[29] T. Penzel, M. Hirshkowitz, J. Harsh, R. D. Chervin, N. Butkov, M. Kryger, B. Malow, M. V. Vitiello, M. H. Silber, C. A. Kushida, A. L. Chesson jr, "Digital analysis and technical specifications," J. Clin. Sleep Med., vol. 3, no. 2, pp. 109–120, March 2007.