MONASH University

Asynchronous Breathing and Patient Ventilator Interaction Assessment: A Machine Learning Approach

Loo Nien Loong

BSc. (Honours) in Mechatronics Engineering

A thesis submitted for the degree of Doctor of Philosophy at

Monash University in 2020

School of Engineering

Copyright Notice

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

Abstract

Mechanical ventilation (MV) is a critical treatment to patients with acute respiratory failure (ARF). However, suboptimal MV settings may lead to poor patient-ventilator interaction (PVI) which predisposes to frequent occurrence of asynchronous breathing (AB). Effort in optimising PVI to reduce AB occurrence is often an arduous task, as PVI assessment is typically evaluated retrospectively by computing asynchronous index (AI). Thus, the prevalence and consequences of AB are frequently underestimated. Hence, an approach capable of detecting and evaluating the magnitude of AB could potentially provide better MV management.

This thesis investigates the quality of PVI on patient's outcome by computing AB frequency and its corresponding magnitude of AB using machine learning approach. A clinical observational trial was carried out to collect airway waveform data of mechanical ventilated respiratory failure patients. The trial successfully collected around 2 million breathing cycles. These data were then manually classified into normal and asynchronous breaths and subsequently used for validating and testing of the machine learning approach. Convolutional Neural Network (CNN) models were trained to detect AB using prospective and retrospective mechanical ventilated respiratory failure patient's airway waveform data. The CNN model trained with one-dimension airway waveform data structure was able to detect AB with 90% accuracy, suggesting that CNN model can be used for AB detection.

In addition, a Generative Adversarial Network (GAN) and Convolutional Autoencoder (CAE) algorithms were also developed to quantify the magnitude of each detected AB. These methods were capable of reconstructing asynchronous affected airway waveform data to a

normal airway waveform. The reconstruction enabled the quantification of the magnitude of asynchrony. The ability of these developed machine learning models to reconstruct the airway waveform data were compared with an existing model-based method. We found that among these methods, CAE outperformed the other two in reconstructing AB with least distortion with mean absolute percentage error (MAPE) of 0.21%.

Finally, a metric, Ventilator-Interaction (VI) was developed to assess the quality of patient-ventilation interaction. We investigated the AI and VI in 18 mechanically ventilated respiratory failure patients by performing correlation analysis between patient's AI and PVI with their arterial blood gases (ABG). It was found that patients recruited in this study have varying AI and both metrics demonstrated different magnitudes. However, it also showed that, despite patient's AI was high, the magnitude of AB induced could be minimal, suggesting that AI alone may not be sufficient to monitor the quality of patient interaction. The developed VI metric may potentially help clinicians to assess the quality of MV treatment, as VI measures the 'resistivity' between the patients with the ventilatory support.

In summary, the machine learning models presented in this thesis are capable of detecting and quantifying the magnitude of AB. These models can potentially aid clinicians to evaluate the patient's condition during lung treatment; thus, allowing better MV decision making.

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution. To the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Print Name: Loo Nien Loong

Date: 26th of November 2020

Acknowledgements

This work could not have been possible completed without the support and assistance of so many people. Immeasurable appreciation for the help and support to the following individuals who have contributed and supported me throughout the course of this thesis completion.

I would like to thank my supervisor Dr. Chiew Yeong Shiong and co-supervisor Assoc. Prof Tan Chee Pin who have provided valuable advice and expertly guidance throughout this research. I acknowledge their valuable and insightful comments and suggestions that help to improve my work. I also want to express my gratitude for their continuous positive feedback and encouragement motivated me to accomplish the project objectives successfully in a timely manner.

I would also like to thank Prof. Basri and Dr. Azrina for their benevolence and unselfishly act to help us to screen and recruit patients into clinical trial for data collection. Besides that, I would also like to thank nurse, Ms. Fizah, medical assistance, Mr. Kamal and clinicians from International Islamic University Malaysia (IIUM), Kuantan Campus for their unselfishly assistance in patient data collection and equipment set up. This project could not be possibly accomplished without their support and assistance.

My special thanks go to Ganesh, See Long, and Qing Arn for their outstanding assistance, time and effort contributed in assisting me to set-up equipment in intensive care unit at IIUM and analyse the collected data. They were also kindly provided several suggestions and advices to improve the technical aspects of the set-up and enhance the quality of my project.

Besides that, I also wish to express my deepest gratitude by thanking the laboratory technical officers Mr. Bathmanathan and Mr. Tharmaa for providing technical support while I was working on my project in the mechatronics lab. I am indebted to their selfless support for lending me a powerful workstation to run my computation-intensive simulations. My project could not be possibly accomplished in time without their assistance.

This acknowledgement section is not complete without mentioning my family's contribution. While my PhD journey is bumpy and winding, my parents had provided tremendous and continuous moral support and encouragements. To Mum and Dad for their caring and support and for always asking me how progress is going. Thank you very much for being there, I could not reach to this stage without their perpetual inspiration and love.

Last but not least, I must acknowledge and express my deepest gratitude for the financial support from Ministry of Higher Education Malaysia (MOHE) Fundamental research grant scheme (FRGS) (Ref: FRGS/1/2016/TK03/MUSM/03/2), Monash University Malaysia Advance Engineering Platform (AEP) and MedTech Core University of Canterbury. I am extremely grateful and appreciate their support financial throughout my PhD journey.

Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 8

List of Tables

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Abbreviations

AB	Asynchronous Breathing
ABG	Arterial Blood Gases
ABReCA	Asynchronous Breathing Reconstruction Convolutional Autoencoder
ABReGAN	Asynchronous Breathing Reconstruction Generative Adversarial Network
Adam	Adaptive Moment Estimation
AI	Asynchronous Index
ALIEN	Automated Logging of Inspiratory and Expiratory Non-Synchronized Breathing
ANZCTR	Australia New Zealand Clinical Trial Registry
APACHE II	Acute Physiology and Chronic Health Evaluation II
APE	Absolute Percentage Error
ARDS	Acute Respiratory Distress Syndrome
ARF	Acute Respiratory Failure
AUC_{Rec}	Area Under the Curve of Reconstructed Airway Pressure
AUC_{Asyn}	Area Under the Curve of Asynchronous Airway Pressure
b	Trainable Bias
BiPAP	Bilevel Positive Airway Pressure
CAE	Convolutional Autoencoder
CAP	Community-Acquired Pneumonia
CARE Trial	Clinical Application of Respiratory Elastance Trial
CNN	Convolutional Neural Network
CO_2	Carbon Dioxide
CPAP	Continuous Positive Airway Pressure
CURE Soft	Clinical Utilisation of Respiratory Elastance Software
D	Discriminator
E	Elastance
E_{ac}	Elastance of Target Airway Pressure
Ea_{di}	Electrical Activity of The Diaphragm
ECG	Electrocardiography
E_{rec}	Elastance of Reconstructed Airway Pressure
F	Number of Convolutional Filter
FiO_2	Fractional Inspired Oxygen
T	Ground Truth
G	Generator
GAN	Generative Adversarial Network
HAP	Hospital-Acquired Pneumonia
i	Input
i_{noise}	Random Uniform Noise Sample
ICU	Intensive Care Unit
IIUM	International Islamic University Malaysia
IPR	Iterative Pressure Reconstruction
IQR	Interquartile Range
IREC	IIUM Research Ethics Committee
J	Cross-entropy Loss Function
L	Loss Function
LSTM	Long Short-Term Memory
m	Unpooling Temporal Step Size
M	Width of Trainable Weights
M_{asyn}	Magnitude of Asynchrony
MAPE	Mean Absolute Percentage Error

MSE	Mean Squared Error
MV	Mechanical Ventilation
η	Min-Batch Size
N	Height of Trainable Weights
NAVA	Neurally Assisted Ventilatory Assist
NUC	Next Unit of Computing
O	Output of Convolutional Layer
O_2	Oxygen
P	Number of Encoder/Decoder Layer
PaO_2/ FiO_2 ratio	Ratio of Partial Pressure of Oxygen to Fractional Inspired Oxygen
$PaCO_2$	Partial Pressure of Carbon Dioxide
PaO_2	Partial Pressure of Oxygen
Paw	Airway Pressure
PC	Pressure Control
PCA	Principal Component Analysis
PEEP	Positive End Expiratory Pressure
PEEPi	Intrinsic Positive End-Expiratory Pressure
P_e	Patient Effort
P_0	Pressure Offset
PREDATOR	Pressure Reconstruction by Eliminating the Demand Effect of Spontaneous Respiration
PVI	Patient Ventilator Interaction
Q	Flow
R	Resistance
R_{rec}	Resistance of Reconstructed Airway Pressure
R_{ac}	Resistance of Target Airway Pressure
RAM	Random Access Memory
RM	Respiratory Mechanics
RM_{Act}	Respiratory Mechanics of Target Airway Pressure
RM_{Rec}	Respiratory Mechanics of Reconstructed Airway Pressure
ReLU	Rectified Linear Unit
SIMV	Synchronized Intermittent-Mandatory Ventilation
SOFA	Sequential Organ Failure Assessment
SPONT	Spontaneous breathing
SSD	Solid State Drive
SSE	Sum Squared Error
SVC	Superior Vena Cava
t	Time
V	Volume
VC	Volume Control
VILI	Ventilator-Induced Lung Injury
VI	Ventilator-Interaction Index
W	Trainable Weights
$\bar{X}$	Generator Reconstructed Output
$\hat{x}$	Normalized Breathing Cycle
$\bar{Y}$	Encoder Reconstructed Output
$\hat{y}$	Predicted Output from CNN Model
z	Decoder Reconstructed Output
α	Batch Normalization Momentum Value
ß	*Adam*'s Momentum Term
γ	Allowable Negative Values for LeakyReLU
ℓ	Convolutional Layer
σ	Activation Function
θ	Trainable Parameters

Chapter 1
Introduction

1.1 Respiratory Physiology

Human respiratory system is part of our human body system; the function is to maintain adequate supply of oxygen (O_2), an element needed for breaking down sugars and fatty acid to produce energy as well as to ensure sufficient expulsion of carbon dioxide (CO_2), a by-product of metabolism (Bates, 2009). Figure 1.1 shows an overview structure of respiratory system.

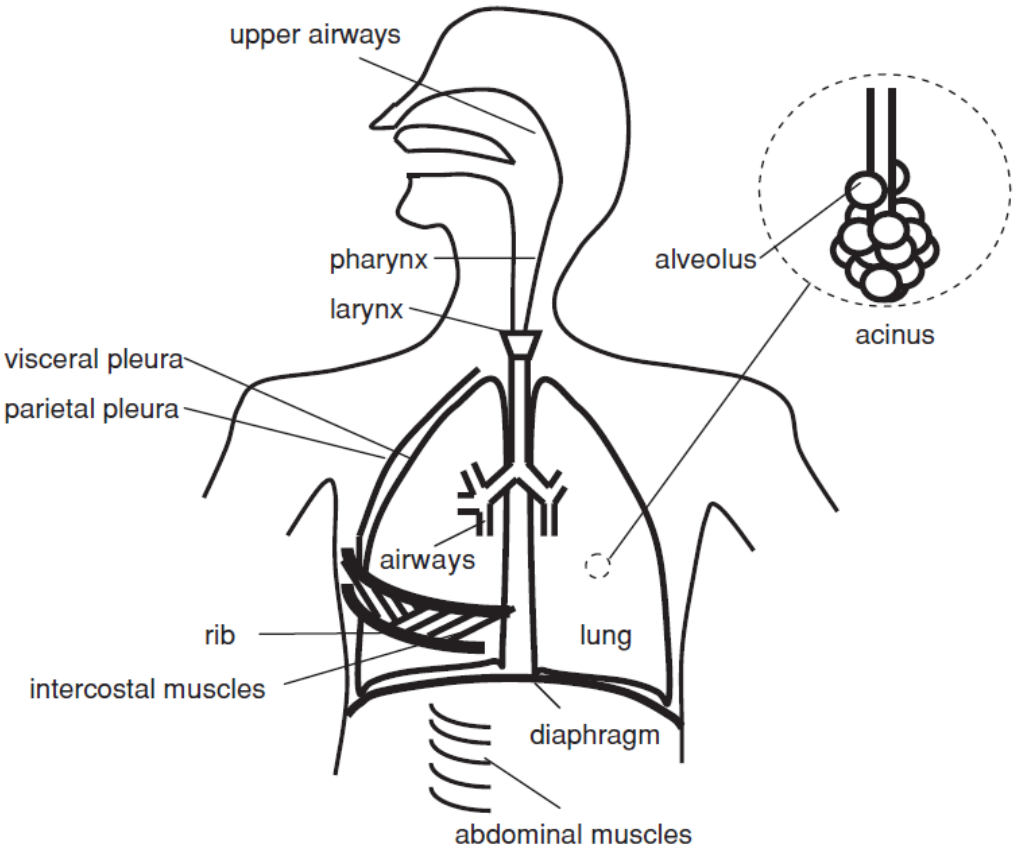

Figure 1.1: Overall structure of the human respiratory system (Bates, 2009)

In order to channel the air into the lungs from the atmosphere, a differential pressure must be created between the surrounding environment and inside the lung, and this can be done

through muscular work. During inspiration, the external intercostal muscles and diaphragm muscles contract to increase the volume of the thoracic cavity. The increment in volume creates negative pressure in the lung; thus, enabling air from atmosphere to flow into the lung. Conversely, during expiration, the inspiratory muscles relax to allow elastic recoil of the chest wall to deflate the lungs until the pressure in the lung is greater than atmospheric pressure; thus, the air flows out of the lung. These processes are known as ventilation, a muscular pumping action to draw air from the atmosphere into the lung and expel air from lungs back to the environment. Figure 1.2 shows the process of inspiration and expiration.

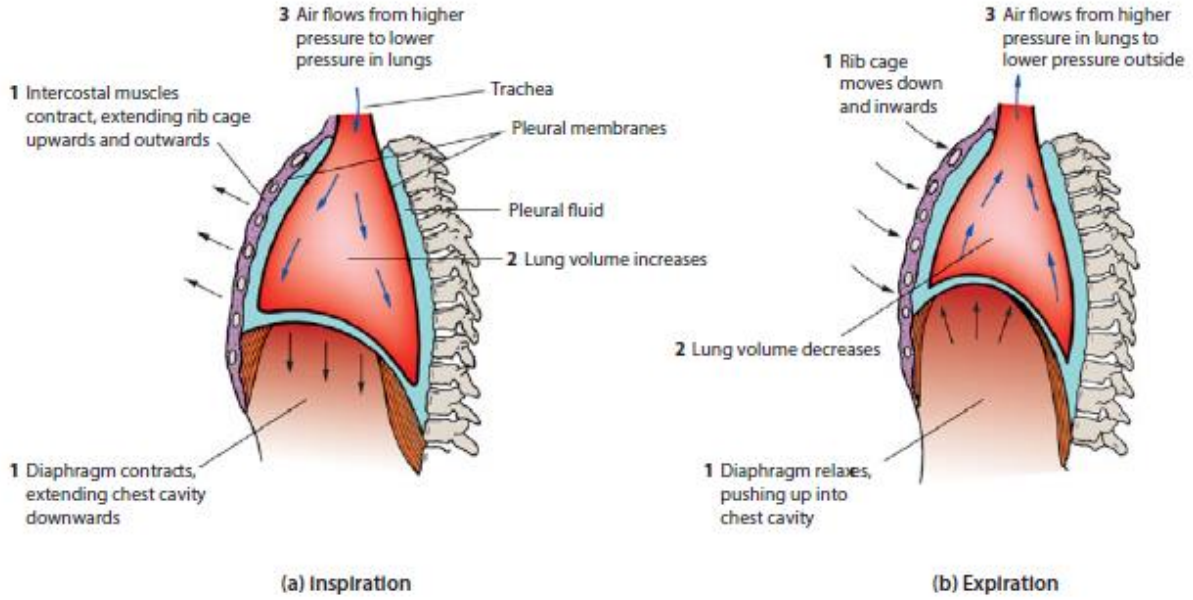

Figure 1.2. During inspiration, the chest moves outwards and upwards to increase volume; conversely, when the chest contracts, air is rushed out from lungs due to the increase of lungs pressure during expiration (Newton & Joyce, 2010).

Lung is adapted for gaseous exchange as the vast surface area consists of millions of alveoli enhances the efficiency to diffuse oxygen into the blood capillaries. This is achieved by possessing a tree-like structure that consists of branches that become progressively abundant as they undergo bifurcations. Trachea and alveoli are bifurcated about 23 times into

an area approximately the size of a tennis court (Barrett et al., 2019; Bates, 2009). Figure 1.3 shows the branching between trachea and alveolar in the human lungs.

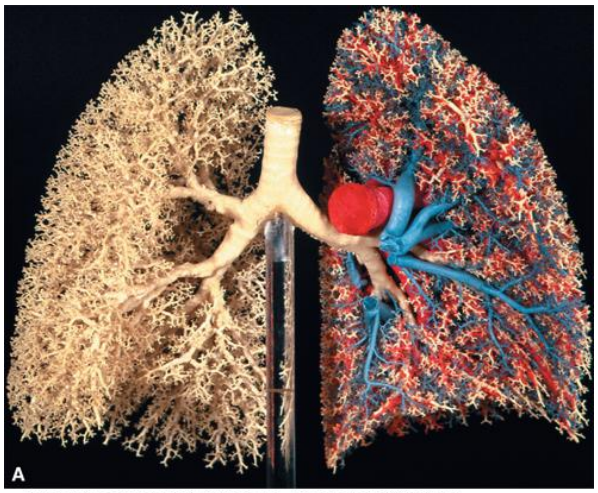

Figure 1.3: A dichotomous branching between trachea and alveolar in human lung (Barrett et al., 2019).

Perfusion or gaseous exchange process occurs at alveolus where the oxygen is diffused into blood capillaries and binds with haemoglobin to form oxyhaemogoblin; whereas, carbon dioxide diffuses out of blood capillaries and expels into the atmosphere during expiration. Typical adult human lung contains around 480 million alveoli with an average diameter of 200 μm (Ochs et al., 2004). Figure 1.4 shows the cross section of capillaries. Type I and Type II cells are major cells forming the alveolar wall (pneumocytes). Type I cells play a crucial role in forming the alveolar wall; whereas, Type II cells secrete proteins and lipids to reduce the surface tension of alveolar to prevent collapse. The membrane separating the alveolar air and the blood has an average thickness of 2μm (Knudsen & Ochs, 2018). Hence, huge respiratory surface area and thin alveolar promote rapid gases exchange; thus, allowing the body to maintain an adequate level of arterial partial pressure of oxygen (PaO_2) and arterial pressure of carbon dioxide ($PaCO_2$). Furthermore, satisfactory elimination of carbon dioxide helps the body to maintain pH in a normal range.

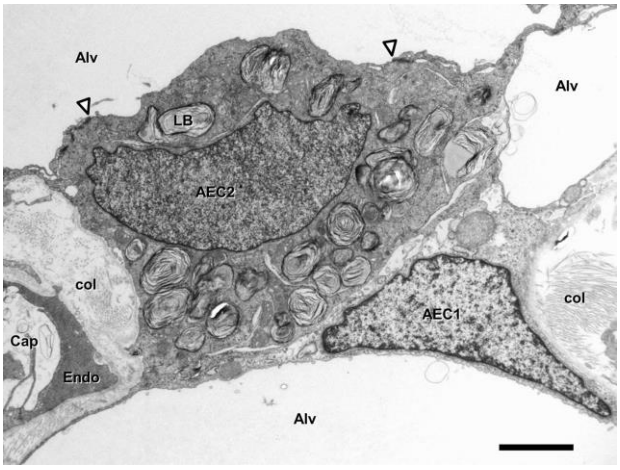

Figure 1.4: A cross section of capillaries under elector microscope showing Type I (AEC 1), surfactant-storing lamellar bodies (LB) in Type II (AEC 2) cell, collagen fibrils (col), alveolar lumen (Alv), capillary lumen (Cap) and (Endo) capillary endothelial cell (Knudsen & Ochs, 2018).

1.2 Respiratory Failure

Respiratory failure is a condition where the lung fails in gas exchange function (Pappert et al., 1994). This condition leads to CO_2 retention or O_2 inadequacy in the blood system. The respiratory system consists of two parts: the lung governs gas-exchanging perfusion process, and the pump controls ventilation in the lungs (Roussos & Koutsoukou, 2003). Hypoventilation or poor ventilation occurs when air fails to enter the alveoli; while, hypoxemia or poor perfusion results from ventilation/perfusion mismatching. These are the main factors of respiratory failure.

The pump that controls ventilation consists of chest wall, respiratory muscles and central nervous system. Damage to any organ in the system potentially lead to hypoventilation. Hypoventilation usually caused by central nervous system defection, respiratory muscles fatigue and neuromuscular transmission impairment. Unsuccessful to create negative intrapleural pressure by respiratory muscles will result in respiratory failure as the air fails to enter the lungs. Besides that, the reduction of effective or healthy surface area for gas

exchange perfusion due to blockage in conducting airways or fluid accumulation in alveolar is another reason for respiratory failure. Presence of fluid in the lung or known as pulmonary oedema impedes the oxygen from diffusing into the capillaries; thus, reduces the oxygen in the body and causes hypoxemia. Causes of pulmonary oedema include pneumonia, near drowning, and chemical fumes inhalation.

1.2.1 Acute Respiratory Distress Syndrome

Acute respiratory distress syndrome (ARDS) is a form of severe respiratory failure and was first characterized by Ashbaugh et al. (Ashbaugh et al., 1967). The definition of ARDS has evolved over the years due the lack of diagnostic biomarkers especially on the inflammation and permeability of lungs; hence, the definition often relies on clinical features and chest imaging as surrogate (Thompson et al., 2017). The Berlin definition of ARDS proposed in 2012 includes the use of computed tomography for qualifying opacities diagnostics to achieve a more reliable definition to improve case recognition (Ferguson et al., 2012). ARDS is defined as follows:

- The acute disorder onset time is within 1 week of clinical insult.
- Bilateral opacities are consistent with pulmonary oedema on chest radiographs and not fully explained by effusions, lobar collapse or nodules.

ARDS can be categorized into mild, moderate and severe by computing the PaO_2/FiO_2 ratio with a minimum level of 5 cmH_2O positive end expiratory pressure (PEEP) or non-invasive continuous positive airway pressure (CPAP) for mild ARDS. The categories are shown as below:

- Mild ARDS ($200 < PaO_2/FiO_2 \leq 300$ *mmHg*)
- Moderate ARDS ($100 < PaO_2/FiO_2 \leq 200$ *mmHg*)
- Severe ARDS ($PaO_2/FiO_2 \leq 100$ *mmHg*)

ARDS can occur due to widespread inflammation in the lung and potentially induce injuries to alveolar or cause pulmonary oedema and atelectasis (Bernard et al., 1994; Thompson et al., 2017). This further impairs the capability and effectiveness of ventilation and perfusion process in the lung. Common etiologic risk factors to ARDS development include pneumonia, pancreatitis, pulmonary contusion, drowning, drug overdose, major trauma, severe burns and pulmonary vasculitis (Ferguson et al., 2012). Despite advancement in supportive care; studies show that ARDS still imposes high mortality rate at range from 30% to 60% (Bellani et al., 2016; Phua et al., 2009; Villar et al., 2011). Furthermore, reports also suggested that sub-optimal administrated treatment could induce lung injuries (Brower et al., 2000; Meade et al., 2008; Mercat et al., 2008) resulting in risk of death and intensive care unit (ICU) stay increment. This implies that, death due to ARDS is sometimes caused by injuries during support treatment.

1.3 Mechanical Ventilation

Mechanical ventilation (MV) is a vital treatment for respiratory failure patient with the purpose of supporting their work of breathing to reinstate or maintain oxygenation. The use of ventilatory assistance can be traced back to 19th century where ventilatory assistance was provided via negative pressure ventilator (Kacmarek, 2011; Arthur S. Slutsky, 2015). The device was an air-tight box with the patient maintained in a sitting position. The ventilation was achieved via sub-atmospheric pressure generated by pumping air into and out of the device manually. In the 1930s, negative pressure ventilation became more clinical realistic when the use of Drinker iron lung negative pressure ventilator was deployed to treat polio patients (Drinker & Shaw, 1929). However, the effort to avoid device leakage as well as difficulty in maintaining effective ventilation leads to negative pressure ventilator elimination

and replacement by positive pressure ventilators. Positive-pressure ventilators allow PEEP establishment and able to sustain high airway pressure; this increases the lung's ability to recruit and retain alveoli for better ventilation-perfusion matching, atelectasis decrement and alveolar ventilation improvement (Luciano Gattinoni et al., 2010; Rossi et al., 1994).

Rapid advancement in technology and more recent research has improved the functions and reliability of the positive-pressure ventilators (Kacmarek, 2011) such as introducing various MV modes and treatment strategies to improve the chance for lung recovery (Downs et al., 1974; Singh et al., 2014; Stock et al., 1987; van der Staay & Chatburn, 2018). Despite these methods are introduced to improve treatment; different modes and approaches pose different functions and advantages. Subsequently, this complicates clinical decision making. Importantly, there is a lack of guidelines, consensus and standardisation on MV management. This limitation often leads to varying MV settings or, in some cases, fails to optimise ventilator settings to fit patient-specific condition (L. Gattinoni et al., 2003; Hubmayr, 2011; Kostic et al., 2011). Ironically, a lifesaving MV treatment has the potential to impose ventilator-induced lung injuries (VILI) which further worsens the lung condition if use incorrectly (L. Gattinoni et al., 2003; Kallet & Branson, 2007).

A classic example of different in MV management is the contradiction in setting high or low ventilation airway pressure or positive end expiratory pressure (PEEP). There is no specific standard protocol to determine optimum PEEP level (Chase et al., 2014; Sundaresan et al., 2011). PEEP is used to recruit collapsed lung units and prevent them from closing during expiration. Collapsed lung units will lower the blood oxygen due to their inability to perform gas exchange process. However, when determining an optimum PEEP level, a delicate trade-off is required between maximising gas exchange (Tusman et al., 1999) while

preventing further damaging the lungs (Ricard et al., 2003; Arthur S Slutsky & Ranieri, 2013) due to the lack of non-invasive medical equipment to diagnose patient's condition. Thus, finding a balance to prevent further injuries while assuring optimal treatment is challenging due to the nature of heterogeneity of patient's condition towards treatment. Hence, MV settings selection are often based on clinical intuition and medical experience (V. J. Major et al., 2018).

1.3.1 Problem Statement

The goal of MV treatment is to improve gas exchange and assist work of breathing. Establishment of optimal interaction between ventilator and the patient is thus essential and necessary for optimal MV performance. This patient-ventilator interaction (PVI) relies on the interplay between patient pathophysiology and the ventilatory support; in other words, PVI describes the ventilator response to the breath delivered by the ventilator, consecutively, the patient responds to the mechanical breath (Epstein & Chatburn, 2011; Gilstrap & MacIntyre, 2013; Sassoon & Foster, 2001). The quality of patient ventilation interaction implies how well a patient is interacting with the ventilator. A better quality means patient is synchronised with the ventilator, whereas a low quality means patient is experiencing AB and there is a significant mismatch between patient demand and ventilator supply (Moorhead et al., 2013). It is evident that atrocious PVI is associated with dire outcomes; therefore, achieving and upholding quality of PVI is utmost important. However, optimum PVI is difficult to achieve due to the complex interaction of patient pathophysiology and mechanical breath delivery. Imperfect PVI occurs because the interplay of both the patient and ventilator is confounded by multifaceted factors related to patient or ventilator (Karen G Mellott et al., 2009). A patient-related factor often associates with highly dynamic behavioural, mechanical, chemical, as well as disease-states. While a ventilator-related factor often involves suboptimal MV settings delivered by clinicians or due to presence of noise in the ventilator circuit.

Nonetheless, these factors induce poor quality of PVI which then predispose to asynchronous breathing (AB) occurrence. AB occurrence is often unpredictable and spontaneous; moreover, lack of automatic real-time AB detection technique is still limited at clinical bedside (Dres et al., 2016; Georgopoulos et al., 2006). Henceforth, visual inspection which is the standard approach to detect AB events is arduous and inapplicable in a time critical clinical environment (Colombo et al., 2011; I. I. Ramirez et al., 2017). As a result, the frequency and impact of AB towards patient's condition are frequently undervalued; thus, effort to eradicate AB to improve PVI becomes a challenging task. While asynchrony index (AI) is the standard approach to quantify the quality of PVI during treatment. Although the correlation between AI with patient's outcomes have been investigated extensively, the actual causative mechanism of AB predisposing to adverse patient's condition still remain unknown and sometimes contradictive. For example Blanch et al. (Blanch et al., 2015) and Martos-Benítez et al. (Martos-Benítez et al., 2020) found that severe AB occurrence (AI>10%) is associated with high mortality and lower PaO2/FiO2 ratio, but a study conducted by Rolland et al. (Rolland-Debord et al., 2017) found that severe AB occurrence is not associated with adverse outcome. Such contradiction may imply that AI computation alone may fail to reflect the actual impact of AB. Hence, it is speculated that, the magnitude of patient-effort induced in AB might play a deterministic role in affecting patient's outcome. Hence, there is a need to develop methods to detect and quantify AB to reflect patient's PVI quality during treatment.

In short, mechanical ventilation treatment is necessary to reinstate oxygenation as well as to provide supportive work of breathing to respiratory failure patients. But methods in providing optimal patient-ventilator interaction treatment remain obscure due to lack of understanding and real-time detection of the problem. Furthermore, it is evident that, poor PVI can cause irreversible dire impact towards patient's condition if leave untreated (Epstein

& Chatburn, 2011; E. Kondili et al., 2003). Hence, there is a need for a real-time non-invasive method to provide insights to patient's PVI condition that could potentially improve MV treatment.

1.4 Research Objective and Scope

The aim of this research is to investigate how the degree of 'mismatching' due to suboptimal PVI (Christer A. Sinderby & Beck, 2012) impacts patient's impacts patient's mechanical ventilation treatment outcome via developing machine learning models to monitor and quantify asynchronous breathing. Machine learning approach potentially provides a unique solution to automatically quantify patient-specific PVI condition.

The scope of this project comprises of 4 major components as noted below:

1. Conduct clinical observational trial and setting up data acquisition system for mechanical ventilation patient data collection in intensive care unit (ICU).

2. Development of an automated detection of patient asynchronous breathing method via machine learning convolutional neural network (CNN).

3. Development of machine learning models, a) generative adversarial network (GAN) and b) convolutional autoencoder (CAE) for quantifying mechanical ventilation asynchrony magnitude.

4. Investigation of the patient-ventilator interaction via quantification of asynchrony index and asynchrony magnitude assessment on actual clinical data.

The first work focuses on setting-up clinical trial to collect data for model development and validation. The second work revolves around developing a model to detect AB in real time. The third work focuses on the development of a model to quantify severity of AB. Lastly, the final research scope involves investigation of the effect of AB on patient's outcome by using the developed methods and a new developed metric to assess PVI.

1.5 Summary of thesis layout

The thesis is outlined as follows:

- **Chapter 2** provides literature review on causative factors and asynchronous breathing trait identification via waveform analysis. It also covers overview of the related work in patient-ventilator interaction monitoring and assessment. The related work is separated into two sections: asynchronous breathing monitoring and magnitude of asynchronous breathing quantification.

- **Chapter 3** describes the details of clinical trial setup that provides clinical data for model development. This chapter also provides an overview of data acquisition system used to collect continuous mechanical ventilation airway waveform data during the clinical trial.

- **Chapter 4** presents the development of a machine learning model for asynchronous breathing detection. The model is tested with different data composites and data quantities. The performance of each model in differentiating asynchronous breathing and non-asynchronous breathing is presented.

- **Chapter 5** introduces two different machine learning models developed to quantify the magnitude of AB via reconstructing airway pressure. A comparison study between the developed machine learning models with an existing method is performed.

- **Chapter 6** presents the investigation on how quality of patient-ventilator interaction impacts the patient's outcome using the developed models in Chapters 4 and 5. A ventilation index metric for assessing the quality of patient-ventilator interaction is introduced in this chapter. This metric and asynchronous index are studied in our patient cohort.

- **Chapters 7** and **8** concludes the finding of the research and describes the future work of improving models to better assess patient-ventilator interaction during mechanical ventilation treatment.

1.6 Contributions

This thesis investigates the application of machine learning models to monitor and assess the quality of patient-ventilator interaction during mechanical ventilation treatment. The machine learning models can be trained and applied at patient bedside to monitor patient-ventilator interaction in real-time. These models potentially provide unique information to clinicians to understand and assess patient's condition in real time. It is essential that the developed models are robust to noises and erratic shapes. The major contributions of this thesis are:

- A data-driven model to distinguish between normal or asynchronous breathing in real time.
- A data-driven model capable of quantifying the magnitude of asynchronous breathing.

- A novel metric to assess patient-ventilator interaction.
- A correlation study on the quality of patient-ventilator interaction affecting patient's outcome.

Chapter 2
Background

There are two main unresolved conundrums in asynchronous breathing: the frequency of asynchronous breathing occurrence and the magnitude of asynchronous breathing. This chapter provides an overview of related work in the effort of detecting and quantifying the magnitude of asynchronous breathing. This chapter begins by describing the causal mechanisms and patterns of each type of asynchronous breathing. In section 2.2, we summarize the existing methods at bedside to detect asynchronous breathing and measure the magnitude of asynchronous breathing. This is followed, in Section 2.3, by an overview of existing and proposed machine learning approaches to identify or measure the magnitude of asynchronous breathing.

2.1 Asynchronous breathing

Patient-ventilator asynchrony or asynchronous breathing (AB) is a phenomenon when the timing is not matching with the timing of ventilator cycle (Bulleri et al., 2018; Karen G Mellott et al., 2009; K. G. Mellott et al., 2014; Gastón Murias et al., 2016; Nilsestuen & Hargett, 2005; Sassoon & Foster, 2001; Souza Leite et al., 2020). However, given the complexity and heterogeneity of patient's response to treatment, AB incidence is common during treatment and it can occur spontaneously in any ventilation modes. Studies have shown that approximately 25% of patients may experience AB during mechanical ventilation (MV) treatment (Blanch et al., 2015; Robinson et al., 2013; A. W. Thille et al., 2006). It is speculated that, the occurrence of AB is mainly due to poor patient ventilator interaction (PVI) because of excessive sedation (B. M. De Wit et al., 2009; Holanda et al., 2018; Vaschetto et al., 2014) or suboptimal mechanical ventilation (MV) settings selection (de Wit, 2011;

Epstein, 2011; K. G. Mellott et al., 2014; A. W. Thille et al., 2006). Frequent AB occurrence may cause dire impacts such as prolonging patient's dependency on MV, increasing sedative drugs usage, and worsening mortality rate (Blanch et al., 2015; Branson et al., 2013; M. de Wit et al., 2009).

There are seven known types of asynchronies namely: ineffective triggering, auto-triggering, double triggering, reverse triggering, flow asynchrony, premature cycling and delayed triggering. These asynchronies can be identified via airway pressure-time and flow-time waveform interpretation (Nilsestuen & Hargett, 2005; I. Ramirez & Arellano, 2018). In the following section, we provide a brief description of each type of AB.

2.1.1 Ineffective triggering

Ineffective triggering or miss-trigger occurs when ventilator fails to detect the presence of patient's breathing effort and often occurred in partial support ventilation mode (Blanch et al., 2015; K. G. Mellott et al., 2014). A study conducted by Thille et al. (A. W. Thille et al., 2006) stated that ineffective triggering is accounted for more than 98% of all asynchronies. Consequently, frequent ineffective triggering occurrence may predispose to patient's ventilatory demand unsatisfactory (B. M. De Wit et al., 2009); thus resulting in respiratory fatigue or injury due to work of breathing increment. Besides, ineffective triggering may build up intrinsic positive end-expiratory pressure (PEEPi) in the lung, causing the patient's effort insufficient to overcome the trigger threshold (Branson et al., 2013; G. Murias et al., 2016; A. W. Thille et al., 2006). PEEPi occurs when the patient's inspiratory time is shorter than ventilator inspiratory time, the ventilator continues to provide flow to patient leaving exhalation time to be reduced (Sassoon & Foster, 2001).

It is believed that ineffective triggering occurs because of diminished respiratory drive due to long term MV treatment or disease condition which weakens patient's effort to trigger ventilator sensor (Branson et al., 2013; Bulleri et al., 2018). Furthermore, deep sedation or unsuitable threshold sensitivity ventilator settings by clinicians may also provoke ineffective triggering event (Branson et al., 2013; Gurevitch & Gelmont, 1989; Sassoon & Foster, 2001).

Ineffective triggering (A. W. Thille et al., 2006) can be described as a simultaneous decrease in airway pressure and an increase in air flow without assisted cycle. Figure 2.1 depicts the occurrence of ineffective triggering (red arrows), where a small positive inflection will appear at the expiration of the flow waveform; while, a small negative inflection will appear at the expiration of the pressure waveform. The subsequent breathing cycle fails to be initiated as the deflection does not surpass the threshold. Oesophageal monitoring or measuring diaphragm electrical activity is also one of the ways to detect ineffective triggering during MV besides visual inspection on pressure or flow waveform (Ashutosh et al., 1975; Branson et al., 2013; Georgopoulos et al., 2006; I. Ramirez & Arellano, 2018); however, oesophageal monitoring requires invasive placement of oesophageal catheter which is not a standard routine care in ICU.

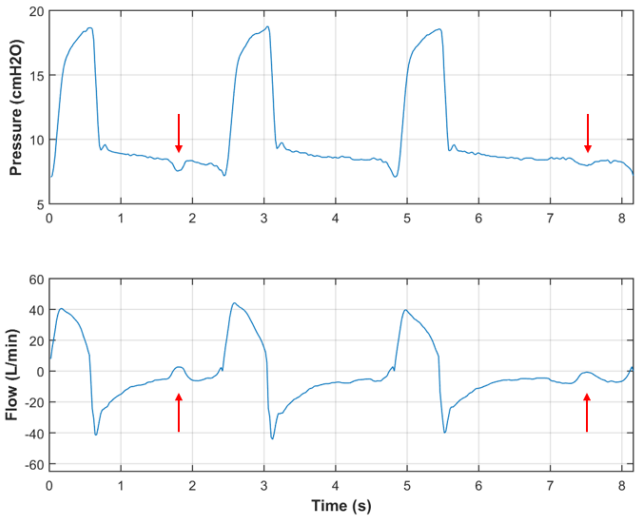

Figure 2.1: Red down arrows point towards the failed triggered effort in pressure/time waveform; whereas, red up arrows show ineffective efforts in flow/time waveform.

2.1.2 Auto-triggering

Auto-triggering occurs when the ventilator delivers airflow in the absence of patient's initiation (A. W. Thille et al., 2006) or cardiogenic oscillations (Imanaka et al., 2000). Auto-triggering event can be described as the absence of negative airway pressure prior to machine-delivered drop; thus, delivering unscheduled breathing cycles. Figure 2.2 depicts the scenario where breathing cycles were initiated automatically due to the absence of patient induced pressure drops in the beginning of inhalation. Similar to detecting ineffective triggering, oesophageal monitoring is the gold standard of detecting the presence of auto-triggering (Branson et al., 2013).

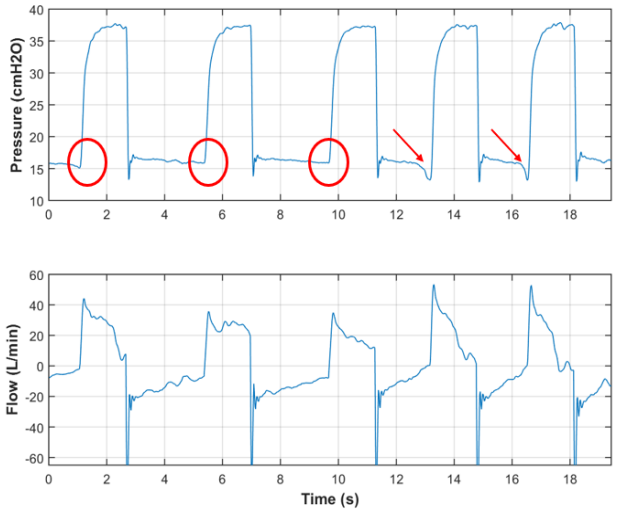

Figure 2.2: As the breathing cycles were triggered due leakage; thus, resulting in absence of airway pressure drop in pressure/time waveform (red circle) as oppose to pressure drop due to patient effort (red arrow)

Auto-triggering event arises mainly due to introduction of noise into the circuit because of leakage, presence of liquid or even hiccups which affects the sensitivity of the sensor; thus causing flow or pressure profile distortion (Georgopoulos et al., 2006; Gilstrap & MacIntyre, 2013; Eumorfia Kondili et al., 2007). The occurrence of auto-triggering can be associated with several factors such as inappropriate ventilator trigger setting (Henry et al., 2016) or excessively low trigger sensitivity (too sensitive) causing difficulty in precisely detecting

patient's effort. Auto-triggering often leads to severe consequences such as PEEPi increment, hyperventilation and promote ventilator-induced-lung-injury (VILI) (Henry et al., 2016).

2.1.3 Double triggering

Double triggering refers to the development of 2 consecutives breath cycles separated by abbreviated expiratory time, half or lesser of the inspiratory time (A. W. Thille et al., 2006). Double triggering is common during volume control continuous mandatory ventilation and pressure support ventilation (Blanch et al., 2015); however, double triggering may also occur during pressure support ventilation with high flow termination criterion (Sassoon & Foster, 2001). Inappropriate ventilator settings may cause excessive demand for flow or volume due to short ventilator inspiratory time; resulting in mandatory breath delivery followed by a second breath as the trigger threshold is surpassed. Consequently, pulmonary barotrauma or volutrauma will occur (Karen G Mellott et al., 2009; A. W. Thille et al., 2006). Furthermore, sighs, coughs while breathing, or inappropriate settings may also trigger such phenomenon.

Double triggering can be spotted at the bedside via flow and pressure waveforms inspection (Liao et al., 2011). Figure 2.3 shows that the expiratory time of the second breathing cycle (black arrow) is shorter than the others (red arrow). Double triggering occurs when neural inspiratory time is longer than the ventilatory inspiratory time.

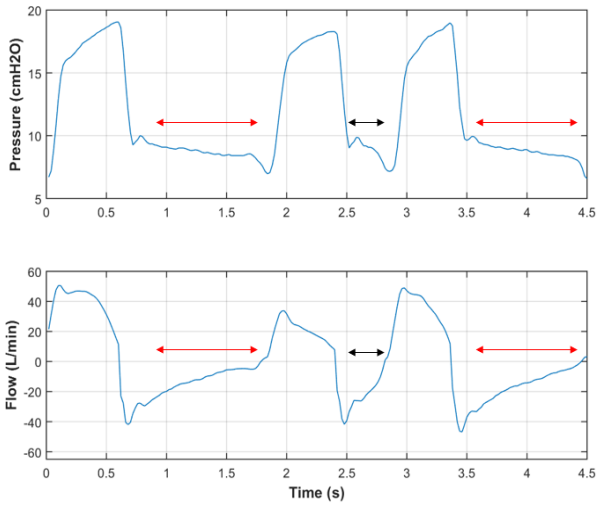

Figure 2.3. The expiration time is shorter during double triggering (black arrow) than expiration of normal breathing cycles (red arrow).

2.1.4 Reverse triggering

Reverse triggering was originally reported in 1980s, yet, it is still a frequently under-recognized form and unique type of neuromechanically asynchrony. (Akoumianaki et al., 2013; Branson et al., 2013). It was described as a form of diaphragmatic muscle contractions triggered by ventilator insufflations and referred such phenomenon as "entrainment", in which patient's respiratory center is activated in response to a passive insufflation of the lungs. There are several hypotheses explaining the occurrence of reverse triggering. Murias et al. (Gastón Murias et al., 2016) stated the occurrence of reverse triggering maybe due to the increase of compliance as previously closed alveoli and airways have been successfully recruited during MV; as for spontaneous breathing patients, the increase of compliance maybe due to activation of patient's inspiratory muscle. The compliance increment causes a drop in pressure during inspiratory plateau. However, the exact causal mechanism is unknown as the occurrence of reverse triggering was reported in brain-death patients (Delisle et al., 2016) and heavily sedated ARDS patients (Akoumianaki et al., 2013). Despite the unknown cause, reverse triggering might result in inducing double-triggering, increasing oxygen consumption and plateau pressure instability (Branson et al., 2013).

Similarly, reverse triggering can be observed through waveforms analysis, where, reverse triggering occurs when a drop in pressure is apparent in breathing cycle (I. Ramirez & Arellano, 2018). Figure 2.4 depicts two 'reversely triggered' breathing cycles in pressure/time waveform. The pressure waveforms exhibit apparent drop in pressure followed by increment of pressure during reverse triggering.

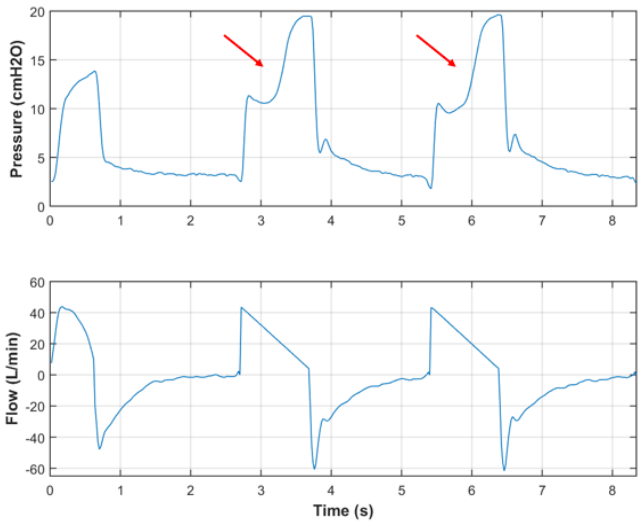

Figure 2.4: The red arrows show negative deflections deforming an airway pressure profile during reverse triggering.

2.1.5 Flow Asynchrony

Flow asynchrony occurs when the patient demand is not matching with the ventilator flow (Blokpoel et al., 2016) causing patient's breathing workload to increase. Marini and colleagues (Branson et al., 2013) found out that when flow delivery is reduced, the patient's work of breathing increases by more than 50% in order to satisfy their ventilatory demand. Flow asynchrony can lead to the creation to PEEPi which is the key factor contributing to barotrauma or volutrauma. Clinicians are able to detect flow asynchrony via ventilator waveform analysis; where a concave shape negative dip with a concomitant constant flow pattern occurs on pressure waveform (I. Ramirez & Arellano, 2018). Figure 2.5 shows the pressure waveform has been "scoped out" during flow asynchrony. As the peak flow set on

the ventilator no longer cater the patient's flow demand due to increment of patient effort, the airway pressure waveform will be gradually dished out (Nilsestuen & Hargett, 2005).

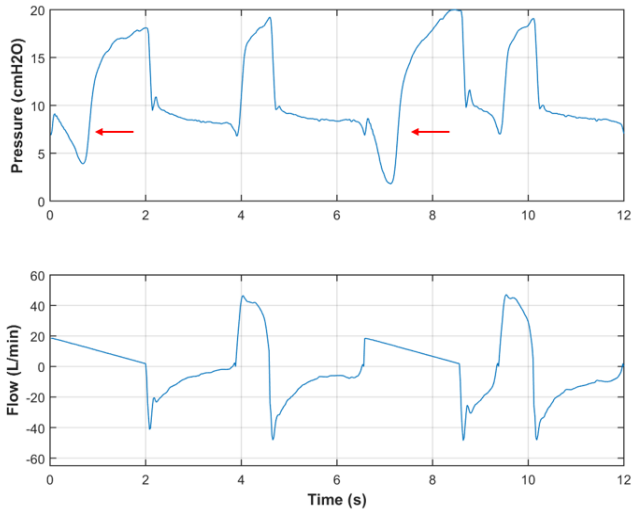

Figure 2.5: Two examples of breathing cycles experience concave shape negative drop on pressure waveform when flow asynchrony occurs.

2.1.6 Premature cycling

Premature cycling or early termination occurs when patient's inspiratory time is lesser than the ventilator inspiratory time (K. G. Mellott et al., 2014). Such event often occurs due to poor satisfying of patient's ventilatory demand and potentially associates with double triggering asynchrony when patient continues to contract inspiratory muscle; thus, resulting in triggering threshold and initiating new breath (Branson et al., 2013; Karen G Mellott et al., 2009). There are 2 known mechanism which could possibly trigger premature termination: early in flow delivery or late in flow delivery (Nilsestuen & Hargett, 2005). Early in flow delivery will initiate an overshoot of pressure threshold (breath-termination criteria) at the initial inspiration due to rapid valve-opening causing the ventilator breath to be terminated prematurely. On the other hand, when the flow delivery to the patient is delayed, flow-termination point may occur earlier due to higher peak flow; thus, resulting in flow-termination criterion being met sooner during inspiration. Premature cycling will cause excessive inspiratory work into and during expiratory phase as well as an overestimation of

respiratory rate (Karen G Mellott et al., 2009). Premature cycling will induce a positive deflection at the end of expiration on flow waveform; whereas a negative deflection will appear at the end of expiration on pressure waveform (Nilsestuen & Hargett, 2005). Figure 2.6 shows the first breath is normal whereas the second and third breath encounters premature cycling. It can be observed that in flow/time waveform, the flow spikes while the airway pressure drops immediately after the end of inspiratory cycle, indicating the patient's effort endures.

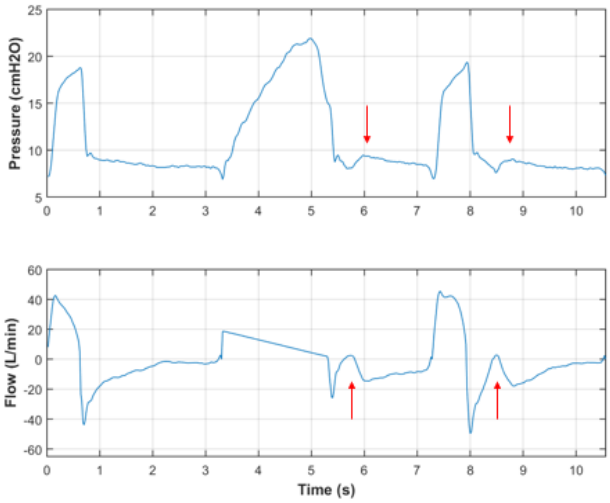

Figure 2.6: The red down arrows show the continuation of inspiratory effort after inspiratory cycle is ended in pressure/time waveform. The red up arrows show a sudden spike in the expiratory flow due to patient's inspiratory effort.

2.1.7 Delayed triggering

Delayed triggering or delayed termination occurs when ventilator continues to provide mechanical breath even when patient's muscular inspiration is complete (Holanda et al., 2018). Delayed triggering might due to presence of air leakage in the system (Bulleri et al., 2018; Calderini et al., 1999) or inappropriate timing in cycling setting (Subira et al., 2018). In addition, patients with chronic obstructive pulmonary disease (COPD) and asthma are speculated at risk for delayed triggering occurrence due to presence of airway resistance (Chiumello et al., 2007). Delayed triggering will cause severe consequences such as

promoting PEEPi or potentially causes ineffective triggering on the next breath due to limited exhalation time. The accumulation of pressure in the lung will result in increased of expiratory load and potentially damage the lung due to barotrauma (Karen G Mellott et al., 2009). Pressure spike can be observed on pressure waveform prior the beginning of exhalation when delayed triggering occurs. On the other hand, a rapid decline in respiratory flow can be observed when delayed triggering occurs (I. Ramirez & Arellano, 2018). Figure 2.7 shows the first and second breaths are delayed triggering, while the 3rd breath is a normal breath. It can be observed from pressure/time waveform that, the breathing cycles affected by delayed triggering experience spikes before the start of exhalation phrase; similarly, a sudden decline can be detected at the end of inspiration cycle in flow waveform.

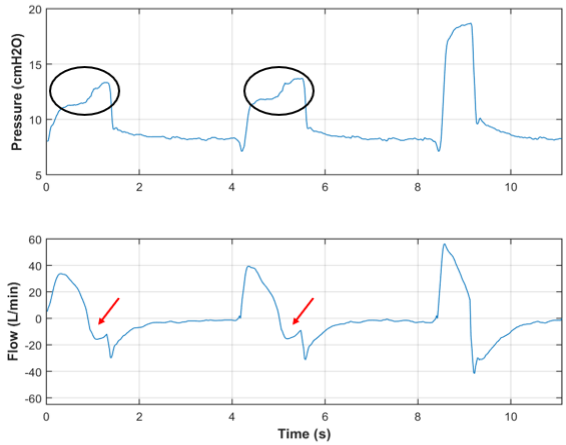

Figure 2.7: The black circle shows spikes occur near the end of inspiration; the red arrows show a rapid decline in flow due to limited exhalation time.

2.2 Existing Methods to AB Detection and Quantification

Despite the prevalence of AB during MV, the frequency of AB occurrence or impact of AB events towards treatment's outcome is still unknown (Blanch et al., 2015; K. G. Mellott et al., 2014; W. A. Thille & Brochard, 2007). Unexplained causal factors to AB events and low recognizing ability of the professional often confound clinicians to correct AB (I. I. Ramirez

et al., 2017). Thus, delivering patient-tailored MV treatment to improve PVI or correct AB is often a daunting task.

However, due to the lack of autonomous AB detection method on the bedside, investigations on the exact causal mechanism and severity of AB on patient's outcome are impeded (Dres et al., 2016). Moreover, the conventional method of quantifying AB involves trained personnel to visually inspect airway pressure and flow waveforms to compute asynchronous index (AI) to describe the quality of PVI (Chao et al., 1997). This process is arduous and time consuming. Henceforth, there remains a huge research interest in developing an autonomous methods or models to detect AB and a metric system to measure the degree of affected patient-specific AB effort. These approaches will potentially aid clinicians to determine the causal mechanism to eradicate AB occurrence (de Haro et al., 2019; Subira et al., 2018). The following outlines the existing methods or models to detect the presence of AB or to quantify the magnitude of AB.

2.2.1 Measurement of Oesophageal Pressure

Akoumianaki et al. (Akoumianaki et al., 2014) proposed the use of oesophageal pressure to detect AB. This method measures the difference between airway pressure and oesophageal pressure to estimate the trans-pulmonary pressure. Oesophageal pressure is useful to determine the amount of airway pressure from the ventilator needed to overcome lung and chest wall elastance. The oesophageal pressure can be used to detect the presence of AB by knowing the level of muscle effort. However, this method is not common in ICU as it requires invasive probes which imposes additional cost and might induce further injuries.

2.2.2 NAVA

Electrical activity of the diaphragm can be measured by applying electrodes incorporated into a nasogastric tube (C. A. Sinderby et al., 1997). The magnitude of the electrical activity of the diaphragm (EA_{di}) signal correlates well with respiratory drive of the patient. Neurally assisted ventilatory assist (NAVA) (C. Sinderby et al., 1999) is a mode of assisted mechanical ventilation which measures the EA_{di} to satisfy patient's ventilatory demand by matching spontaneous breathing demand from patient. Besides that, the timing and intensity of the EA_{di} signal from the patient is also useful to assess the interaction between patient and the ventilator which may able to provide additional information and assessment for the clinicians (C. Sinderby et al., 2013). However, NAVA is uncommon in ICU due to costly equipment such as NAVA catheter (Hjelmgren et al., 2016).

2.2.3 Spectral Analysis of Airway Flow

Gutierrez et al. (Gutierrez et al., 2011) proposed a method of assessing AI by using spectral analysis of airway flow waveform. This method implements Fourier transformation to discrete time dependant signal into infinite number of sine and cosine waves. The frequencies and amplitudes of the waves are then displayed as a frequency spectrum with the amplitude ratio of the first harmonic peak (H_1) to that zero frequency is taken as a measure of spectral organization. The results show that the frequency of an asynchronous patient has irregular spectrum with lower and wider H_1 peaks and disappearance of higher frequency harmonics as compared with spectrum of a synchronized patient. Although this method is less invasive and achieved 83% of sensitivity and specificity, but, only taking airway flow changes into consideration, neglects the mismatch of both flow and pressure that defines an AB, which could lead to lower AB detection accuracy. Moreover, this method is limited to retrospective analysis; which is not feasible to deploy in real time critical clinical decision.

2.2.4 ALIEN

Automated Logging of Inspiratory and Expiratory Non-synchronized Breathing (ALIEN) is designed to detect AB automatically (Chiew et al., 2015). This method uses a combination of different approaches to detect AB by detecting irregular or anomalous deflections in airway pressure and flow compared to typical waveforms. ALIEN algorithm separates the waveforms into segments and checks the gradient sign. The segmented signals will be discarded if the net change in flow or pressure is lesser than threshold. Under pressure and flow control, ALIEN algorithm can achieve high specificity and sensitivity; however, during volume-controlled mode, flow-based inspiratory AB detection becomes less reliable, thus forcing ALIEN to rely only on the pressure waveform to detect AB. Besides, the lack of robustness to anomalies present in breathing cycles might deteriorate the performance to detect AB.

2.2.5 Iterative Pressure Reconstruction

Iterative pressure reconstruction (IPR) is a mathematical model developed by Chiew et al. (Chiew et al., 2018) to estimate respiratory mechanics of AB. The model incorporates a single compartment model to reconstruct to an asynchrony free airway pressure waveform. The difference in area under the graph between the reconstructed breathing cycle and original AB is defined as the magnitude of AB. Despite the model proves its effectiveness in restoring AB to asynchrony free airway pressure profile; the lack of robustness to eliminate patient effort occurring in different location impinges on the performance of the model.

2.2.6 Iterative Flow Reconstruction

Iterative interpolative flow reconstruction method was proposed by (Kannangara et al., 2016) with the purpose to better estimate respiratory mechanics via reconstructing asynchrony

affected airway profile. In addition to respiratory mechanics estimation, the method can quantify the magnitude of AB as indicator of PVI. However, the author stated that the efficiency of the method maybe affected when reconstructing aberrant breathing cycles.

2.2.7 PREDATOR

Pressure Reconstruction by Eliminating the Demand Effect of Spontaneous Respiration (PREDATOR) (Redmond et al., 2014) is an approach designed to reconstruct pressure profiles to assess underlying respiratory mechanics. While the timing of diaphragmatic contraction only affects the position of asynchrony occurrence in a breathing cycle; this leads the rest of the breathing cycle are the 'correct' datapoints. PREDATOR reconstructs breathing cycles affected by reverse-triggering via shadowing the 'correct' data from previous breathing cycles. The reconstructed breathing cycles allow better estimation of elastance and resistance values. Similarly, the difference in area under the curve between reconstructed airway profile waveform using PREDATOR and original AB can be used to describe the magnitude of AB. The performance of PREDATOR can be affected if the reverse-triggering exhibits early asynchrony, as less 'correct' data can be adopted to fit the model.

2.3 Machine Learning Approach

The developed methods in literature review is uncommon at clinical bedside as some can be invasive (Akoumianaki et al., 2014), costly (C. Sinderby et al., 1999) or lack of robustness (Chiew et al., 2018; Kannangara et al., 2016) when dealing with aberrant breathing cycles. Therefore, the model developed should avoid introducing additional costs or imposing further injuries to patients; while, ensuring the developed model is robust as well as computationally

simple. For these reasons, we believe that machine learning can potentially transcend the existing models in detecting and quantifying the magnitude AB due to its reliability, robustness and the ability to learn from data without explicitly programming.

Machine learning is a subfield of artificial intelligence and is closely related with statistics (João & André, 2012). The main objective is to optimally realize embeddings via translating high-dimensional vectors into a relatively low-dimensional data, as well as to cluster and classify different categories. The ability to extract meaningful features and group them by learning from past-experienced data makes machine learning adept in building computational model. For such, machine learning techniques have been extensively deployed to solve problems in various fields. For business applications, machine learning is often deployed to perform predictions (Wei-Yang et al., 2012). In autonomous robotics, machine learning is commonly developed for visual processing (Giusti et al., 2016) and tactile sensing (Bandyopadhyaya et al., 2014).

There are two main types of machine learning: supervised and unsupervised learning. In the former, supervised learning involves a model to learn to map the input examples with the output labelled datasets. The performance of a trained model is usually validated with withheld target variables. Once ready, the model will be used for categorization and classification with new data. On the other hand, in unsupervised learning, a model learns to draw inferences from datasets consisting of input examples without corresponding output labelled data. The purpose of unsupervised learning is to discover intrinsic properties or patterns in the data embedded in the low-rank subspaces. This allows hidden or undetected inherent information to be exploited for data analysis. Therefore, the objective or training approach of machine learning may vary depending on the application.

Nonetheless, although studies on the feasibility and the implementation of machine learning in respiratory related field such as respiratory mechanics estimation (Perchiazzi et al., 2002; Perchiazzi et al., 2017), lung injuries detection (Räsänen & León, 1998) or MV modes identification (Leon & Lorini, 1997) have already been investigated. However, the deployment of machine learning approach in AB detection and quantification is still limited. The following outlines several existing or proposed machine learning techniques to detect and quantify AB.

2.3.1 LSTM

Long short-term memory (LSTM) (Hochreiter & Schmidhuber, 1997) is another type of recurrent neural network, in which, the architecture is distinctive from the conventional neural networks. In LSTM, the output of the current input is influenced by the previous computation steps. Henceforth, LSTM is widely deployed to solve temporal sequence problems such as speech recognition in natural language processing (Kaisheng et al., 2014) due to its ability to memorize data to process sequences of inputs.

Zhang et al (L. Zhang et al., 2020), proposed a two-layer LSTM machine learning technique to detect double triggering and ineffective triggering. The model extracts the temporal features from flow and pressure ventilator waveform simultaneous and determine the occurrence of asynchronies. Despite the method achieved remarkable performance; the method is limited to two types of asynchronies detection. Besides, the feasibility of deploying different machine learning algorithms in other types of AB has not been examined.

2.3.2 CNN

Convolutional Neural Network (CNN) is a biologically-inspired subclass of supervised machine learning models and it was first introduced by LeCun (Lecun et al., 1998) in 1994 to perform digits recognition from handwritten images. CNN was designed to extract useful spatial features while reducing computational resources dependency in which powerful central processing unit is absence in the 90s. The sharing of CNN's trainable weights in the kernels in convolutional layer momentously reduce the total number of parameters in the model as the "information" or weights are distributed in the model. This is in contrasts with neurons in multilayer perceptron (MLP) where each neuron has individual parameters vector. Hence, CNN greatly decreases the parameters to train; thus, saving resources and time needed for convergence to minima while ensuring spatially correlated inputs are not affected.

In recent years, CNNs have achieved impressive and astonishing outcomes in critical tasks such as object classification (Redmon et al., 2016; Szegedy et al., 2015) and audio recognition (Abdel-Hamid et al., 2013). CNN is able to achieve remarkable accomplishments due to its ability to capture and learn complex features or patterns automatically by adopting its convolutional layers. Hence, we hypothesise that CNN can detect AB with high accuracy due to its robustness to anomalies in breathing cycles and ability to extract essential and intrinsic AB features. In Chapter 4, we present the implementation of CNN to identify AB.

2.3.3 CAE

Autoencoder is a type of unsupervised machine learning algorithm that learns to imitate its inputs to its outputs (Baldi, 2012; Y. Zhang, 2018). An autoencoder model constitutes two sections namely encoder and decoder. Interestingly, the operation of these sections serves the opposite functionalities. The encoder extracts and compresses the critical information in the

inputs to reduce the dimensionality while preserving the essential details. On the other hand, the decoder learns to map and reconstruct the output to match the input with least distortion. This process involves the optimization of parameters in the convolutional layers to filter unwanted features as well as 'remember' the intrinsic properties in the inputs.

While convolutional autoencoder (CAE) is a traditional autoencoder integrated with convolutional layers (Y. Zhang, 2018). The supremacy of convolutional layers in anomalies removal and features extraction makes CAE adept in capturing essential intrinsic patterns. CAE are commonly applied to filter noise (Ali et al., 2018) or to restore image (Mao et al., 2016). In Chapter 5, we develop a CAE model to reconstruct AB to asynchrony free breathing cycle. This allows us to quantify AB by measuring the difference of area under the curve between AB with reconstructed asynchrony free breathing cycle.

2.3.4 GAN

Generative adversarial network or GAN is a class of unsupervised machine learning technique (Goodfellow et al., 2014) which involves two neural networks: generator and discriminator competing with each other in a zero-sum game framework. The generator learns to produce synthetic data which look as close as real images given a random distribution; whereas the discriminator learns to distinguish whether the generated data from generator is close to the real data. The discriminator will penalize the generator by tuning the parameters if the generated data is distinct from actual data. Ultimately, the generator will generate realistic data after trained; while, the discriminator will fail to distinguish between actual data and generated synthetic data.

The ability of GAN to generate new data by learning the distribution in the training datasets automatically makes GAN adept in realistic synthetic photographs creation from text (Bodnar, 2018) and images restoration (Yeh et al., 2017). In Chapter 5, we deploy GAN technique to restore a presuming normal breathing given AB as input. Similarly, the difference between the generated normal breathing and AB can be computed to quantify the magnitude of AB.

2.4 Summary

It is well established that AB occurrence is spontaneous and often concomitant with random magnitude and erratic asynchronies patterns. The standard method to identify AB is through observation of the ventilator waveforms at bedside, but, long-term visual inspection is an onerous and dauting task. Therefore, the prevalence and consequences of AB are often underestimated because the absence or lack of monitoring tools to identify them (Dres et al., 2016). Table 2.1 summarizes the development of models to detect and quantify AB over years. It is clear that, the performance of several mathematical models may be affected due to aberrant breathing cycles or the occurrence of patient effort in different timing in AB; henceforth, a high robustness to noises model is favourable. Moreover, a non-invasive model but computational simple and without the usage of additional probes may benefit the popularity of monitoring tools at bedside.

Last but not least, autonomous dedicated techniques or software to monitor AB is essential as it enables the extension and expansion of research on elucidating the harmful physiological effects of AB (de Haro et al., 2019). In the next chapter, we will be presenting the clinical trial setup in intensive care unit to collect ventilator waveform data for model development and validation.

Table 2.1: Summary of the development of AB detection and quantification

Authors and methods	Approach	Advantages	Disadvantages
AB Detection			
Oesophageal measurement (Akoumianaki et al., 2014)	Sensor Measurement	• High accuracy as it measures muscle activity	• Requires invasive probes
NAVA (C. Sinderby et al., 2013)	Sensor Measurement	• High accuracy as it measures muscle activity • Able to quantify the patient ventilator-interaction	• Costly as it requires additional instruments
Spectral Analysis (Gutierrez et al., 2011)	Fast Fourier Transformation	• Automatic • High Accuracy • Uses existing flow waveform from ventilator	• Only measures flow waveform • Only able to be conducted retrospectively.
ALIEN (Chiew et al., 2015)	Mathematical Model	• High accuracy • Real time • Uses both pressure and flow waveform	• Lack of robustness
LSTM (L. Zhang et al., 2020)	Machine Learning	• High accuracy • Real time • Uses both pressure and flow waveform	• Only able to detect two types of AB
AB Quantification			
Iterative Pressure Reconstruction (Chiew et al., 2018)	Mathematical Model	• Real time • Allows respiratory mechanics estimation	• Lack of robustness • Time of patient effort occurrence in a breathing cycle may affect performance
Iterative Flow Reconstruction (Kannangara et al., 2016)	Mathematical Model	• Real time • Allows respiratory mechanics estimation	• Lack of robustness to anomalies or noises
PREDATOR (Redmond et al., 2014)	Mathematical Model	• Uses previous information for better AB reconstruction	• Performance may be affected if early asynchrony occurs

Chapter 3
Clinical Trial

This chapter presents the design and setup of a clinical observational trial conducted in this research to collect data for model development. Clinical trial is essential as it documents the data from cohort studies, treatment process and patient's outcomes. The data collected from this trial are imperative for model development, testing, as well as to study the impact of patient-ventilator interaction towards treatment in this research.

3.1 Introduction

In this study, Clinical Application of Respiratory Elastance (CARE), a clinical observational trial is conducted to investigate the quality of AB towards patient's outcome. The trial is conducted in intensive care unit (ICU) in International Islamic University Malaysia (IIUM), Kuantan, Malaysia. The trial collects mechanical ventilator airway pressure and flow waveform data of respiratory failure patients together with their clinical chart. The trial is approved by IIUM Research Ethics Committee (IREC) with trial number IREC-666 and it is registered with the Australia New Zealand Clinical Trial Registry (ANZCTR 12618000468224) (Chiew et al., 2018). In this chapter, the study design, data collection process as well as initial processing of the data are presented.

3.2 Study Design

3.2.1 Patients

The patients in the ICU, diagnosed with acute respiratory failure and ventilated using Puritan Bennett PB980 ventilator (Covidien, Boulder, CO, USA) are screened for inclusion into the

study. Written consent is obtained from patient's family members recruited into the trial. The inclusion and exclusion criteria are as follows:

Inclusion Criteria

1. Patients requiring invasive mechanical ventilation (MV) (Intubation or tracheotomy).

2. Patients with PaO_2/ FiO_2 [oxygen partial pressure to fraction of inspired oxygen] ratio < 300 mmHg).

3. Arterial line in situ.

Exclusion Criteria

1. Patients who are likely to be discontinued from MV within 24 hours.

2. Patients with age < 16.

3. Any medical condition associated with a clinical suspicion of raised intracranial pressure and/or a measured intracranial pressure ≥ 20 cmH_2O.

4. Patients who have a high spinal cord injury with loss of motor function and/ or have significant weakness from any neurological disease.

5. Patients who are moribund and/or not expected to survive for > 72 hours.

6. Lack of clinical equipoise by intensive care unit (ICU) medical staff managing the patient.

3.2.2 Data acquisition system

In this trial, the CURE software (Szlavecz et al., 2014) is implemented to collect data from the ventilator. CURE software was programmed in JAVA, a general-purpose programming language which allows CURE software to run on any platform ranging from mobile devices to personal computers. This study uses the Intel Next Unit of Computing (NUC) and a

GeChic On-Lap 1502i Touch Screen Monitor to house the CURE Software. The data acquisition system specifications are listed in Table 3.1:

Table 3.1: Data acquisition system hardware specification.

Specifications	General Descriptions
Processing Unit	Intel NUC Processor Intel core i5 (5th generation) RAM: 8GB Storage: 256 GB SSD Operating System: Windows 10
Monitor	GeChic On-Lap 1502i Resistive Touch Screen Resolution 1920x1080

The data acquisition system is mounted on the arm holder of Puritan Bennett PB980 ventilator. Figure 3.1 shows the processing unit and touch screen monitor emplaced on CURE mounting device.

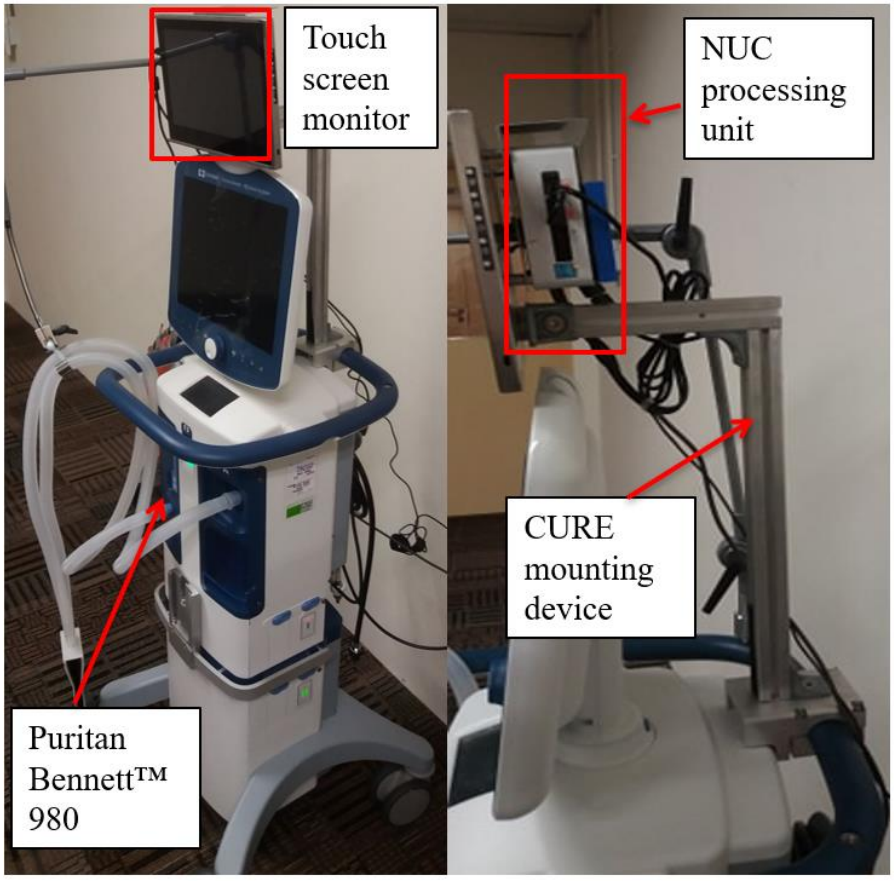

Figure 3.1: The CURE mounting device (red box) is designed to rest on the arm of Puritan Bennet 980, while holding the touch screen and processing unit (black circle) in place.

An automatic backup system operates by periodically duplicating and transferring data to another storage medium such as external drive or cloud storage automatically. The backup system consists of a portable modem to provide wireless internet connection, an external hard disc and Rclone, an open-source software to encrypt and upload patient's data to cloud storage services. The software is programmed to execute upon Windows system bootup to upload data to both cloud and local storage periodically. Besides, the external hard disc acts as additional backup in case of unavailable internet access. Figure 3.2 shows the external hard disc attached to NUC processing unit to backup data automatically.

Figure 3.2: The NUC processing unit saves patient data to external hard disc automatically and periodically to prevent data loss.

3.3 Recruited Patient

A total of 24 patients were recruited into CARE trial within 2 years duration (14 August 2017 – 6 February 2020). The demographics of the recruited patients are shown in Table 3.2 and Table 3.3. The collected data consists of more than 1500 hours or equivalent to 2.1 million breathing cycles ventilated using different MV modes such as synchronized intermittent mandatory ventilation (SIMV), bilevel positive airway pressure (BiPAP), continuous positive airway pressure (CPAP), and spontaneous ventilation (SPONT) modes. In addition to ventilator waveform data collection, patient's bed charts and National Audit on Intensive

Care Unit were collected. These data comprise of valuable patient's condition such as arterial blood gas (ABG), MV length, duration of ICU stay as well as patient's basic information such age and gender. A few samples of mechanical ventilation waveform are shown in Figure 3.3. As the breathing waveform patterns may alter due to the change of MV settings (Hess, 2005); therefore, huge diversities of MV settings may benefit model development and validation.

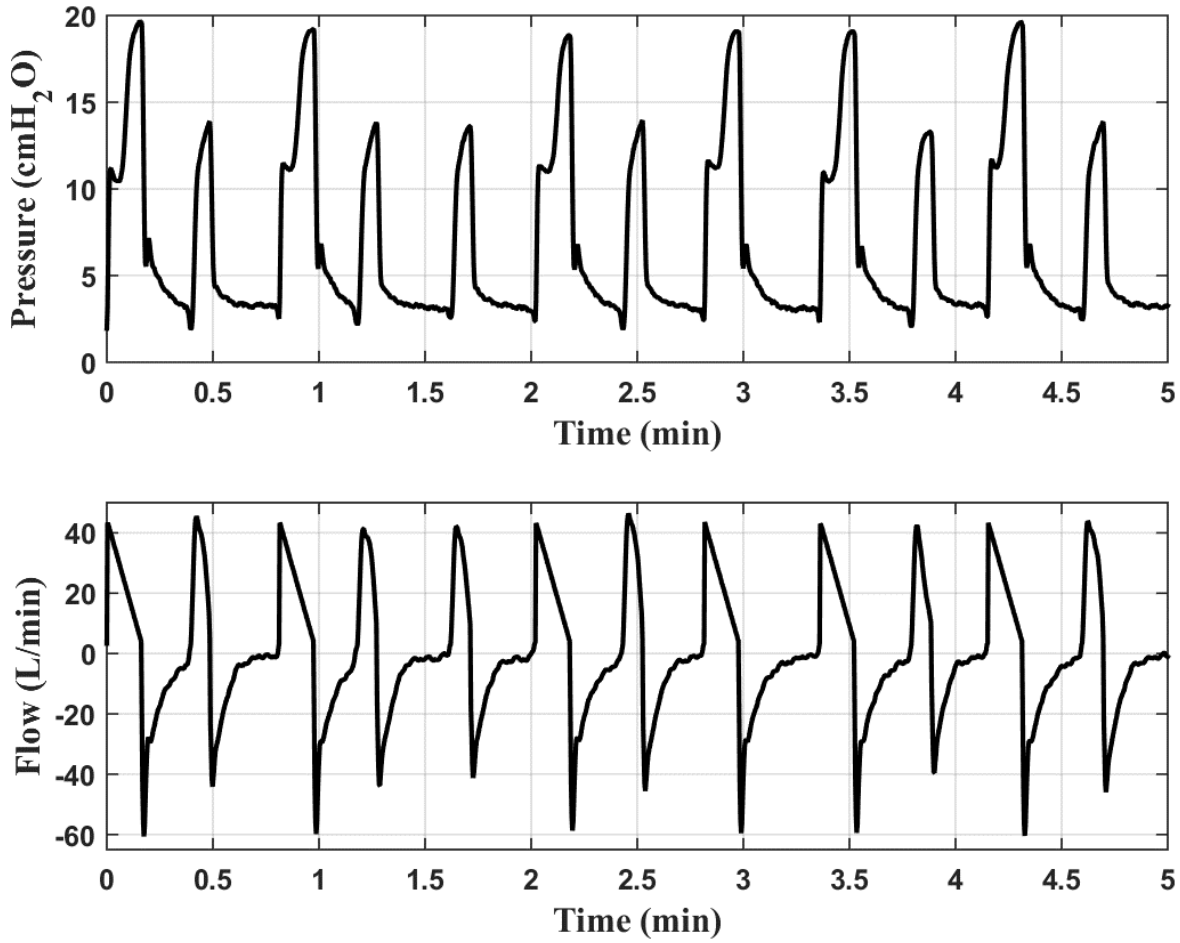

Figure 3.3: The top plot displays the airway pressure waveforms; while the bottom plot shows the airway flow waveforms.

Table 3.2: Patient demographics, admission reasons and ventilation mode.

No.	Gender	Age	Hours of Data Collected	No. of Breath	Clinical Diagnosis	Initial P/F Ratio	Ventilation Modes
1	Female	43	168	270,126	Thyroid Carcinoma	150	SIMV
2	Male	54	59	61,832	Pneumonia	202	SIMV/ SPONT
3	Male	52	44	46,067	SVC Obstruction	95	SIMV
4	Male	64	20	21,286	HAP	238	SIMV
5	Female	63	139	152,229	Klebsiella Sepsis	146	SIMV
6	Female	73	5	5,989	Pneumonia	298	SIMV/ SPONT
7	Female	64	34	38,846	Pneumonia	117	SIMV / BiPAP
8	Male	48	35	36,566	CAP	128	SIMV/ SPONT
9	Male	42	5	6,953	Severe Pneumonia	97	CPAP/ SIMV
10	Female	60	3	2,488	Acute Coronary Syndrome with Cardiac Asthma	80	SIMV/ SPONT
11	Male	64	3	502	HAP	289	SIMV/ BiPAP
12	Male	74	3	3,444	HAP	253	SIMV/ BiPAP
13	Male	63	14	6,075	HAP	104	SIMV/ BiPAP
14	Male	53	49	38,908	HAP	133	SIMV
15	Female	62	55	83,385	HAP	143	SIMV/ SPONT
16	Female	34	112	120,172	Pneumonia	155	SIMV/ SPONT
17	Male	43	42	52,960	Acute Pancreatitis	157	SIMV/ SPONT
18	Female	61	64	62,145	Right Lobar Pneumonia	92	SIMV/ SPONT
19	Male	48	303	368,689	CAP	350	SIMV/ SPONT
20	Female	53	156	216,984	CAP	241	BiPAP /SIMV/ SPONT
21	Female	65	119	169,651	Recurrent Multifocal Infarct with Poor Neurological Recovery	106	SIMV/ BiPAP
22	Male	48	113	128,573	Partially Treated Pneumonia	202	SIMV/ SPONT
23	Male	66	114	149,158	HAP	119	SIMV
24	Male	53	49	56,227	HAP	246	SIMV/ SPONT

HAP – Hospital acquire pneumonia; CAP – Community acquire pneumonia; SVC – Superior vena cava

Table 3.3. Summarized patients' demographic and clinical characteristics. Figures are presented as median [25^{th}-75^{th} percentiles] or percentages.

Patients	n = 24
Age (years)	57 [48-64]
Gender (% men)	58%
Reason for admission n (%)	
Acute respiratory failure	22 (91.7%)
- Sepsis	1 (4.2%)
- Pneumonia	18 (75%)
- Others	3 (12.5%)
Postsurgical	1 (4.2%)
Obstruction	1 (4.2%)
APACHE II	15.5 [12.5–21.5]
SOFA at admission	6.5 [4.0-10.0]
Length of data collection (hours)	49.0 [17.0-113.5]
Number of data (breathing cycles *100,000)	0.55 [0.14-1.39]

APACHE II – Acute physiology and chronic health evaluation II; SOFA – Sequential organ failure assessment.

3.3.1 Manual AB Inspection

Most of the MV modes administrated to the recruited patients during treatment were similar; henceforth, Patients 1 to 7 comprise of more than 600,000 breathing cycles treated with prevalent MV modes (SIMV, BiPAP and SPONT) during treatment were manually categorized into two categories namely: normal and asynchrony. The data from the initial 7 patients comprises of various AB shapes and features and the data will be useful for model development. To ensure impartiality, these data used for model development and validation were cross-checked by three trained researches. During the process of manual AB inspection, multiple computers were deployed to process the recorded ventilator waveform data and saved them individually in Matlab binary file format (.mat) and standard compressed image format (.jpeg) using Matlab 2017b (Natick,MA). The researchers then categorized them through manual inspection on pressure data. The researchers took more than two weeks to complete the analysis. Asynchrony index (AI) (M. de Wit et al., 2009) was calculated using the following equation:

$$AI = \frac{No.\,of\ Asynchronous\ Breaths}{Total\ No.\,of\ Breaths} \times 100\% \tag{3.1}$$

Table 3.4 shows the composition of manually classified breathing cycles in Patients 1 to 7. It shows that every recruited patient experienced different AI with median 39.5% [interquartile range (IQR): 11.5-53.2]. Patient 3 experienced lowest AI at 1.37%; whereas, the highest AI was at 64.8% experienced by Patient 4. This result shows that the AI varies among patients.

Table 3.4: Composition of types of breathing cycle experienced by the first seven patients.

	Number of breaths		Percentage
Patient	**Asynchrony**	**Normal**	**AI (%)**
1	69,620	200,506	25.77
2	33,426	28,406	54.06
3	633	45,434	1.37
4	13,793	7,493	64.80
5	60,066	92,163	39.46
6	403	5,586	6.73
7	19,581	19,265	50.41
Sum	197,522	398,853	-
Median	19,581	28,406	39.46
IQR	[3,923–53,406]	[10,436–80,481]	[11.49–53.15]

In this observational trial, patients were ventilated with different MV settings despite similar clinical diagnosis. This finding shows patients intra- and inter-variability and information variation is useful for developing a robust asynchronous breathing detection model. While the study design is set to collect data continuously for 72 hours, a number of patient data had less than 72 hours. This shorter data period is mainly due to the discontinuation of CURE soft from ventilator as ventilator maintenance is required. In addition, MV patients are often scheduled for other clinical procedure, where the patient is connected to portable MVs. Therefore, some of the data acquisition is not continued after clinical procedure.

In this trial, when determining asynchronous breathing via manual inspection, the lack of taxonomy to consistently identify AB (Chanques et al., 2013; Colombo et al., 2011; Gogineni et al., 2012) may predispose to wrong AB evaluation. It is a challenge where manual inspection of breathing asynchrony is subjective. Thus, the manual classification of asynchronous breathing cycles carried out in this research were compared with two other independent researchers trained to identify AB. The researchers studied the literature on definition of AB and established an agreement among each other on identifying the traits of AB. The ‘ground truth’ is the breathing cycle classified by the researchers as AB or normal

breathing cycle. Specifically, 500 randomly chosen breathing cycles from the collected data to categorize into two categories: normal and asynchrony. The breathing cycles were annotated manually between researchers and their accuracy between them are recorded. Table 3.5 shows that the researchers reached more than 80% accuracy computed using Equation 3.2 when categorizing 500 breathing cycles. The accuracy found in this study was similar to other literature finding (Epstein, 2011; I. I. Ramirez et al., 2017), where they have found variation of accuracy among observers. This result suggested that determining AB through manual inspection may be subjective and there is a need to automate an objective method in identifying AB.

Table 3.5: Manual inspection accuracy attained by three independent researchers.

Researcher	Accuracy
Researcher 1	85.2%
Researcher 2	83.2%
Researcher 3	84.6%
Avg. Accuracy	84.3%

$$Accuracy = \frac{Breathing\ cycle\ denoted\ by\ researchers}{Total\ sample\ breathing\ cycles} \times 100\% \quad (3.2)$$

3.4 Summary

CARE trial was setup to collect data for model development and validation. 24 patients from different demographics were recruited via the trial. We found out manual inspection of asynchronous breathing is subjective and may differ between researchers. In the next chapter, we present the development of machine learning model to detect AB using the manually classified breathing cycles carried out in this study.

Chapter 4
Asynchrony Detection

Abundant ventilator waveform data collected from the clinical trial presented in Chapter 3 allow us to use data driven model approach to classify breathing cycles. The ability of detecting asynchronies plays a vital role in evaluating the quality of ventilator treatment. Henceforth, in this chapter, we present investigation on the application of a machine learning algorithm namely convolutional neural network to detect asynchronous breathing. We examined the optimal number and data composite type for convolutional neural network training to achieve good accuracy using Monte-Carlo simulation. The result on the performance of the trained model on clinical data validation is presented in Section 4.4.

4.1 Introduction

Asynchronous Breathing (AB) is an erratic event that occurs due to poor patient-ventilator interaction (PVI) during mechanical ventilation (MV) treatment (Blokpoel et al., 2016). The impact of AB can be critical to patients as frequent occurrence of AB may impose further injuries to the lungs, lengthen patient's dependency on the MV and worsen mortality rate (Beitler et al., 2016; Branson et al., 2013). Poor PVI occurs due to the mismatching of the timing between patient's natural breathing pattern and the support from the mechanical ventilator. Hence, promoting and maintaining harmonious PVI is the key to reduce AB occurrence rate.

While AB is usually spontaneous and manifests erratic and abnormal ventilator waveforms; but, the standard approach of detecting AB is via long-term visual inspection of the ventilator

waveforms at bedsides by the clinicians (Colombo et al., 2011; I. I. Ramirez et al., 2017; A. W. Thille et al., 2006). As a result, identifying AB becomes arduous and challenging as it requires continuous and consistent ventilator waveform respiratory signals interpretation. Hence, an automated computerized and robust algorithm is necessary for AB detection to better MV management.

In this chapter, we deploy Convolutional Neural Network (CNN) machine learning algorithm to detect the presence of AB automatically. In order to understand the feasibility of and performance of CNN, we trained the CNN model with different data structure and data quantity. We hypothesized that CNN is robust to anomalies and able to detect AB with high accuracy.

4.2 Convolutional Neural Network

CNN was introduced to recognize digits from handwritten images (Lecun et al., 1998) in the 90s. The adaptable kernels in CNN accelerate the learning process to capture complex essential features (Krizhevsky et al., 2017) and decrease the dependency on computational resources. Henceforth, CNN technique is applied in various fields and has shown positive outcomes specifically in object and audio recognition (Krizhevsky et al., 2017; Lee et al., 2009).

Typically, a CNN model comprises of an input and output layer, convolutional layer, activation layer, pooling layer and fully connected layer. The fully connected layer and convolutional layer comprise of trainable parameters which improves to optimize prediction accuracy during training process. The following section describes the operation of each layer in CNN.

4.2.1 Convolutional Layer

The filters or kernels in convolutional layer are adept in features extraction from input, i with the aid from the adaptable weights. The weights, W with size of width, M by height, N, and bias, b in each convolution layer, ℓ are adjusted throughout the training process in order to improve the ability of capturing the critical features (O'Shea & Nash, 2015). A nonlinear activation function, σ will be applied to the summation of the convolutions to obtain output, O ; where O is a tensor size of width, x and height, y. Equation (4.1) represents the operation.

$$O^{\ell}_{x^{\ell+1},y^{\ell+1}} = \sigma\left(\sum_{M=0}^{W}\sum_{N=0}^{H} W_{M,N} * i^{\ell-1}_{(x+M)(y+N)} + b\right) \tag{4.1}$$

4.2.2 Activation Layer

Activation function plays a role in capping the input elements within a manageable range via normalizing and rectifying unacceptable range of values (Nwankpa et al., 2018). There are different variants of activation functions which yield different output such as Linear, Tanh, Sigmoid, Softmax and ReLU functions. Equation (4.2) shows an activation function namely Rectified Linear Units or ReLU. The superior performance of ReLU in computation time reduction as compared with other types of activations has been proven (Krizhevsky et al., 2017); hence, ReLU is commonly applied to the output of convolutional layer. On the other hand, Softmax function, a type of activation function is applied to the output of the fully connected layer to ensure the predicted outcome is constrained within 0 and 1. Softmax function can expressed in Equation (4.3).

$$f(x) = \max\left(i^{\ell}_{x,y}, 0\right) \tag{4.2}$$

$$\sigma\left(i_{x,y}\right) = \frac{e^{i_{x,y_a}}}{\sum_{k=1}^{k} e^{i_{x,y_k}}} \quad for\ a = \ 0, 1, 2, \dots, k \tag{4.3}$$

4.2.3 Pooling layer

The pooling layer reduces the computational complexity by reducing the dimension of the feature maps from previous layer gradually. This greatly decreases the number of parameters required to be trained and controls the overfitting during training (O'Shea & Nash, 2015). Max pooling (Scherer et al., 2010) is a pooling operations which applies a filter size of (M, N) to partition the inputs into multiple non-overlapping sub-regions. The maximum element of the sub-regions of size $\frac{i_x}{M}$, $\frac{i_y}{N}$ is selected as output. Max-pooling operation can be expressed as:

$$O_{x,y}^{\ell-1} = \max\left(i_{x\times M, y\times N}^{\ell}\right) \tag{4.4}$$

4.2.4 Fully Connected Layer

Fully connected layer is usually introduced at the end of the network which unites the entire activations in the previous layer. The trainable weights, W and bias, b , in neurons in the fully connected layers are connected to each other and the connectivity resembles the biological neurons in animal brain (Hebb, 2005; Langille & Brown, 2018). The output, O of the function of a neuron given, n inputs can be defined as:

$$O_{x,y}^{\ell-1} = \sum_{i=1}^{n} W_i i_{x,y_i} + b_i \tag{4.5}$$

4.2.5 Objective Function

An objective function or loss function is a measurement to evaluate the deviation between the model's predicted outcomes with the actual results (Janocha & Czarnecki, 2017). Objective function will compute a large error value to 'punish' the model if predictions diverge too much from actual results. Gradually, the model will learn to reduce the error and improve the

prediction accuracy during backpropagation where the weights and biases are tuned proportional to error magnitude. Cross-entropy loss, or log-loss is a common loss function implemented on machine learning to quantify the dissimilarity between two probability distributions. Similarly, cross-entropy loss will increase if the predicted outcome diverges from actual results. Cross-entropy loss function, $J(\theta)$ can be expressed as:

$$J(\theta) = -\sum_{i}^{N} t_i \log(\hat{y}) \tag{4.6}$$

where N indicates the total number of classes (Normal and AB), label i is the correct classification and $\hat{y}$ is the predicted output from the model.

4.2.6 Optimizer

An optimizer is core to machine learning training as it helps to control or adjust the weights and learning rate to optimize the objective function (Sun et al., 2020). Mini-batch stochastic gradient-based optimization is one of the commonly applied optimization approaches in machine learning field. The advantage of mini-batch stochastic gradient descent (SGD) to compute gradient of a randomly sampled dataset with the size of mini-batch, η instead of entire dataset leads to cheap computation of the gradient on each iteration (M. Li et al., 2014; Robbins & Monro, 1951). The efficiency and feasibility of SGD is proven in notable machine learning applications such as (Graves et al., 2013; D. Li et al., 2013). Examples of optimization algorithm that uses the concept of SGD are *Adadelta* (M. Zeiler, 2012) and *Adam* (Kingma & Ba, 2017). Equations to update the weights in the model using *Adadelta* are expressed in Equations 4.7 to 4.9; whereas *Adam* equations are described in Equations 4.10 to 4.12.

$$D_t = \beta D_{t-1} + (1-\beta)[\theta_t - \theta_{t-1}]^2 \tag{4.7}$$

$$v_t = \beta v_{t-1} + (1-\beta)[\frac{\partial}{\partial \theta_t} J(\theta)]^2 \tag{4.8}$$

$$\theta_t = \theta_{t-1} - \frac{\sqrt{D_{t-1} + \epsilon}}{\sqrt{v_t + \epsilon}} * \frac{\partial}{\partial \theta_t} J(\theta) \tag{4.9}$$

where t is time step, θ represents the trainable weight or parameter, ϵ equals to 10e-6, a constant to ensure equation stability and β is the momentum value.

$$\theta_1 = \theta_{t-1} - \frac{\alpha}{\sqrt{\frac{v_t}{1-\beta_2^t}} + \epsilon} \cdot (\frac{m_t}{1-\beta_1^t}) \tag{4.10}$$

$$m_t = \beta_1 m_{t-1} + (1-\beta_1)\frac{\partial}{\partial \theta_t} J(\theta) \tag{4.11}$$

$$v_t = \beta_2 m_{t-1} + (1-\beta_2)\left[\frac{\partial}{\partial \theta_t} J(\theta)\right]^2 \tag{4.12}$$

where α is the learning rate, β_1 and β_2 are hyper parameters with 0.9 and 0.999 respectively, m and v are moving average of gradient and squared gradient respectively.

4.3 Methodology

In this chapter, our focus is to investigate the feasibility of implementing CNN to detect AB occurrence as well as to examine the optimal configurations (data composite type and training data quantity) for CNN model training. Therefore, the CNN model is not trained to categorize AB. It is hypothesized that CNN is able to detect AB, despite the heterogeneity of AB patterns and shapes. Such robust and automated machine learning model can eradicate the need to programming explicitly to detect AB and provide critical patient's condition information to clinicians in real time.

4.3.1 Patient Data for CNN Development

The 600,000 manually categorized patient's airway pressure waveforms from CARE trial as described in Chapter 3, are used for model training and performance evaluation. 10,000 AB and 10,000 of normal breathing (NB) cycles are selected for training and randomly choose non-repeated 3000 AB and 3000 NB from the remaining breathing cycles for training validation. Finally, the performance of the trained model is tested with about 600,000 classified datasets. These validation data are obtained from the first seven manually annotated patient's breathing cycles. Despite the validation dataset comprise only seven patients, we believe that, these data contain abundance apparent erratic and abnormal AB features ventilated using different MV modes and settings. Thus, it is sufficient for CNN to 'learn' capture them and identify them in high accuracy.

4.3.2 Data Composite Type

To understand the performance of CNN when trained with different data composite types; we train the CNN model with four different data structures. The patient's data are processed using Matlab 2017a (The MathWorks, Natick, MA). Details of the data composite type are shown as follows:

- *1-Dimension Data*

  The discrete airway pressure waveform profile from ventilator are arranged in a sequence and fed to CNN algorithm. The magnitude of each breathing cycle is normalized to 1 and the total of number of data points per breathing cycle is resampled to 150 data points to match with the CNN algorithm input size.

- *Line Graph*

  The airway pressure profile is mapped and plotted in a white background with solid black line. The linewidth is set as 0.5 during plotting on Matlab and saved as portable network graphics (.png) image format at 150x150 pixels as shown in Figure. 4.1. (Left).

- *Area Graph*

  The configuration of *Area Graph* is similar to *Line Graph* except the area under the curve is shaded; similarly, *Area Graph* is saved as portable network graphics (.png) image format at 150x150 pixels. Figure. 4.1. (Middle) shows an example of a normal airway pressure plotted and shaded.

- *Array Data*

  The configuration of *Array Data* is similar to *Area Graph*, apart from that the data format is a Python array. Similar to *Area Graph*, the graph is assigned with integer 1 to represent black pixels. The advantage of this method is avoiding the necessity to save the waveform into an image file; while preserving the data structure in 150 width and 150 height configurations. Figure. 4.2. (Right) shows the configuration of Array Data.

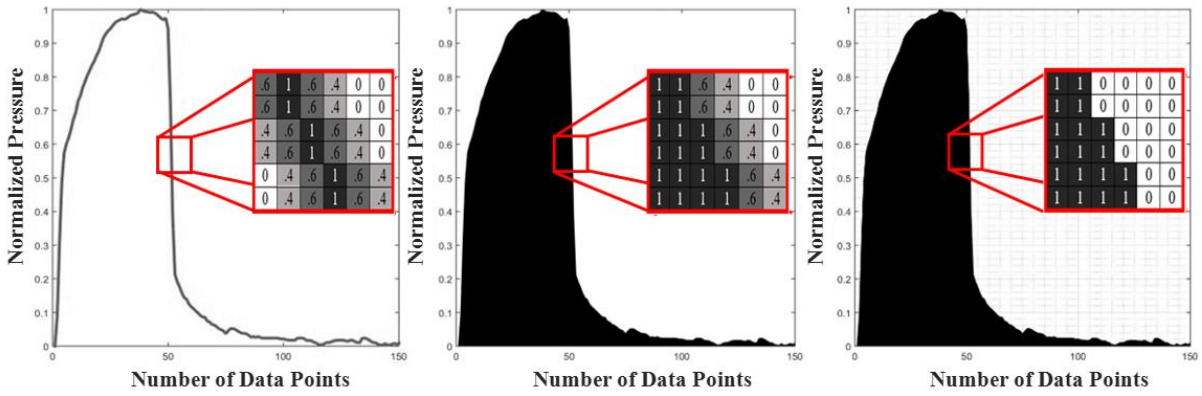

Figure 4.1: Examples of normal breath plotted in Line configuration (Left), Area configuration (Middle) and Array configuration (Right) with zoom-out box showing the pixel information of each configurations. As shown in the zoom-out box, Line and Area configuration underwent anti-aliasing when saving into (.png) image format; hence, grey pixels are produced due to interpolation at the edges or boundaries. Array configuration omits the need of undergoing anti-aliasing, as the waveform pattern is stored in python array.

As the aim of this study is to examine the feasibility of CNN to detect AB; thus, the types of asynchrony are not specified, and we only included asynchronies that are identifiable via visual observation on airway pressure profile for training and validation. AB such as reverse triggering, double triggering, flow asynchrony and delayed triggering (A. W. Thille et al., 2006) are included in the training dataset. Examples of AB present in the pressure waveform are shown in Figure. 4.2. A configuration of 150 × 150 pixels is chosen to enable manual annotation by the researchers. We found that reducing a size lesser than 150 × 150 pixels will reduce the image sharpness and clarity, whereas increasing the size greater will consume additional computation resources without significant return. A min-max normalisation to cap airway pressure within 0 to 1 was applied to ensure faster convergence of model during training (Jayalakshmi & A, 2011; Jin et al., 2015; Sola & Sevilla, 1997) as well as to ensure equal comparison of breathing cycles from different MV settings.

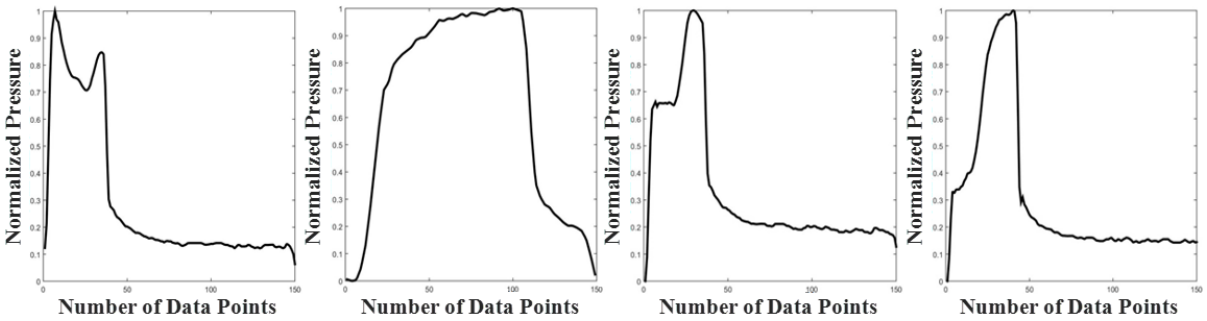

Figure 4.2: Normalised asynchronous pressure. 1st column: Pressure waveform of a breath cycle with reverse triggering. A negative deflection is present at the breath cycle. 2nd column: Pressure waveform of a breath cycle with double triggering. The breathing cycle will appear to have shorter expiration as compared with the others. 3rd column: Pressure waveform of a breath cycle with delayed triggering. A spike will appear on the pressure waveform just before exhalation. 4th column: Pressure waveform of a breath cycle with flow asynchrony. A concave shape negative dip will appear on the pressure waveform.

4.3.3 CNN Architecture

The CNN architecture applied in this study is similar to LeNet-5 (Lecun et al., 1998); apart from we apply some modifications to accommodate different inputs configuration. The architecture comprises of two convolutional layer, two max pooling layer and two fully

connected layers. Furthermore, we apply ReLU activation to every output from convolutional layer and Max pooling layer is applied to the outputs from convolutional layer. Besides, two fully connected layers with a size of 128 neurons and 2 output neurons are introduced to our proposed model. To prevent overfitting, dropout layer (Srivastava et al., 2014) is introduced to the layer after the flattened layer prior the first fully connected layer and in between fully connected layer. The output neurons yield a score ranging from 0 to 1 for each category: Asynchronous Breathing (AB) and Normal Breathing (NB). The highest score yields the category of the breathing cycle. In addition, binary crossentropy loss function is deployed along with '*Adadelta*' optimizer set to default ($\beta = 0$) to update the parameters during training this two-class classification model. All the models are trained 50 epochs with 32 mini-batch size training samples. Figure 4.3 shows the sequence of the CNN process where the input passes through multiple layers for feature learning and classification.

Table 4.1: Architecture of CNN according to data type.

Layers	**1-Dimension Data**	**Line Graph, Area Graph, Array Data**
	Settings/ Name	**Settings/ Name**
Input	150x1	150x150x1
Convolution	32x3	32x3x3
Activation	ReLU	ReLU
Convolution	32x3	32x3x3
Activation	ReLU	ReLU
Max Pooling	2	2x2
Dropout	0.5	0.5
Fully Connected	128	128
Activation	ReLU	ReLU
Dropout	0.5	0.5
Fully Connected	2	2
Activation	Softmax	Softmax

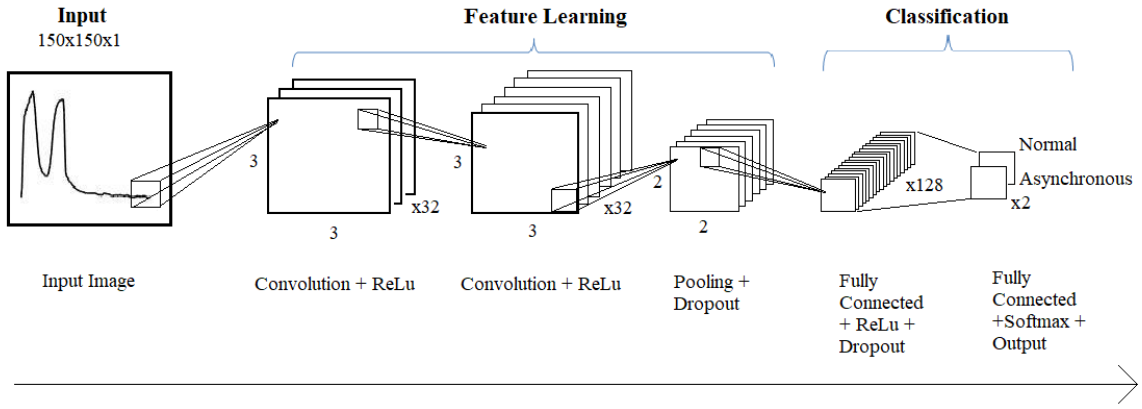

Figure 4.3: A generalised CNN training process flow.

4.3.4 Computational Setup and Training Process

All the CNN models are programmed in Python programming language (Python Software Foundation) and trained offline on a desktop computer (Core i5-7400 CPU, 8GB DDR4 RAM, NVIDIA GTX 1050Ti 4GB GPU). Theano (The Theano Development et al., 2016), a library written in python is implemented throughout the research to enable Graphical Processing Unit (GPU) usage to train machine learning models. As machine learning is computationally intensive, the deployment of GPU in our model training allows parallelism which significantly reduces our resources and time required in tuning the parameters (Chetlur et al., 2014).

4.3.5 Sensitivity and Specificity Analysis

True Positive (TP) is the number of times when the CNN model can correctly identify AB, and True Negative (TN) is the number of times when the model identifies NB. False Negative (FN) denotes the total number of times when the model has incorrectly classified normal breathing cycle as AB, and False Positive (FP) is the number of times when normal breathing cycles have been identified as AB.

4.3.6 Study I: CNN Models' Performance and Evaluation

Monte-Carlo simulation is conducted for each CNN model trained with different data composite type and different number of training datasets to study the performance of the CNN models. Specifically, the training sets per category for each CNN models were, *n*: 300, 1000, 5000 and 10,000. The training datasets are randomly chosen from 20,000 manually classified breathing cycles comprising both AB and NB.

The performance of the models is evaluated using sensitivity and specificity analysis. The Monte-Carlo simulation is repeated for 100 iterations, with each iteration's sensitivity and specificity recorded. To avoid arbitrary results, all CNN models are rebooted with a new model every iteration during Monte-Carlo simulation. During Monte-Carlo simulation, all models are initialised with a random number generator by seedi$_n$g the same value to ensure equal comparison. Figure 4.4 shows a schematic flow of the Monte-Carlo simulation. Nonetheless, in order to ensure impartiality, the validation dataset of 6000 breathing cycles are isolated from the training data.

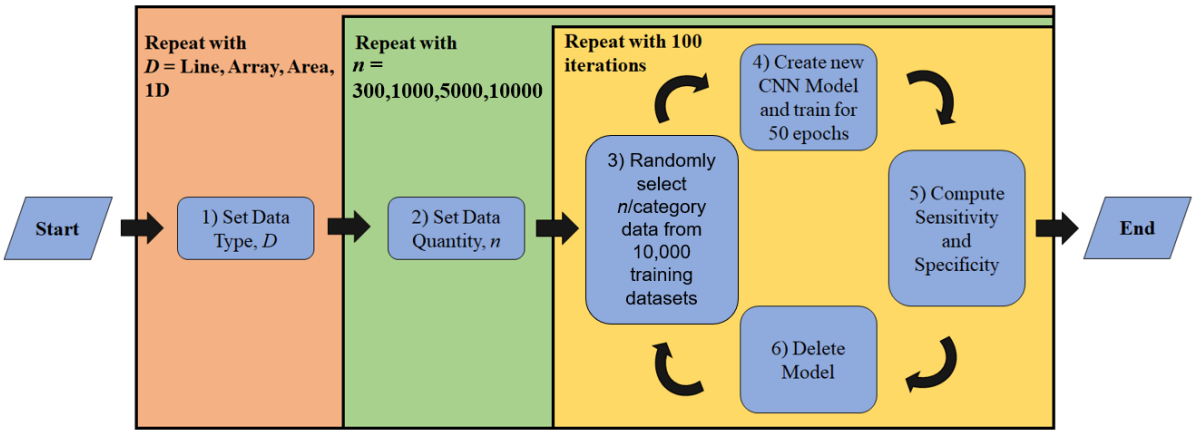

Figure 4.4: Monte-Carlo simulation process to test robustness of CNN models.

4.3.7 Study II: Performance Evaluation using Clinical Data

This study extends the previous study, where we investigate the performance of highest accuracy trained in 1-dimensional and 2-dimensional data in Monte-Carlo simulation CNN model to validate with clinical data. The 600,000 manually categorised breathing cycles from 7 mechanically ventilated patients are used for model evaluation.

4.4 Results

4.4.1 Study I: Robustness Test

In this study, four CNN models are developed and trained with different data composite type and data quantity. The CNN models are named according to the data structure they are trained in. The four CNN models are: a) CNN_1D (one dimension data), b) CNN_Line (Line Graph), c) CNN_Area (Area Graph) and d) CNN_Array (Array Data).

Figure 4.5 shows the sensitivity and specificity plot of Monte-Carlo simulation when trained with different number of training data from each category. The results show that, for each CNN model, when trained with randomly chosen datasets, the performance of CNN models improves proportionally to the training data quantity. Interestingly, when CNN_Area was trained with *n* lesser than 5000, CNN_Area achieved relatively higher consistency as it clustered together; despite low sensitivity and specificity performance. On the contrary, CNN_Line attained the least consistency when trained with *n* lesser than 5000, as the performance is scattered.

The distribution plot in Figure 4.6 shows the summary of the performance of CNN models from each data structure when trained with 10,000 training data. All CNN data structure type

achieved high performance with near 100% sensitivity and specificity when trained with 10,000 training datasets per category and validated with 3000 datasets from each category.

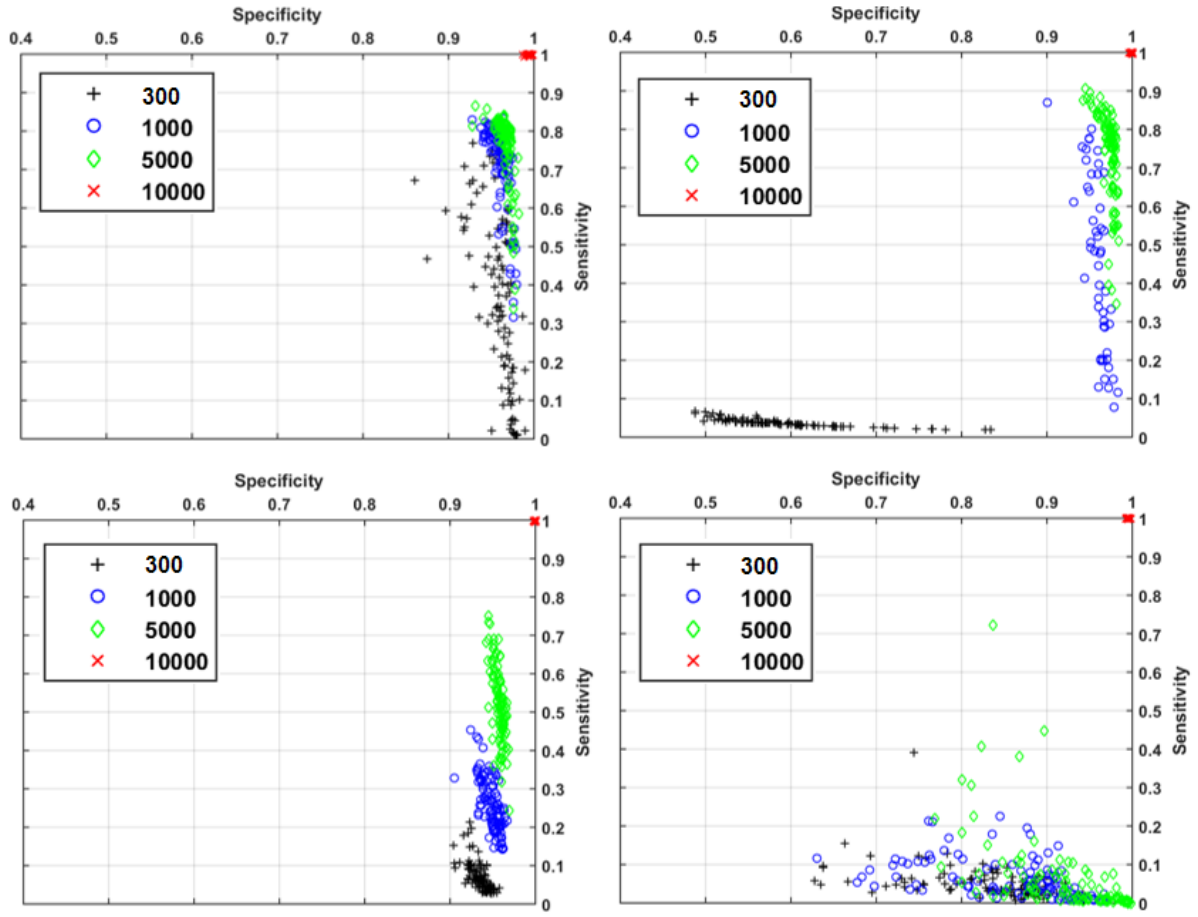

Figure 4.5: Monte-Carlo simulation of each data structure. Top left: Monte-Carlo results of CNN 1D. Bottom left: Monte-Carlo results of CNN_Area. Top Right: Monte-Carlo results of CNN_Array. Bottom Right: Monte-Carlo results of CNN_Line.

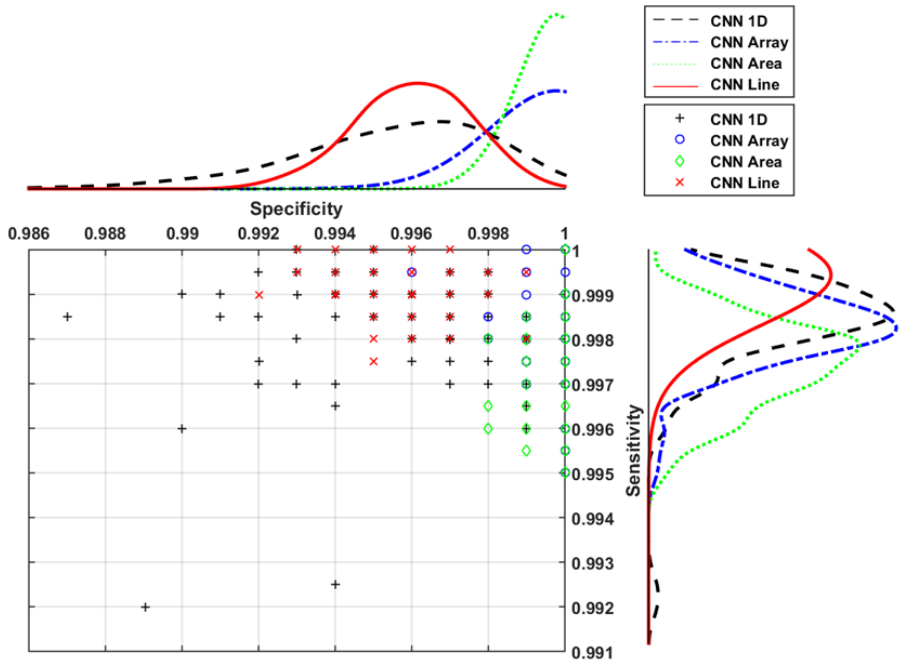

Figure 4.6: Performance of CNN models trained with 10,000 training data from different data structures.

Table 4.2 shows the summary of total training time of Monte-Carlo simulation for each data structure and quantity. The results show that the time required to complete the training increases proportionally as the data quantity increases; except, CNN_1D required the least training time regardless of training data quantity as compared CNN trained with 2-dimensional data. CNN required 5 hours to complete the Monte-Carlo simulation when trained with merely 300 image training datasets; whereas, CNN_1D only required 10 times lesser.

Table 4.2: Total simulation run time of Monte-Carlo simulation for each type of data configuration and data quantity.

	Data Type	
Data Quantity /Category	**1-Dimension Data**	**Line Graph, Area Graph, Array Data**
300	0.5 hour	5 hours
1000	1 hour	45 hours
5000	2 hours	75 hours
10,000	4 hours	150 hours

Table 4.3 shows the summary of the CNNs performance from Monte-Carlo simulation. It was found out that when trained with 300 data, CNN_1D achieved lowest consistency with

95th percentile and 5th percentile difference of 70% sensitivity when compared with other 2-demensional data structure. Similarly, CNN_1D had highest specificity 95th percentile and 5th percentile difference when trained with 10,000 when compared with other data composite type. On the contrary, 2-dimensional data structures (Array, Area and Line) achieved relatively lower median sensitivity and specificity than 1-demensional data structure type when trained with 5000 datasets or lesser.

Table 4.3: The summary of sensitivity and specificity of the CNN models during Monte Carlo simulation.

Data	CNN_1D		CNN_Array		CNN_Area		CNN_Line	
Quantity	Sen	Spe	Sen	Spe	Sen	Spe	Sen	Spe
300	37.8 [1.9-72.3]	96.3 [91.8-98.0]	3.7 [2.1-5.9]	57.4 [50.4-76.5]	6.3 [2.9-15.2]	93.5 [91.7-95.1]	4.2 [1.1-12.0]	85.1 [66.5-93.8]
1000	74.4 [43.5-81.7]	96.2 [94.1-97.7]	48.0 [12.7-77.6]	96.1 [94.2-97.7]	24.6 [15.3-36.2]	95.1 [93.2-96.2]	4.7 [0.8-17.8]	88.0 [70.8-95.8]
5000	78.8 [51.2-83.5]	96.7 [95.3-97.7]	76.5 [52.0-87.7]	95.2 [97.5-98.2]	51.6 [36.6-68.9]	95.8 [94.5-96.6]	2.1 [0.3-31.3]	93.3 [80.6-99.4]
10,000	99.9 [99.7-100]	99.6 [99.1-99.9]	99.9 [99.7-100]	100 [99.9-100]	99.8 [99.6-99.9]	100 [99.9 -100]	100 [99.8-100]	99.6 [99.4-99.8]

Sen = Sensitivity, Spe = Specificity, Median [5th – 95th percentile]

Table 4.4 shows the sensitivity, specificity and accuracy of the best CNN model based on the Monte-Carlo simulation. Results show that, all the models achieved an average accuracy of more than 99%; while, CNN_Area and CNN_Array achieved 100% accuracy when validated with another set of validation data.

Table 4.4: Sensitivity analysis of the robustness of the best CNN models From the Monte-Carlo simulation.

CNN_1D			CNN_Array			CNN_Area			CNN_Line		
Sen	Spe	Acc	Sen	Spe	Acc	Sen	Spe	Acc	Sen	Spe	Acc
100.00	99.90	99.95	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.90	99.95

Sen = Sensitivity, Spe = Specificity, Acc = Accuracy

4.4.2 Study II: Performance on Clinical Data

Table 4.5 shows the performance of CNN_1D and CNN_Array when tested on clinical data. Results showed that CNN_1D attained high average specificity of 91.44% and sensitivity of 73.37% and CNN_Array only yielded specificity of 90.34% and sensitivity of 68.27%. Both CNN models showed good accuracy in AB detection; but both CNN models performed poorly in Patient 3 who was ventilated using SIMV pressure control mode. CNN_1D and CNN_Array had sensitivity of 24.01% and 22.42%, respectively. CNN_1D and CNN_Array also achieved relatively poor performance in detecting normal breathing cycles in Patient 4 with specificity of 54.77% and 57.54% respectively. Overall, despite the performance of CNN_1D and CNN_Array degraded in some cases; their overall performance achieved an average of near 90.0% accuracy with more than 500,000 validation clinical data.

Table 4.6 shows the comparison between the AI computed manually and CNN computed AI. The results show that there are some discrepancies in performance between CNN_1D and CNN_Array when detecting the presence of AB. It is observed that the performance of CNN_1D deteriorated when identifying AB in patients in PC or SPONT modes such as Patients 2, 3 and 6.

Table 4.5: Patient demographic along with CNN_1D and CNN_Array performance.

				CNN_1D			CNN_Array		
Patient No.	**Diagnostic**	**MV Mode**	**Total Breathing Cycle**	**Sen**	**Spe**	**Acc**	**Sen**	**Spe**	**Acc**
P1	Thyroid carcinoma	SIMV (VC mode)	270,126	92.81	99.15	97.43	83.18	98.59	94.00
P2	Pneumonia	SIMV (VC mode)/ SPONT	61,832	84.03	93.46	87.67	76.82	94.16	82.38
P3	SVC Obstruction	SIMV (PC mode)	46,067	22.83	99.80	95.81	20.84	99.38	96.46
P4	HAP	SIMV (VC mode)	21,286	97.84	54.77	70.79	94.02	57.54	73.35
P5	Klebsiella Sepsis	SIMV (VC mode)	152,229	86.68	94.26	91.10	83.31	88.62	86.55
P6	Pneumonia	SIMV (VC mode)/ SPONT	5,989	40.99	99.55	90.48	24.84	99.75	80.02
P7	Pneumonia	SIMV (VC mode)/ BiPAP	38,846	96.42	94.89	95.65	94.88	94.33	94.60
		Total/Average	596,675	73.37	91.44	89.34	68.27	90.34	86.76
		25^{th} Percentile	25,676	51.75	93.66	88.37	37.84	90.01	80.61
		Median	46,067	86.68	94.89	91.10	83.18	94.33	86.55
		75^{th} Percentile	129,630	95.52	99.45	95.77	91.34	99.18	94.45

*SVC – Superior vena cava; HAP – Hospital acquired pneumonia; SIMV – Synchronous intermittent mandatory ventilation; SPONT – Spontaneous; BiPAP – Bilevel Positive Airway Pressure; VC – Volume mode; PC – Pressure Control; Sen – Sensitivity; Spe – Specificity; Acc – Accuracy

Table 4.6: Comparison between manual and CNN computed AI.

	AI (%)		
Patient No.	**Manual**	**CNN_1D**	**CNN_Array**
P1	25.77	27.10	29.80
P2	54.06	62.62	67.93
P3	1.37	5.17	3.72
P4	64.80	37.19	43.33
P5	39.46	41.66	39.04
P6	6.73	15.48	26.34
P7	50.41	49.61	50.15
Median	39.46	37.19	39.04
IQR	[11.49–53.15]	[18.34–47.62]	[29.21–48.45]

IQR – Interquartile Range

4.5 Discussion

4.5.1 Study I: Robustness Test

The distribution plot in Figure 4.5 shows that when trained with 300 breathing cycles, CNN_1D and CNN_Line had the lowest performance compared to others, with their sensitivity and specificity scattered in Monte-Carlo simulation. When CNN_1D was trained with more data, the precision improved significantly, and its performance begin to be consistent. However, CNN_1D still struggled to maintain consistency even when trained with 10,000 training datasets. Table 4.3 shows that CNN_1D had the highest specificity 5^{th} and 95^{th} percentile difference. We speculate that the inconsistency CNN's performance was due to the outlier data points or anomalies in the training data. The presence of irregularities in the training data may decrease the CNN performance (Khamis et al., 2005) as non-essential features are extracted for training. Similarly, possible reason behind CNN_1D lower performance when trained with 10,000 is due to model overfitting as repetitive insignificant features and irregularities were introduced into CNN. In this case, dropout is an effective method to prevent overfitting. Henceforth, by introducing different dropout rate or batch normalization method in the model may help to improve the model's performance (Ioffe & Szegedy, 2015; Srivastava et al., 2014).

CNN_Array's performances improved when trained with more data and began to cluster when trained with more than 5000 datasets. Inconsistency CNN_Array's performance when trained with lesser amount of data shows that CNN struggled to capture the vast 2-dimensional latent information from the data structure; however, both median sensitivity and specificity achieved by CNN_Array is still greater than CNN_Area and CNN_Line when trained with 5000 data or lesser. Furthermore, CNN_Area had sensitivity and specificity clustered close together when trained with only 300 data points. This result shows that, CNN_Area had consistency in extracting and learning essential features; although, the

performance was still far behind from CNN_1D. The superiority of CNN_Area performance as compared with CNN_Line showed that shading area under the curve provides additional information; thus, enhanced CNN performance.

Furthermore, Table 4.3 shows that the median of CNN_Area and CNN_Line were lower than CNN_1D and CNN_Array when trained with training datasets lesser than 5000. We believed that the ability of CNN_Area and CNN_Line to extract features drop noticeably when analysing anti-aliased images due to the presence of redundant anti-aliased pixel information. By removing redundant pixel information, the dimension of the data is altered leaving only binary data; thus, CNN does not require additional training dataset to extract and distinguish these anti-aliased features.

Table 4.4 shows that all the CNN models achieved an average of 99% sensitivity and specificity or more than 99.95% accuracy when conducted with validation data to test the robustness of CNN models. This result showed that CNN is a highly robust machine learning algorithm that is capable of capturing AB regardless of input data type (1D or Image) when trained with adequate amount of data. However, the vast amount of information in 2 dimensional data structure abates the performance of CNN to extract and learn the features; thus, requiring additional amount of training data for CNN to learn and understand the underlying patterns. In contrast, the inconsistent CNN_1D performance had led to relatively higher specificity 5^{th} and 95^{th} percentile gap as compared with 2-dimensional data structure type when trained with 10,000. It is believed that, the abundance of pixel information in 2-dimensional data structure benefits CNN during training; thus, CNN_Area, CNN_Line and CNN_Array achieved relatively consistent performance compared with CNN_1D.

Furthermore, Table 4.2 shows that CNN_1D required least training time to complete Monte-Carlo simulation. Simulation completion time is affected by the training data sizes. 2-dimensional data structure consists of 150 by 150 pixels or equivalent to 22,500 pixels information; whereas 1-dimensional data only consists of 150 datapoints. Moreover, additional pixels information indicates additional memory consumption in the system. Hence, additional resources and time is required to train 2-dimensional data composite types.

Nonetheless, despite both CNN_Area and CNN_Array achieved similar performances and outperform CNN_1D when trained with 10,000 datasets, CNN_1D data structure type is still favourable due to the remarkable performance and consistency when trained with lower training dataset quantity. Besides, low training time could potentially conserve more resources as CNN_1D consistently attains the least amount of training time among all models.

4.5.2 Study II: Performance on Clinical Data

The results in Table 4.5 show that CNN_1D and CNN_Array are capable to distinguish AB and NB effectively. However, the lack of training data from different ventilation mode could compromise the performance of CNN. As the amount of training dataset with pressure-controlled (PC) ventilation mode may be relatively lesser than the breathing cycles under the different modes; hence, CNN_1D and CNN_Array failed to identify AB correctly in these patients. In addition, the atrocious performance of the models to identify normal breath in Patient 4 could suggest that 10,000 normal or AB training data ventilated using different MV modes is insufficient. As breathing cycles can be altered due to the change of MV modes (Baedorf Kassis et al., 2021; Daoud et al., 2012; Leon & Lorini, 1997), assuring training data comprising balance amount of AB and normal breathing cycle ventilated using different MV

will potentially prevent biasness in the model. training data is insufficient; thus, causing the model to underfit during Monte-Carlo simulation analysis.

Furthermore, despite Monte-Carlo simulation shows that CNN_Array achieved high sensitivity and specificity; its performance is relatively poor when compared to CNN_1D in the clinical data validation. It is possible that the performance deterioration is due to the feature alteration or loss during conversion from 1-dimensional to 2-dimensional data structure. Furthermore, poor CNN_Array performance may also be due to the lack of trainable parameters in fully connected layer, impeding the ability to cater large quantity pixel information in 2-dimensional data structure (Bansal et al., 2017; Basha et al., 2020).

Besides that, Table 4.6 shows there are discrepancies between manually computed AI and AI attained by CNN models. Similarly, the difference in performance could indicate that CNN_1D and CNN_Array models are underfitting due to the lack of AB features ventilated under different MV modes. For example, Patient 6 was ventilated using a mixture between SPONT and SIMV. The lack of training data ventilated using SPONT could be the reason behind the deterioration of CNN_1D to compute AI closer to the AI computed manually for Patient 6. Nonetheless, despite the difference in performance, CNN_1D achieved better average sensitivity and specificity than CNN_Array. Hence, the CNN model trained with 1-dimensional data structure is preferred in AB detection task in this study.

4.5.3 Limitation

In summary, the 1-dimensional data format in this study is the optimal data type to perform AB detection. There are a few limitations in the developed models that we wish to highlight. The first limitation is that this study only detects the present of asynchronies in airway

pressure. The absence of the ability of detect flow waveform could deprecate the assessment of patient-ventilator interaction. One of the concerns in this research is that the effect of normalisation of the training data is not studied thoroughly. In particular, normalisation may alter the data, potentially lose its original form (KumarSingh et al., 2015; U POR, 2011). It is however, normalisation offer opportunities to compare inter- and intra- variability of AB, reducing the amount of data needed for model training and faster convergence (Jin et al., 2015).

Moreover, the MV settings or patient-specific breathing pattern may limit the performances of the CNN models. In this study, the training data mainly consists of SIMV and VC , where patient-effort is obvious in airway pressure (Ramirez & Arellano, 2018). This allows us to investigate the effect of asynchrony induced in AB in correlation with patient's oxygenation during MV in the following chapter. Other settings such as neurally adjusted ventilator assist, proportional assisted ventilation or airway pressure release ventilation were not included (V. J. Major et al., 2018). The variabilities in pressure or flow profile in different MV modes thus requires a wider range of AB and NB data for the development of a one size-fit-all CNN model.

Another limitation is that the CNN models developed here may not be able to detect all forms of asynchrony. CNN may fail to identify them due to the subtle features defining these asynchronies. For example, ineffective triggering is described as a simultaneous decrease in airway pressure and an increase in airflow without assisted cycle (A. W. Thille et al., 2006). Figure 4.7 shows an example of ineffective triggering; however, as the magnitude of the deflection is small, CNN may fail to capture the deflection due to unclear features. Similarly, asynchronies that were dependent of previous breathing cycles and not present in the current

cycle may be detected using different machine learning approach (L. Zhang et al., 2020) More studies are needed in order to determine the appropriate settings to capture the pattern of AB with small deflection magnitude.

Finally, all CNN models in this study are only capable of distinguishing AB and NB breathing cycles but unable to quantify magnitude of asynchronies or recognize types of AB. Further studies on quantifying AB are required to provide additional information on patient-ventilator interaction to the clinicians.

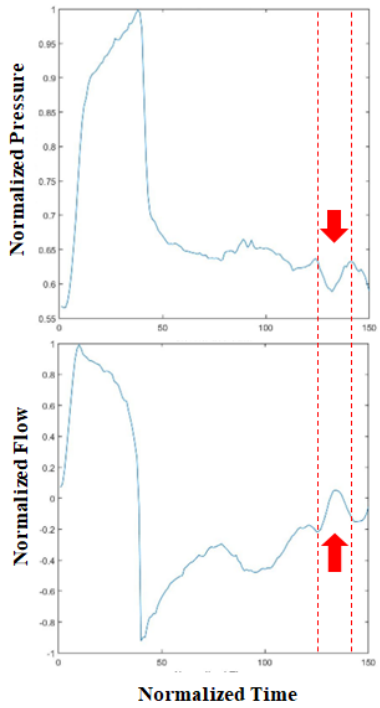

Figure 4.7: Example of ineffective triggering with both pressure and flow waveform.

4.6 Summary

In this chapter, we have presented the development of CNN models to detect AB using airway pressure data from the MV. It was found that performance of CNN varies with different data structure and quantity. In this study which involved airway pressure waveform, the 1-dimension data structure is the preferred data type for CNN training due to accuracy and consistency, besides rapid training completion time could potentially save more

resources. Furthermore, we conclude that the performance of each CNN model regardless of data composite type can achieve high accuracy when trained with additional data.

In summary, we show that the CNN's ability in detecting AB is promising and it can be implemented at bedside to monitor the quality of PVI during MV treatment automatically. As CNN model is only able to detect the presence of AB during MV; the ability to assess the magnitude of AB is still absent. In the next chapter, we present two machine learning models to quantify the magnitude of AB. The chapter includes a comparison study between the developed machine learning models with an existing mathematical model.

Chapter 5
Asynchrony Magnitude Quantification

This chapter presents the development of machine learning algorithms in quantifying the magnitude of asynchronous breathing. The ability to assess the degree of severity of asynchronous breathing may help to further understand the effects of asynchronous breathing towards patient's condition. We present two different machine learning algorithms that potentially compute the magnitude of asynchrony. The machine learning algorithms are *ABReGAN* and *ABReCA*. The performance of these methods is then compared with an existing mathematical model.

5.1 Introduction

The shape of a breathing cycle may be altered due to the presence of patient effort. For example, reverse triggering is a neuromechanically triggered AB when the diaphragmatic muscle contraction is triggered by ventilator insufflations (Akoumianaki et al., 2013). The presence of reverse triggering imposes anomalies to the mechanical ventilation airway pressure waveform. Consequently, the patient's effort (asynchrony) masks the underlying physiological properties; (Brochard et al., 2012; Talmor et al., 2008); therefore leading to mis-identification of respiratory mechanics (Chiew et al., 2015; Lucangelo et al., 2007; V. Major et al., 2015). Figure 5.1 shows an example of poor fitting of using single compartment model due to the presence of patient effort. Hence poor estimation of respiratory mechanics using conventional model severely hinders the usage of model-based respiratory mechanics in MV monitoring and guidance.

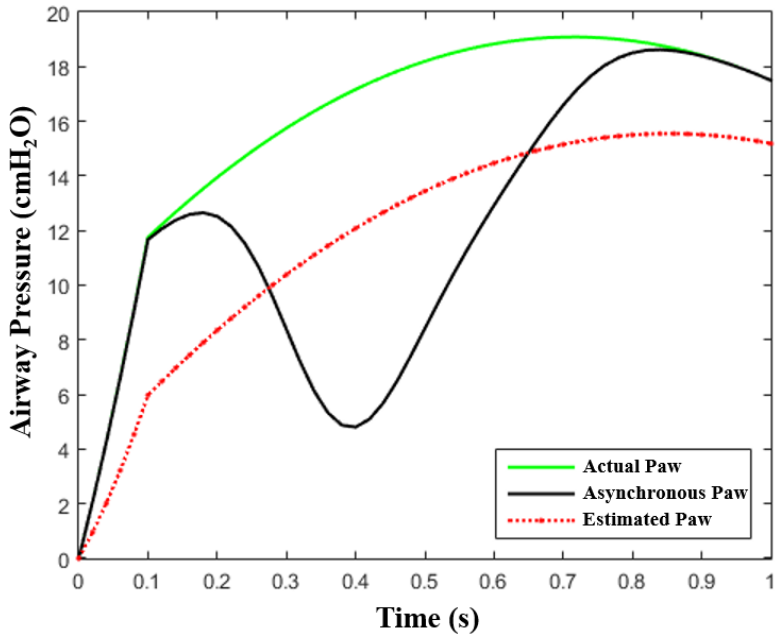

Figure 5.1: Linear regression yields elastance of 30 cmH_2O/L and resistance of 4 cmH_2Os/L; however, the actual airway pressure has elastance of 35 cmH_2O/L and resistance of 10 cmH_2Os/L. Presence of patient effort leads to miscalculation of elastance and resistance values.

Airway pressure or flow profile reconstruction is a technique to reconstruct or 'restore' AB to an asynchrony free breathing cycle in the effort to better estimation of the underlying respiratory mechanics (Chiew et al., 2018; Kannangara et al., 2016). Once reconstructed, the airway profile will manifest a normal breathing cycle. As a result, the difference in area under the curve between the reconstructed breathing cycle and AB can be defined as the magnitude of an AB. Chiew et al (Chiew et al., 2018) proposed a method namely iterative pressure reconstruction (*IPR*), a mathematical model which uses single compartment model to reconstruct AB to normal breathing cycle for better respiratory mechanics estimation. However, *IPR* lacks robustness when reconstructing asynchronies that present at different location, as a result, this impinges on the performance to reconstruct AB. Therefore, high robustness machine learning approach is proposed to improve the performance to reconstruct AB with least error.

In this chapter, two different machine learning techniques to reconstruct airway pressure are presented. The first method uses Generative Adversarial Network (GAN), an unsupervised machine learning technique that utilizes two neural networks namely discriminator, *D* and generator, *G* competing each other in a zero-sum game framework (Goodfellow et al., 2014). Ultimately, the trained *G* model will learn to restore AB to a clean airway pressure. The second approach utilizes another unsupervised machine learning algorithm, namely convolutional autoencoder (CAE) (Y. Zhang, 2018) to reconstruct AB to asynchrony free breathing cycle. These models are trained and validated with simulated patient effort imposed breathing cycles.

This chapter further extends the study by including a comparison study between the performance of developed machine learning models and a mathematical model to reconstruct breathing cycles. We hypothesise that the exemplary high robustness of machine learning to 'learn' and 'understand' essential and intrinsic AB properties could outperform conventional mathematical model to better reconstruct AB. Ultimately, the best performing model is selected for further testing with actual clinical data.

5.2 Methods

5.2.1 Simulated AB for Training and Validation

In order to prepare data for machine learning training and assessment, a forward simulation was performed. The forward simulation incorporates both single compartment model, SCM, and gaussian effort model, GEM (Arunachalam et al., 2020), to generate normal breathing cycle and patient's effort respectively.

SCM can be extended to simulate different magnitude and patterns of AB by imposing patient effort generated using GEM into the function. The single compartment model with patient effort, is defined:

$$P_{aw}(t) = EV(t) + RQ(t) + P_0 + P_e \tag{5.1}$$

where P_{aw} (cmH_2O) is the airway pressure at time t (s), consists volume, V (L), the integral of airway flow, Q (L/s), the respiratory (resistance, R (cmH_2Os/L) and elastance, E (cmH_2O/L)) and P_e (cmH_2O) is the patient effort.

On the other hand, GEM model utilizes three Gaussian basis functions to model the nonlinear patient effort. These basis functions allow GEM to simulate different shapes and sizes of patient effort present in reverse triggering. GEM can be expressed as follow:

$$P_e = \sum_{i=1}^{3} A_i e^{-\left(\frac{f-\mu_i}{\sigma_i}\right)^2} \tag{5.2}$$

The gaussian effort model consists peak effort of pressure of i^{th} Gaussian curve, A_i(cmH_2O), normalized patient effort timing, f, center of i^{th} Gaussian curve, μ_i and width of i^{th} Gaussian curve σ_i.

In this study, a total of 1,000 normal breathing cycle with their corresponding 420,000 unique AB were simulated using the parameter range shown in Table 5.1. Each simulated pressure waveform is normalized to 0 and 1 and resampled to 64 data points. Normalization and resampling are necessary to reduce variability of the signal magnitude and duration. As a result, the normalized breathing cycle will yield a range of respiratory mechanics *E* from 1 to 1.77 and *R* from 0.08 to 0.62 during models' performance validation. These values are relatively smaller than the actual *E* and *R* used to simulate breathing cycles. Furthermore, as

the volume and flow settings were kept constant, therefore, only volume control (VC) mode breathing cycles were generated in this study.

Table 5.1: Parameters used for synthetic data generation simulation.

Parameters (Unit)	**Range**	**Resolution**
Elastance, E (cmH_2O/L)	25-60	5
Resistance, R (cmH_2Os/L)	1-20	1
Magnitude of Pressure, A_1 (cmH_2O)	1-10	1
Magnitude of Pressure, A_2 (cmH_2O)	1-10	1
Magnitude of Pressure, A_3 (cmH_2O)	1-10	1
Centroid of Gaussian curve, μ	-3 – 3	3
Width of Gaussian curve, σ	1	Constant

5.2.2 Model Training and Development

The main objective of the developed machine learning models (GAN and CAE) is to reconstruct AB with least distortion to compute the magnitude of AB. We begin training the models using the simulated AB as input and the simulated normal breathing cycles as output. The trained models' performance is evaluated by determining the difference of the ground truth respiratory mechanics (RM) with the RM of the reconstructed AB. Figure 5.2 shows the overall flow of models' performance evaluation in this study. Detailed information of GAN and CAE models training and development process are discussed in Section 5.2.5 and 5.2.6 respectively.

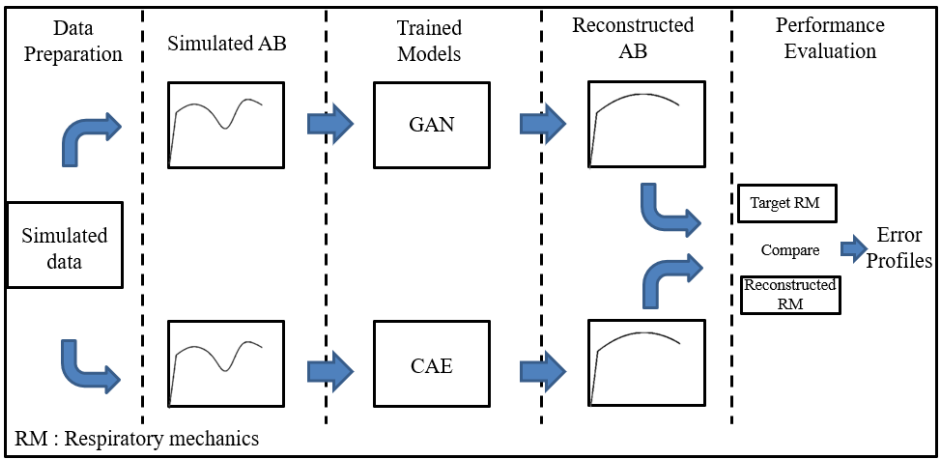

Figure 5.2: Overview machine learning models performance evaluation process.

5.2.3 Models' Performance Evaluation

The RM parameters (resistance, R and elastance, E) value of each model reconstructed airway pressure, P_{aw} will be estimated by fitting the single compartment model equation using linear regression (Bates, 2009):

$$P_{aw}(t) = EV(t) + RQ(t) + P_0 \tag{5.3}$$

where airway flow, Q, and lung volume, V were kept constant and pressure offset P_0 was kept as zero.

The performance of the developed models is validated using error profiles such as Mean Absolute Percentage Error (MAPE), Sum Squared Error (SSE) and Absolute Percentage Error (APE). These errors measure the disparity between the reconstructed elastance, E_{rec} and resistance, $R_{rec}$with the actual elastance, $E_{ac}$and resistance, R_{ac}.

APE computes error by measuring the absolute distance between actual and predicted points; while, MAPE extends APE by measuring the mean error. Formula to compute APE and MAPE are shown as below:

$$APE = \left[\frac{(RM_{Act} - RM_{Rec})}{RM_{Act}}\right] \times 100\% \tag{5.4}$$

$$MAPE = Mean\left[\frac{(RM_{Act} - RM_{Rec})}{RM_{Act}}\right] \times 100\% \tag{5.5}$$

where RM_{Act} is the actual respiratory mechanics and RM_{Rec} is the reconstructed respiratory mechanics. While, the sum of squared error, SSE formula is shown as below:

$$SSE = \sum (RM_{Act} - RM_{Rec})^2 \tag{5.6}$$

5.2.4 Computational Setup

All the machine learning models are trained offline on a computer running on Windows 10 with Intel Core i5-7400 CPU (4 cores), 32 GB DDR4 RAM and NVIDIA GTX 1050Ti 4GB GPU. The models are developed in Python 3.5 (Python Software Foundation) and Python library, Theano (The Theano Development et al., 2016). Theano allows parallel computation using Graphical Processing Unit (GPU) to reduce resources and training time consumption.

5.2.5 Generative Adversarial Network (GAN)

Generative Adversarial Network (GAN) is an unsupervised deep learning architecture that involves two neural network namely generator, $G(i_{noise})$ and discriminator, $D(\bar{X}, T)$ contesting each other in a zero-sum game framework (Goodfellow et al., 2014). Ultimately, the generator learns to generate a synthetic data, $\bar{X}$ similar to actual data points by mapping the random uniform noise sample, i_{noise} inputs. On the other hand, the D receives synthetic data, $\bar{X}$ from G to determine whether the generator produces output close to the target, T. During training, both G and D will contest with each other, where G will minimize its error by generating a synthetic data which is close to real data distribution, T; whereas, D will maximize its ability to determine the authenticity (real or fake) of the generated synthetic data. Eventually, G will learn to generate synthetic data similar to real data distribution by mapping the inputs. The objective function of GAN can be expressed as:

$$\frac{min}{G}\frac{max}{D}J(D,G) = Er[logD(\bar{X})] + Er\left[\log\left(1 - D\left(G(i_{noise})\right)\right)\right] \tag{5.7}$$

where Er is the error computed by the discriminator or the generator. The output result of the discriminator is regarded as optimal when the distribution of the generated data is close or equivalent to the real data (Goodfellow et al., 2014).

AB Reconstruction Generative Adversarial Network (ABReGAN)

AB Reconstruction Generative Adversarial Network *(ABReGAN)* uses the GAN algorithm to reconstruct AB to a normal breathing cycle for the use of quantifying the magnitude of AB. The G in *ABReGAN* is designed to accept simulated AB as inputs, I and generate synthetic data, $\bar{X}$. The flow of *ABReGAN* reconstruction training process is shown in Figure 5.3. During the training process, G will learn the distribution of AB patterns and generate synthetic normal breathing cycle with least distortion. On the other hand, D will update the parameters according to the difference between the generated and actual normal breathing. Ultimately, once *ABReGAN* is trained, the G will learn to reconstruct AB to an asynchrony free breathing cycle when given AB as input.

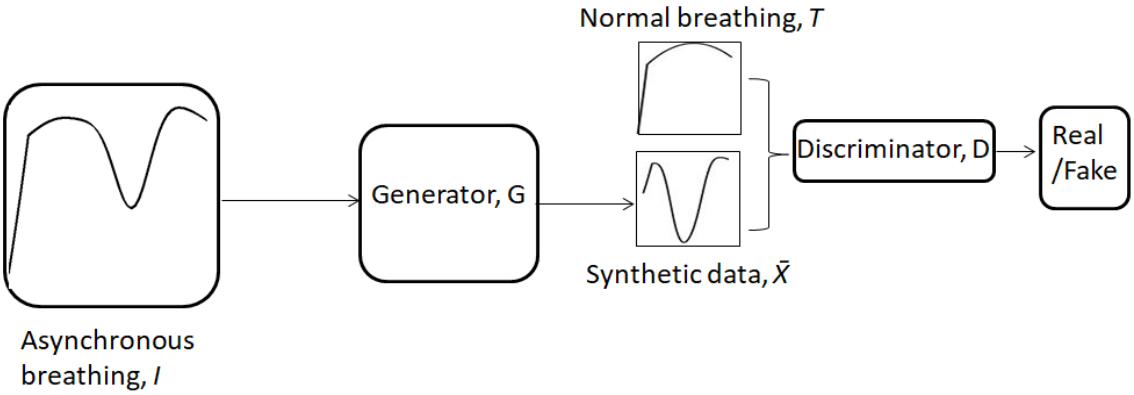

Figure 5.3: *ABReGAN* architecture during training. Asynchronous breathing will be the input of generator and discriminator will compare the generated $\bar{X}$ data with ground truth, T.

ABReGAN Architecture and performance evaluation

Figures 5.5, 5.6 and Table 5.2 show overall *ABReGAN* architecture respectively. *ABReGAN* is trained 128 mini-batch size with adaptive moment estimation (*Adam*) (Kingma & Ba, 2017) as optimizer as denoted in Equation 4.10 to 4.12 and trained for 10,000 iterations. Batch normalization momentum is set as 0.8 and alpha value for LeakyRelu set as 0.2. Once *ABReGAN* is trained, it is validated with 420,000 of its own training datasets to evaluate the

performance of *ABReGAN* in recognizing its 'observed' data. In order to evaluate the robustness of *ABReGAN*, k-fold cross validation technique is applied (Berrar, 2018). K-fold cross-validation technique is a common practice to evaluate the robustness of a machine learning model. K-fold process involves shuffling and partitioning of the whole training data into different training and testing groups with *k* number of times. In this study, the performance of *ABReGAN* is evaluated by a 10-fold cross validation on the 420,000 simulated breathing cycles. The flowchart shown in Figure 5.4 depicts the k-fold evaluation process. Ten percent of the training dataset are reserved for model testing while the rest are used for training. The following outlines the operations implemented in *ABReGAN*.

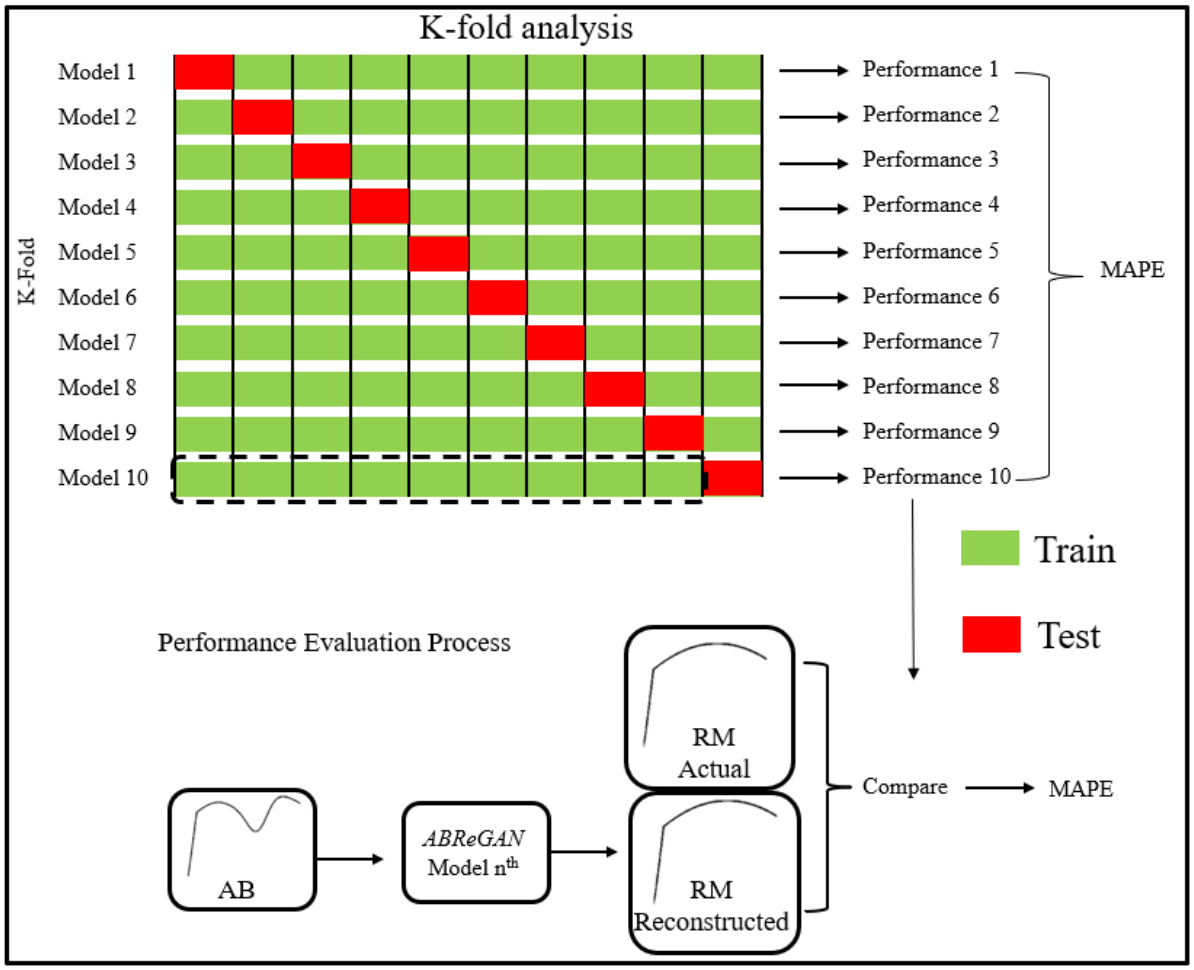

Figure 5.4. Evaluation process of the 10-fold cross validation for *ABReGAN* to test its robustness. All of the data are randomly shuffled before 10-fold analysis.

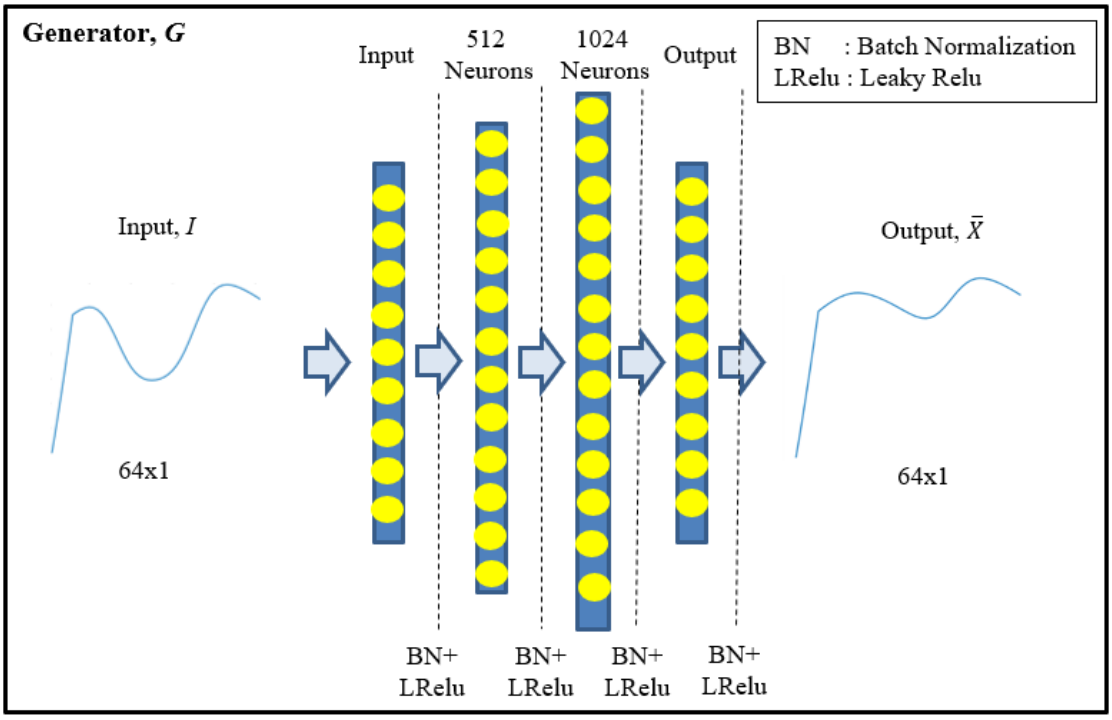

Figure 5.5: The generator accepts AB as input and learns to output normal breathing cycles.

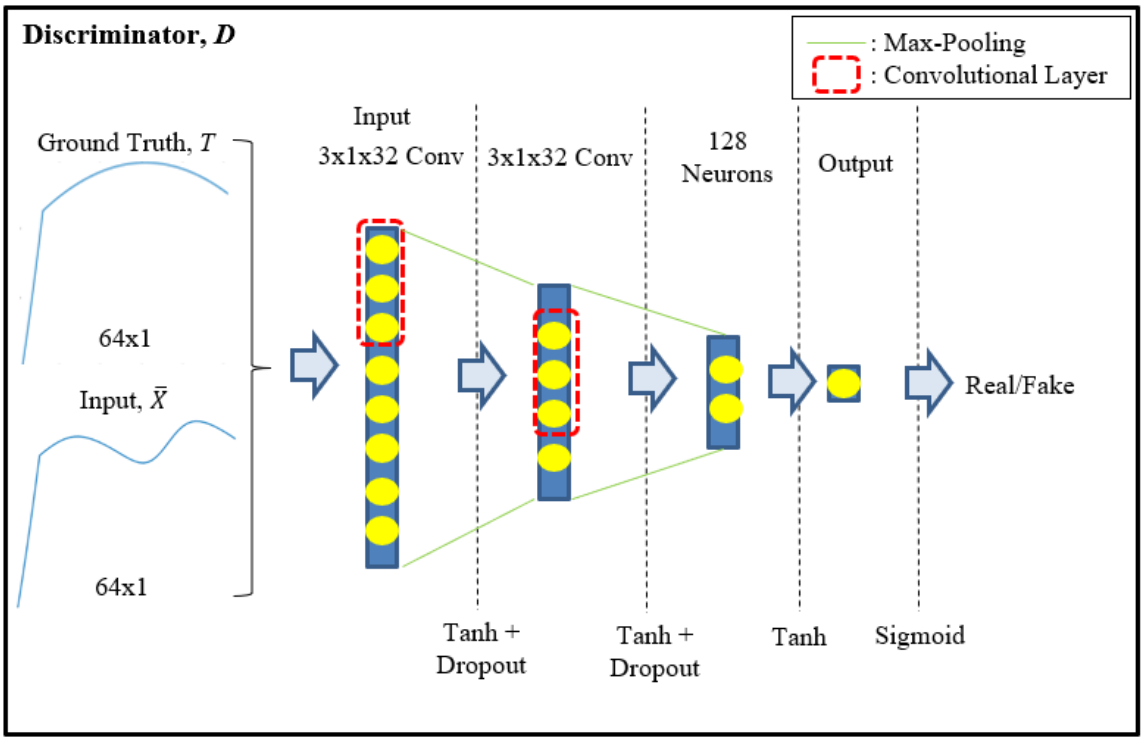

Figure 5.6: The discriminator compares the difference between the ground truth and synthetic output from generator and determines if they look alike.

Table 5.2: Architecture of *ABReGAN*.

Discriminator		**Generator**	
Layer	**Settings /Name**	**Layer**	**Settings /Name**
Input	64x1	Input	64x1
Convolution	3x1x32	Activation	LeakyReLU (β = 0.2)
Activation	Tanh	Normalization	Batch Normalization (α = 0.8)
Dropout	0.5	Hidden Layer	512 neurons
Max Pooling	2	Activation	LeakyReLU (β = 0.2)
Convolution	3x1x32	Normalization	Batch Normalization (α = 0.8)
Activation	Tanh	Hidden Layer	1024 neurons
Dropout	0.5	Activation	LeakyReLU (β = 0.2)
Max Pooling	2	Normalization	Batch Normalization (α = 0.8)
Hidden layer	128 neurons	Output layer	64x1
Activation	Tanh	Activation	LeakyReLU (β = 0.2)
Output layer	1	Normalization	Batch Normalization (α = 0.8)
Activation	Sigmoid		

Generator, G

The G in *ABReGAN* contains 4 hidden layers, 4 activation layers followed by batch normalization layer and it is compiled with binary cross-entropy as loss function as expressed in Equation (4.6) and shown in Figure 5.5. Batch normalization is implemented on the outputs of all fully connected layers to reduce overfitting (Ioffe & Szegedy, 2015). Batch-Normalization is a technique to improve stability and computational speed during training by normalizing and scaling the inputs, i_i for machine learning. Batch-Normalization normalized output $\hat{x}$ can be expressed as follows:

$$Minibatch\ Input\ Mean, \mu_b = \frac{1}{\eta}\sum_{i=1}^{\eta} i_i \tag{5.8}$$

$$Minibatch\ Input\ Variance, {\sigma_b}^2 = \frac{1}{\eta}\sum_{i=1}^{\eta}(i_i - \mu_b)^2 \tag{5.9}$$

$$Running\ Mean, \mu_{mov} = \alpha(\mu_{mov}) + (1-\alpha)\mu_b \tag{5.10}$$

$$Running\ Variance, {\sigma_{mov}}^2 = \alpha({\sigma_{mov}}^2) + (1-\alpha){\sigma_b}^2 \tag{5.11}$$

$$Normalize\ Output, \hat{x} = \frac{i_i - \mu_{mov}}{\sqrt{{\sigma_{mov}}^2 + 0.001}} \tag{5.12}$$

where η is the total number of mini-batch and α is the momentum value.

In addition, LeakyReLU (Maas, 2013) activation layer is applied in G to improve GAN stability and training (Radford et al., 2015; Xu et al., 2015). Unlike ReLU activation layer Equation (4.2), LeakyReLU allows certain range of negative values as inputs, i; thus, enables the model to learn additional information. LeakyReLU can be denoted as follows:

$$f(i) = \max(i, \gamma(i)) \tag{5.13}$$

where γ is the allowable negative value to pass through the LeakyReLU.

Discriminator, D

Figure 5.6 shows the architecture of discriminator. The D architecture consists 2 convolutional layers with 32 filters, 4 activation layers, 2 hidden layers and 2 pooling layers and it is compiled using binary cross-entropy as loss function as described in Equation (4.6). While the convolution layer, max pooling and hidden layer operations in D are denoted in Equation (4.1), Equation (4.4), and Equation (4.5) respectively; Tanh and sigmoid activation functions (Nwankpa et al., 2018) in D can be expressed as follows:

$$Tanh, \sigma(i) = \frac{e^i - e^{-i}}{e^i + e^{-i}} \tag{5.14}$$

$$Sigmoid, \sigma(i) = \frac{1}{1 + e^{-i}} \tag{5.15}$$

where i is the input to the activation function.

Results

Figure 5.7 contains two subplots. The top graph depicts the generator losses, while the bottom graph describes the losses of discriminator for real data (Dis-real) and losses of discriminator for reconstructed data (Dis-fake) during training. It can be observed that, the generator losses were erratic in the earlier runs before stabilized at approximately 1.7 error between iteration 1700 to iteration 6000; The variance of the loss was found to increase beyond iteration 6000. On the other hand, the discriminator attained high fluctuation of losses throughout the entire training process, but they remained stable. Furthermore, the Dis-real and Dis-fake losses stabilized between iteration 1000 to 6000 before Dis-fake losses surpassed Dis-real losses. Despite both Dis-real and Dis-fake losses exhibited high variance; the losses remained stable approximately at 0.7 loss. Table 5.3 shows the results of k-fold validation for reconstructing AB. *ABReGAN* achieved mean MAPE of 26.35% and 189.04% when estimating elastance values and resistance values respectively.

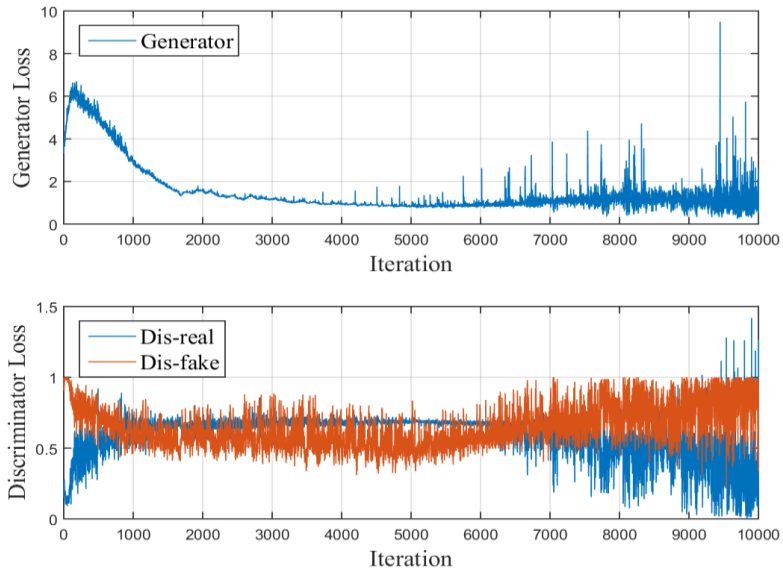

Figure 5.7: Subplots of generator (top) and discriminator (bottom) losses obtained from one of *ABReGAN* model from k-fold analysis.

Table 5.3: Performance of *ABReGAN* obtaining normalized elastance and resistance values in the 10-fold validation analysis.

Respiratory mechanics	MAPE (%)
Elastance, E	26.35 [25.30 – 26.90]
Resistance, R	189.04 [184.28 – 199.77]

Median [25th -75th percentile]

Figure 5.8 top left and right figures show two successfully reconstructed AB as the reconstructed airway pressure intimately resembles the actual airway pressure with minimal error (SSE = 0.04 (Top left), 0.09 (Top right)). The absolute error between E_{rec} and E_{ac} were 0.01 (Top left) and 0.01 (Top right) cmH_2O/L and R_{rec} and R_{ac}were 0.04 (Top left) and 0.07 (Top right) cmH_2Os/L respectively. Figure 5.8 bottom left and right graphs denote examples of failed reconstruction as *ABReGAN* generated normal breathing cycle with huge discrepancy (SSE = 4.452 (Bottom left), 0.88 (Bottom right)). The absolute error between E_{rec} and E_{ac} were 0.46 (Bottom Left) and 0.40 (Bottom Right) cmH_2O/L. Absolute error between $R_{rec}$and R_{ac}were 0.57 (Bottom left) and 0.13 (Bottom Right) cmH_2Os/L.

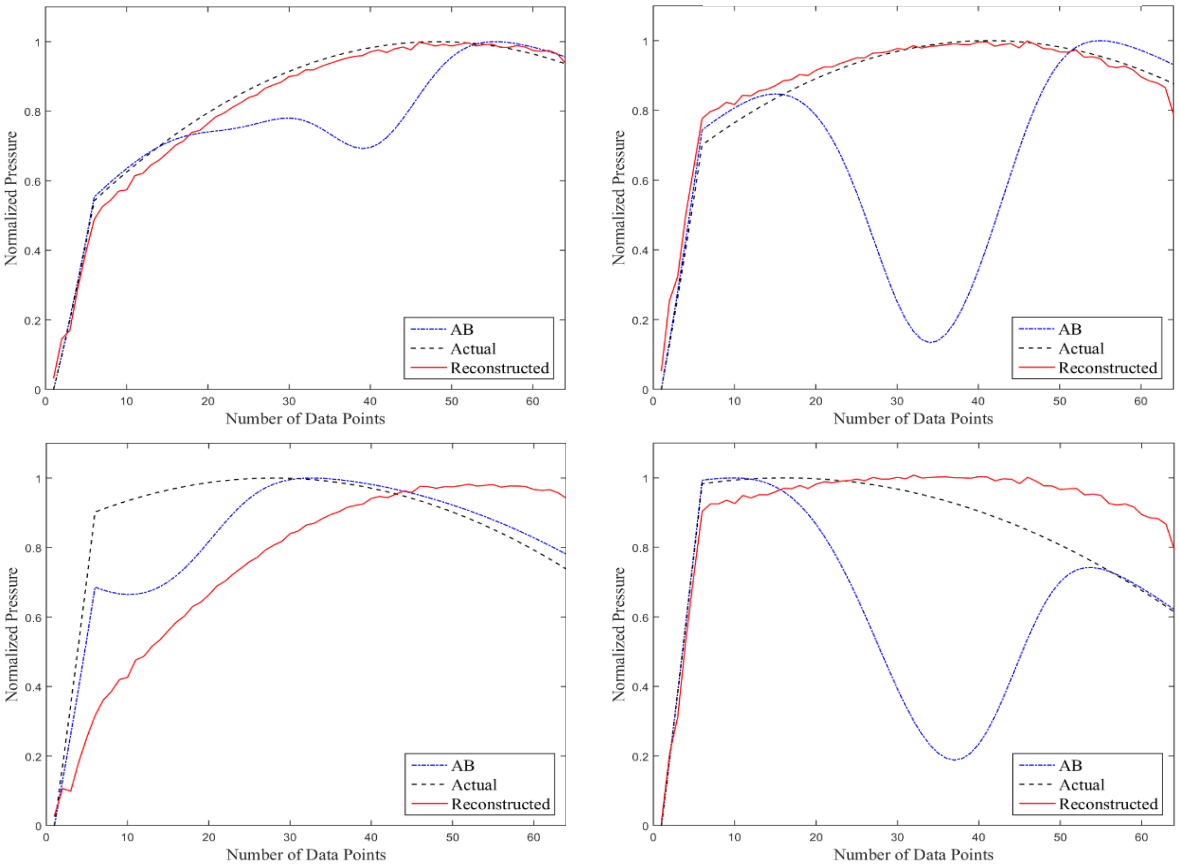

Figure 5.8: Comparison of *ABReGAN* reconstructed airway pressure, target normal airway pressure and AB. Top left and right figures show that the generated airway pressure agrees the features defining a normal breathing cycle but with the presence of minor magnitude of oscillations. Bottom left and right figures show *ABReGAN* fails to reconstruct to actual normal breathing.

Discussion

Non-convergence issues (mode collapse and vanishing gradient) are common during GAN training due to the enormous difficulties in finding the equilibrium between the generator and discriminator (Arjovsky & Bottou, 2017; Salimans et al., 2016; Thanh-Tung & Tran, 2020). Ideally, both neural networks should update proportionally; however, if one of the networks outperforms another, they could repeatedly diminish the progress one another and cause non-convergence. In this study, although *ABReGAN* was trained successfully as the training process of *ABReGAN* was stable and steady as shown in Figure 5.7; the k-fold validation results show that *ABReGAN* attained high error when computing both elastance and resistance. Atrocious elastance and resistance estimation may due to the failure of convergence of *ABReGAN*. However, achieving or maintaining stability in GAN during training is difficult as stopping criteria to GAN training is unclear (Mescheder et al., 2018); therefore, it is not ideal to obtain the least validation error during training and additional studies are required to improve *ABReGAN*. Detail discussion and further analysis on the performance of *ABReGAN* when evaluated with 420,000 datasets continues in Section 5.2.7.

In summary, although *ABReGAN* was able to reconstruct, its effectiveness and capability may be limited due to the error attained in Figure 5.7. Henceforth, we develop another machine learning model, in the effort to investigate alternative approach to better reconstruct AB. The following section describes an alternative machine learning algorithm to reconstruct AB to quantify the magnitude of AB.

5.2.6 Convolutional Autoencoder

Autoencoder (Baldi, 2012) is a neural network model that is designed and trained to reproduce its input in an unsupervised manner. An autoencoder model comprises two main

blocks namely encoder and decoder and employs a symmetric structure. The encoder is trained to compress input and extracts critical information into a lower dimensional representation; whereas, the decoder learns to reproduce the input using the given latent representation extracted by encoder.

Convolutional autoencoders (CAE) (Y. Zhang, 2018) are conventional autoencoders composes with convolutional layers in encoder and decoder layers. The efficiency and superiority of the adaptable kernels in convolutional layer helps the crucial spatial information extraction from the inputs; thus, accelerates the convergence process. Similar to autoencoder, CAE comprises an encoder to input features selection and compression as well as a decoder to decompress and reassemble original inputs based on the feature representation. The compressed representation contains essential information from the input representations at the lowest possible reconstruction error; this information is critical for the decoder to reassemble back to its inputs with least distortion. Representation of the feature map from encoder with convolutional layer can be computed using the following equation:

$$Encoder, \bar{Y} = f(x) = \sigma(i * W + b) \tag{5.7}$$

where i is the input, σ is an activation, * denotes convolution operation, W and b represents weight and bias respectively.

On the other hand, the decoder will map the output from encoder, $\bar{Y}$ and perform a convolution operation. The objective of the decoder is to ensure the output from decoder, z will have the exact same features and dimensions with the input. The decoder can be expressed using the following equation:

$$Decoder, z = f'^{(Encoder)} = \sigma(\bar{Y} * W + b) \tag{5.8}$$

where σ is an activation, * denotes the convolution operation, and W and b represents weight and bias.

AB Reconstruction Convolutional Autoencoder (ABReCA)

AB Reconstruction Convolutional Autoencoder (*ABReCA*) consists multiple convolutional layers with each layer contains trainable filters and every output from each layer undergoes batch normalization and max pooling operation. Figure 5.9 shows the overall flow of *ABReCA*, where it reconstructs the normal breathing cycle from an AB input. During training phase, the encoder learns to capture and identify the crucial patterns of simulated AB to assemble them into a compressed representation; while, the decoder expands the compressed representation and reconstructs it into a normal breathing. Although the concept to reconstruct AB using CAE has been realised, the optimal architecture of *ABReCA* is still undetermined. Henceforth, we perform k-fold cross validation to investigate the performance of *ABReCA* trained with different number of convolution filters, *F*, and encoder/decoder layer, *P*.

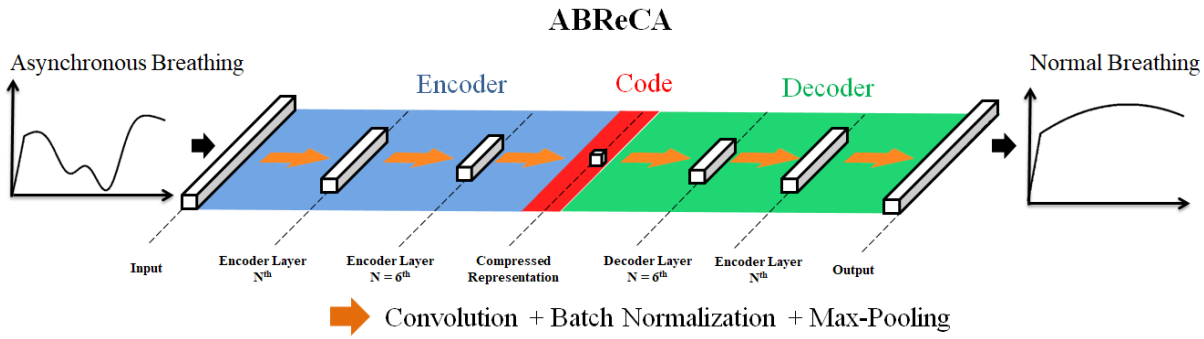

Figure 5.9: A simple flow diagram describing the flow of *ABReCA* to reconstruct an AB.

ABReCA Architecture and performance evaluation

In this study, we perform 10-fold cross validation to study the performance of *ABReCA* trained with various combination of F = 1,2,4,8,16,32 and P = 1,2,3,4,5,6. Figure 5.10 shows the flowchart of the 10-fold cross-validation. For every k-folds iteration, the training set consists 378,000 training datasets, whereas the testing set comprises 42,000 unique breathing cycles which then be used for performance evaluation. 30% of the training data are reserved for validation to select the model with lowest validation error during training to be evaluated using performance metrics (MAPE). The behaviour of *ABReCA* when trained with different combinations F and P is recorded for investigation. Each *ABReCA* model is trained for 100 epochs with 128 batch size and assessed by computing the difference between elastance of reconstructed AB, E_{rec} and, resistance of reconstructed AB, R_{rec} with the, elastance of actual breathing cycle E_{ac} and, resistance of actual breathing cycle R_{ac} using the performance metrics MAPE.

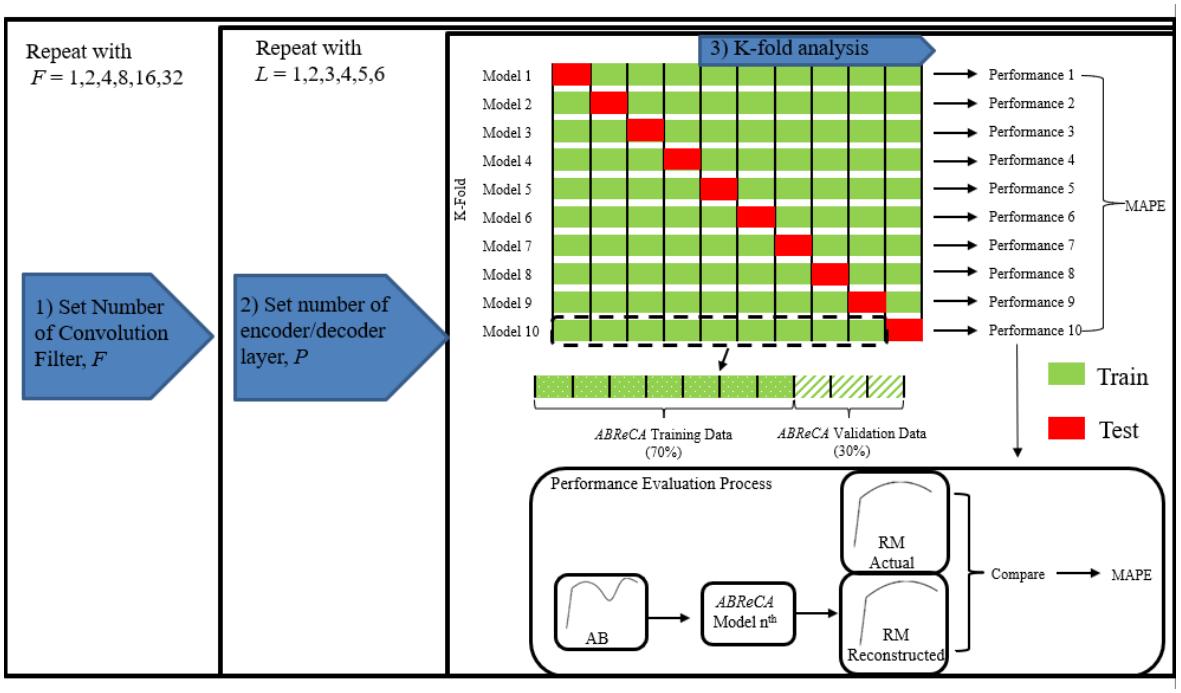

Figure 5.10: 420,000 dataset were segregated into 10 different training/validation sets when k is folded 10 times. The training dataset were shuffled prior segregation.

The main architecture of *ABReCA* consists convolutional layer, batch normalization layer, activation function (ReLU), pooling (Max-pooling) and unpooling layers. Several operations are repeated from other models for example: the operation of convolutional layer is described in Equation (4.1), ReLU activation function is denoted in Equation (4.2), Max-pooling function is described in Equation (4.4); whereas. batch normalization is mentioned in Equation (5.8 – 5.12) in this chapter. Unpooling layer is newly introduced in *ABReCA* with the purpose to revert the downsampled outcomes of pooling operation (M. D. Zeiler & Fergus, 2014). Essentially, unpooling operation upsamples small inputs into a larger size by repeating each temporal step with size, m times along the axis. Figure 5.11 depicts the unpooling operation.

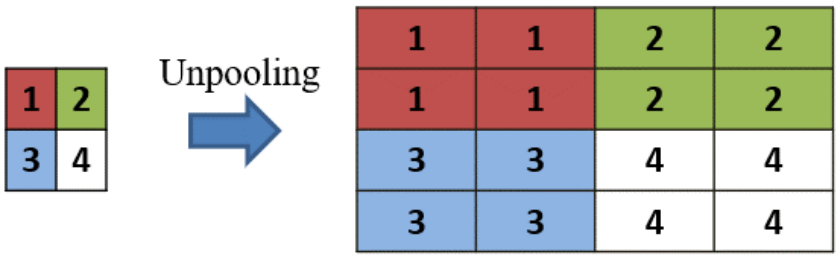

Figure 5.11: Unpooling operation doubles the size of the input by repeating the values in each cell.

On the other hand, we apply Mean Squared Error, MSE as the training loss function to evaluate the performance for parameters adjustment during training (Janocha & Czarnecki, 2017). MSE computes the loss, L by calculating the difference of the ground truth, T and reconstructed output from decoder, z over batch size, η. MSE can be expressed as follows:

$$MSE, L = \frac{1}{\eta}\sum_{i=1}^{\eta}(T - z)^2 \tag{5.9}$$

Similarly, *ABReCA* uses Adaptive moment estimation, *Adam* as optimizer to compute the objective function (Kingma & Ba, 2017). The hyper-parameters of *Adam* are set to default initially with learning rate equals to 0.001 as well as momentum term $ß_1$ and $ß_2$ equals to 0.9 and 0.999 respectively. Equations to express *Adam* are described in Equation 4.10 to 4.12.

Results

We performed 10-Fold analysis to determine the optimal number of encoder/decoder layers, *P* and the amount of convolution filter, *F*, applied to every convolutional layer. Tables 5.4 and 5.5 show that as the *F* and *P* increase, the error to obtain accurate RM reduces. However, when *ABReCA* was trained with only 1 CNN filter, its performance deteriorated regardless of the number of encoder and decoder layers. Figure 5.12 shows the relation between the number of CNN filters and number of encoder/decoder layers. The results show that when *ABReCA* was trained with 32 CNN filters and 6 encoder and decoder layers, *ABReCA* reached its peak performance by achieving an average of MAPE median 0.9% resistance and 7.8% elastance mean error when trained. In contrast, when trained with 3 encoder/decoder layers and 1 convolution filter, *ABReCA* achieved median MAPE of 138.1% and 21.3% when computing resistance and elastance values respectively. In fact, *ABReCA* achieves median error of more than 80% when trained with 1 CNN filter regardless of number of encoder/decoder layers. Nonetheless, *ABReCA* achieved relatively better performance in computing elastance values than resistance values as shown in Tables 5.4 and 5.5.

Table 5.4: Mean absolute percentage error for elastance from K-Fold analysis.

Number of Filters	Number of Encoder/Decoder Layers (MAPE Elastance %)					
	1	2	3	4	5	6
1	11.2 [10.9-28.4]	14.5 [11.8-82.4]	21.3 [11.4-24.6]	10.5 [10.0-15.5]	11.5 [8.7-17.0]	14.4 [10.9-17.7]
2	11.8 [10.0-18.7]	10.1 [8.7-11.3]	8.9 [7.1-14.8]	8.6 [5.2-12.5]	5.2 [3.2-14.6]	6.1 [4.8-11.1]
4	9.0 [8.2-12.9]	9.5 [7.4-18.0]	5.4 [3.8-6.9]	5.3 [4.2-6.3]	2.3 [1.5-4.2]	2.4 [1.9-2.7]
8	12.9 [11.0-14.7]	7.6 [5.7-15.8]	3.7 [3.3-5.5]	3.5 [2.2-4.2]	1.6 [1.0-2.3]	1.3 [0.8-1.7]
16	11.7 [9.7-19.5]	6.6 [4.5-11.2]	4.6 [3.4-5.2]	2.9 [2.7-3.7]	0.9 [0.6-1.2]	1.5 [1.2-1.7]
32	15.7 [8.2-16.4]	7.2 [6.5-11.5]	5.0 [3.5-11.0]	2.3 [2.2-3.0]	1.1 [0.8-1.6]	0.9 [0.6-1.3]

Median [Interquartile Range]

Table 5.5: Mean absolute percentage error for resistance from K-Fold analysis

Number of Filters	Number of Encoder/Decoder Layers (MAPE Resistance %)					
	1	2	3	4	5	6
1	85.8 [78.6-123.4]	126.6 [100.0-191.9]	138.1 [94.9-168.9]	125.2 [91.9-137.0]	119.4 [88.7-156.1]	131.2 [124.5-152.7]
2	79.1 [57.4-91.1]	78.5 [64.1-102.4]	73.3 [45.6-100.0]	60.2 [47.3-97.4]	48.3 [24.7-130.0]	40.8 [25.5-119.5]
4	74.2 [57.6-78.3]	58.0 [49.0-119.8]	28.3 [27.4-48.2]	31.5 [25.6-34.1]	17.2 [10.0-28.0]	20.6 [15.1-21.9]
8	70.6 [61.9-111.4]	39.8 [30.6-79.5]	28.4 [18.4-41.7]	20.4 [16.4-28.1]	9.8 [8.6-13.1]	12.1 [6.0-14.0]
16	78.6 [57.3-88.1]	39.6 [28.6-71.2]	20.9 [16.7-26.2]	20.4 [15.1-24.9]	8.2 [4.8-9.6]	9.1 [6.5-11.6]
32	68.4 [57.2-110.9]	33.4 [29.9-39.3]	19.6 [16.7-27.2]	22.3 [15.6-29.6]	8.7 [7.0-12.0]	7.8 [6.0-10.3]

Median [Interquartile Range]

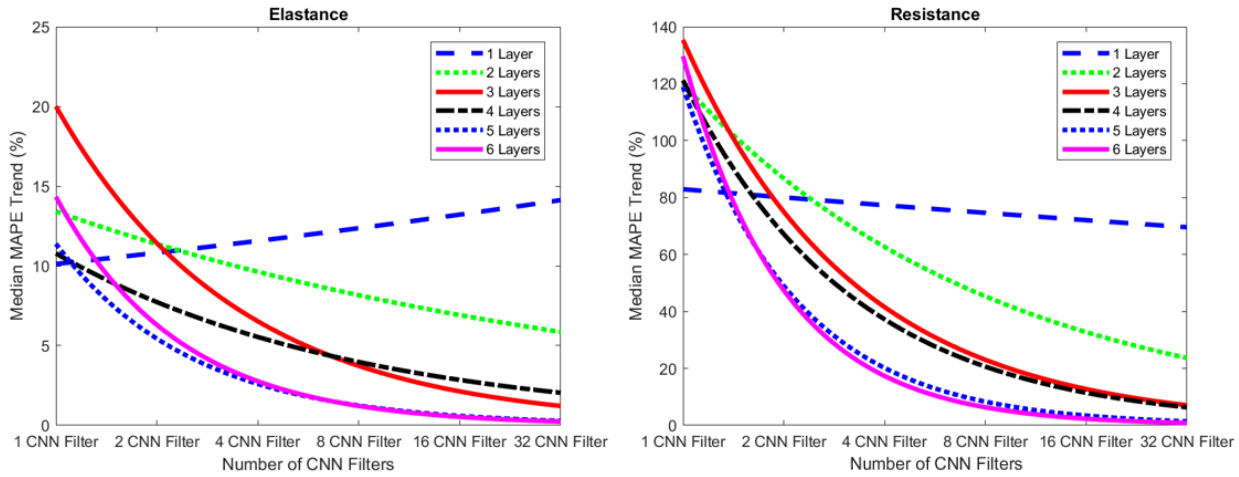

Figure 5.12: The trends show the performance of *ABReCA* computing correct elastance (left) and resistance (right) when trained with different filter numbers and number of encoder/decoder layers

Figure 5.13 shows the optimal configuration of *ABReCA* applied in this study. The encoder comprises 6 convolutional layers, followed by batch normalization layers and max pooling operation. By doing so, the inputs are compressed down to tensor size of 1 by 32 convolution filters. The decoder mirrors the architecture of encoder by repeatedly upsampling, batch normalizing and applying convolutional layers to restore to the size of inputs.

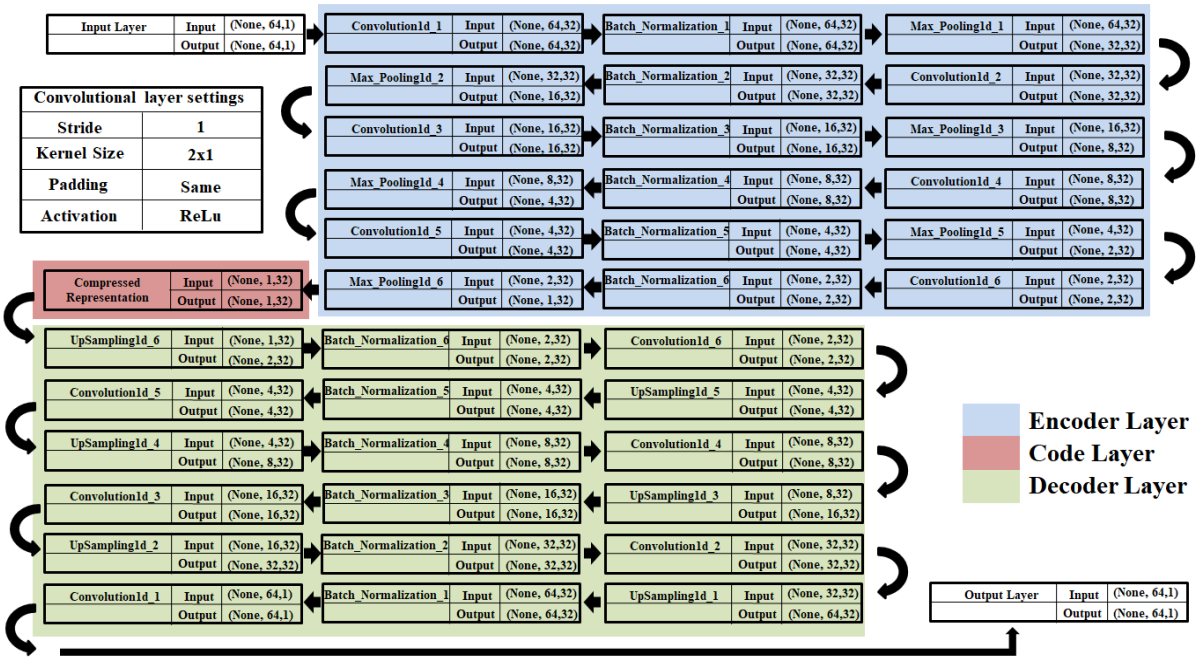

Figure 5.13: The optimal architecture of *ABReCA* comprises 32 convolution filters and 6 layers depth of encoder and decoder layers. Each convolutional layer is followed by batch normalization and max-pooling.

Figure 5.14 shows the training and validation loss when *ABReCA* was trained with optimal sets of encoder/decoder layer and convolutional filters. Among 10 k-fold analysis, the model achieves the lowest validation error with 1.7e-05 at epoch 85 during training and it is selected for further analysis in the subsequent performance comparison.

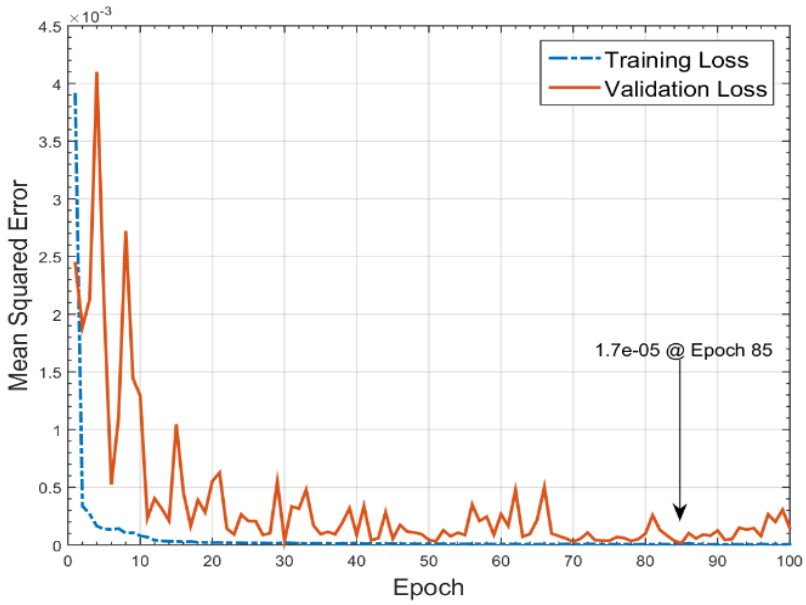

Figure 5.14: Training and validation loss of *ABReCA* with 32 CNN filters and 6 encoder/decoder layers. The model achieving lowest validation error (1.7e-05 MSE) during training was selected.

Figure 5.15 shows four examples of *ABReCA* reconstructed AB. Top left and right figures show two successfully reconstructed AB as the reconstructed airway pressure intimately resembles the actual airway pressure with minimal error (SSE = 0.001 (Top left), 0.002 (Top right)). The absolute error between E_{rec} and E_{ac} were 0.003 (Top left) and 0.009 (Top right) cmH_2O/L and R_{rec} and R_{ac} were 0.096 (Top left) and 0.001 (Top right) cmH_2Os/L respectively. The bottom left and right plots show examples of failed reconstructions. As shown in the Figure 5.15 *ABReCA* reconstructed normal breathing cycle with errors (SSE = 0.074 (Bottom left) and 0.517 (Bottom right)). The absolute error between E_{rec} and E_{ac} were 0.102 (Bottom left) and 0.259 (Bottom right) cmH_2O/L. Absolute error between R_{rec} and R_{ac} were 0.068 (Bottom left) and 0.185 (Bottom right) cmH_2Os/L.

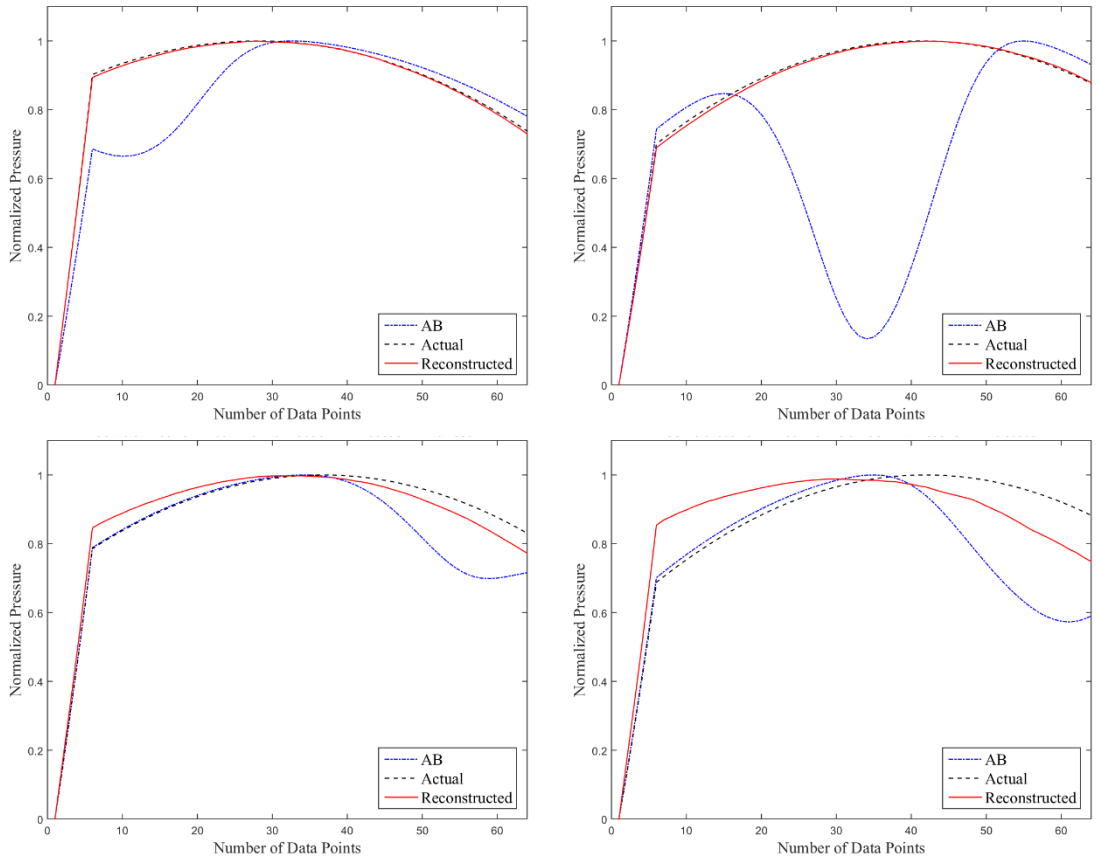

Figure 5.15: Three different types of AB were selected for testing. Top left and right figures show that the *ABReCA* reconstructs airway pressure with minor distortion. Bottom left and right figures show *ABReCA* fails to reconstruct to actual normal breathing.

Discussion

Tables 5.4 and 5.5 and the trend shown in Figure 5.12 show that, when *ABReCA* was trained with only 1 CNN filter, its performance deteriorated regardless of the number of encoder and decoder layers. This shows that, the model requires greater number of trainable parameters in convolution filter to capture and learn the intricate and erratic AB features to restore AB to asynchrony free breathing cycle (Basha et al., 2020; Krizhevsky et al., 2017). Furthermore, the difference in median MAPE was almost similar when *ABReCA* was trained with F = 16 and 32; despite the increment in encoder and decoder layers. This indicates that the *ABReCA* with convolution filter size of 16 is sufficient to AB reconstruction. Nonetheless, the model trained with F = 32 and P = 6 is still preferred due its achievement in lowest median MAPE among all.

It is hypothesized that, *ABReCA* is able to achieve low error is due to the encoder repeatedly removing trivial AB features in the inputs; ultimately, those essential AB features are compressed into one size-tensor through convolutional layers and max-pooling operation. Similarly, the decoder repeatedly expands and learns the information in compressed representation to reconstruct AB through convolutional layers. Moreover, introducing multiple convolutional layers grows the number of trainable neurons; thus, allowing *ABReCA* to 'memorize' large quantity of AB intrinsic properties (O'Shea & Nash, 2015) and enables better AB reconstruction.

The training and validation losses shown in Figure 5.14 suggest that *ABReCA*'s learning and reconstruction capability increases during training; despite, the validation loss fluctuates even though *ABReCA* has reached its maximum performance. This fluctuation is likely due to variability of the learning rate regularized by *Adam* during training. Although high variance in validation loss could indicate model overfitting and introducing dropout could improve the performance (Gal & Ghahramani, 2015; Srivastava et al., 2014); but this can be ignored as the validation loss is low with an average *MSE* of 3.6e-04.

Figure 5.15 shows four reconstructed breathing cycles using *ABReCA*. It is observed that *ABReCA* reconstructed with errors when tested with breathing cycles with late asynchrony triggering. This could indicate that *ABReCA* is not proficient in reconstructing breathing cycles with late patient effort triggering; this could possibly due to overfitting as mentioned earlier. Overall, the error obtained by *ABReCA* is relatively lower as compared with *ABReGAN*. Besides that, *ABReCA* has slight advantage in reconstructing pressure waveform, as the oscillation is absent from the *ABReCA* reconstructed breathing cycles compared with

the *ABReGAN* generated breathing cycles. Thus, error difference between reconstructed and actual normal breathing cycle by *ABReCA* is minimal.

5.2.7 Model comparison

This section presents a comparison study between the developed machine learning models (*ABReGAN* and *ABReCA*) with *IPR*. To ensure impartiality and unbiased comparison, all the models are tested and evaluated using same validation dataset using four different performance metrics namely SSE, APE, MSE and MAPE. The validation dataset comprises entire 420,000 simulated AB that were generated for machine learning training.

Results

Figure 5.16 and Table 5.6 shows that *ABReCA* achieved the best performance followed by *IPR* then *ABReGAN*. It successfully achieved MSE with only 5e-6 and 1e-5 or MAPE 0.2 % and 2.64 % when estimating elastance and resistance respectively. This shows that *ABReCA* performs relatively better when estimating elastance values than estimating resistance values. On the other hand, *ABReGAN* attained the highest error among all with MAPE of 28.91 % and 224.47 % or MSE of 7.80 and 1.44 when computing elastance and resistance respectively. *IPR* achieved MAPE of 7.34% and 19.59% as well as MSE of 0.02 and 0.005 when computing elastance and resistance respectively.

Table 5.6: Performance of each type of models tested with 420,000 validation datasets.

Model	RM	SSE	MSE	APE (%)	MAPE (%)
ABReGAN	*E*	0.013 [0.003 – 0.042]	7.80	7.38 [3.45 – 13.75]	28.91%
	R	0.005 [0.001 – 0.015]	1.44	54.35 [19.45 – 90.11]	224.47%
ABReCA	*E*	2e-6 [5e-7 – 6e-6]	5e-6	0.18 [0.08 – 0.26]	0.20 %
	R	8e-6 [2e-6 – 2e-5]	1e-5	1.45 [0.56 – 2.37]	2.64 %
IPR	*E*	0.013 [0.003 – 0.030]	0.02	6.56 [2.99 – 10.98]	7.34%
	R	0.003 [0.001 – 0.006]	0.005	11.35 [6.13 – 21.61]	19.59%

RM – Respiratory mechanics; *E* – Elastance; *R* – Resistance

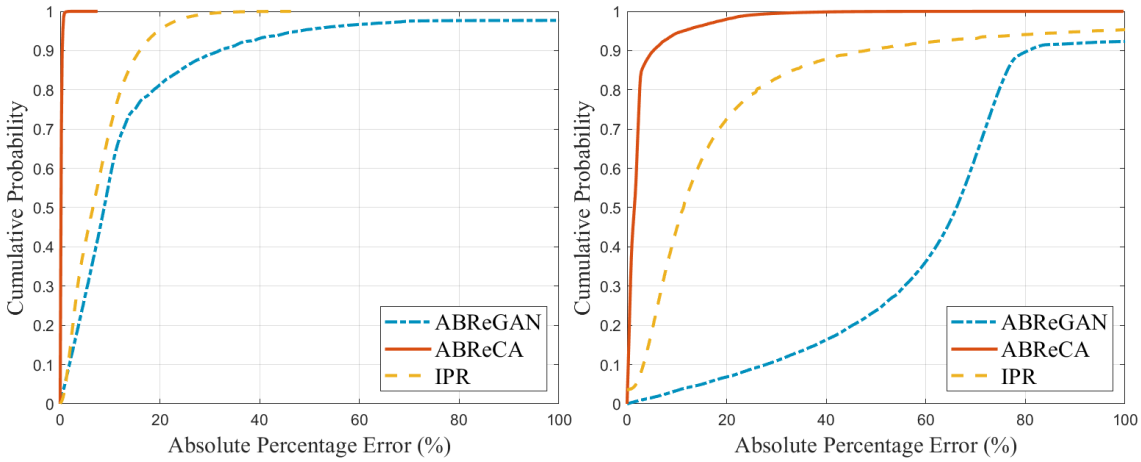

Figure 5.16: Empirical cumulative distribution plot of resistance (Left) and elastance (Right) APE of different models when tested with 420,000 of unique simulated AB.

Discussion

Figure 5.16 shows the empirical cumulative distribution function of APE attained by *ABReGAN*, *ABReCA* and *IPR* to compute elastance (Left figure) and resistance (Right figure) when validated with all 420,000 simulated data. Despite Figure 5.8 shows that *ABReGAN* is able to reconstruct AB with minimal error; the inconsistent performance and high error rate as shown in Figure 5.16 may hinder the implementation of *ABReGAN* at clinical bedside. Furthermore, although *ABReGAN* was trained and validated using the exact dataset as *ABReCA*; the former model achieved lowest accuracy in RM estimation. It is speculated that poor architecture impedes the efficiency and ability of *ABReGAN* in extracting features; thus, modifying architecture or optimization approach may potentially help in performance improvement (Arjovsky et al., 2017; Bi et al., 2020).

Although the results suggested that *ABReCA* is able to eliminate patient-effort via reconstructing AB with least distortion; the estimation of resistance value by *ABReCA* is relatively lower when estimating the elastance value. This is mainly due to the poor reconstruction of AB with late patient effort triggering as shown in Figure 5.15. One of the reasons behind this deterioration of performance is likely related to model fitting during

training as mentioned earlier. Similarly, introducing dropout layer could potentially overcome model overfitting or the lack of training data with different asynchrony patterns and features. Nonetheless, further investigation is necessary to improve the performance of *ABReCA*.

Figure 5.16 shows that *IPR* is able to capture and estimate the true underlying respiratory mechanics in some cases; however, the performance of *IPR* deteriorates as the *IPR* fails to estimate both elastance and resistance values. It is speculated that timing of patient effort occurrence in asynchronous breathing cycle may affect the performance of *IPR*; thus, impeding the performance of *IPR* to restore AB to normal breathing cycle. In this case, machine learning eliminates the needs to explicitly design mathematical models to cater every possible AB shapes and sizes.

Overall, both *ABReGAN* and *ABReCA* are able to perform pressure reconstruction when given AB as input. In the following section, we investigate the performance of *ABReCA* when tested using clinical data randomly selected from the CARE trial recruited patients.

5.2.8 Magnitude of Asynchrony

The difference in area under the curve of *ABReCA* reconstructed airway pressure and original airway pressure can be defined as the magnitude of AB. Figure 5.17 shows examples of AB magnitude, M_{asyn}. The magnitude of AB can be quantified using Equation 5.10:

$$Asynchrony\ Magnitude, M_{asyn} = \frac{\left|AUC_{Rec} - AUC_{Asyn}\right|}{AUC_{Rec}} \times 100\% \quad (5.10)$$

where M_{asyn} denotes the magnitude of asynchrony for each breathing cycle. AUC_{Rec} is the area under the curve of the reconstructed breathing cycle and AUC_{Asyn} is the area under the curve of the asynchrony airway pressure. The computed magnitude of asynchrony can be

used as a metric to measure the degree of MV controlled breath is affected patient-specific AB effort (Kannangara et al., 2016).

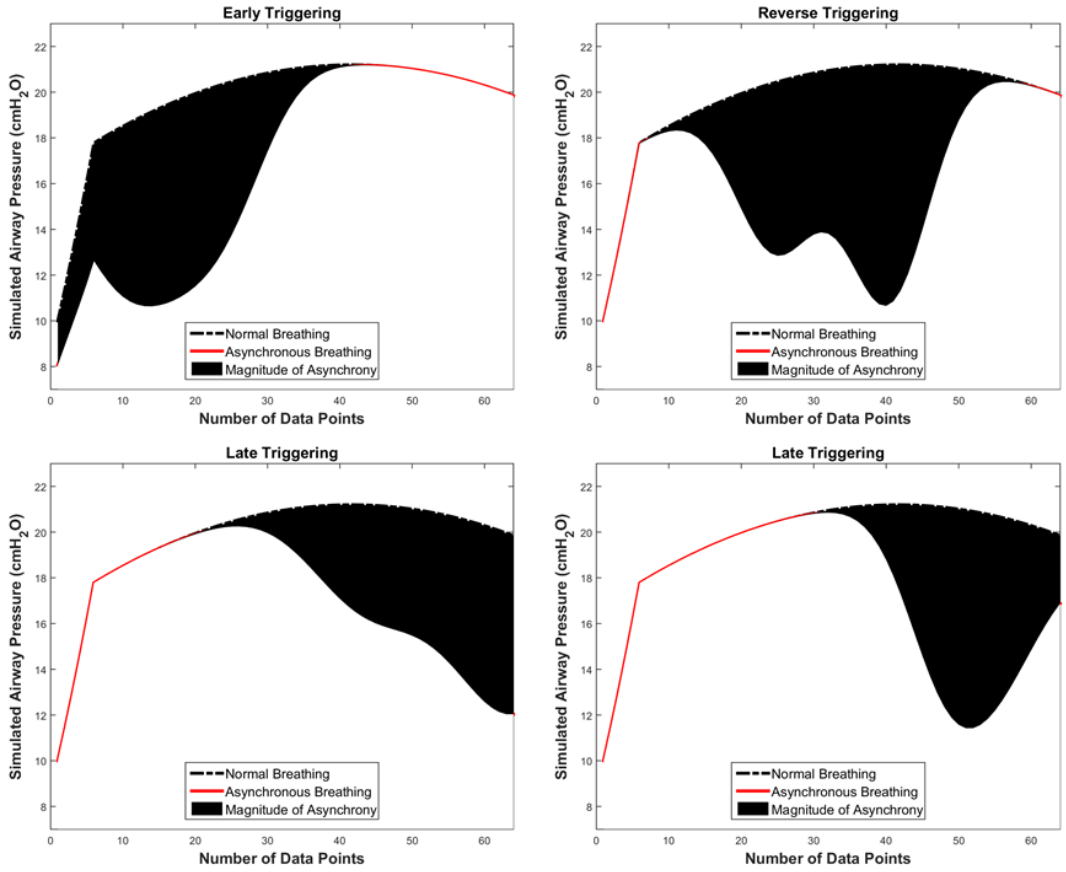

Figure 5.17: Examples of 4 different AB with different M_{asyn}. 16.3% and 20.1% (Top Left to right) 13.5% and 13.9% (Bottom Left to right). Difference in area under the curve of breathing cycles indicates the magnitude of the asynchrony.

Results and discussion

Figure 5.18 shows examples of 6 randomly chosen and different breathing cycles with asynchrony experienced by CARE trial recruited patient reconstructed using *ABReCA*. The top four figures denote successful reconstructions whereas the bottom left and right denote failed reconstructions.

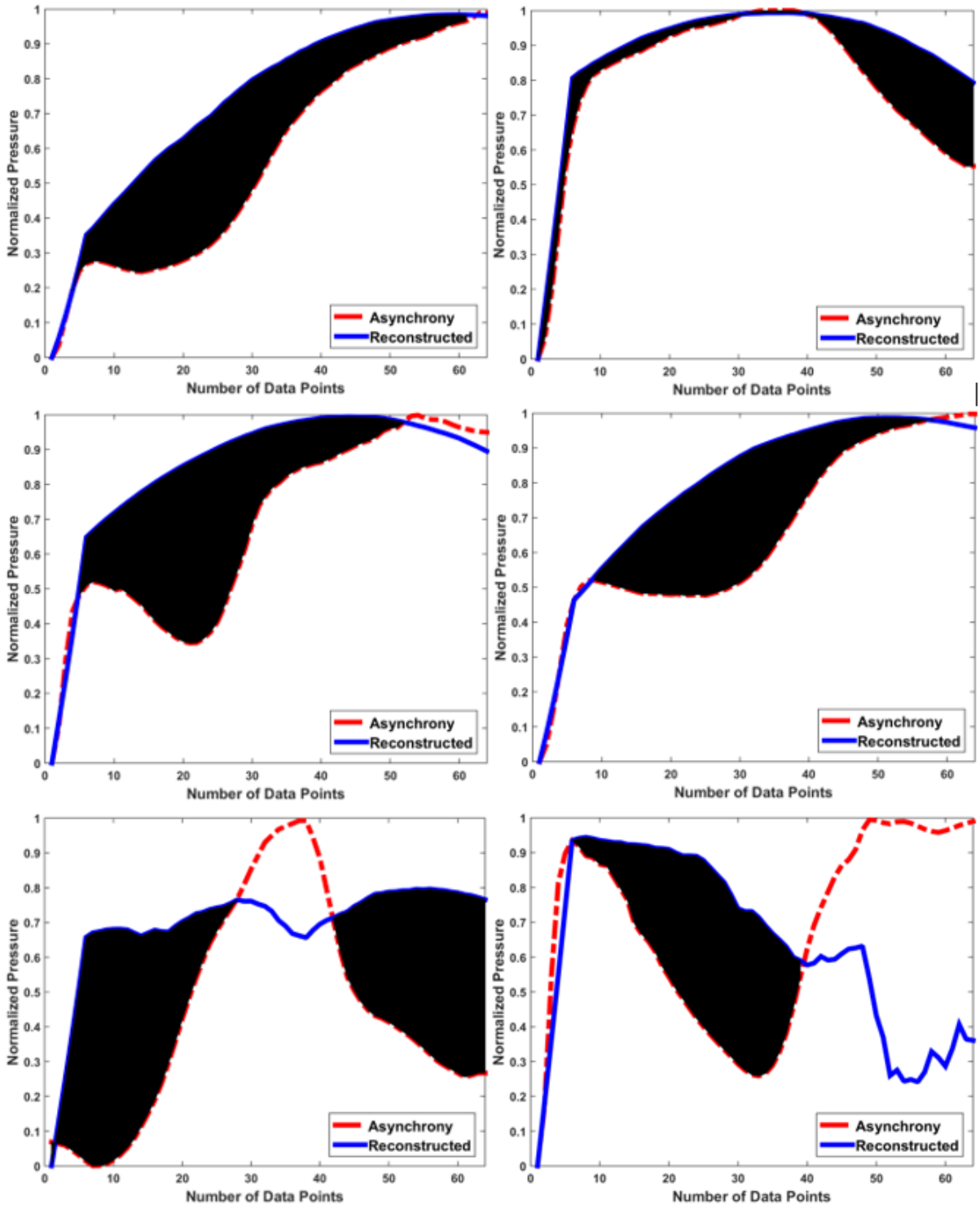

Figure 5.18: 6 different AB experienced by patient with different M_{asyn}. 23.1% and 9.1% (Top Left to right), 20.2% and 17.4% (Middle Left to right) and 36.1% and 6.46% (Bottom Left to right). Magnitude of asynchrony is defined as the difference in area under the curve of reconstructed breathing cycles with original asynchrony breathing.

Results from Figure 5.18 show that *ABReCA* is capable of reconstructing any types of AB collected from actual patient despite *ABReCA* was trained with simulated data. The results

show that *ABReCA* captures the shape of ABs and reconstructs them with 'reasonable' normal shape and pattern. It is however, imperfect reconstruction may occur due to the lack of actual patient data inclusion into training data. As patients breathing pattern are variable, it may be difficult to reconstruct AB into a erfect normal breathing. Hence, additional simulated data is required for training to improve the performance (Ying, 2019), but this may also lead to unwanted overfitting due to similar feature distribution and variation (Vezhnevets & Barinova, 2007). Therefore, more studies are required to generate quality dataset or determine optimal settings to improve *ABReCA*'s performance.

5.3 Limitations

The presence of patient effort muscular effort may alter the airway pressure or/and flow waveform profile; however, *ABReCA* method has only demonstrated the ability to reconstruct AB present in pressure waveform. Hence, further investigations are required to develop a model to reconstruct and quantify the magnitude of asynchronies in airway flow profile. Besides that, the simulated data might be insufficient to train a robust model as the model was only trained and validated with total of 420,000 simulated breathing cycles. Furthermore, patient effort may occur in other ventilation mode, but, *ABReCA* is only developed using VC simulated data. Hence, future work should extend the number of simulated data to generate every possible shapes and aspects of AB experienced by patients as well as breathing data from various ventilator modes to improve the performance of *ABReCA*. The respiratory mechanics of respiratory failure patients maybe altered (Chiew et al., 2011; Chiew et al., 2015). Hence, covering additional range of respiratory mechanics of the patients will allow ABReCA to capture AB in different patient-specific respiratory mechanics; potentially perform a better and more accurate reconstruction. However, setting a higher range of respiratory mechanics and smaller resolution (resolution =1) will also include significant

additional data. It is estimated that more than 5 million breathing cycles can be generated if using the settings in Table 5.7.

Table 5.7: Number of combinations required to cover wider range of respiratory mechanics.

Parameters (Unit)	**Range**	**Resolution**	**Combination(s)**
Elastance, E (cmH_2O/L)	15-60	1	45
Resistance, R (cmH_2Os/L)	1-20	1	20
Magnitude of Pressure, A_1 (cmH_2O)	1-10	1	10
Magnitude of Pressure, A_2 (cmH_2O)	1-10	1	10
Magnitude of Pressure, A_3 (cmH_2O)	1-10	1	10
Centroid of Gaussian curve, μ	-3 – 3	1	6
Width of Gaussian curve, σ	1	Constant	1
	Total combination		5,400,000

5.4 Summary

Two different machine learning models namely *ABReGAN* and *ABReCA* to reconstruct airway pressure waveform to assess the magnitude of AB are presented in this chapter. Despite the models were trained and validated using same dataset, *ABReCA* achieved remarkable low error. Moreover, when tested with clinical data, *ABReCA* is able to reconstruct patient's breathing cycle that agrees the patterns and shapes of a normal breathing cycle. In the next chapter, both CNN model from Chapter 4 and *ABReCA* are deployed simultaneously along with a novel metric system developed to study the correlation between PVI and patient's arterial blood gases.

Chapter 6

Application of CNN and Autoencoder in Patient Asynchronous Index and Asynchrony Magnitude Quantification

In this chapter, we present a preliminary study of implementing the developed models from Chapter 4 and Chapter 5 to establish correlations between the effect of patient-ventilator quality and patient's oxygenation level. This chapter will also investigate into the impact of patient-ventilator interaction quality towards patient's outcomes. As the frequency and the impact of asynchronous breathing occurrence towards patient's condition during treatment is still unknown; it is imperative to measure and evaluate the effect asynchronous breathing.

6.1 Introduction

Partial pressure of oxygen, PaO_2 and partial pressure of carbon dioxide, $PaCO_2$ measurements are part of the routine when evaluation of various disease states such as respiratory failure (Collins et al., 2015). PaO_2 and $PaCO_2$ values can be measured via conducting blood gas analysis, a diagnostic tool generally used to measure the partial pressure of gas in blood by taking blood from circulatory system (capillary, vein, or artery) (Luciano Gattinoni et al., 2018). An arterial blood gas (ABG) is a blood sample extracted from an artery and it provides valuable information such as respiratory and metabolic disorder (Sood et al., 2010). While patient-specific oxygenation and ventilation can be assessed non-invasively through pulse oximetry measurement, ABG measurement is still the standard in ICU (Pretto et al., 2014). Hence, in this study, we investigate the relationship between the quality of PVI and patient's oxygenation status, via ABG measurement to interpret patient's outcome during MV

treatment. A novel metric namely ventilator-interaction index (VI) is introduced to assess the quality of PVI.

Despite multiple cohorts or studies have been conducted in the effort of elucidating the impact and frequency of AB occurrence; the direct effects of AB are still unknown (de Haro et al., 2019). Moreover, most of the studies conducted are limited to a short evaluation period (ranging from several minutes to 24 hours), neglecting the magnitude of AB and taking the patient's oxygenation state into account (Blanch et al., 2015; B. M. De Wit et al., 2009; M. de Wit et al., 2009; Rué et al., 2017; A. W. Thille et al., 2006). For example: Martos-Benítez et al. (Martos-Benítez et al., 2020) conducted a study to identify the relationship of AB with sedation level, hemogasometric and clinical results; they showed that severe AB occurrence (AI>10%) was associated with a lower PaO_2/FiO_2 ratio. However, the study conducted only observed 30 minutes of ventilatory waveforms and neglected asynchronies with patient effort induced such as reverse triggering. Thus, we speculate that computing the frequency of AB occurrence alone may not be sufficient to reflect the actual patient's quality of patient-ventilator interaction. Therefore, we sought to determine if the magnitude of AB plays a role in affecting patient's PaO_2 and $PaCO_2$ level. This is achieved by conducting correlation analysis between ABG with the VI metric of collected ABG at that particular hour.

The following section presents the studies of the impact of quality of PVI towards patient's oxygenation level during MV treatment. In Section 6.2, we describe our approach to investigate the relationship between VI and patient's condition. This is followed, in Section 6.3, the results and discussion of our investigation. The limitations are discussed in Section 6.4 then we conclude our findings in Section 6.5.

6.2 Models and Algorithms

An algorithm is written to incorporate both machine learning models (CNN model – described in Chapter 4) and (*ABReCA* – described in Chapter 5) to study the outcome of mechanically ventilated patients from CARE trial (Chapter 3). Figure 6.1 shows an overview of the algorithm to access AI and VI hourly. The algorithm is divided into two layers: pre-processing layer and analytics layer.

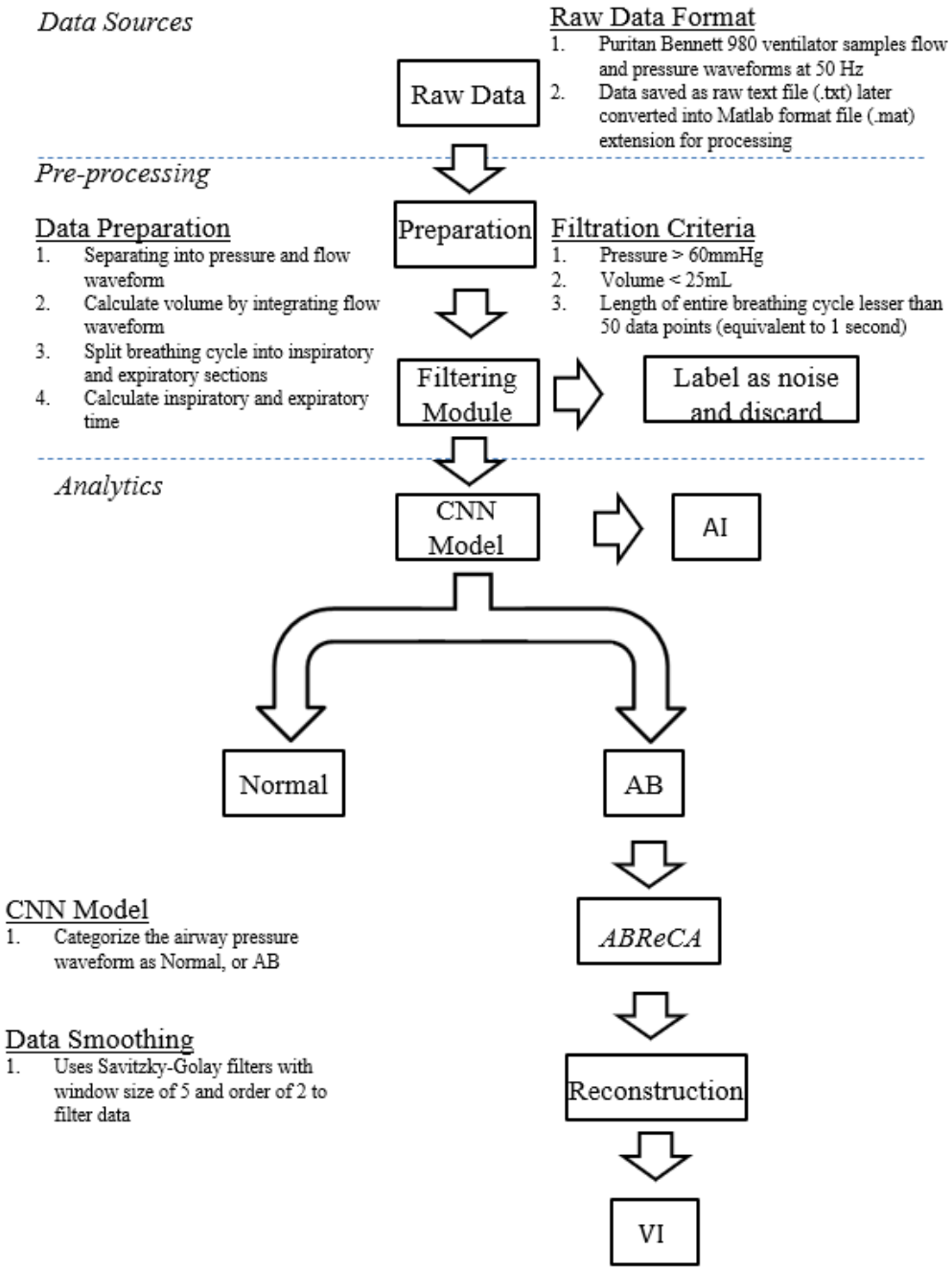

Figure 6.1: Flow diagram of algorithm to assess AI and VI hourly.

In the pre-processing layer, the patient's data are pre-processed prior to any analysis. This layer prepares the raw ventilator waveform by performing several tasks such as data filtering, resampling, and normalizing. Besides that, this layer keeps track of the time of airway breathing occurrence and organize them in hourly manner.

The analytics layer handles the pre-processed airway breathing cycles to analyse the patient's condition via AI and VI computation. This layer comprises of CNN and *ABReCA* models to quantify the AB frequency and magnitude; while, statistical tests are conducted to the outcomes to establish findings using Matlab 2017b (Natick, MA).

Prior the execution of the algorithm, the machine learning models are initialized with the same random seed during model development to ensure equality and reproducibility. The models and algorithms implemented are written and executed in Python 3.6 running on an Intel Core i5 7th generation 3GHz central processing unit, a NVIDIA GTX 1050Ti graphics processing unit and Windows 10 with 32 GB random access memory. The following describes the operation of each layer.

6.2.1 Data Source - Patient data

The CARE trial recruited 24 patients of ventilator waveform data, but 6 patients with total of number of breathing cycles collected lesser than 10,000 data or total hours of data collected lesser than 12 hours are excluded from this study. Patient excluded are P6, P9, P10, P11, P12 and P13. The final study cohort consists of 18 patients (10 males and 8 females).

6.2.2 Pre-processing layer

The raw ventilator pressure waveforms are subjected to noise (Arnal, 2018; Hess, 2005); thus, data pre-processing is required prior to feeding in the developed asynchrony detection and asynchrony magnitude quantification algorithm.

Data Conversion: The first step in data processing involves file conversion. We convert the ventilator raw waveform data in (.txt) file format to floating points formatting (.mat) Matlab file format using Matlab 2017b (Natick, MA). During format conversion, the Matlab algorithm replaces the annotations 'BS' and 'BE' in the raw file with the number 0 in both pressure and flow data to indicate the position of the start or end of a breathing cycle.

Segmentation: After data conversion, the raw data is structured to contain flow in first column and pressure in second column. We segment the raw data into individual breathing cycle via locating the position of 0 present in both pressure and flow waveform. Subsequently, the breathing cycle is articulated into inhalation and expiration section.

Filtering: The presence of noise due to ventilator circuitry leakage or presence of oscillations because of liquid accumulation in the circuit is common during MV (Stauffer et al., 1981). As the collected airway breathing cycles contain copious amount of noises; hence, a Savitzky-Golay filter with window size of 5 and order of 2 is applied to remove the outliers. In addition, airway pressure that contains peak pressure more than 60 cmH_2O, inspiratory elastance lesser than 0, inspiratory volume lesser than 25ml and total number of data points of a breathing cycle lesser than 50 are discarded and marked as noise (Kim et al., 2019).

Normalisation: Normalisation is necessary as the machine learning models were trained with normalised data as well as to reduce variability. Each breathing cycle and inspiration section of airway pressure is normalized to 0 and 1 by using the following equation:

$$normalized\ breathing\ cycle, \hat{x} = \frac{i}{max(i)} - \min(i) \tag{6.1}$$

Where i is the breathing cycle and $\hat{x}$ is the normalized output.

Resampling: Data resampling is necessary to ensure consistency in the length of signal. This can be achieved by introducing new data points within the signal via interpolation method. This method ensures datapoints uniformity to match with model's input configuration, while preserving the valuable information in the signal. CNN model accepts 150 number of data points, while *ABReCA* accepts 64 number of data points.

Arterial Blood Gasses: A total of 114 patient's bed charts with ABG information are collected for analysis.

6.2.3 Analytics layer

The analytics layer is the core layer. This layer applies the developed machine learning models to identify and quantify the magnitude of asynchrony as well as to compute the patient's AI and PVI. The following outlines the processes involved.

Breathing Classification: In Chapter 4, the CNN 1D model which achieved highest accuracy is selected to perform AB detection. CNN 1D accepts the normalised and resampled airway pressure waveform to determine if the breathing cycle is normal or asynchronous. If the breathing cycle is AB, the magnitude of AB will be assessed. The frequency and time of AB

and normal breathing cycles occurrence are recorded. Asynchrony index (AI) is computed hourly using Equation 3.1.

AB magnitude Quantification: If AB is detected, *ABReCA* model will accept the inspiration section of airway pressure waveform and reconstructs it. The magnitude of asynchrony is computed by computing the ratio of the difference in area of reconstructed airway pressure to the under the curve of original AB airway pressure. Magnitude of AB can be calculated using the following equation:

$$Asynchrony\ Magnitude, M_{asyn} = \frac{\left|AUC_{Rec} - AUC_{Asyn}\right|}{AUC_{Rec}} \times 100\% \tag{6.2}$$

where AUC_{Rec} is the area under the curve of reconstructed airway pressure and AUC_{Asyn} is the area under the curve of original AB airway pressure (Chatburn & Mireles-Cabodevila, 2020; Chiew et al., 2018).

Patient-ventilator Interaction Assessment: Equation 6.2 is then extended to describe the quality of patient-ventilator interaction by simply subtracted by 100% as described in Equation 6.3. While Equation 6.4 computes the average ventilator-interaction index (VI) within an hour by summing the VI values within an hour and divided by the total number of breathing cycles occurrence in an hour.

$$Ventilator\text{-}Interaction\ Index, VI = \ 100\% - M_{asyn} \tag{6.3}$$

$$Average\ VI\ in\ a\ hour, VI_{hour} = \frac{\sum VI\ \ in\ An\ Hour}{Total\ Breathing\ cycle\ in\ An\ Hour} \tag{6.4}$$

This novel VI metric enables better reflection of the ‘resistivity’ between the patient with the ventilatory support. In other words, the higher the VI value the better the quality of

interaction between patient and ventilator. The AI and VI for the patients recruited in the CARE trial are calculated and studied.

6.3 Results and Discussion

6.3.1 Relationship between VI and AI

Figure 6.2 shows the VI versus AI plot for this patient cohort. It is observed that the CNN detected AI ranges from 0% to 99.4%, showing inter- and intra-patient variation of AI in different hours. On the other hand, when VI is measured, it is found that VI is not a direct inverse of AI. A Pearson's correlation $R_{VI/AI}$ of -0.49 is found between VI and AI, showing a negative relationship between VI and AI metrics, as expected. In other words, as AI increases the VI decreases, because patient's 'resistivity' towards ventilatory support intensifies as the patient experience more frequent of AB occurrence. But higher AI does not lead to exact lower VI because the VI metric takes into consideration of the magnitude of each asynchronous breathing cycle within an hour.

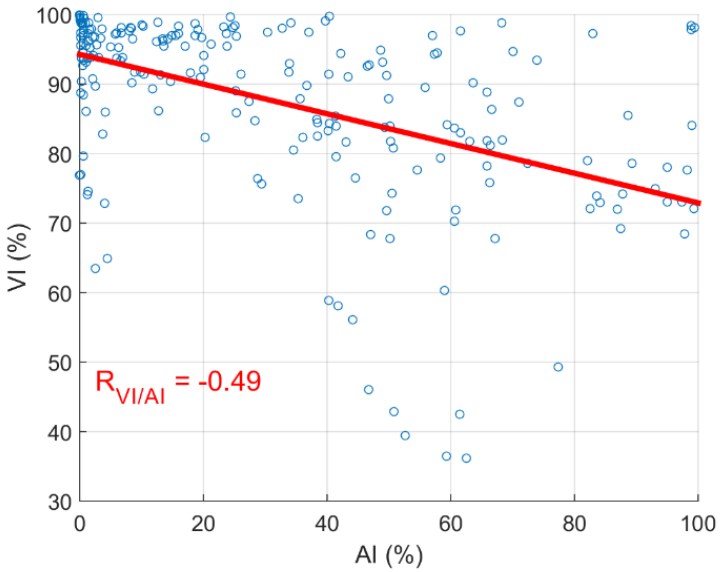

Figure 6.2: Patients' VI versus AI scatter plot.

Figure 6.3 shows plots of 10 consecutive breathing cycles comprising both normal and asynchronous breathing cycles experienced by Patients 1 (Top), 24 (Middle) and 21 (Bottom)

respectively. A red background in a breathing cycle shows a breathing cycle classified as asynchrony by the CNN algorithm whereas green background indicates a normal breathing cycle. The black coloured region is the airway profile reconstruction performed by the *ABReCA* model for quantification of the VI.

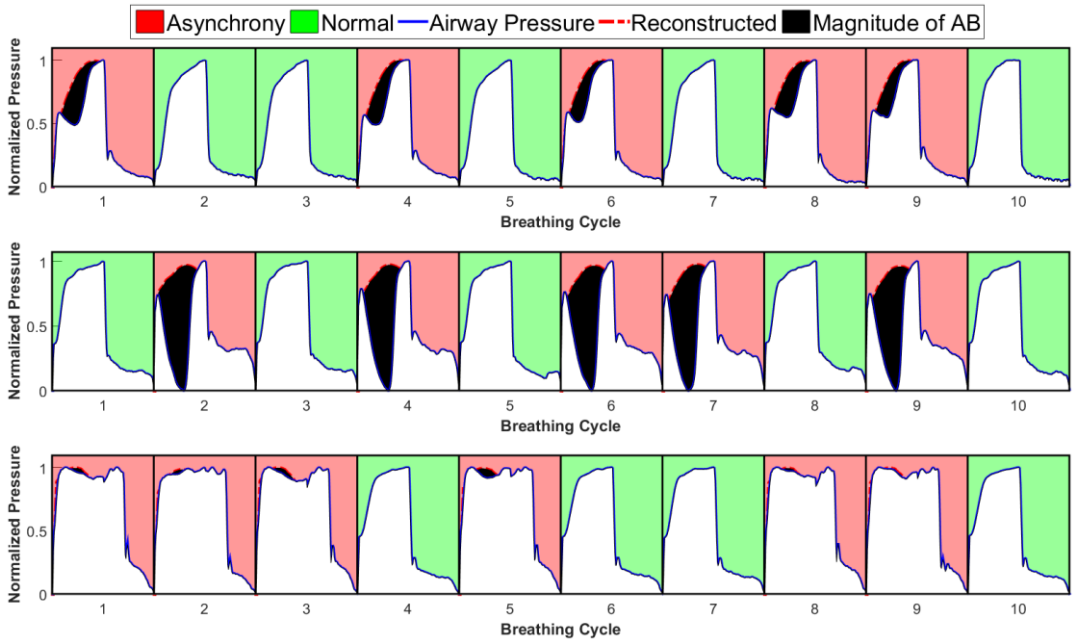

Figure 6.3: Patient 1 (Top) experienced AI of 50% with average magnitude of asynchrony of 14.78% and VI of 85.23% within 10 breathing cycles. Patient 24 (Middle) experienced AI of 50% with average magnitude of asynchrony of 46.26% and VI of 53.73% within 10 breathing cycles. Patient 21 (Bottom) experienced AI of 60% with average magnitude of asynchrony of 3.16% and VI of 96.83% within 10 breathing cycles.

In these 30 breathing cycles as shown in Figure 6.3, the breathing cycles denoted as AB have some form of fluctuations or anomalies in their airway pressure profile. Using the *ABReCA*, the magnitude of these anomalies can be quantified as magnitude of asynchrony. Among these 3 patients' AI and VI, Patients 1 and 24 had AI of 50% while Patient 21 experienced AI of 60%. While the AI for Patient 21 was high, the patient experienced relatively lesser average magnitude of asynchrony of 3.16% and VI of 96.83% as compared with Patient 1 and Patient 24 with magnitude of asynchrony of 14.78% and 46.26% respectively. As a result, Patient 1 and Patient 21 attained VI more than 80%; whereas, Patient 24 experienced lowest VI of 53.73%. Thus, this result clearly shows that measuring

AI alone may not be sufficient to access the quality of PVI. A VI metric that correlates with the AI metric while taking in consideration of the magnitude of AI could be a better assessment of the quality of MV treatment.

6.3.2 Correlation between AI and VI versus Patient's ABG during Treatment

Table 6.1 shows the patient demographics and the summary of the data analysis for this study. The patients' VI, AI, PaO_2, $PaCO_2$, are also shown in Figure 6.4. Most patients (n=14 of 18) were diagnosed with pneumonia. The MV modes for these patients were SIMV, SPONT and BiPAP. The patients' PaO_2 was median 91.0 mmHg (IQR: 75.2-117.8)) and $PaCO_2$ was 39.1 mmHg (IQR: 35.0-44.2).

As shown in Figure 6.4 and Table 6.1, both AI and VI are very different for each patient. The AI for this patient cohort is median 23.5% [IQR: 2.5.7–50.6] with minimum and maximum AI at 0% and 99.4% respectively. The AI for this patient cohort is similar to AI reported in other literatures (Chiew et al., 2015; K. G. Mellott et al., 2014) but AI as high as 99.4% is also observed during treatment; this could represent poor management of patient-ventilator interaction (Akoumianaki et al., 2013).

Most patients had high AI index (AI>10%) (A. W. Thille et al., 2006), however, they maintained more than 70% VI most of the time during MV treatment, despite the patients experienced frequent AB occurrence. For example, Patients 8 and 15 had AI value of more than 70% using CNN classification. This AI value in theory, indicates a suboptimal ventilation delivery; however, these patients also have high VI achieving about 80%. These results suggest that in this patient cohort, patient's desynchrony towards MV assistance was common, but the magnitude of asynchrony was quite low for most of the time, where they did

not resist with the ventilatory support. One reason behind such phenomenon could be the MV supply may not exactly match patient's demand but the degree of dyssynchronous effect is insufficient to alter the patient's airway waveform that are usually used as an indication of asynchrony (Akoumianaki et al., 2013; Yoshida et al., 2018).

Table 6.1: Patients' demographic included in this study.

No.	Clinical Diagnosis	Initial P/F Ratio	MV mode	PaO_2 (mmHg)	$PaCO_2$ (mmHg)	AI (%)	VI (%)
P1	Thyroid Carcinoma	150	SIMV	89.4 [81.2–93.6]	38.4 [32.9–41.0]	29.4 [3.6–42.9]	76.0 [70.0–81.0]
P2	Pneumonia	202	SIMV/ SPONT	105.0 [86.2–118.3]	39.6 [37.7–42]	30.4 [12.8–61.6]	93.2 [82.0–98.1]
P3	SVC Obstruction	95	SIMV	94.3 [81.0–137.5]	49.0 [46.1–52.7]	0.4 [0.1–1.1]	98.8 [98.4–99.4]
P4	HAP	238	SIMV	133.0 [130.8–142.8]	31.3 [30.0–34.1]	26.1 [19.3–56.3]	91.4 [83.8–94.5]
P5	Klebsiella Sepsis	146	SIMV	175.0 [141.5–186.3]	38.8 [31.0–41.6]	38.4 [24.6–49.8]	89.5 [84.6–97.1]
P7	Pneumonia	117	SIMV / BiPAP	97.2 [79.7–126.0]	32.0 [31.4–34.3]	45.8 [40.8–58.4]	76.6 [63.3–85.9]
P8	CAP	128	SIMV/ SPONT	95.8 [87.9–109.0]	44.9 [42.3–45.4]	74.2 [58.4–87.6]	75.4 [74.6–77.4]
P14	HAP	133	SIMV	107.0 [103.0–120.3]	47.1 [45.9–48.3]	25.3 [4.5–44.4]	92.7 [87.3–99.4]
P15	HAP	143	SIMV/ SPONT	109.0 [97.2–114.0]	43.7 [41.8–47.3]	70.5 [33.8–87.4]	82.9 [73.0–94.3]
P16	Pneumonia	155	SIMV/ SPONT	86.7 [73.3–95.1]	37.9 [36.7–40.7]	41.4 [5.4–60.6]	86.0 [82.5–93.4]
P17	Acute Pancreatitis	157	SIMV/ SPONT	117.0 [103.0–130.8]	36.4 [34.2–37.3]	40.3 [9.03–56.5]	83.7 [71.0–87.6]
P18	Right Lobar Pneumonia	92	SIMV/ SPONT	66.6 [62.2–79.3]	63.5 [59.1–81.7]	34.2 [9.5–67.6]	93.0 [79.1–97.8]
P19	CAP	350	SIMV/ SPONT	81.3 [67.1–102.0]	35.0 [29.5–38.3]	46.5 [13.3–60.8]	90.2 [71.7–95.3]
P20	CAP	241	BiPAP /SIMV/ SPONT	129.0 [109.0–154.0]	38.4 [34.4–40.3]	2.9 [0.2–14.7]	94.6 [90.4–99.6]
P21	Recurrent Multifocal Infarct with Poor Neurological Recovery	106	SIMV/ BiPAP	73.8 [68.2–90.7]	46.6 [41.0–48.1]	18.0 [10.8–47.3]	98.2 [97.1–98.9]
P22	Partially Treated Pneumonia	202	SIMV/ SPONT	86.3 [64.4–90.8]	37.8 [34.4–38.4]	1.6 [0.6–2.7]	96.8 [93.7–98.1]
P23	HAP	119	SIMV	81.3 [69.7–89.1]	44.2 [40.2–48.2]	14.9 [2.8–19.5]	96.9 [95.5–98.3]
P24	HAP	246	SIMV/ SPONT	89.2 [86.1–105.0]	39.6 [38.1–41.5]	1.3 [0.6–2.6]	72.9 [83.6–93.8]
Median		148		91.0	39.1	23.5	92.7
IQR		[119–202]		[75.2–117.8]	[35.0–44.2]	[2.5–50.6]	[81.8–97.3]
Median of medians				39.2	95.1	29.9	90.8
IQR of medians				[37.8–44.9]	[86.3–109.0]	[14.9–41.4]	[82.9–94.6]

CAP – Community Acquired Pneumonia; HAP – Hospital Acquired Pneumonia; SVC – Superior Vena Cava; SIMV – Synchronized Intermittent-Mandatory Ventilation; BiPAP – Bilevel positive airway pressure; SPONT – Spontaneous breathing

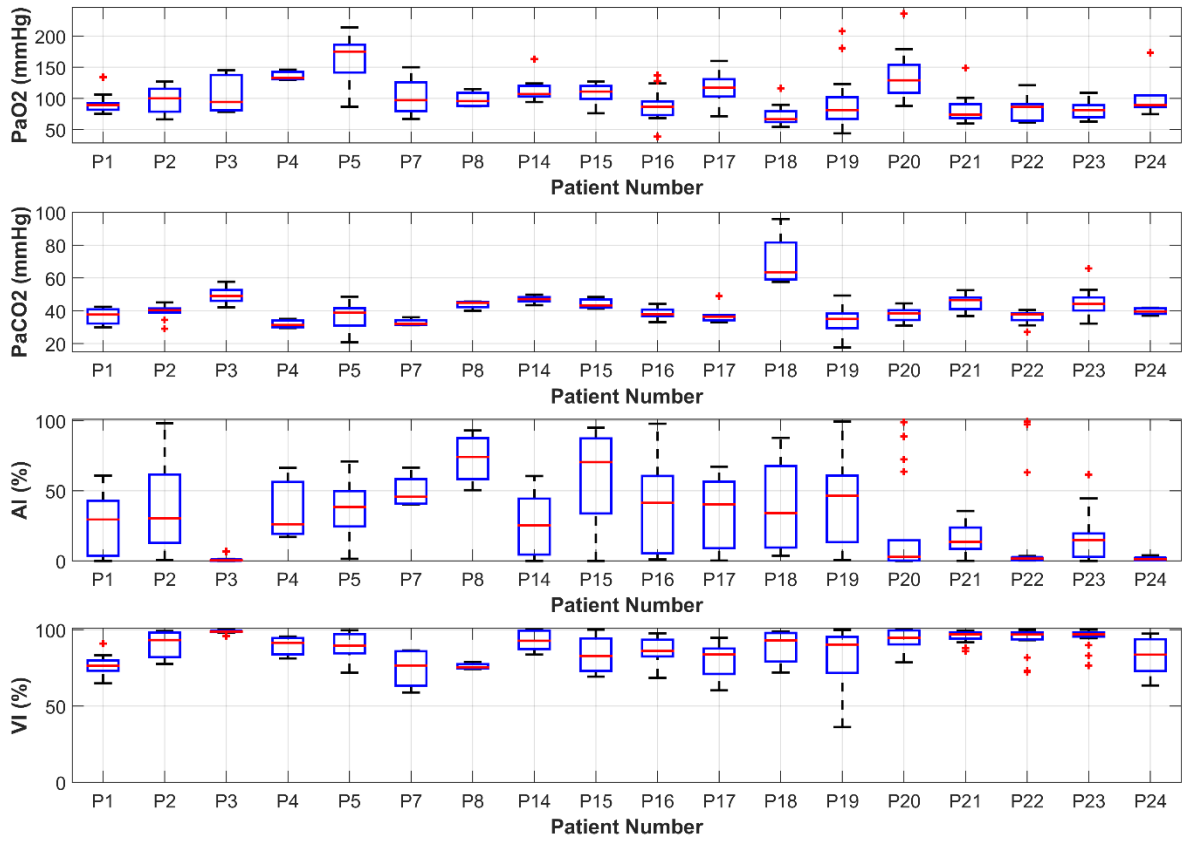

Figure 6.4: Patients' PaO_2, $PaCO_2$, AI and VI presented in boxplot.

Table 6.2 shows the Pearson's correlation coefficient and the significance values when testing AI, VI with PaO_2 and $PaCO_2$. A positive correlation (R>0) between VI versus PaO_2 and negative correlation (R<0) between VI versus $PaCO_2$ suggest a quality MV management may enhance patient's oxygenation intake and carbon dioxide expulsion. This trend can be observed in 7 patients (Patients 5, 7, 8, 14, 19, 20, and 23). However, there are three patients, Patients 1, 3, and 21 exhibited the opposite trend.

A negative relationship can be seen in Patients 17, 22 and 24 when conducting the correlation test in VI versus PaO_2 and VI versus $PaCO_2$. Similarly, a positive relationship between VI and PaO_2 and $PaCO_2$ can be observed in Patients 2, 4, 15, 16, and 18 in which both patient's PaO_2 and $PaCO_2$ level increased. These results suggested that while high VI can be viewed as good patient-ventilator interaction, it does not truly reflect MV patients

quality. There are other cofounding factors such as optimal ventilator settings and sedation used that influence the quality of MV delivery in a patient (Aragón et al., 2019; Guo et al., 2018).

Table 6.2: The correlation coefficient (R value) and significance (P value) when tested with AI or VI with PaO_2 or $PaCO_2$.

		VI				AI			
		PaO_2		$PaCO_2$		PaO_2		$PaCO_2$	
Patient	**ABG Data**	**R Value**	**P Value**	**R Value**	**P Value**	**R Value**	**P Value**	**R Value**	**P Value**
P1^	13	-0.48	0.09	0.18	0.55	0.44	0.14	-0.17	0.59
P2	11	0.13	0.71	0.23	0.50	-0.43	0.19	0.14	0.69
P3^	8	-0.53	0.17	0.18	0.66	0.40	0.33	-0.10	0.81
P4	3	0.41	0.73	0.91	0.28	-0.50	0.66	-0.85	0.35
P5*	15	0.17	0.54	-0.57	0.03	-0.25	0.38	0.57	0.03
P7*	4	0.84	0.16	-0.25	0.75	0.90	0.10	0.27	0.73
P8*	4	0.65	0.35	-1.0	0.01	-0.11	0.89	-0.36	0.64
P14*	7	0.32	0.49	-0.52	0.23	-0.49	0.26	0.32	0.48
P15	6	0.82	0.04	0.08	0.88	-0.75	0.09	-0.39	0.45
P16	19	0.16	0.50	0.24	0.31	0.04	0.86	0.23	0.35
P17	7	-0.37	0.41	-0.04	0.93	0.57	0.18	0.08	0.86
P18	11	0.42	0.19	0.53	0.09	-0.41	0.21	-0.47	0.15
P19*	41	0.35	0.02	-0.27	0.09	0.43	0.00	-0.26	0.11
P20*	18	0.49	0.04	-0.06	0.80	-0.50	0.03	-0.23	0.36
P21^	14	-0.18	0.54	0.28	0.33	-0.05	0.85	0.05	0.86
P22	17	-0.19	0.48	-0.26	0.31	0.12	0.64	0.20	0.44
P23*	19	0.10	0.68	-0.30	0.21	-0.09	0.72	0.11	0.65
P24	6	-0.28	0.58	-0.40	0.43	0.79	0.06	0.71	0.11
25^{th}	6.3	-0.19	0.16	-0.30	0.21	-0.43	0.10	-0.26	0.35
Median	11	0.17	0.45	-0.05	0.32	-0.07	0.24	0.07	0.47
75^{th}	16.5	0.42	0.54	0.23	0.66	0.43	0.66	0.23	0.69

Patients with * indicate attaining positive VI correlation with PaO_2 but negative correlation with $PaCO_2$. Patients with ^ indicate attaining negative VI correlation with PaO_2 but positive correlation with $PaCO_2$.

6.3.3 Patients' VI Correlation versus ABG (Positive, Neutral and Negative)

Figure 6.5 first column shows the scatter plots of VI versus PaO_2 and $PaCO_2$ of Patient 19 (Top), Patient 23 (Middle) and Patient 1 (Bottom). Patient 19 exhibited positive VI versus PaO_2 correlation with R = 0.35, Patient 23 showed a weak correlation of VI against PaO_2 with R = 0.10, and Patient 1 showed a negative correlation of VI against PaO_2 with R = -0.48.

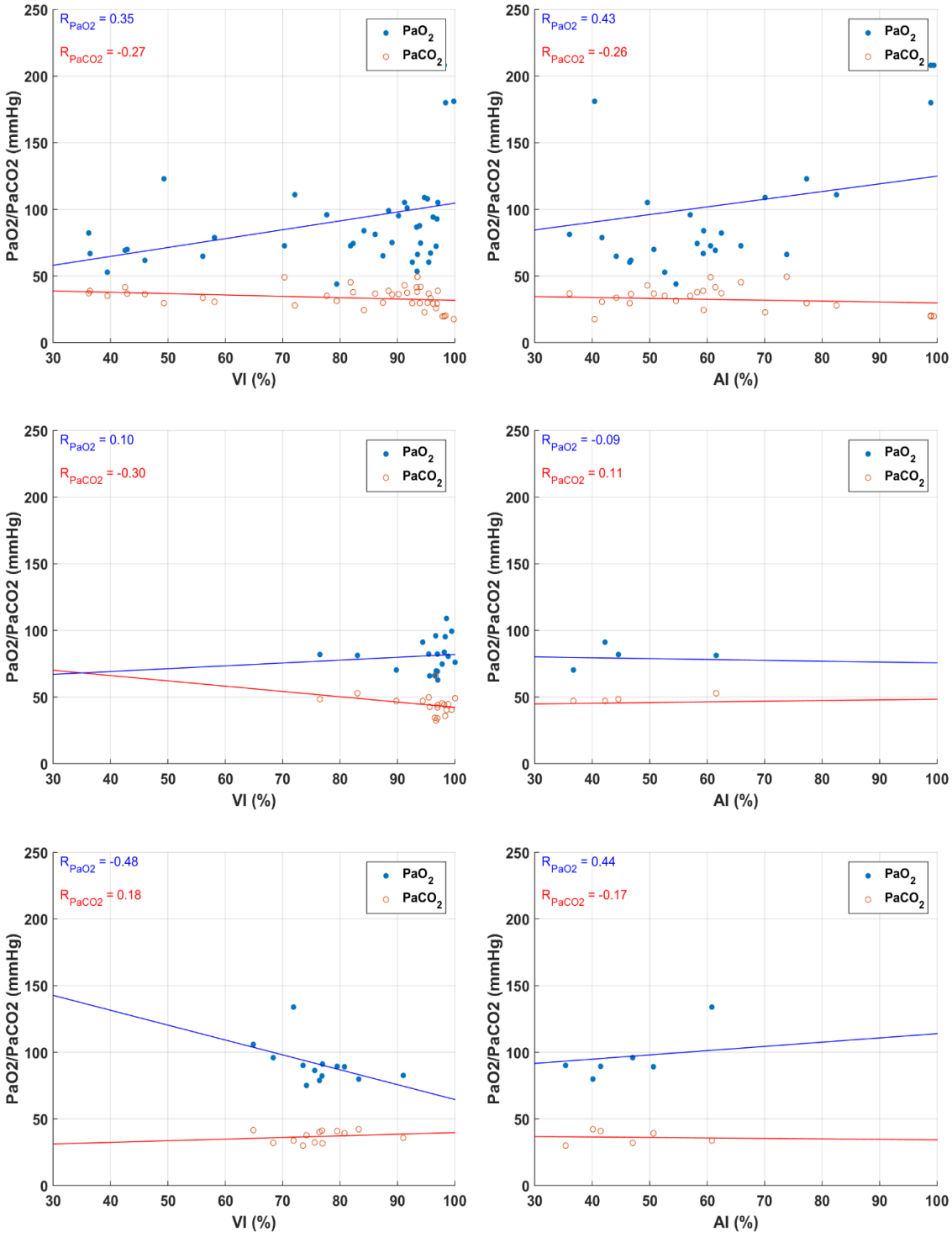

Figure 6.5: Illustration of the strength of relationship of VI (Left column) and AI (Right column) when tested against PaO_2 and $PaCO_2$ displayed in scatter plot. Patient 19 (Top row), Patient 23 (Middle row) and Patient 1 (Bottom row) exhibited positive (R_{P19} = 0.35), neutral (R_{P23} = 0.10) and negative (R_{P1} = -0.48) correlation with tested VI against PaO_2 respectively.

Patient 19's PaO_2 was found to be higher at higher AI; and lower $PaCO_2$ value with lower AI value. This finding is different from literature (Epstein, 2011; Martos-Benítez et al., 2020; Yonis et al., 2015), where a higher AI should result in poor oxygenation which may predispose to lower PaO_2 and higher $PaCO_2$. However, for the same patient, the VI showed positive correlation with PaO_2 and vice versa when tested against $PaCO_2$. This finding indicates that a VI metric can be high despite of higher AI as it accounts for the magnitude of AB. A weak correlation of VI against PaO_2 and $PaCO_2$ observed in Patient 23 and a negative VI trend was observed in Patient 1. These results show that an obvious trend in PaO_2 and $PaCO_2$ cannot be observed despite of high VI and low AI. It also suggests that there are other factors that contributes to the PaO_2 and $PaCO_2$ trend (Park et al., 2015).

Figure 6.6 shows the box-whiskers plot of VI, AI, PaO_2 and $PaCO_2$ for Patient 19. It is observed that Patient 19 experienced high PaO_2 variation in Day 1, suggesting acute onset of patient condition. The PaO_2 values reduced sharply from 120 mmHg to around 70 mmHg on Day 2 and remains consistent at around median 75 mmHg throughout the MV treatment. $PaCO_2$ increased gradually during the initial 5 days of lung treatment but declined steadily starting from Day 6. The $PaCO_2$ values had relatively lower variability as compared with PaO_2. The AI metric showed relatively higher variance on Days 1, 2, 3, 6, 7, 8 and 11 during treatment as compared with VI metric which showed little deviation throughout the treatment except on Day 5. The VI metric showed that the patient's lower ventilator interaction from Day 2 to Day 5 before reaching 80%. This could indicate that patients not adapting to MV support in early ventilation period. The VI started stabilises from Day 6 onwards, suggesting that the MV settings fulfilled the patient's ventilatory demand. The VI trend observation also is supported by the rise of $PaCO_2$ in the first five days and before seeing a decreasing trend after Day 6.

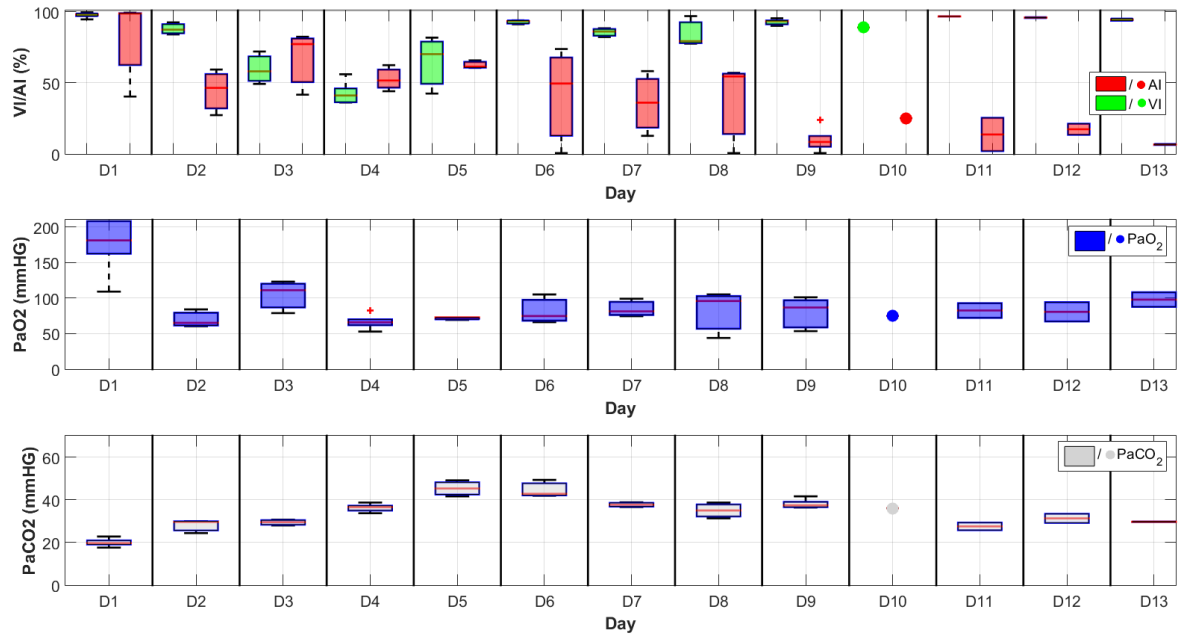

Figure 6.6: Subplots of box plot showing the variance in VI, AI, PaO_2 and $PaCO_2$ experienced by Patient 19. Box plot is replaced with scatter plot if only one ABG is collected on that day.

The VI, AI, PaO_2 and $PaCO_2$ for Patient 1 is shown in Figure 6.7 in box plots. As compared with Patient 19, Patient 1's condition was stable as low fluctuation in PaO_2 and $PaCO_2$ during treatment is observed. Despite high variability of AI was attained from Day 2 to 5, VI indicated consistent and stable patient's 'resistivity' with MV delivery. In fact, VI showed steady increment throughout the patient's MV treatment. However, the effect of higher AI and VI was not obvious in Patient 1 PaO_2 or $PaCO_2$ level. Such incident could indicate that there are multiple factors which could cause deterioration or improvement of PaO_2 (Blackwood et al., 2006) aside from MV patient-ventilator interaction.

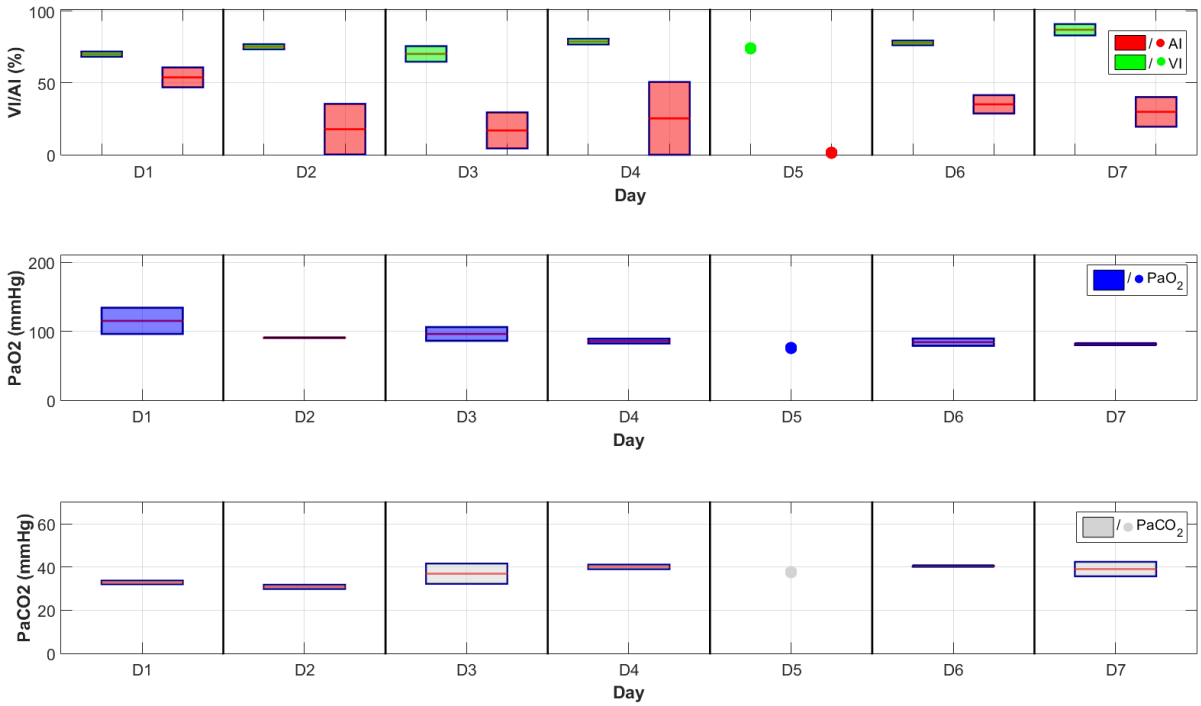

Figure 6.7: Subplots of box plot tracking the change in VI, AI, PaO_2 and $PaCO_2$ experienced by Patient 1. Box plot is replaced with scatter plot if only one ABG is collected on that day.

Ideally, the AI should exhibit positive correlation with $PaCO_2$ level in the blood due to frequent AB occurrence. Conversely, a positive correlation is expected when tested VI with PaO_2 because good VI indicates optimal MV treatment as patient exhibits lesser 'resistivity'. However, this trend is only observed in some of the patients in this study cohort. This result suggests that there are other factors leading to the deterioration or improvement of patient's PaO_2 and $PaCO_2$ during MV besides AB occurrence. As a result, VI or AI metric alone may not truly reflect the patient's condition when establishing the relationship between PVI and ABG.

6.4 Summary

The CNN AI classification and VI metric calculated using *ABReCA* were used to assess the quality of MV management. These two metrics were able to provide additional insight to patient's response to MV treatment that was previously unavailable, potentially benefits the

clinicians to promote or to achieve optimal patient-ventilator interaction. Despite both VI and AI metric do not exhibit strong correlation with the patient's outcome; we speculate that AB occurrence is one of many factors predisposing to poor oxygenation during MV, and more research investigating into this problem is required.

Chapter 7
Conclusions

7.1 Conclusions

In this thesis, we have presented machine learning models to detect asynchronous breathing (AB) and to quantify the magnitude of AB and a metric to evaluate patient-ventilator interaction (PVI). The developed models and metrics were tested with simulated and clinical data. Finally, these machine learning models were used to investigate the quality of PVI towards patient's clinical outcome.

In Chapter 3, a clinical study (CARE trial) to collect patient's ventilator waveform data and patient's arterial blood gases (ABG) was presented. CARE trial recruited 24 mechanically ventilated respiratory failure patients with total of more than 2 million breathing cycles. The different mechanical ventilation (MV) modes and settings among the recruited patients help to create a database for machine learning training and performance validation. Besides that, the heterogeneity of patient's condition towards different MV treatment may aid in elucidating the impact of frequent AB occurrence towards patient's ABG. The manual categorised breathing cycles showed that the first seven recruited patients experienced asynchronous index (AI) of median 39.46% [interquartile range (IQR):11.49%–53.15%]. The results and findings showed that the patients experienced different AB index as it is patient-specific and depends on MV settings. Thus, there is a need to develop models and metric to assess PVI automatically in order to explicate patient's condition and outcomes.

Chapter 4 highlighted the development and implementation of Convolutional Neural Network (CNN), a machine learning algorithm to detect AB automatically. Monte-Carlo simulation was conducted to identify the optimal dataset quantity and data composite type for

CNN training. The results showed that 1-dimensional data structure with 10,000 training datasets from each breathing cycle category achieved relatively better performance as compared with other configurations. Overall, CNN was able to detect AB occurrence with sensitivity of 73.37% and specificity of 91.44% when validated with more than 500,000 breathing cycles. The ability to detect AB automatically enables the potential to monitor the quality of PVI in large.

Chapter 5 focused on the development of two machine learning models, namely generative adversarial network (GAN) and convolutional autoencoder (CAE) to quantify the magnitude of AB through reconstructing AB to asynchrony free breathing cycles. The magnitude of AB is defined as the difference in area under the curve between original AB and reconstructed breathing cycle. The performance of GAN and CAE was compared with a mathematical model to reconstruct AB with least error. The results showed that, CAE outperformed other models by attaining lowest mean absolute percentage error (MAPE) of 0.21% when evaluated with validation dataset.

Chapter 6 investigated the effect of quality of MV treatment on patient's outcome (ABG). This was done by deploying the trained models to classify and quantify AB to evaluate PVI of CARE trial recruited patients. Ventilator-interaction index (VI), a metric to measure the magnitude of AB was developed to better describe PVI. Correlation analysis was conducted to establish the relationship between VI and patient's pressure of oxygen, PaO_2 and partial pressure of carbon dioxide, $PaCO_2$. Results showed that, the developed metric was able to describe or capture the changes in PaO_2 or $PaCO_2$ values when there is a change in VI amidst several patients. Surprisingly, when tested with clinical data, we found that the magnitude of AB experienced by patients is minimal as compared with AI values. This shows that, AI

computation alone maybe insufficient to describe patient's actual condition during treatment. However, the relationships between AI or VI metric systems and PaO_2 or $PaCO_2$ in most patients are vague. Hence, we speculated that other possibilities could contribute to deterioration or improvement of oxygenation and carbon dioxide level besides the quality of MV treatment. Therefore, additional studies to expound causal mechanism or the nature of AB to patient's outcome are still required. Nonetheless, this metric offers a novel approach to quantify PVI which could potentially provide additional insight to patient's condition for the clinicians.

In conclusion, this thesis has shown machine learning models are able to identify and possibly, quantify the magnitude of AB in real time. While these models have shown promising results in clinical application; there is necessity to improve the ability of machine learning to categorize different types of AB. Thus, the ability to detect AB in real time and large scale enables better understanding of MV quality during treatment.

7.2 Communication of Research

This following are the communication of results for this research:

Conferences:

- **Loo, N. L.**, Chiew, Y. S., Tan, C. P., Arunachalam, G., Ralib, A. M., & Mat-Nor, M. B. (2018). A MACHINE LEARNING MODEL FOR REAL-TIME ASYNCHRONOUS BREATHING MONITORING. *IFAC-PapersOnLine*, *51*(27), 378-383. doi:https://doi.org/10.1016/j.ifacol.2018.11.610 **(Presented)**

- Chiew, Y. S., Chase, J. G., Arunachalam, G., Tan, C. P., **Loo, N. L.**, Chiew, Y. W., Mat Nor, M. B. (2018). Clinical Application of Respiratory Elastance (CARE Trial) for Mechanically Ventilated Respiratory Failure Patients: A Model-based Study. *IFAC PapersOnLine, 51*(27), 209-214. doi:10.1016/j.ifacol.2018.11.641 **(Presented)**

- **Loo, N. L.**, Chiew, Y. S., Tan, C. P., Arunachalam, G., Ralib, A. M., & Mat-Nor, M. B. (2018). A MACHINE LEARNING MODEL FOR REAL-TIME ASYNCHRONOUS BREATHING MONITORING. *IFAC-PapersOnLine*, *51*(27), 378-383. doi: https://doi.org/10.1016/j.ifacol.2018.11.610 **(Presented)**

Journals:

- **Loo, N. L.**, Chiew, Y. S., Tan, C. P., Ralib, A. M., & Mat-Nor, M. B. Convolutional Neural Network in Monitoring Asynchronous Breathing Waveform **(Submitted)**

- **Loo, N. L.**, Chiew, Y. S., Tan, C. P., Ralib, A. M., & Mat-Nor, M. B. A Machine Learning Approach to Assess Magnitude of Asynchrony Breathing **(Submitted)**

Chapter 8
Future Work

There are many experiments, improvements and adaptations have been remained for future due to the lack of time (i.e. clinical trials setup, patient's recruitment and patient's data collation are very time consuming.). This chapter provides several suggestions and proposals to investigate into different potential approach or model consideration to improve the research. The first section gives some suggestions on improving data quality collected from clinical trial. This is followed, in Section 8.2 by discussions on improving the models to provide the insight of patient's condition.

8.1 Clinical Trial

8.1.1 Real-Time Clinical Data Collection

The clinical trial (CARE trial) as described in Chapter 3 was designed to provide patient's breathing pattern as well as patient's bed chart which holds vital signs such as blood pressure, arterial blood gasses, MV settings and many more. Although these data help to reflect patient's condition during treatment, these data are often recorded retrospectively and manually. Hence, it is often difficult to pair with the collected real time data from the ventilator due to inconsistency. For example, the blood pressure may not be measured and recorded hourly consistently; whereas, patient's vital signs information can be recorded by clinicians early or later within the hour. The lack of consistency in clinical data may predispose to analysis imprecision. Therefore, if available and feasible, clinical data should be collected in real time and in parallel to ventilatory data or ensure the patient's data collection is in consistent manner.

8.1.2 Additional Clinical Collection

Currently, only arterial blood gases are collected and extracted from patient's bed charts for biomarker analysis. The analysis can be improved by collecting additional clinical data such as thermodilution cardiac output and sedation level as these are proven to cause AB occurrence (M. de Wit et al., 2009; Imanaka et al., 2000). These data can help to elucidate the causative mechanism of AB in order to better manage the quality of MV.

8.2 Asynchronies Detection and Quantification

8.2.1 CNN Model

Currently, the developed model as presented in Chapter 4 is only capable of differentiate between AB and normal breathing. While the model can identify double triggering, reverse triggering, flow asynchrony and delayed-triggering; but the lack of competency in categorizing different AB may impede elucidation the causal mechanism or impact of AB. Besides, the absence of taking flow waveform into consideration may affect the performance of the model as flow waveform may accommodate crucial information defining an asynchrony. Henceforth, a machine learning model that takes both pressure and flow waveforms into account is necessary to classify different types of AB in high accuracy. Besides that, additional studies are required to understand and avoid the effect of unbalance training dataset towards CNN performance due the lack of equal sample AB during training. Moreover, the presence of noise due to leakage or presence of secretion (Arnal, 2018) is prevalent during MV treatment. An additional category to detect the presence of secretion maybe useful as indicator for clinicians to remove secretion. Besides that, additional work is still required to train CNN model to identify AB in different MV modes aside from VC mode. Hence, setting up additional clinical trial to collect more clinical data with different MV modes is necessary. The ability to detect different AB in different MV modes will enable us

to study in-depth on the AB aetiology and its effect towards patient's recovery in a larger scale.

The lack of agreement between researchers during AB classification may impede the performance of CNN to detect AB. Therefore, additional work and effort can be carried out to standardise AB classification. In particular, involving more trained clinicians, and annotating more AB data may improve the model's performance and consistency (Nagendran et al., 2020). This process to refine the performance of CNN model can be repeated until the detected breathing cycles are agreed or similar with the consented AB traits (Sidey-Gibbons & Sidey-Gibbons, 2019).

Moreover, the absence of ability to detect AB occurrence due to leakage or noise such as double triggering or auto-triggering (A. W. Thille et al., 2006) may also impede the implementation of VI as a metric to assess MV quality. The PVI equation computes the magnitude of AB by calculating the area under the curve of asynchrony affected region. However, the occurrence of AB due to leakage are often not altered by patient's effort; thus, the breathing cycles are often look 'normal' as it does not involve any muscles contraction or 'entrainment'. Therefore, expanding an additional category in CNN to detect double triggering and auto-triggering may help to improve the VI metric by simply excluding the calculation of area under the curve of those breathing cycles.

Another potential future work in AB detection is to investigate the effect of normalised breathing data towards machine learning model development. It is known normalisation of breath could potentially alter the data form and thus the machine learning model may not capture these unique breaths. Thus, with more non-subjective clinical data, we can investigate

the effects of normalisation or any data processing method to better understand the machine learning model development process, its advantages and disadvantages.

Lastly, CNN model can be integrated with CURE software which runs natively on JAVA environment by using deep learning 4 JAVA (DL4J) library (Team, 2016). DL4J enables the deployment and execution of Python developed machine learning models on JAVA program by simply loading the trained model. This will allow ease of deployment of model at clinical bedside without the need to re-program CURE software in Python to deploy CNN to detect AB.

8.2.2 AB Model Quantification

The performance of *ABReCA* has shown robustness to reconstruct AB to asynchrony free breathing cycle with minimal error, despite it was trained using simulated data. However, the lack of additional training dataset which caters the patterns of other MV modes may limit its potential to be deployed in ICU environment. Therefore, future work should diversify the patterns of simulated breathing cycles.

Furthermore, as autoencoder is adept in dimensionality reduction, the latent representation in *ABReCA*'s code layer can be exploited to explore the undiscovered intrinsic properties of breathing cycles (Wang et al., 2015). Principal component analysis (PCA) to emphasize variation or patterns in the data (Moore, 1981) to discover unexplored underlying information in the breathing cycles. These critical features can be used to track or monitor patient's condition as the disease progresses; thus, allowing better MV monitoring.

8.2.3 Integrating the Two Models to Classify and Quantify breathing cycles

We have presented two machine learning models namely CNN and *ABReCA* to classify and quantify respectively. These models were trained independently and requires two different training datasets and configuration. As a result, training these models to improve performance separately may consume additional time and resources. Currently, we have learned the optimal configuration to train *ABReCA* to reconstruct AB with least error. Therefore, we can develop an integrated machine learning model which is able classify and quantify the magnitude of AB concurrently by using the latent information from *ABReCA*. Figure 8.1 shows the overview of the integrated machine learning model to quantify and classify simultaneously. The integrated system can help to accelerate the deployment of machine learning model in ICU to monitor patient's condition.

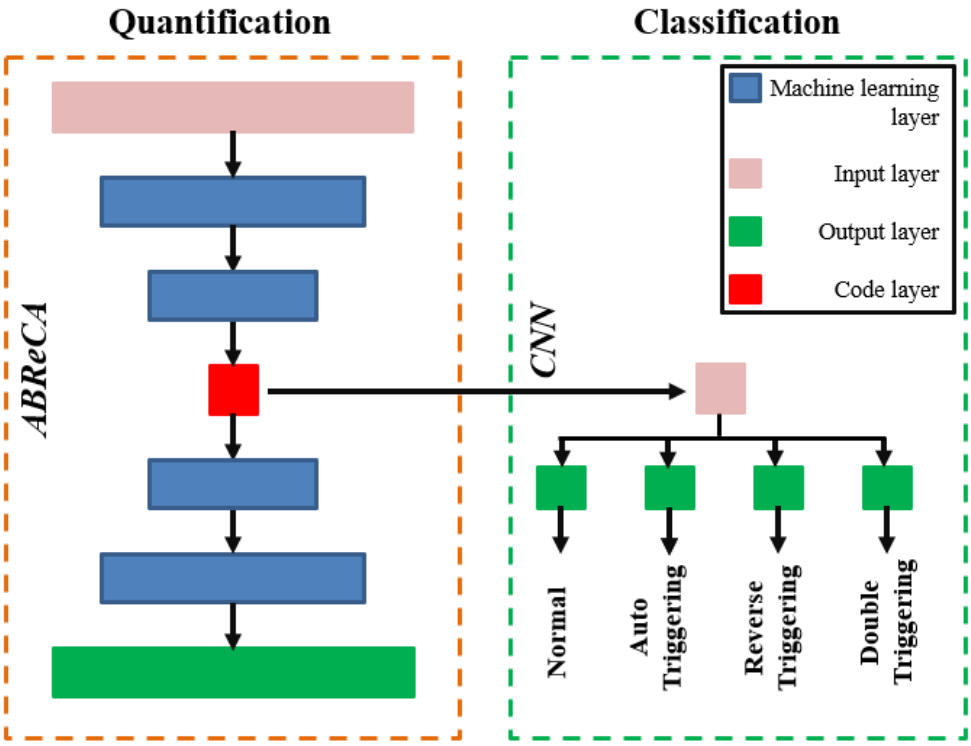

Figure 8.1: An overview of integrated system able to classify different types of AB and reconstruct to quantify the magnitude of AB.

As *ABReCA*'s neurons are well trained; re-training the integrated system is not necessary. Transfer learning is a technique to re-purpose the existing model to perform other objectives

without the need to develop a model from scratch (Sinno Jialin & Qiang, 2010; Tan et al., 2018). By doing so, time and resources to re-train the model can be avoided; thus, encourages and accelerates the model's performance evaluation and deployment. It is however, extra investigation is needed to prepare additional training dataset which comprise different types of AB such as auto triggering, reverse triggering and double triggering may be required in order to realise this concept.

8.2.4 Detection and Quantification of AB in other MV Modes

In this thesis, we have demonstrated the methods to detect and quantify the patient-effort present in AB during VC MV mode. However, this work needs to be extended to different MV modes, especially, patients maybe ventilated in different MV setting. In particular, we believe that development of an extended *ABReCA* that enables flow reconstruction is also important, as AB occurrence in Pressure controlled mode is equally prevalent. For this purpose, additional clinical data obtained from MV PC mode is required to train a model to detect and perform flow reconstruction. Futuremore, a mathematical model capable of simulating synthetic AB during pressure controlled could also be helpful in training the extended model.

Bibliography

Abdel-Hamid, O., Deng, L., & Yu, D. (2013). *Exploring convolutional neural network structures and optimization techniques for speech recognition.* Paper presented at the INTERSPEECH.

Akoumianaki, E., Lyazidi, A., Rey, N., Matamis, D., Perez-Martinez, N., Giraud, R., Mancebo, J., Brochard, L., & Richard, J. M. (2013). Mechanical ventilation-induced reverse-triggered breaths: a frequently unrecognized form of neuromechanical coupling. *CHEST*, *143*(4), 927-938. doi:10.1378/chest.12-1817

Akoumianaki, E., Maggiore, S. M., Valenza, F., Bellani, G., Jubran, A., Loring, S. H., Pelosi, P., Talmor, D., Grasso, S., Chiumello, D., Guerin, C., Patroniti, N., Ranieri, V. M., Gattinoni, L., Nava, S., Terragni, P. P., Pesenti, A., Tobin, M., Mancebo, J., Brochard, L., & Group, P. W. (2014). The application of esophageal pressure measurement in patients with respiratory failure. *Am J Respir Crit Care Med*, *189*(5), 520-531. doi:10.1164/rccm.201312-2193CI

Ali, I., Lashari, H., Hassan, S., Maitlo, A., & Qureshi, B. (2018). Image Denoising with Color Scheme by Using Autoencoders. *18*, 158-161.

Aragón, R. E., Proaño, A., Mongilardi, N., de Ferrari, A., Herrera, P., Roldan, R., Paz, E., Jaymez, A. A., Chirinos, E., Portugal, J., Quispe, R., Brower, R. G., & Checkley, W. (2019). Sedation practices and clinical outcomes in mechanically ventilated patients in a prospective multicenter cohort. *Critical Care*, *23*(1), 130. doi:10.1186/s13054-019-2394-9

Arjovsky, M., & Bottou, L. (2017). Towards Principled Methods for Training Generative Adversarial Networks. *arXiv.org*.

Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein GAN.

Arnal, J.-M. (2018). Monitoring During Expiration. In J.-M. Arnal (Ed.), *Monitoring Mechanical Ventilation Using Ventilator Waveforms* (pp. 59-80). Cham: Springer International Publishing.

Arunachalam, G. R., Chiew, Y. S., Tan, C. P., Ralib, A. M., & Nor, M. B. M. (2020). Patient asynchrony modelling during controlled mechanical ventilation therapy. *computer methods and programs in biomedicine*, *183*, 105103. doi:10.1016/j.cmpb.2019.105103

Ashbaugh, D. G., Bigelow, D. B., Petty, T. L., & Levine, B. E. (1967). Acute respiratory distress in adults. *Lancet (London, England)*, *2*(7511), 319. doi:10.1016/s0140-6736(67)90168-7

Ashutosh, K., Gilbert, R., Auchincloss, J. H., Jr., & Peppi, D. (1975). Asynchronous Breathing Movements in Patients with Chronic Obstructive Pulmonary Disease. *CHEST, 67*(5), 553-557. doi:10.1378/chest.67.5.553

Baedorf Kassis, E., Su, H. K., Graham, A. R., Novack, V., Loring, S. H., & Talmor, D. S. (2021). Reverse Trigger Phenotypes in Acute Respiratory Distress Syndrome. *Am J Respir Crit Care Med, 203*(1), 67-77. doi:10.1164/rccm.201907-1427OC

Baldi, P. (2012). *Autoencoders, unsupervised learning, and deep architectures.* Paper presented at the Proceedings of ICML workshop on unsupervised and transfer learning.

Bandyopadhyaya, I., Babu, D., Kumar, A., & Roychowdhury, J. (2014). Tactile sensing based softness classification using machine learning. In (pp. 1231-1236).

Bansal, A., Castillo, C., Ranjan, R., & Chellappa, R. (2017). The Do's and Don'ts for CNN-based Face Verification. *arXiv.org*.

Barrett, K. E., Barman, S. M., Brooks, H. L., & Yuan, J. X. J. (2019). Introduction to Pulmonary Structure & Mechanics. In *Ganong's Review of Medical Physiology, 26e*. New York, NY: McGraw-Hill Education.

Basha, S. H. S., Dubey, S. R., Pulabaigari, V., & Mukherjee, S. (2020). Impact of fully connected layers on performance of convolutional neural networks for image classification. *Neurocomputing, 378*, 112-119. doi:10.1016/j.neucom.2019.10.008

Bates, J. H. T. (2009). *Lung Mechanics: An Inverse Modeling Approach*. Cambridge: Cambridge University Press.

Beitler, J. R., Malhotra, A., & Thompson, B. T. (2016). Ventilator-induced Lung Injury. *Clinics in chest medicine, 37*(4), 633-646. doi:10.1016/j.ccm.2016.07.004

Bellani, G., Laffey, J. G., Pham, T., Fan, E., Brochard, L., Esteban, A., Gattinoni, L., van Haren, F., Larsson, A., McAuley, D. F., Ranieri, M., Rubenfeld, G., Thompson, B. T., Wrigge, H., Slutsky, A. S., & Pesenti, A. (2016). Epidemiology, Patterns of Care, and Mortality for Patients With Acute Respiratory Distress Syndrome in Intensive Care Units in 50 Countries. *Jama, 315*(8), 788-800. doi:10.1001/jama.2016.0291

Bernard, G., Artigas, A., Brigham, K., Carlet, J., Falke, K., Hudson, L., Lamy, M., LeGall, J., Morris, A., & Spragg, R. (1994). Report of the American-European consensus conference on ARDS: Definitions, mechanisms, relevant outcomes and clinical trial coordination. *Intensive Care Medicine, 20*(3), 225-232. doi:10.1007/BF01704707

Berrar, D. (2018). Cross-Validation. In *Encyclopedia of Bioinformatics and Computational Biology* (pp. 542-545).

Bi, F., Man, Z., Xia, Y., Liu, W., Yang, W., Fu, X., & Gao, L. (2020). Improvement and Application of Generative Adversarial Networks Algorithm Based on Transfer Learning. *Mathematical Problems in Engineering, 2020*, 9453586. doi:10.1155/2020/9453586

Blackwood, B., Wilson-Barnett, J., Patterson, C. C., Trinder, T. J., & Lavery, G. G. (2006). An evaluation of protocolised weaning on the duration of mechanical ventilation*. *Anaesthesia, 61*(11), 1079-1086. doi:https://doi.org/10.1111/j.1365-2044.2006.04830.x

Blanch, L., Villagra, A., Sales, B., Montanya, J., Lucangelo, U., Lujan, M., Garcia-Esquirol, O., Chacon, E., Estruga, A., Oliva, J. C., Hernandez-Abadia, A., Albaiceta, G. M., Fernandez-Mondejar, E., Fernandez, R., Lopez-Aguilar, J., Villar, J., Murias, G., & Kacmarek, R. M. (2015). Asynchronies during mechanical ventilation are associated with mortality. *Intensive Care Med, 41*(4), 633-641. doi:10.1007/s00134-015-3692-6

Blokpoel, R. G., Burgerhof, J. G., Markhorst, D. G., & Kneyber, M. C. (2016). Patient-Ventilator Asynchrony During Assisted Ventilation in Children. *Pediatr Crit Care Med, 17*(5), e204-211. doi:10.1097/PCC.0000000000000669

Bodnar, C. (2018). *Text to Image Synthesis Using Generative Adversarial Networks.*

Branson, R. D., Blakeman, T. C., & Robinson, B. R. H. (2013). Asynchrony and dyspnea. *Respiratory Care, 58*(6), 973. doi:10.4187/respcare.02507

Brochard, L., Martin, G. S., Blanch, L., Pelosi, P., Belda, F. J., Jubran, A., Gattinoni, L., Mancebo, J., Ranieri, V. M., Richard, J.-C. M., Gommers, D., Vieillard-Baron, A., Pesenti, A., Jaber, S., Stenqvist, O., & Vincent, J.-L. (2012). Clinical review: Respiratory monitoring in the ICU - a consensus of 16. *Critical care (London, England), 16*(2), 219. doi:10.1186/cc11146

Brower, R. G., Matthay, M. A., Morris, A., Schoenfeld, D., Thompson, B. T., & Wheeler, A. (2000). Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *N Engl J Med, 342*(18), 1301-1308. doi:10.1056/nejm200005043421801

Bulleri, E., Fusi, C., Bambi, S., & Pisani, L. (2018). Patient-ventilator asynchronies: types, outcomes and nursing detection skills. *Acta bio-medica : Atenei Parmensis, 89*(7-S), 6-18. doi:10.23750/abm.v89i7-S.7737

Calderini, E., Confalonieri, M., Puccio, P. G., Francavilla, N., Stella, L., & Gregoretti, C. (1999). Patient-ventilator asynchrony during noninvasive ventilation: the role of expiratory trigger. *Intensive Care Med, 25*(7), 662-667. doi:10.1007/s001340050927

Chanques, P. G., Kress, B. J., Pohlman, B. A., Patel, B. S., Poston, B. J., Jaber, B. S., & Hall, B. J. (2013). Impact of Ventilator Adjustment and Sedation–Analgesia Practices on

Severe Asynchrony in Patients Ventilated in Assist-Control Mode*. *Critical care medicine, 41*(9), 2177-2187. doi:10.1097/CCM.0b013e31828c2d7a

Chao, D. C., Scheinhorn, D. J., & Stearn-Hassenpflug, M. (1997). Patient-ventilator trigger asynchrony in prolonged mechanical ventilation. *CHEST, 112*(6), 1592-1599. doi:10.1378/chest.112.6.1592

Chase, J. G., Moeller, K., Shaw, G. M., Schranz, C., Chiew, Y. S., & Desaive, T. (2014). When the value of gold is zero. *BMC Research Notes, 7*(1), 404. doi:10.1186/1756-0500-7-404

Chatburn, R. L., & Mireles-Cabodevila, E. (2020). 2019 Year in Review: Patient-Ventilator Synchrony. *Respiratory Care, 65*(4), 558. doi:10.4187/respcare.07635

Chetlur, S., Woolley, C., Vandermersch, P., Cohen, J., Tran, J., Catanzaro, B., & Shelhamer, E. (2014). cudnn: Efficient primitives for deep learning. *arXiv preprint arXiv:1410.0759*.

Chiew, Y. S., Chase, J. G., Arunachalam, G., Tan, C. P., Loo, N. L., Chiew, Y. W., Ralib, A. M., & Mat Nor, M. B. (2018). Clinical Application of Respiratory Elastance (CARE Trial) for Mechanically Ventilated Respiratory Failure Patients: A Model-based Study. *IFAC PapersOnLine, 51*(27), 209-214. doi:10.1016/j.ifacol.2018.11.641

Chiew, Y. S., Chase, J. G., Shaw, G. M., Sundaresan, A., & Desaive, T. (2011). Model-based PEEP optimisation in mechanical ventilation. *BioMedical Engineering OnLine, 10*(1), 111. doi:10.1186/1475-925X-10-111

Chiew, Y. S., G. Pretty, C., Beatson, A., Glassenbury, D., Major, V., Corbett, S., Redmond, D., Szlávecz, Á., Shaw, G., & Chase, J. (2015). *Automated Logging of Inspiratory and Expiratory Non-Synchronized Breathing (ALIEN) for Mechanical Ventilation* (Vol. 2015).

Chiew, Y. S., Pretty, C., Docherty, P. D., Lambermont, B., Shaw, G. M., Desaive, T., & Chase, J. G. (2015). Time-Varying Respiratory System Elastance: A Physiological Model for Patients Who Are Spontaneously Breathing. *PLoS One, 10*(1), e0114847. doi:10.1371/journal.pone.0114847

Chiew, Y. S., Tan, C. P., Chase, J. G., Chiew, Y. W., Desaive, T., Ralib, A. M., & Nor, M. B. M. (2018). Assessing mechanical ventilation asynchrony through iterative airway pressure reconstruction. *computer methods and programs in biomedicine, 157*, 217-224.

Chiumello, D., Polli, F., Tallarini, F., Chierichetti, M., Motta, G., Azzari, S., Colombo, R., Rech, R., Pelosi, P., Raimondi, F., & Gattinoni, L. (2007). Effect of different cycling-off criteria and positive end-expiratory pressure during pressure support ventilation in patients with chronic obstructive pulmonary disease. *Critical care medicine, 35*(11), 2547-2552. doi:10.1097/01.CCM.0000287594.80110.34

Collins, J.-A., Rudenski, A., Gibson, J., Howard, L., & O'Driscoll, R. (2015). Relating oxygen partial pressure, saturation and content: the haemoglobin-oxygen dissociation curve. *Breathe (Sheffield, England), 11*(3), 194-201. doi:10.1183/20734735.001415

Colombo, D., Cammarota G Fau - Alemani, M., Alemani M Fau - Carenzo, L., Carenzo L Fau - Barra, F. L., Barra Fl Fau - Vaschetto, R., Vaschetto R Fau - Slutsky, A. S., Slutsky As Fau - Della Corte, F., Della Corte F Fau - Navalesi, P., & Navalesi, P. (2011). Efficacy of ventilator waveforms observation in detecting patient-ventilator asynchrony. *Crit Care Med, 39*(1530-0293 (Electronic)), 2452-2457. doi:10.1097/CCM.0b013e318225753c

Daoud, E. G., Farag, H. L., & Chatburn, R. L. (2012). Airway pressure release ventilation: What do we know? *Respir Care, 57*(2), 282-292. doi:10.4187/respcare.01238

de Haro, C., Ochagavia, A., López-Aguilar, J., Fernandez-Gonzalo, S., Navarra-Ventura, G., Magrans, R., Montanyà, J., & Blanch, L. (2019). Patient-ventilator asynchronies during mechanical ventilation: current knowledge and research priorities. *Intensive Care Medicine Experimental, 7*(Supplement 1), 1-14. doi:10.1186/s40635-019-0234-5

De Wit, B. M., Miller, A. K., Green, E. D., Ostman, K. H., Gennings, K. C., & Epstein, K. S. (2009). Ineffective triggering predicts increased duration of mechanical ventilation. *Critical care medicine, 37*(10), 2740-2745. doi:10.1097/CCM.0b013e3181a98a05

de Wit, M. (2011). Monitoring of Patient-Ventilator Interaction at the Bedside. *Respiratory Care, 56*, 61-72.

de Wit, M., Pedram, S., Best, A. M., & Epstein, S. K. (2009). Observational study of patient-ventilator asynchrony and relationship to sedation level. *Journal of Critical Care, 24*(1), 74-80. doi:10.1016/j.jcrc.2008.08.011

Delisle, S., Charbonney, E., Albert, M., Ouellet, P., Marsolais, P., Rigollot, M., Savary, D., Richard, J.-C. M., & Serri, K. (2016). Patient-Ventilator Asynchrony due to Reverse Triggering Occurring in Brain-Dead Patients: Clinical Implications and Physiological Meaning. *American journal of respiratory and critical care medicine, 194*(9), 1166. doi:10.1164/rccm.201603-0483LE

Downs, B. J., Klein, F. E., Sautels, H. D., Modell, R. J., & Kirby, R. R. (1974). INTERMITTENT MANDATORY VENTILATION: A NEW APPROACH TO WEANING PATIENTS FROM MECHANICAL VENTILATORS. *Survey of Anesthesiology, 18*(3), 261-261. doi:10.1097/00132586-197406000-00033

Dres, M., Rittayamai, N., & Brochard, L. (2016). Monitoring patient–ventilator asynchrony. *Current Opinion in Critical Care, 22*(3), 246-253. doi:10.1097/MCC.0000000000000307

Drinker, P., & Shaw, L. A. (1929). AN APPARATUS FOR THE PROLONGED ADMINISTRATION OF ARTIFICIAL RESPIRATION: I. A Design for Adults and Children. *The Journal of clinical investigation, 7*(2), 229-247. doi:10.1172/JCI100226

Epstein, S. K. (2011). How Often Does Patient-Ventilator Asynchrony Occur and What Are the Consequences? *Respiratory Care, 56*(1), 25. Retrieved from http://rc.rcjournal.com/content/56/1/25.abstract

Epstein, S. K., & Chatburn, R. L. (2011). Patient-Ventilator Interaction. *Respiratory Care, 56*(1), 13. doi:10.4187/respcare.01150

Ferguson, N. D., Fan, E., Camporota, L., Antonelli, M., Anzueto, A., Beale, R., Brochard, L., Brower, R., Esteban, A., Gattinoni, L., Rhodes, A., Slutsky, A. S., Vincent, J. L., Rubenfeld, G. D., Thompson, B. T., & Ranieri, V. M. (2012). The Berlin definition of ARDS: an expanded rationale, justification, and supplementary material. *Intensive Care Med, 38*(10), 1573-1582. doi:10.1007/s00134-012-2682-1

Gal, Y., & Ghahramani, Z. (2015). Bayesian convolutional neural networks with Bernoulli approximate variational inference. *arXiv preprint arXiv:1506.02158*.

Gattinoni, L., Carlesso, E., Brazzi, L., & Caironi, P. (2010). Positive end-expiratory pressure. *Current Opinion in Critical Care, 16*(1), 39-44. doi:10.1097/MCC.0b013e3283354723

Gattinoni, L., Carlesso E Fau - Cadringher, P., Cadringher P Fau - Valenza, F., Valenza F Fau - Vagginelli, F., Vagginelli F Fau - Chiumello, D., & Chiumello, D. (2003). Physical and biological triggers of ventilator-induced lung injury and its prevention. *Eur Respir J Suppl, 47*(0904-1850 (Print)), 15s-25s. doi:10.1183/09031936.03.00021303

Gattinoni, L., Pesenti, A., & Matthay, M. (2018). Understanding blood gas analysis. *Intensive Care Medicine, 44*(1), 91-93. doi:10.1007/s00134-017-4824-y

Georgopoulos, D., Prinianakis, G., & Kondili, E. (2006). Bedside waveforms interpretation as a tool to identify patient-ventilator asynchronies. *Intensive Care Medicine, 32*(1), 34-47. doi:10.1007/s00134-005-2828-5

Gilstrap, D., & MacIntyre, N. (2013). Patient-ventilator interactions. Implications for clinical management. *Am J Respir Crit Care Med, 188*(9), 1058-1068. doi:10.1164/rccm.201212-2214CI

Giusti, A., Guzzi, J., Ciresan, D. C., He, F. L., Rodriguez, J. P., Fontana, F., Faessler, M., Forster, C., Schmidhuber, J., Caro, G. D., Scaramuzza, D., & Gambardella, L. M. (2016). A Machine Learning Approach to Visual Perception of Forest Trails for Mobile Robots. *IEEE Robotics and Automation Letters, 1*(2), 661-667. doi:10.1109/LRA.2015.2509024

Gogineni, V. K., Brimeyer, R., & Modrykamien, A. (2012). Patterns of patient-ventilator asynchrony as predictors of prolonged mechanical ventilation. *Anaesth Intensive Care, 40*(6), 964-970. doi:10.1177/0310057x1204000607

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). *Generative adversarial nets.* Paper presented at the Advances in neural information processing systems.

Graves, A., Mohamed, A.-R., & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. In (pp. 6645-6649).

Guo, L., Xie, J., Huang, Y., Pan, C., Yang, Y., Qiu, H., & Liu, L. (2018). Higher PEEP improves outcomes in ARDS patients with clinically objective positive oxygenation response to PEEP: a systematic review and meta-analysis. *BMC anesthesiology, 18*(1), 172. doi:10.1186/s12871-018-0631-4

Gurevitch, M. J., & Gelmont, D. (1989). Importance of trigger sensitivity to ventilator response delay in advanced chronic obstructive pulmonary disease with respiratory failure. *Critical care medicine, 17*(4), 354-359. doi:10.1097/00003246-198904000-00011

Gutierrez, G., Ballarino, G. J., Turkan, H., Abril, J., De La Cruz, L., Edsall, C., George, B., Gutierrez, S., Jha, V., & Ahari, J. (2011). Automatic detection of patient-ventilator asynchrony by spectral analysis of airway flow. *Crit Care, 15*(4), R167. doi:10.1186/cc10309

Hebb, D. O. (2005). *The organization of behavior: A neuropsychological theory*: Psychology Press.

Henry, N. R., Russian, C. J., & Nespral, J. (2016). Identifying Potential Ventilator Auto-Triggering Among Organ Procurement Organization Referrals. *Progress in Transplantation, 26*(2), 129-134. doi:10.1177/1526924816640649

Hess, D. R. (2005). Ventilator waveforms and the physiology of pressure support ventilation. *Respir Care, 50*(2), 166-186; discussion 183-166. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/15691390

Hjelmgren, J., Bruce Wirta, S., Huetson, P., Myrén, K.-J., & Göthberg, S. (2016). Health economic modeling of the potential cost saving effects of Neurally Adjusted Ventilator Assist. *Therapeutic advances in respiratory disease, 10*(1), 3-17. doi:10.1177/1753465815603659

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation, 9*(8), 1735-1780. doi:10.1162/neco.1997.9.8.1735

Holanda, M. A., Vasconcelos, R. D. S., Ferreira, J. C., & Pinheiro, B. V. (2018). Patient-ventilator asynchrony. *Jornal brasileiro de pneumologia : publicacao oficial da*

Sociedade Brasileira de Pneumologia e Tisilogia, 44(4), 321-333. doi:10.1590/S1806-37562017000000185

Hubmayr, R. D. (2011). Point: Is low tidal volume mechanical ventilation preferred for all patients on ventilation? Yes. *CHEST, 140*(1), 9-11. doi:10.1378/chest.11-0825

Imanaka, R. H., Nishimura, R. M., Takeuchi, R. M., Kimball, R. W., Yahagi, R. N., & Kumon, R. K. (2000). Autotriggering caused by cardiogenic oscillation during flow-triggered mechanical ventilation. *Critical care medicine, 28*(2), 402-407. doi:10.1097/00003246-200002000-00019

Ioffe, S., & Szegedy, C. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *arXiv.org*.

Janocha, K., & Czarnecki, W. (2017). On Loss Functions for Deep Neural Networks in Classification. *Schedae Informaticae, 25*. doi:10.4467/20838476SI.16.004.6185

Jayalakshmi, T., & A, S. (2011). Statistical normalization and back propagation for classification. *International Journal Computer Theory Engineering (IJCTE), 3*, 89-93.

Jin, J., Li, M., & Jin, L. (2015). Data Normalization to Accelerate Training for Linear Neural Net to Predict Tropical Cyclone Tracks. *Mathematical Problems in Engineering, 2015*, 1-8. doi:10.1155/2015/931629

João, G., & André, C. P. L. F. d. C. (2012). Machine Learning. In A. Information Resources Management (Ed.), *Machine Learning: Concepts, Methodologies, Tools and Applications* (pp. 13-22). Hershey, PA, USA: IGI Global.

Kacmarek, R. M. (2011). The Mechanical Ventilator: Past, Present, and Future. *Respiratory Care, 56*(8), 1170. doi:10.4187/respcare.01420

Kaisheng, Y., Baolin, P., Yu, Z., Dong, Y., Zweig, G., & Yangyang, S. (2014). Spoken language understanding using long short-term memory neural networks. In (pp. 189-194).

Kallet, R. H., & Branson, R. D. (2007). Do the NIH ARDS Clinical Trials Network PEEP/F_{IO₂} Tables Provide the Best Evidence-Based Guide to Balancing PEEP and F_{IO₂} Settings in Adults? *Respiratory Care, 52*(4), 461. Retrieved from http://rc.rcjournal.com/content/52/4/461.abstract

Kannangara, D. O., Newberry, F., Howe, S., Major, V., Redmond, D., Szlavecs, A., Chiew, Y. S., Pretty, C., Benyo, B., Shaw, G. M., & Chase, J. G. (2016). Estimating the true respiratory mechanics during asynchronous pressure controlled ventilation. *Biomedical Signal Processing and Control, 30*, 70-78. doi:10.1016/j.bspc.2016.06.014

Khamis, A., Ismail, Z., Khalid, H., & Tarmizi Mohammed, A. (2005). The Effects of Outliers Data on Neural Network Performance. *Journal of Applied Sciences, 5*(8), 1394-1398.

Kim, K. T., Knopp, J., Dixon, B., & Chase, G. (2019). Quantifying neonatal pulmonary mechanics in mechanical ventilation. *Biomedical Signal Processing and Control, 52*, 206-217. doi:https://doi.org/10.1016/j.bspc.2019.04.015

Kingma, D., & Ba, J. (2017). Adam: A Method for Stochastic Optimization. *arXiv.org*.

Knudsen, L., & Ochs, M. (2018). The micromechanics of lung alveoli: structure and function of surfactant and tissue components. *Histochemistry and Cell Biology, 150*(6), 661-676. doi:10.1007/s00418-018-1747-9

Kondili, E., Prinianakis, G., & Georgopoulos, D. (2003). Patient–ventilator interaction. *BJA: British Journal of Anaesthesia, 91*(1), 106-119. doi:10.1093/bja/aeg129

Kondili, E., Xirouchaki, N., & Georgopoulos, D. (2007). Modulation and treatment of patient–ventilator dyssynchrony. *Current Opinion in Critical Care, 13*(1), 84-89. doi:10.1097/MCC.0b013e328011278d

Kostic, P., Zannin E Fau - Andersson Olerud, M., Andersson Olerud M Fau - Pompilio, P. P., Pompilio Pp Fau - Hedenstierna, G., Hedenstierna G Fau - Pedotti, A., Pedotti A Fau - Larsson, A., Larsson A Fau - Frykholm, P., Frykholm P Fau - Dellaca, R. L., & Dellaca, R. L. (2011). Positive end-expiratory pressure optimization with forced oscillation technique reduces ventilator induced lung injury: a controlled experimental study in pigs with saline lavage lung injury. *Crit Care, 15*(1466-609X (Electronic)), R126. doi:D - NLM: PMC3218989 EDAT- 2011/05/18 06:00 MHDA- 2011/11/16 06:00 CRDT- 2011/05/18 06:00 PHST- 2011/02/20 00:00 [received] PHST- 2011/04/09 00:00 [revised] PHST- 2011/04/28 00:00 [accepted] PHST- 2011/05/18 06:00 [entrez] PHST- 2011/05/18 06:00 [pubmed] PHST- 2011/11/16 06:00 [medline] AID - cc10236 [pii] AID - 10.1186/cc10236 [doi] PST - ppublish

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet Classification with Deep Convolutional Neural Networks. *Association for Computing Machinery. Communications of the ACM, 60*(6), 84. doi:10.1145/3065386

KumarSingh, B., Verma, K., & Thoke, A. (2015). Investigations on Impact of Feature Normalization Techniques on Classifier's Performance in Breast Tumor Classification. *International Journal of Computer Applications, 116*, 11-15. doi:10.5120/20443-2793

Langille, J. J., & Brown, R. E. (2018). The Synaptic Theory of Memory: A Historical Survey and Reconciliation of Recent Opposition. *Frontiers in systems neuroscience, 12*, 52-52. doi:10.3389/fnsys.2018.00052

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278-2324. doi:10.1109/5.726791

Lee, H., Pham, P., Largman, Y., & Ng, A. Y. (2009). *Unsupervised feature learning for audio classification using convolutional deep belief networks.* Paper presented at the Advances in neural information processing systems.

Leon, M., & Lorini, L. (1997). Ventilation Mode Recognition Using Artificial Neural Networks. *Computers and biomedical research, an international journal, 30*(5), 373-378. doi:10.1006/cbmr.1997.1452

Li, D., Jinyu, L., Jui-Ting, H., Kaisheng, Y., Dong, Y., Seide, F., Seltzer, M., Zweig, G., Xiaodong, H., Williams, J., Yifan, G., & Acero, A. (2013). Recent advances in deep learning for speech research at Microsoft. In (pp. 8604-8608).

Li, M., Zhang, T., Chen, Y., & Smola, A. J. (2014). *Efficient mini-batch training for stochastic optimization*. Paper presented at the Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, New York, New York, USA. https://doi.org/10.1145/2623330.2623612

Liao, K.-M., Ou, C.-Y., & Chen, C.-W. (2011). Classifying Different Types of Double Triggering Based on Airway Pressure and Flow Deflection in Mechanically Ventilated Patients. *Respiratory Care, 56*(4), 460. doi:10.4187/respcare.00731

Lucangelo, U., Bernabè, F., & Blanch, L. (2007). Lung mechanics at the bedside: make it simple. *Current Opinion in Critical Care, 13*(1), 64-72. doi:10.1097/MCC.0b013e32801162df

Maas, A. L. (2013). *Rectifier Nonlinearities Improve Neural Network Acoustic Models*.

Major, V., Simon, C., Redmond, D., Beatson, A., Glassenbury, D., Chiew, Y. S., Pretty, C., Desaive, T., Szlavecz, A., Benyo, B., Shaw, G. M., & Chase, J. G. (2015). Assessing Respiratory Mechanics of Reverse-Triggered Breathing Cycles - Case Study of Two Mechanically Ventilated Patients. *IFAC-PapersOnLine, 48*(20), 505-510. doi:https://doi.org/10.1016/j.ifacol.2015.10.191

Major, V. J., Chiew, Y. S., Shaw, G. M., & Chase, J. G. (2018). Biomedical engineer's guide to the clinical aspects of intensive care mechanical ventilation. *BioMedical Engineering OnLine, 17*(1), 169. doi:10.1186/s12938-018-0599-9

Mao, X.-J., Shen, C., & Yang, Y.-B. (2016). Image Restoration Using Convolutional Auto-encoders with Symmetric Skip Connections.

Martos-Benítez, F. D., Domínguez-Valdés, Y., Burgos-Aragüez, D., Larrondo-Muguercia, H., Orama-Requejo, V., Lara-Ponce, K. X., & González-Martínez, I. (2020). Outcomes of

ventilatory asynchrony in patients with inspiratory effort. *Revista Brasileira de terapia intensiva, 32*(2), 284. doi:10.5935/0103-507x.20200045

Meade, M. O., Cook, D. J., Guyatt, G. H., Slutsky, A. S., Arabi, Y. M., Cooper, D. J., Davies, A. R., Hand, L. E., Zhou, Q., Thabane, L., Austin, P., Lapinsky, S., Baxter, A., Russell, J., Skrobik, Y., Ronco, J. J., & Stewart, T. E. (2008). Ventilation strategy using low tidal volumes, recruitment maneuvers, and high positive end-expiratory pressure for acute lung injury and acute respiratory distress syndrome: a randomized controlled trial. *Jama, 299*(6), 637-645. doi:10.1001/jama.299.6.637

Mellott, K. G., Grap, M. J., Munro, C. L., Sessler, C. N., & Wetzel, P. A. (2009). Patient-ventilator dyssynchrony: clinical significance and implications for practice. *Critical care nurse, 29*(6), 41-55. doi:10.4037/ccn2009612

Mellott, K. G., Grap, M. J., Munro, C. L., Sessler, C. N., Wetzel, P. A., Nilsestuen, J. O., & Ketchum, J. M. (2014). Patient ventilator asynchrony in critically ill adults: frequency and types. *Heart & lung : the journal of critical care, 43*(3), 231-243. doi:10.1016/j.hrtlng.2014.02.002

Mercat, A., Richard, J.-C. M., Vielle, B., Jaber, S., Osman, D., Diehl, J.-L., Lefrant, J.-Y., Prat, G., Richecoeur, J., Nieszkowska, A., Gervais, C., Baudot, J., Bouadma, L., Brochard, L., & Expiratory Pressure Study Group, f. t. (2008). Positive End-Expiratory Pressure Setting in Adults With Acute Lung Injury and Acute Respiratory Distress Syndrome: A Randomized Controlled Trial. *Jama, 299*(6), 646-655. doi:10.1001/jama.299.6.646

Mescheder, L., Geiger, A., & Nowozin, S. (2018). Which Training Methods for GANs do actually Converge? *arXiv.org*.

Moore, B. (1981). Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control, 26*(1), 17-32. doi:10.1109/TAC.1981.1102568

Moorhead, K., Piquilloud, L., Lambermont, B., Roeseler, J., Chase, J., Vignaux, L., Bialais, E., Tassaux, D., Jolliet, P., & Desaive, T. (2013). Patient-Ventilator Synchrony and Tidal Volume Variability using NAVA and Pressure Support Mechanical Ventilation Modes. *IFAC Proceedings Volumes, 44*(1), 569-574.

Murias, G., de Haro, C., & Blanch, L. (2016). Does this ventilated patient have asynchronies? Recognizing reverse triggering and entrainment at the bedside. *Intensive Care Medicine, 42*(6), 1058-1061. doi:10.1007/s00134-015-4177-3

Murias, G., Lucangelo U Fau - Blanch, L., & Blanch, L. (2016). Patient-ventilator asynchrony. *Curr Opin Crit Care, 22*(1531-7072 (Electronic)), 53-59. doi:10.1097/MCC.0000000000000270

Nagendran, M., Chen, Y., Lovejoy, C. A., Gordon, A. C., Komorowski, M., Harvey, H., Topol, E. J., Ioannidis, J. P. A., Collins, G. S., & Maruthappu, M. (2020). Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ, 368*, m689. doi:10.1136/bmj.m689

Newton, T. J., & Joyce, A. P. (2010). *Human Perspectives (5th ed.)*. South Melbourne, Australia: Cengage Learning Australia.

Nilsestuen, J. O., & Hargett, K. D. (2005). Using ventilator graphics to identify patient-ventilator asynchrony. *Respir Care, 50*(2), 202-234; discussion 232-204. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/15691392

Nwankpa, C., Ijomah, W., Gachagan, A., & Marshall, S. (2018). *Activation Functions: Comparison of trends in Practice and Research for Deep Learning*.

O'Shea, K., & Nash, R. (2015). *An Introduction to Convolutional Neural Networks*.

Ochs, M., Nyengaard, J. R., Jung, A., Knudsen, L., Voigt, M., Wahlers, T., Richter, J., & Gundersen, H. J. G. (2004). The number of alveoli in the human lung. *American journal of respiratory and critical care medicine, 169*(1), 120. doi:10.1164/rccm.200308-1107OC

Pappert, D., Rossaint, R., Slama, K., Grüning, T., & Falke, K. J. (1994). Influence of Positioning on Ventilation-Perfusion Relationships in Severe Adult Respiratory Distress Syndrome. *CHEST, 106*(5), 1511-1516. doi:10.1378/chest.106.5.1511

Park, J. J., Choi, D. J., Yoon, C. H., Oh, I. Y., Lee, J. H., Ahn, S., Yoo, B. S., Kang, S. M., Kim, J. J., Baek, S. H., Cho, M. C., Jeon, E. S., Chae, S. C., Ryu, K. H., & Oh, B. H. (2015). The prognostic value of arterial blood gas analysis in high‐risk acute heart failure patients: an analysis of the Korean Heart Failure (KorHF) registry. *European Journal of Heart Failure, 17*(6), 601-611. doi:10.1002/ejhf.276

Perchiazzi, G., Hedenstierna, G., Vena, A., Ruggiero, L., Giuliani, R., & Fiore, T. (2002). *Monitoring respiratory mechanics using artificial neural networks*.

Perchiazzi, G., Rylander, C., Pellegrini, M., Larsson, A., & Hedenstierna, G. (2017). Robustness of two different methods of monitoring respiratory system compliance during mechanical ventilation. *Medical & Biological Engineering & Computing, 55*(10), 1819-1828. doi:10.1007/s11517-017-1631-0

Phua, J., Badia, J. R., Adhikari, N. K. J., Friedrich, J. O., Fowler, R. A., Singh, J. M., Scales, D. C., Stather, D. R., Li, A., Jones, A., Gattas, D. J., Hallett, D., Tomlinson, G., Stewart, T. E., & Ferguson, N. D. (2009). Has mortality from acute respiratory distress syndrome decreased over time?: A systematic review. *American journal of respiratory and critical care medicine, 179*(3), 220. doi:10.1164/rccm.200805-722OC

Pretto, J. J., Roebuck, T., Beckert, L., & Hamilton, G. (2014). Clinical use of pulse oximetry: Official guidelines from the Thoracic Society of Australia and New Zealand. *Respirology, 19*(1), 38. doi:10.1111/resp.12204

Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks.

Ramirez, I., & Arellano, D. (2018). Identifying Patient-Ventilator Asynchrony Using Waveform Analysis. *Palliative Medicine & Care: Open Access, 4*(4), 1-4. doi:10.15226/2374-8362/4/4/00147

Ramirez, I. I., Arellano, D. H., Adasme, R. S., Landeros, J. M., Salinas, F. A., Vargas, A. G., Vasquez, F. J., Lobos, I. A., Oyarzun, M. L., & Restrepo, R. D. (2017, 2017/02//). Ability of ICU health-care professionals to identify patient-ventilator asynchrony using waveform analysis. *Respiratory Care, 62(2),* 144+. Retrieved from https://link.gale.com/apps/doc/A480594296/AONE?u=monash&sid=AONE&xid=6df8a787

Räsänen, J., & León, M. A. (1998). Detection of lung injury with conventional and neural network-based analysis of continuous data. *J Clin Monit Comput, 14*(6), 433-440.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016, 27-30 June 2016). *You Only Look Once: Unified, Real-Time Object Detection.* Paper presented at the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Redmond, D. P., Major, V., Corbett, S., Glassenbury, D., Beatson, A., Á, S., Chiew, Y. S., Shaw, G. M., & Chase, J. G. (2014, 8-10 Dec. 2014). *Pressure reconstruction by eliminating the demand effect of spontaneous respiration (PREDATOR) method for assessing respiratory mechanics of reverse-triggered breathing cycles.* Paper presented at the 2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES).

Ricard, J. D., Dreyfuss, D., & Saumon, G. (2003). Ventilator-induced lung injury. *European Respiratory Journal, 22*(42 suppl), 2s-9s. doi:10.1183/09031936.03.00420103

Robbins, H., & Monro, S. (1951). A Stochastic Approximation Method. *The Annals of Mathematical Statistics, 22*(3), 400-407.

Robinson, B. R. H., Blakeman, T. C., Toth, P., Hanseman, D. J., Mueller, E., & Branson, R. D. (2013). Patient-Ventilator Asynchrony in a Traumatically Injured Population. *Respiratory Care, 58*(11), 1847. doi:10.4187/respcare.02237

Rolland-Debord, C., Bureau, C., Poitou, T., Belin, L., Clavel, M., Perbet, S., Terzi, N., Kouatchet, A., Similowski, T., & Demoule, A. (2017). Prevalence and Prognosis Impact of Patient–Ventilator Asynchrony in Early Phase of Weaning according to Two Detection Methods. *Anesthesiology, 127*(6), 989-997. doi:10.1097/aln.0000000000001886

Rossi, A., Santos, C., Roca, J., Torres, A., Félez, M. A., & Rodriguez-Roisin, R. (1994). Effects of PEEP on VA/Q mismatching in ventilated patients with chronic airflow obstruction. *Am J Respir Crit Care Med, 149*(5), 1077-1084. doi:10.1164/ajrccm.149.5.8173744

Roussos, C., & Koutsoukou, A. (2003). Respiratory failure. *European Respiratory Journal, 22*(47 suppl), 3s. doi:10.1183/09031936.03.00038503

Rué, M., Andrinopoulou, E. R., Alvares, D., Armero, C., Forte, A., & Blanch, L. (2017). Bayesian joint modeling of bivariate longitudinal and competing risks data: An application to study patient‐ventilator asynchronies in critical care patients. *Biometrical Journal, 59*(6), 1184-1203. doi:10.1002/bimj.201600221

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016). Improved Techniques for Training GANs. *arXiv.org*.

Sassoon, C. S., & Foster, G. T. (2001). Patient-ventilator asynchrony. *Curr Opin Crit Care, 7*(1), 28-33. doi:10.1097/00075198-200102000-00005

Scherer, D., Müller, A., & Behnke, S. (2010). Evaluation of pooling operations in convolutional architectures for object recognition. *Artificial Neural Networks–ICANN 2010*, 92-101.

Sidey-Gibbons, J. A. M., & Sidey-Gibbons, C. J. (2019). Machine learning in medicine: a practical introduction. *BMC Medical Research Methodology, 19*(1), 64. doi:10.1186/s12874-019-0681-4

Sinderby, C., Liu, S., Colombo, D., Camarotta, G., Slutsky, A. S., Navalesi, P., & Beck, J. (2013). An automated and standardized neural index to quantify patient-ventilator interaction. *Crit Care, 17*(5), R239. doi:10.1186/cc13063

Sinderby, C., Navalesi, P., Beck, J., Skrobik, Y., Comtois, N., Friberg, S., Gottfried, S. B., & Lindstrom, L. (1999). Neural control of mechanical ventilation in respiratory failure. *Nat Med, 5*(12), 1433-1436. doi:10.1038/71012

Sinderby, C. A., & Beck, J. (2012). Patient–Ventilator Interactions. In J.-L. Vincent & J. B. Hall (Eds.), *Encyclopedia of Intensive Care Medicine* (pp. 1676-1690). Berlin, Heidelberg: Springer Berlin Heidelberg.

Sinderby, C. A., Beck, J. C., Lindstrom, L. H., & Grassino, A. E. (1997). Enhancement of signal quality in esophageal recordings of diaphragm EMG. *J Appl Physiol (1985), 82*(4), 1370-1377. doi:10.1152/jappl.1997.82.4.1370

Singh, P., Borle, A., & Trikha, A. (2014). Newer nonconventional modes of mechanical ventilation.(Review Article)(Report). *Journal of Emergencies, Trauma, and Shock, 7*(3), 222. doi:10.4103/0974-2700.136869

Sinno Jialin, P., & Qiang, Y. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering, 22*(10), 1345-1359. doi:10.1109/TKDE.2009.191

Slutsky, A. S. (2015). History of Mechanical Ventilation. From Vesalius to Ventilator-induced Lung Injury. *American journal of respiratory and critical care medicine, 191*(10), 1106-1115. doi:10.1164/rccm.201503-0421PP

Slutsky, A. S., & Ranieri, V. M. (2013). Ventilator-induced lung injury. *New England Journal of Medicine, 369*(22), 2126-2136. doi:10.1056/NEJMra1208707

Sola, J., & Sevilla, J. (1997). Importance of input data normalization for the application of neural networks to complex industrial problems. *IEEE transactions on nuclear science, 44*(3), 1464-1468. doi:10.1109/23.589532

Sood, P., Paul, G., & Puri, S. (2010). Interpretation of arterial blood gas. *Indian journal of critical care medicine : peer-reviewed, official publication of Indian Society of Critical Care Medicine, 14*(2), 57-64. doi:10.4103/0972-5229.68215

Souza Leite, W., Novaes, A., Bandeira, M., Olympia Ribeiro, E., Dos Santos, A. M., de Moura, P. H., Morais, C. C., Rattes, C., Richtrmoc, M. K., Souza, J., Correia de Lima, G. H., Pinheiro Modolo, N. S., Gonçalves, A. C. E., Ramirez Gonzalez, C. A., Do Amparo Andrade, M., Dornelas De Andrade, A., Cunha Brandão, D., & Lima Campos, S. (2020). Patient-ventilator asynchrony in conventional ventilation modes during short-term mechanical ventilation after cardiac surgery: randomized clinical trial. *Multidisciplinary respiratory medicine, 15*(1), 650. doi:10.4081/mrm.2020.650

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). *Dropout: A Simple Way to Prevent Neural Networks from Overfitting* (Vol. 15).

Stauffer, J. L., Olson, D. E., & Petty, T. L. (1981). Complications and consequences of endotracheal intubation and tracheotomy: A prospective study of 150 critically ill adult patients. *The American journal of medicine, 70*(1), 65-76. doi:10.1016/0002-9343(81)90413-7

Stock, C. M., Downs, B. J., & Frolicher, A. D. (1987). Airway pressure release ventilation. *Critical care medicine, 15*(5), 462-466. doi:10.1097/00003246-198705000-00002

Subira, C., de Haro, C., Magrans, R., Fernandez, R., & Blanch, L. (2018). Minimizing Asynchronies in Mechanical Ventilation: Current and Future Trends.(Report). *Respiratory Care, 63*(4), 464. doi:10.4187/respcare.05949

Sun, S., Cao, Z., Zhu, H., & Zhao, J. (2020). A Survey of Optimization Methods From a Machine Learning Perspective. *IEEE Transactions on Cybernetics, 50*(8), 3668-3681. doi:10.1109/TCYB.2019.2950779

Sundaresan, A., Chase, J. G., Shaw, G. M., Chiew, Y. S., & Desaive, T. (2011). Model-based optimal PEEP in mechanically ventilated ARDS patients in the Intensive Care Unit. *BioMedical Engineering OnLine, 10*(1), 64. doi:10.1186/1475-925x-10-64

Szegedy, C., Wei, L., Yangqing, J., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015, 7-12 June 2015). *Going deeper with convolutions.* Paper presented at the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Szlavecz, A., Chiew, Y. S., Redmond, D., Beatson, A., Glassenbury, D., Corbett, S., Major, V., Pretty, C., Shaw, G. M., Benyo, B., Desaive, T., & Chase, J. G. (2014). The Clinical Utilisation of Respiratory Elastance Software (CURE Soft): a bedside software for real-time respiratory mechanics monitoring and mechanical ventilation management. *BioMedical Engineering OnLine, 13*, 140-140. doi:10.1186/1475-925X-13-140

Talmor, D., Sarge, T., Malhotra, A., O'Donnell, C. R., Ritz, R., Lisbon, A., Novack, V., & Loring, S. H. (2008). Mechanical ventilation guided by esophageal pressure in acute lung injury.(Clinical report). *The New England journal of medicine, 359*(20), 2095. doi:10.1056/NEJMoa0708638

Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). A Survey on Deep Transfer Learning. *arXiv.org*.

Team, E. D. j. D. (2016). DL4J: Deep Learning for Java. Retrieved from https://github.com/eclipse/deeplearning4j

Thanh-Tung, H., & Tran, T. (2020). On Catastrophic Forgetting and Mode Collapse in Generative Adversarial Networks. *arXiv.org*.

The Theano Development, T., Al-Rfou, R., Alain, G., Almahairi, A., Angermueller, C., Bahdanau, D., Ballas, N., Bastien, F., Bayer, J., Belikov, A., Belopolsky, A., Bengio, Y., Bergeron, A., Bergstra, J., Bisson, V., Snyder, J. B., Bouchard, N., Boulanger-Lewandowski, N., Bouthillier, X., de Brébisson, A., Breuleux, O., Carrier, P.-L., Cho, K., Chorowski, J., Christiano, P., Cooijmans, T., Côté, M.-A., Côté, M., Courville, A., Dauphin, Y. N., Delalleau, O., Demouth, J., Desjardins, G., Dieleman, S., Dinh, L., Ducoffe, M., Dumoulin, V., Kahou, S. E., Erhan, D., Fan, Z., Firat, O., Germain, M., Glorot, X., Goodfellow, I., Graham, M., Gulcehre, C., Hamel, P., Harlouchet, I., Heng, J.-P., Hidasi, B., Honari, S., Jain, A., Jean, S., Jia, K., Korobov, M., Kulkarni, V., Lamb, A., Lamblin, P., Larsen, E., Laurent, C., Lee, S., Lefrancois, S., Lemieux, S., Léonard, N., Lin, Z., Livezey, J. A., Lorenz, C., Lowin, J., Ma, Q., Manzagol, P.-A., Mastropietro, O., McGibbon, R. T., Memisevic, R., van Merriënboer, B., Michalski, V., Mirza, M., Orlandi, A., Pal, C., Pascanu, R., Pezeshki, M., Raffel, C., Renshaw, D., Rocklin, M., Romero, A., Roth, M., Sadowski, P., Salvatier, J., Savard, F., Schlüter, J., Schulman, J., Schwartz, G., Serban, I. V., Serdyuk, D., Shabanian, S., Simon, É., Spieckermann, S., Subramanyam, S. R., Sygnowski, J., Tanguay, J., & van Tulder, G. (2016). *Theano: A Python framework for fast computation of mathematical expressions*.

Thille, A. W., Rodriguez, P., Cabello, B., Lellouche, F., & Brochard, L. (2006). Patient-ventilator asynchrony during assisted mechanical ventilation. *Intensive Care Med, 32*(10), 1515-1522. doi:10.1007/s00134-006-0301-8

Thille, W. A., & Brochard, W. L. (2007). Promoting Patient-Ventilator Synchrony. *Clinical Pulmonary Medicine, 14*(6), 350-359. doi:10.1097/CPM.0b013e31815b1234

Thompson, B. T., Chambers, R. C., & Liu, K. D. (2017). Acute Respiratory Distress Syndrome. *The New England journal of medicine, 377*(6), 562. doi:10.1056/NEJMra1608077

Tusman, G., Bohm, S. H., Vazquez de Anda, G. F., do Campo, J. L., & Lachmann, B. (1999). 'Alveolar recruitment strategy' improves arterial oxygenation during general anaesthesia. *Br J Anaesth, 82*(1), 8-13. doi:10.1093/bja/82.1.8

U POR, K. A. S. (2011). *Investigation of Normalization Techniques and Their Impact on a Recognition Rate in Handwritten Numeral Recognition*.

van der Staay, M., & Chatburn, R. L. (2018). Advanced modes of mechanical ventilation and optimal targeting schemes. *Intensive Care Medicine Experimental, 6*(1), 30-30. doi:10.1186/s40635-018-0195-0

Vaschetto, R., Cammarota, G., Colombo, D., Longhini, F., Grossi, F., Giovanniello, A., Della Corte, F., & Navalesi, P. (2014). Effects of Propofol on Patient-Ventilator Synchrony and Interaction During Pressure Support Ventilation and Neurally Adjusted Ventilatory Assist*. *Critical care medicine, 42*(1), 74-82. doi:10.1097/CCM.0b013e31829e53dc

Vezhnevets, A., & Barinova, O. (2007). *Avoiding Boosting Overfitting by Removing Confusing Samples*, Berlin, Heidelberg.

Villar, J., Blanco, J., Añón, J., Santos-Bouza, A., Blanch, L., Ambrós, A., Gandía, F., Carriedo, D., Mosteiro, F., Basaldúa, S., Fernández, R., & Kacmarek, R. (2011). The ALIEN study: incidence and outcome of acute respiratory distress syndrome in the era of lung protective ventilation. *Intensive Care Medicine, 37*(12), 1932-1941. doi:10.1007/s00134-011-2380-4

Wang, Y., Yao, H., & Zhao, S. (2015). Auto-Encoder Based Dimensionality Reduction. *Neurocomputing, 184*. doi:10.1016/j.neucom.2015.08.104

Wei-Yang, L., Ya-Han, H., & Chih-Fong, T. (2012). Machine Learning in Financial Crisis Prediction: A Survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 42*(4), 421-436. doi:10.1109/TSMCC.2011.2170420

Xu, B., Wang, N., Chen, T., & Li, M. (2015). Empirical Evaluation of Rectified Activations in Convolutional Network.

Yeh, R. A., Chen, C., Yian Lim, T., Schwing, A. G., Hasegawa-Johnson, M., & Do, M. N. (2017). *Semantic image inpainting with deep generative models.* Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Ying, X. (2019). *An Overview of Overfitting and its Solutions.* Paper presented at the Journal of Physics: Conference Series.

Yonis, H., Crognier, L., Conil, J.-M., Serres, I., Rouget, A., Virtos, M., Cougot, P., Minville, V., Fourcade, O., & Georges, B. (2015). Patient-ventilator synchrony in Neurally Adjusted Ventilatory Assist (NAVA) and Pressure Support Ventilation (PSV): a prospective observational study. *BMC anesthesiology, 15*(1), 117. doi:10.1186/s12871-015-0091-z

Yoshida, T., Nakamura, M. A. M., Morais, C. C. A., Amato, M. B. P., & Kavanagh, B. P. (2018). Reverse Triggering Causes an Injurious Inflation Pattern during Mechanical Ventilation. *American journal of respiratory and critical care medicine, 198*(8), 1096-1099. doi:10.1164/rccm.201804-0649LE

Zeiler, M. (2012). ADADELTA: An Adaptive Learning Rate Method. *arXiv.org*.

Zeiler, M. D., & Fergus, R. (2014, 2014//). *Visualizing and Understanding Convolutional Networks.* Paper presented at the Computer Vision – ECCV 2014, Cham.

Zhang, L., Mao, K., Duan, K., Fang, S., Lu, Y., Gong, Q., Lu, F., Jiang, Y., Jiang, L., Fang, W., Zhou, X., Wang, J., Fang, L., Ge, H., & Pan, Q. (2020). Detection of patient-ventilator asynchrony from mechanical ventilation waveforms using a two-layer long short-term memory neural network. *Computers in Biology and Medicine, 120*, 103721. doi:https://doi.org/10.1016/j.compbiomed.2020.103721

Zhang, Y. (2018). *A Better Autoencoder for Image: Convolutional Autoencoder.* Paper presented at the ICONIP17-DCEC. Available online: http://users. cecs. anu. edu. au/Tom. Gedeon/conf/ABCs2018/paper/ABCs2018_paper_58. pdf (accessed on 23 March 2017).