Ph.D. Dissertation of Engineering

Exploring the Application Studies of Artificial Intelligence for Tabular and Image Data - Insights from Radiology, Taxation, and Security Detection Research Fields -

정형 및 이미지 데이터에 관한 인공지능 응용 연구
- 영상 의학, 세금 징수, 보안 탐지 분야를 중심으로 -

August 2023

Program in Dept. of Intelligence and Information
Graduate School of Convergence Science and Technology
Seoul National University

Eunchan Kim

Exploring the Application Studies of Artificial Intelligence for Tabular and Image Data - Insights from Radiology, Taxation, and Security Detection Research Fields -

Advisor Byungjoon Yoo

Submitting a Ph.D. Dissertation of Engineering

August 2023

Program in Dept. of Intelligence and Information
Graduate School of Convergence Science and Technology
Seoul National University

Eunchan Kim

Confirming the Ph.D. Dissertation written by
Eunchan Kim

August 2023

Chair	Jong Woo Choi	(Seal)
Vice Chair	Byungjoon Yoo	(Seal)
Examiner	Gahgene Gweon	(Seal)
Examiner	Jiwoong Choi	(Seal)
Examiner	Yeunwoong Kyung	(Seal)

Abstract

In recent years, the application and implementation of artificial intelligence (AI) has been rapidly growing and expanding in various industrial and research fields. Many scholars have confirmed the active adoption of AI technology in agriculture, manufacturing, education, finance, administration, medicine, security, and other study areas. This dissertation aims to validate these trends and present three studies on the application of AI in the fields of radiology, taxation, and security detection, which are the sub-fields of medical, financial, and security industries respectively. The three AI application studies presented in the dissertation are conducted using tabular and image data. The tabular data-based studies include 1) chronic obstructive pulmonary disease (COPD)-related pulmonary function prediction and 2) local tax delinquents prediction research, while 3) object detection research on airport baggage X-ray images is presented as image data-based research. The goal of this dissertation is to identify research gaps and opportunities in AI applied research by conducting a literature review of the current state of artificial intelligence applied research in both tabular and image data research, as well as specific fields of each of these three studies.

Ultimately, this dissertation provides insights into the methodologies for AI applied research and comparatively analyzes the AI applied research identified through conducting a broad literature review and presenting three research findings. The research verifies inconsistencies and differences in AI application research conducted across each field, and suggests that there is a relatively greater opportunity for the newer research topics. Furthermore, the research identifies machine learning, deep learning, and improvement of deep learning as the main degrees of AI application techniques, and confirms that there are more research opportunities in the early stages of research, such as, introducing machine learning. The research framework presented in this thesis, derived through these academic distinction processes, would serve as a theoretical basis for establishing AI application research academically in the future. The findings provide significant insights for researchers and industry experts in AI applications, as AI technologies are increasingly used in a variety of fields and are expected to play a more important role in the future. Additionally, this dissertation suggests future research directions in the AI application field by identifying the limitations of research based on these research results, and emphasizes the need for continuous efforts to develop and systematize AI technology that can be used in various fields.

Keywords: Applications of artificial intelligence, information systems engineering and management, tabular and image data, radiology, taxation, security detection

Student Number: 2019-31312

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background and Motivation

Recently, artificial intelligence (AI) has become an increasingly popular research field, with applications in various academic fields and industries. The use of AI in medical, financial, and security-related fields has a potential to improve efficiency, accuracy, and decision-making processes (Karthick and Gopalsamy, 2022; Paul et al., 2023). Radiology, tax arrear, and security detection fields are among the areas where AI applications have yielded promising results.

According to Zhang and Lu (2021), the ability to process large amounts of data and identify patterns through machine learning (ML) algorithms makes AI an essential tool for analysis and decision-making in related tasks. The motivation behind this study is to explore the potential of AI applications for tabular and image data in the radiology, tax arrear, and security detection fields. The aim of this study is to examine the current research, identify gaps, and explore opportunities for future research, thereby contributing to the development of AI applications in these fields.

Accordingly, the purpose of this study is to investigate and explore the potential of AI applications for tabular and image data analysis in the radiology, taxation, and security detection fields. Because of the exhaustive review of existing research, identification of gaps, and exploration of potential future applications of AI, the findings of this research will have an impact on the further development of these technologies and their potential applications. This is of vital importance, as the increased use of AI in these fields has the potential to considerably improve decision-making and analysis accuracy.

For this purpose, the focus of this study is on the application of AI for tabular and image data, analyzing how its implementation affects the development of decision-making in these domains. Additionally, the advantages and disadvantages of AI implementation, as well as the current limitations, are analyzed and discussed. All these elements are discussed in detail and contextualized in the research on radiology, tax arrears, and security detection, to ensure an adequate representation of their corresponding needs and expectations from AI applications. Generally, this research seeks to expand the existing understanding laid out by the current research and create new knowledge regarding the potential applications of AI in these three industries. This would substantially benefit both the AI industry and fields it is applied in, potentially improving the accuracy of decision-making in the fields of radiology, taxation, and security detection.

1.2 Research Questions and Objectives

The research questions of this dissertation are centered around the understanding on the current applications of AI for tabular and image data in the fields of radiology, tax arrears, and security detection. Through addressing these research questions, this dissertation aims to uncover any new possibilities of applying AI in tabular and image data analysis, as well as to suggest possible research areas to better leverage the potential of AI for data processing in the given research fields. Accordingly, the research questions are as follows:

RQ1. What are the current AI applications for tabular and image data in the radiology, tax arrears, and security detection fields?

RQ2. What are the commonalities and differences in these fields?

RQ3. What are the gaps and opportunities for further research?

The purpose of this study is to investigate the current applications of AI for tabular (structured) and image data analysis in the fields of radiology, taxation, and security detection. The commonalities and differences between these areas are revealed, and gaps and opportunities for further research are identified. First, various AI technologies applied in each of the three fields and their corresponding roles are examined. Then, the successes and shortcomings of these implementations are assessed based on their accuracy and effectiveness. An overall understanding on the utilization of AI for tabular and image data is provided. The assessment of commonalities and differences in these fields will further help to identify the unique features of AI in each area. It will also serve as the basis for detecting opportunities for innovation and suggesting possible future AI implementations in these areas. Furthermore, in this study, the differences and similarities between each field are investigated in terms of the data types and techniques used, technological complexity of tasks, accuracy of outcomes, social and ethical implications of applications, and potential bias in the algorithms used. Finally, the gaps in the research in each of these fields, which require further exploration and investigation of potential solutions, are highlighted. The objectives of this study are as follows:

1. To establish an overview of AI applications in the radiology, tax arrear, and security detection fields concerning both tabular and image data.
2. To examine the commonalities and differences between the three research fields regarding their applications of AI for tabular and image data.
3. To investigate the potential opportunities for further AI research in the radiology, tax arrears, and security detection fields.

1.3 Contribution and Significance

This study seeks to examine the current use and potential of AI and ML applications in radiology, tax arrears, and security detection. By conducting a thorough comparison of research findings in these three fields, this study aims to highlight the existing opportunities for applying AI technologies and ML algorithms in various fields. The research reviewed in this study identified several benefits associated with AI and ML applications, such as improved accuracy and efficiency in decision-making. Furthermore, it discussed the potential problems and areas of concern, providing insights into the potential opportunities and possible gaps in AI applications in these fields.

The results are expected to have great significance, both for research and development efforts, as well as in industry and professional practice. By conducting an extensive review of existing literature and identifying existing gaps, this study seeks to add to the knowledge base by highlighting the potential of AI in various domains. Furthermore, it provides a general roadmap for research and industry professionals looking to adopt and deploy AI and ML practices in their respective domains. This study seeks to provide a better understanding of the implications of AI and ML algorithms for radiology, tax arrears, and security detection. In particular, this research focuses on examining the possible strengths and weaknesses of different algorithms and how they can be best leveraged for applications in different domains. Overall, this study contributes to the body of knowledge related to AI and ML algorithms, offering valuable insight into their potential to solve complex problems and improve decision-making processes.

1.4 Scope and Limitations

This study aims to explore the applications of AI in the fields of radiology, tax arrears, and security detection to advance the knowledge and understanding of the respective use cases. Specifically, this study focuses on the use of AI to process tabular and image data, while other aspects of AI, such as natural language processing, robotics, or gaming, are beyond the scope of this research.

The scope and breadth of this research are limited by the availability and quality of the data used to analyze the various applications of AI in these domains. While extensive research is being conducted in the aforementioned fields, much of the existing research does not meet the rigor or completeness that would be desirable for obtaining a comprehensive understanding of AI's potential uses. Consequently, the findings from this research are likely to be more exploratory than definitive.

Another limitation of this research is its scope. While it does not attempt to include every branch of AI applications, it does focus exclusively on the three chosen domains, thus restricting the potential for making broader statements about AI in general. Furthermore, any general conclusions drawn from this research will likely only apply to the limited domains explored and similar areas of application. Given the scope and limitations of this research, it is hoped that its results can help provide greater insight into the potential of AI and contribute to achieving an increased understanding of the many possible use cases for AI in various fields and industries.

1.5 Structure of the Dissertation

This dissertation is structured into five chapters. Chapter 1 provides an introduction to the study, including the background, motivation, research questions and purpose, contribution and significance, research scope and limitations, and structure of the paper. Chapter 2 presents a literature review of AI applications, including the predictive research based on tabular data, research on object detection based on image data, prior research on AI applications, criteria and selection order of applied research cases, and current research in the radiology, tax arrears, and security detection fields. Additionally, this chapter identifies the research gaps and opportunities.

Chapter 3 presents the first study conducted for implementing AI with tabular data, with a focus on the prediction of a key metric named relative regional air volume change (RRAVC) related to respiratory ailments using numerical values extracted from radiographic images. The chapter outlines the study's design, data collection and preprocessing methods, and experimental results. It also discusses the limitations and future directions of the study. Chapter 4 describes the second study conducted for implementing AI with tabular data, to predict tax delinquency among local taxpayers. The chapter discusses the study's design, data collection and preprocessing methods, experimental results, limitations, and future research directions. Chapter 5 presents the third study, which focuses on using AI for image detection in security applications, specifically in identifying objects using X-ray in airport baggage control. The chapter outlines the study, related data collection approaches and methods, experimental results, limitations, and future research directions.

Chapter 6 offers a comparative analysis of the three AI studies, highlighting the commonalities, differences, and attributes in the AI techniques and methodologies implemented in these studies. Then, the chapter concludes the dissertation by summarizing the main findings and contributions of this research. Finally, the chapter discusses the implications and limitations of the study, as well as potential future research directions in the field of AI applications.

Chapter 2

Literature Review

2.1 Overview of AI Application Implemented in Research

The implementation of AI in the computer science field has improved significantly over the years. One concept that cannot be excluded from the discussion is data, which can be split into two broad categories: structured data (or tabular data) and unstructured data, such as images, text, and audio. According to Paschen et al. (2020), while AI research primarily focused on unstructured data, there has recently been a push to also explore the capabilities of AI on structured data. This chapter will comprehensively examine this expansion from structured to unstructured data, and how AI can be used in prediction and detection research in the three main domains, namely, radiology, tax arrears, and security detection. It is important to examine the existing literature in the field to inform and gain insights on current trends and developments in this realm of AI research. To facilitate this understanding, the order and criteria of the selected research cases and research outputs in the aforementioned domains are identified to be subsequently discussed in this review. Through understanding these findings, a clearer perspective on potential gaps, successes, and improvements in current technology can be achieved.

By looking into radiology, tax arrears, and security detection, this chapter will provide an insight into the current uses of AI on tabular data, as well as an understanding of prediction and detection methods used with unstructured data. Thereafter, the limitations and discrepancies in their development stages and corresponding opportunities for research will be analyzed and discussed. The selection order of the research cases, which is tailored to identify the areas of strengths and opportunities, is carefully defined as well. The focus of this literature review is to equip the readers with an overall basic understanding of the main uses and developments in tabular and unstructured data related to AI research, as well as to identify differences in development levels and opportunities for future research in the three domains. In addition, the review will evaluate the research outputs in these fields. As this chapter takes a thorough look at both prediction and detection methods and their current progress, it will conclude with potential improvements and new avenues of research and paths to reach them.

2.1.1 Research on Tabular Data

The application of predictive research based on tabular data is particularly relevant in modern world. By taking advantage of ML algorithms and data analytics, various industries, such as healthcare and finance, can gain valuable insight from vast datasets and make more informed decisions (Stiglic et al., 2020). Currently, many researchers are focusing on the analysis of tabular data and use of various algorithms to facilitate this process, such as tree-based, shallow methods, and deep learning algorithms (Cannarile et al., 2022; Meddage et al., 2022). However, due to the special attributes of tabular datasets, a universal solution to analyze this data type has not been found yet. According to Raschka et al. (2020), tabular data analysis is an important tool for ML and data analytics, allowing for complex predictions to be made from significant amounts of structured data. To achieve this, researchers have used various ML algorithms, including tree-based and shallow methods to deep learning algorithms. Although advances in the field of deep learning have enabled more effective predictions from image and audio data, clear performance improvements in the case of tabular data are lacking (Zhou et al., 2017; Grinsztajn et al., 2022). This suggests that algorithms should be tailored to the unique characteristics of the tabular data to get the most out of the information.

The prevailing opinion is that there is no field that deep learning cannot penetrate. Deep neural network models, including neural networks and their derivatives, have been evaluated to be significantly efficient and effective in various domains involving text, audio, and image (Brown et al., 2020; Cheng et al., 2022; and Zhang et al., 2020). However, tabular data have been evaluated as a field that has not been thoroughly explored (Kadra et al., 2021; Arik and Pfister, 2019). Research on tabular data is ongoing and evolving but many challenges remain in processing and analyzing this type of data at large scales. One reason for this is that tabular data often uses more complex and structured formats with many different categories and values that can be difficult to organize and process effectively. Tabular data can also be more difficult to explore than image data, as it often requires a deeper understanding of the underlying relationships and patterns in the data. ML algorithms can be trained to analyze and classify tabular data but they are not always able to identify the underlying relational and causal factors driving patterns in such data.

Certainly, deep learning is an extremely powerful method and solution for tabular data as well but it cannot be used to solve every problem. According to some researchers, most tabular data models, which have recently received attention from the research community, have exhibited performances inferior to that of XGBoost (Chen and Guestrin, 2016) and LightGBM (Ke et al., 2017) with significantly less tuning (Shwartz-Ziv and Armon, 2022). It has even been argued that deep learning is not an absolute necessity for tabular data (Shwartz-Ziv and Armon, 2022). Several other studies (Prokhorenkova et al., 2018; Chen and Guestrin, 2016; and Friedman, 2001) have supported the claims of Shwartz-Ziv and his colleague. These studies have supported the use and consideration of tree-based ensemble algorithms including XGBoost and LightGBM for cases involving real-life tabular datasets. Therefore, there remain numerous challenges and opportunities for further development and innovation in tabular data analysis.

Tree-based methods have become popular owing to their ability to determine relationships between data points quickly and easily. For example, Dabiri et al. (2022) stated that random forests are often used for tabular data analysis because of their effectiveness in data mining. Moreover, Dabiri et al. (2022) showed that tree-based methods function by training decision trees from various sub-samples of the dataset and then combining the individual trees into a single predictive model. Despite their wide usage, research has shown that tree-based methods are not always the most reliable solution for tabular data analysis (Chen and Guestrin, 2016; Santos et al., 2022). DeLancey et al. (2019) also showed that shallow learning algorithms have been employed in various ML tasks. One popular algorithm for tabular data analysis, according to Chui et al. (2020), is the support vector machine (SVM), which works by creating an optimal hyperplane that separates different categories of data points. Therefore, this algorithm is known to be effective at discovering complex patterns in data. However, its accuracy can be limited by its lack of a deep structure.

Deep learning algorithms have become a prominent tool in the field of ML in recent years. Zhang et al. (2021) indicated that deep learning algorithms are suited for image recognition and text analysis; however, research has demonstrated that they are less effective for tabular data analysis. One reason for this is that the data entropy is usually too low to exploit the full potential of deep learning algorithms. Research also suggests that there may be several factors preventing deep learning algorithms from performing at the highest level when dealing with tabular data (Fang et al., 2020). Fang et al. indicated that the limited progress made in the area of tabular data analysis, lack of adequate data, and lack of sufficient entropy in the data all seemed to contribute to the limited performance of deep learning algorithms in this context.

To help researchers understand the nuances of tabular data, research has explored this field from various perspectives. This has included using regression models for the analysis of predictive models, making comparisons between supervised and unsupervised learning algorithms, and exploring how transfer learning can improve tabular data prediction (Ahmad and Chen, 2020). These efforts help researchers better understand the unique requirements of tabular data and further advance the application of AI in tabular data analysis. Research into tabular data and its associated predictive techniques has yielded a better understanding of the data and various algorithms that can be used to extract insights from them. By further exploring the field and potential for transfer learning, researchers can help unlock the potential of deep learning algorithms to gain more insights related to tabular data analysis (Loey et al., 2021). However, to truly unlock this potential, there is a need for further research with domain knowledge and greater understanding of the nuances of tabular data and its associated algorithms.

Currently, various algorithms can be used for tabular data analysis. Tree-based methods, shallow algorithms, and deep learning models are all common approaches used to handle tabular datasets. Unfortunately, Zhang et al. (2021) indicated deep learning algorithms have not yet proven to be more effective in tabular data analysis than their shallow counterparts. One reason for this could be the limited progress made by deep learning in this field. The properties of tabular data require that the algorithms used to process such data must be tailored to the data characteristics. As such, a generalized algorithm that could effectively handle all tabular data types is yet to be developed. Furthermore, some experts claim that the lack of data is another key factor that affects the performance of deep learning models in tabular data analysis. This is because tabular data requires a high amount of entropy to be effectively leveraged by the models (Santos et al., 2022). Therefore, if there is insufficient data or data is not of high quality, deep learning models are not likely to yield good results. For these reasons, further research is needed to gain a better understanding of the characteristics of tabular data and how to effectively use deep learning algorithms in this field. This would improve the accuracy and efficiency of tabular data analysis and help advance the field of deep learning in general.

Generally, the literature on predictive research based on tabular data is extensive, and researchers continue to investigate various approaches to effectively utilize ML algorithms for tabular data analysis. Although tree-based methods and shallow learning algorithms are currently the most effective approaches for tabular data analysis, deep learning algorithms are also making significant progress and have significant potential for future applications. It is hoped that with further research, a universal solution for tabular data analysis can be obtained, improving the accuracy of ML models and advancing the field in general.

2.1.2 Research on Image Data

1) Overview

Research on object detection based on image data has advanced significantly recently with the implementation of deep learning algorithms such as convolutional neural networks (CNNs). According to Krizhevsky et al. (2017) and Song et al. (2020), CNNs can process images quickly and accurately by learning from visual information. This is an important advancement as it allows for real-time detection and tracking of objects. In addition, Masita et al. (2020) indicated that, owing to their capability of extracting information from large amounts of data, CNNs are a useful tool for developing advanced image recognition applications. In security, surveillance, and autonomous driving, CNNs have been utilized to improve object detection accuracy (Sanil et al., 2020). To identify objects, deep learning models use convolutional layers that process images in a pixel-by-pixel manner and extract the necessary features. Consequently, the model understands and correctly classifies the object. Therefore, different strategies, such as transfer learning, can be used to enhance the detection accuracy by fine-tuning the model with existing training data.

Recently, using 3D vision systems combined with deep learning models has been proposed as an improved method for object detection. Bai et al. (2020) stated that 3D systems capture information about the spatial relationship between objects contained in the image, allowing for better object recognition and localization. This technique can be further enhanced by fusing depth data with CNNs for improving the detection performance. In addition to the CNN-based detection methods, using other deep learning algorithms, such as recurrent neural networks (RNNs) and generative adversarial networks (GANs) are becoming increasingly popular. These models have been found to achieve better performance on certain tasks, such as detecting objects from complex images (Al-Dhabi and Zhang, 2021; Bahmei et al., 2022). Generally, RNNs can learn from past and current data, making them better equipped for temporal detection, while GANs allow for establishing a high level of control when training a deep learning model for image recognition tasks.

Most research on object detection based on image data involves utilizing ML algorithms for recognizing objects contained in images and videos (Gupta et al., 2021). Gupta et al. indicated that this type of AI application has numerous potential use cases, including those in the fields of security, surveillance, and autonomous driving. These fields have received extensive attention in recent years, with scientists focusing on harnessing the power of deep learning algorithms to create systems that are more accurate and efficient in analyzing image data. According to Saleh et al. (2022), CNNs have shown great potential recently, as they have been able to learn sophisticated features from images without requiring a manual input from a user. CNNs can take an image as input, break it down into small patches, and generate output maps, indicating the locations of certain objects (Saleh et al., 2022). They have become particularly useful for analyzing and identifying objects in complicated and large-scale environments. CNNs can also be trained to identify objects based on previous learning outcomes, giving the system the ability to adjust itself for more complex and unpredictable environments. Recent research in this field includes developing systems that can adjust themselves as new data becomes available and identify objects from low-resolution images (Masita et al., 2020; Saleh et al., 2022). By leveraging on modern image data acquisition techniques, such as deep learning and transfer learning, the field of image object detection has become even more robust and efficient.

In addition to deep learning techniques, scientists have experimented with combining ML algorithms with conventional object detection algorithms to increase accuracy and speed (Bai et al., 2020). According to Bai et al., some of the more popular object detection techniques used in this context include boosting and Haar cascading. The combination of these methods has been shown to provide accurate and fast results when recognizing objects in large databases. As the field continues to advance, scientists will be able to identify even more sophisticated and nuanced objects from images. Moreover, object detection based on image data has significantly improved in the past few years. Based on the powerful deep learning models, such as CNNs, 3D vision systems, RNNs, and GANs, it is now possible to accurately detect and identify objects in images and videos. This research provides invaluable insight into the potential applications of deep learning algorithms in security, surveillance, and autonomous driving fields.

2) Types of Object Detection Technology

Object detection locates an object contained in an image for verification and automatically categorizes its type (Hättenschwiler et al., 2018). Object detection algorithms can be categorized into two types: two-stage detector algorithms, which perform region proposal and object classification separately, and one-stage detector algorithms, which perform these two processes concurrently (Morera et al., 2020). The former can be transformed into the latter with enhanced speed and lower calculation cost. The most representative two-stage detector algorithms, listed in order of increasing speed, are 1) region-based convolutional neural network (R-CNN), 2) fast R-CNN, and 3) faster R-CNN. Two-stage detector algorithms propose a region of interest where the object for detection might be located, extract the object features, and perform learning to mark and categorize the bounding box of the object. Contrastingly, one-stage detectors can be categorized into You Only Look Once (YOLO) and single-shot multibox detector (SSD) algorithms. As can be inferred from its name, a one-stage detector proposes and categorizes the bounding box simultaneously, which saves time and reduces the cost of calculation and inference.

2.2 Previous Studies on AI Application Research

2.2.1 Criteria and Order for Selecting Studies

AI has emerged as a powerful tool in many industries but it has especially made a mark in healthcare-, finance-, and security-related fields. Its ability to process and analyze vast amounts of data and identify patterns using ML algorithms makes it an efficient instrument for decision-making and analysis tasks. In particular, AI has demonstrated its potential to enhance accuracy and precision in radiology, tax arrears, and security detection fields. These achievements have been acknowledged by researchers, such as Lu (2019) and Zhang and Lu (2021). Thus, to select the most relevant studies on AI applications in the radiology, tax delinquency, and security detection fields, a set of criteria and an order of selection have been established. The criteria consider the relevance of the research to the specific field, quality of the research, and novelty of the research. In terms of the order of selection, the studies are sorted based on their recency. In addition, the level and subset of AI research presented by Dwivedi et al. (2021) and Regona et al. (2022) are also considered. Those researchers divided the level of AI research into three categories: 1) application of ML algorithms (mainly tree-based shallow algorithms), 2) application of deep learning algorithms, and 3) deep learning algorithms with improvements.

The selection criteria used for the research studies includes three aspects: the relevance of the research to the respective field, quality of the research, and novelty of the research. First, the research studies must apply to the field in which they are investigating and exploring the potential of AI technology. This ensures that the study yields reliable and accurate results that can be used in real-world scenarios. Secondly, the quality of the research must also be considered. This includes considering the accuracy of the findings and results, as well as the methodology used in collecting and analyzing data. Lastly, the novelty of the research is also important to ensure that the study results in meaningful and new insights or approaches to using AI technology.

The order of selection used in this research focuses on the recency of the studies. This allows the study to identify and analyze the most recent advances in AI application research. By considering the research that was done in the last two to three years, this study can more accurately gauge the most advanced trends and understand what applications AI can have in the given field. This in turn allows for more accurately predicting the future applications of AI in these fields. Establishing the criteria and order for selecting the relevant studies for this research ensures that only the most meaningful, reliable, and novel results are analyzed and discussed. This allows for establishing an accurate and meaningful understanding of the applications of AI in the various fields being studied.

2.2.2 Field of Radiology

Research into the potential of AI for radiology applications has proven the capability of these algorithms in the context of both tabular and image data analyses. One example is the development of an AI-based algorithm for detecting strokes from computerized tomography (CT) scans, which outperformed traditional methods in both sensitivity and specificity (Shafaat et al., 2021). According to Shafaat et al., other AI algorithms have been developed for similar tasks, such as automated lung cancer diagnosis, utilizing CT and magnetic resonance imaging (MRI) scan data. These findings underline the potential of AI-driven image processing to improve the accuracy of diagnosis in clinical radiology, as well as to detect abnormalities more rapidly and with greater confidence compared with current methods. Kaka et al. (2021) assessed the performance of using AI to diagnose cerebral aneurysms from CT scans. Although the AI system was able to accurately detect the aneurysm in some cases, it was found that the performance varied significantly depending on the resolution of the scan. This demonstrated how certain parameters could significantly impact the effectiveness of AI systems, suggesting the need for conducting more research to maximize the accuracy of such systems. However, a key issue that needs to be considered is the variability of the datasets and algorithms used in different studies, as these have a strong influence on the accuracy and generalizability of AI models (Kaka et al., 2021). Generally, studies have demonstrated the potential of combining AI algorithms with the expertise of human radiologists, demonstrating the effectiveness of the synergy between human and machine intelligence.

1) General Studies

The application of AI to medical imaging data has been gaining considerable attention from the research community recently. Studies have suggested that using AI can help radiologists make faster and more accurate diagnoses by automating or improving image analysis. The study conducted by Papanastasopoulos et al. (2020) indicated that AI techniques, such as CNNs, have been successful in detecting lesions, classifying different diseases, and diagnosing cancers in radiological images, such as those produced by computed tomography (CT) and magnetic resonance imaging (MRI). Recent studies conducted by Choi et al. (2013) and Mun et al. (2021) exploring the application of AI in the field of radiology have successfully demonstrated its capabilities for diagnosing various conditions. Mun et al. indicated that AI algorithms are particularly effective in the detection, classification, and diagnosis of lesions with both tabular and image data, such as those produced by CT and MRI scans. Therefore, this form of AI-based image analysis has greatly improved the efficiency and accuracy of diagnosis regarding various diseases, such as breast cancer, lung cancer, and Alzheimer's disease.

Several studies have been conducted to examine the accuracy and potential of AI-based algorithms for the detection and classification of abnormalities in image data. In a study conducted by Bibi et al. (2020), the authors determined that an AI model outperformed conventional imaging techniques for diagnosing breast cancer, achieving a higher accuracy in diagnosing breast lesions. The authors further found that using AI increased the accuracy of conventional imaging, further supporting the use of AI in this domain. In another study, He et al. (2022) investigated the effect of using CNNs to predict the prognosis of stroke patients, achieving relatively higher accuracy. They showed that deep learning algorithms achieved high diagnostic accuracy on an image dataset obtained from patients with reperfusion patterns, outperforming both clinicians and other methods. Similarly, Haq et al. (2022) demonstrated that using deep learning-based algorithms enabled automated segmentation and classification of pulmonary nodules on CT scans, resulting in high accuracy. Generally, these studies demonstrate the potential of AI for automated image analysis, enabling faster and more accurate diagnoses in radiology.

Deep learning algorithms, particularly CNNs, have been proven to be very useful in this context. Using these networks, He et al. (2022) indicated that the performance of the diagnostic process was significantly improved owing to the associated automation, as well as exhibiting improved accuracy and reduced turnaround time for analysis. CNNs are very well-suited for such medical AI research, with studies on stroke prediction models and lung cancer diagnosis models achieving high accuracies in their results. However, Papanastasopoulos et al. (2020) indicated that using AI for radiology research is still relatively new and needs to be further developed to reach its full potential. It has also faced many challenges, with medical professionals expressing some hesitation to fully rely on AI models to make accurate diagnoses. As a result, more research and development need to be conducted. In addition, existing studies lack the scale to explore the full potential of these applications. Future research should focus on investigating further uses of AI-based image analysis in the field of radiology.

In general, AI applications in the radiology field have shown great potential. Moreover, AI-based image analysis techniques are being explored to automate or support the diagnosis of breast cancer, lung cancer, and Alzheimer's disease. While further research is necessary to investigate the clinical utility of these techniques, early results demonstrate that AI can enable radiologists to perform diagnoses and classify medical imaging data with higher accuracy, the current research has found that AI-based image analysis in the radiology field has immense potential, with encouraging results being produced in many areas (Haghighi et al., 2019). Its accuracy and ability to reduce turnaround times in diagnoses can significantly help in supporting medical professionals and improving patient outcomes. Thus, further research should continue to be conducted to reach the full potential this approach.

2) Quantitative Studies

Quantitative research in the field of radiology is mainly a study using data extracted through CT and X-ray images. This is mainly a study in which it is difficult to draw clinical conclusions only with the image itself, or is conducted to confirm more detailed clinical results. In this context, several studies have been conducted that focus on AI algorithms used in analysis, such as those related to the diagnosis of medical conditions. A study conducted by Dong et al. (2020) investigated using ML algorithms to detect COVID-19 cases in a patient population and achieved excellent performance. The same research also showed how the same techniques could be applied to the analysis of quantitative features extracted from MRI images to detect the presence of prostate cancer, demonstrating the versatility of AI algorithms in radiology studies. Another study conducted by Choudhury and Asan (2020) demonstrated how AI can predict the risk of stroke, using a dataset constructed from patients. Through the analysis of features extracted from MRI scans, the researchers demonstrated how the algorithm could accurately identify those at a higher risk for suffering from a stroke. Here, AI was used to classify the probability of lung cancer from patient scans. In both studies, the results were promising, demonstrating how AI can be used to accurately detect and classify various medical conditions.

A study conducted by Chamberlin et al. (2021) applied deep learning to detect lung nodules in CT scans and found that AI was able to accurately detect nodules with higher sensitivity compared to the results obtained by human readers, who achieved am accuracy of 95.24%. Chamberlin et al. evaluated AI algorithms that demonstrated improved accuracy when compared to conventional ML models. Additionally, a study conducted by Voter et al. (2021) investigated the effectiveness of AI applications in identifying pathological features from CT scans, noting the potential to reduce false-negative diagnoses and provide timely and accurate treatment recommendations. These results are encouraging and suggest that AI holds promise for improving accuracy and efficiency in radiology research. Nonetheless, further research is necessary to fully understand the capabilities and limitations of AI-based solutions (Chamberlin et al., 2021). Specifically, efforts are needed to explore the ethical and safety implications of using AI in clinical decision-making, as well as develop more reliable and accurate algorithms tailored to specific medical conditions. Such research is expected to significantly improve the understanding on AI-based medical imaging analysis and provide substantial benefits to the healthcare industry.

The research examining the application of AI algorithms in radiological studies is increasing in numbers and providing valuable insight into the capabilities and limitations of AI systems. Although promising results have been obtained, further research is needed to fully understand the capabilities of these algorithms and how to maximize their effectiveness for different medical conditions. In the long run, such advances have the potential to significantly improve the accuracy and efficiency of medical diagnoses, providing better outcomes for patients (C. Liu et al., 2020). Therefore, it is evident that there is much potential for AI algorithms to make a meaningful contribution to medical diagnosis in the radiology field. However, further research is needed to investigate the potential of AI algorithms and explore methods for addressing their variability. Such research should also consider how AI-based models may be deployed in tandem with human clinicians to better optimize medical diagnoses.

3) Respiratory Studies

ML has been widely used in respiratory research in the medical field, especially in radiology. Representative examples include early prediction of childhood asthma, early prediction of asthma exacerbations, and characterization of asthma and chronic obstructive pulmonary disease (COPD) phenotypes (Howard et al., 2015; Castaldi et al., 2020; Finkelstein and Jeong, 2017; Patel et al., 2022; Haghighi et al, 2019). COPD is the third leading cause of death globally and is expected to become the first (World Health Organization, 2020; OECD, 2018). Pulmonary function test (PFT) results, such as forced expiratory volume in the first second (FEV1) and forced vital capacity (FVC), are commonly used to determine the stages of COPD (Global Initiative for Chronic Obstructive Lung Disease, 2020). However, because the PFT reflects only the entire lung function and not regional lung decline of functional features before the destruction of lung tissue (Galbán et al., 2012; Zou et al., 2021), early detection or self-recognition is difficult until the entire lung function is severely declined. Management and detection of early COPD have recently received significant attention (Laucho-Contreras and Cohen-Todd, 2020). ML approaches with imaging have been capable of characterizing early-stage progression of COPD (Zou et al., 2021). Furthermore, ML has been widely used to investigate various aspects of COVID-19 (Kwekha-Rashid et al., 2021; Rasheed et al., 2021; Alimadadi et al., 2020). Finally, ML is intensively used for the early diagnosis of lung diseases, which is very important in treatment efficacy (Patel et al., 2022; Kadir and Gleeson 2018; Cai et al., 2015).

Recently, in addition to ML, CT scans have been used to quantitatively characterize imaging-based regional lung function in lung diseases (Choi et al., 2017; Amelon et al., 2011; Choi et al., 2013; Bodduluri et al., 2013; Shin et al., 2020; Li et al., 2021; Chae et al., 2021). Chae et al. (2021) first introduced and developed a standardized measure of CT-based local ventilatory capacity called the RRAVC to differentiate COPD patients from healthy subjects.

Due to the characteristic nature of the lungs, where functionally favorable regions for respiration are observed differently in the lungs of healthy subjects compared to those of patient groups, it is difficult to make a clinical diagnosis based solely on lung imaging. Furthermore, even among healthy lungs, there are regional variations in functional distribution and characteristics. As a result, they concluded that the RRAVC serves as a valuable imaging biomarker for assessing regional lung ventilation and quantifying the airflow within specific lung regions based on CT scans. This helps researchers detect progressive airflow limitation in lung disease. The RRAVC also characterizes the effects of supine versus prone body positions in regional lung ventilation distribution of normal lungs (Shin et al., 2020).

From both inspiratory and expiratory CT images of a human subject, first segmentation and image registration are conducted. Then, RRAVC values are calculated at local lung regions based on the mathematical definition (Chae et al., 2021).

2.2.3 Field of Tax Delinquency

In the realm of tax collection, AI has shown considerable potential in identifying and predicting tax evasion. AI algorithms, such as ML and deep learning, are used to detect and analyze trends from large datasets to ultimately predict tax evasion cases or help prioritize certain tax cases for investigation (Abedin et al., 2021). With the potential to improve the effectiveness and cost efficiency of tax collection processes, AI holds great potential for the financial services industry. Recent research has demonstrated the effectiveness of AI for tabular and image data applications in taxation, with various studies being conducted to identify and predict tax evasion (Höglund, 2017). AI has the potential to revolutionize the efficiency and accuracy of taxation and revenue collection procedures, providing a modern alternative to more traditional time-consuming approaches.

Research has found that ML and deep learning can be used to analyze tax evasion patterns and predict the tendencies of tax evaders (Abedin et al., 2022). According to Abedin et al., tax collection-related tasks can be performed faster and more accurately by utilizing AI. Moreover, AI can prioritize cases under investigation and automate the relevant collection processes. However, the ethical implications of using AI in taxation need to be addressed. In particular, it is essential to make sure that algorithms used in tax collection are reliable and unbiased to avoid any potential misidentifications or mislabeling of taxpayers. A recent study conducted by Lai (2020) has demonstrated that AI-assisted tax filing has the potential to increase accuracy and efficiency in the field of taxation. In their experiment, the authors used a deep learning-based neural network to detect and categorize tax documents, as well as to generate customized reminders for users. Their results demonstrated that the AI-assisted system reduced the amount of time required for completing the filing process and improved the accuracy of filing tasks.

Another study conducted by Ali et al. (2021) predicted delinquency on mortgage loans utilizing ML techniques. Their experiment employed an AI system for identifying taxpayers who have been misclassified or are at a higher risk of evading taxes. Moreover, the results presented by Li (2019) demonstrated that the AI-assisted system improved the efficiency of identifying potential tax evasion cases while maintaining a lower false-positive rate compared to manual analysis.

AI is particularly beneficial for identifying patterns and predicting potential problems with taxpayers and revenue collection efforts. AI algorithms have been used to identify tax evaders and predict delinquent cases, enabling financial agencies to identify individuals for further investigation (Huang and Yen, 2019). According to Huang and Yen (2019), automated detection and analysis can also be used to identify taxpayers likely to default in the future. In addition, AI algorithms have been used to prioritize cases for investigation, helping to ensure resources are deployed efficiently and cost-effectively (Bhatore et al., 2020). The implications of utilizing AI in taxation are potentially significant. Not only can AI improve the efficiency of tax collection activities but it also has the potential to save resources and increase the overall tax take. However, the use of AI must be done with a clear understanding of the technology and consideration of the potential consequences of misidentification or mislabeling. Transparency and accuracy are key to any AI application in taxation and financial regulation because failure to accurately identify or prioritize individuals can lead to costly delays or disputes (Lessmann et al., 2015).

However, it is important to be aware of the ethical implications of using AI in tax collection (Munoko et al., 2020). According to Munoko et al., misidentifying or mislabeling taxpayers can have dire consequences; moreover, the algorithms used must be not only reliable but also transparent and unbiased. As such, there is a need for further research and development of AI applications to ensure their compliance with ethical guidelines and codes of conduct. Studies examining the use of AI in tax collection have focused on detecting delinquent cases, predicting future delinquencies, and optimizing tax identification processes (Stahl et al., 2022). Therefore, AI has proven successful in addressing in all these problems, as algorithms can analyze significant amounts of data to quickly identify patterns that could not be done by human operators.

A thorough literature review of studies conducted on the application of AI for tax collection reveals the effectiveness of such approaches. Several studies have underlined the potential of AI in aiding tax collection, through its ability to streamline processes, save costs, and improve accuracy. However, more research is needed to understand how to develop AI algorithms that are reliable and ethically sound while ensuring maximum efficacy. These applications need to consider the broader legal and financial implications, implementing failsafe systems to ensure taxpayers are protected from wrongful actions or identification. AI holds great potential for modernizing and improving tax collection; however, the remaining challenge is to ensure accuracy and accountability for all users of the system.

2.2.4 Field of Security Detection

1) General Studies

The advancements in AI technology have substantially expanded its potential applications in security detection systems. The application of AI to image data analysis has been explored in various research studies related to security detection. These studies have focused on detecting objects in images and videos, as well as analyzing tabular data to identify patterns in large datasets that may indicate the presence of security threats.

To better understand the application of AI in security detection, several key studies were reviewed. A study conducted by Akcay et al. (2018) explored how CNNs could be used to detect security threats in digital images. Their study concluded that using deep learning algorithms provided significantly better results than traditional techniques. In addition, a research conducted by Rueda et al. (2022) proposed a hybrid model combining CNNs and SVM for malware detection and intrusion detection, obtaining a higher accuracy than that of other similar methods. In addition to these studies, there have been numerous others in the field of AI and security detection. Notable among these is a research conducted by Wei et al. (2020)

that proposed a convolutional neural network-based system for detecting phishing web pages based on image data. The study discussed the importance of detecting malicious websites, as phishing attacks have become increasingly sophisticated recently. Another study, conducted by Lee and Cho (2020), focused on a real-time weapon recognition system based on an enhanced deep learning algorithm, noting that the accuracy and performance of the system were superior to other existing approaches. Other researchers have also investigated various approaches, such as artificial neural networks (ANNs), decision trees, and genetic algorithms (Shenfield et al., 2018). These algorithms have been used to detect various anomalies or unusual activities. Other applications have been studied to detect fraudulent financial activities and identify fake customer profile images (Gillen and Morrison, 2015).

AI applications for image data analysis focus on object detection, such as weapons, suspicious packages, and persons of interest. Deep learning algorithms, specifically CNNs, have been used extensively to identify and classify objects in images and videos (Rabbi et al., 2020). With the assistance of AI technology, computer vision-based security systems can better detect suspicious activity and minimize the risk of unauthorized access to sensitive information. CNN, RNN, YOLO, and other deep learning algorithms have been applied to image and video datasets to detect objects of a security interest, such as weapons, suspicious packages, and persons of interest (Akcay et al., 2018). For instance, a CNN-based image recognition system was applied to identify human trafficking victims. Similarly, deep learning algorithms have been used to detect signs of terrorism and threats in large datasets of communication and other information. Apart from these object-based security detection applications, AI algorithms have been used to identify anomalies in communication traffic that may indicate an attack on a computer system (Tong et al., 2020). Additionally, AI-based malware detection systems have been developed to protect companies and governments from malicious attacks. Studies have also focused on leveraging AI algorithms for the early detection of potential security threats. For example, ML algorithms have been applied to recognize certain behavioral patterns of human and computer activities that could potentially signal a security breach.

Generally, the application of AI technology to the field of security detection has shown immense potential. This technology has enabled organizations and governments to detect security threats more accurately and efficiently. In the future, AI-based security detection systems are expected to become increasingly sophisticated and accurate. Thus, this technology has the potential to revolutionize the security detection industry. Research into the application of AI technology for tabular and image datasets for security detection purposes has experienced significant developments recently. ANNs, decision trees, genetic algorithms, and CNNs have been studied and used to detect various anomalies and objects, improving the accuracy and efficiency of security detection systems. As security threats become increasingly sophisticated, it is expected that AI technology will play an increasingly important role in ensuring security and protecting sensitive information.

2) Transport Studies

Over several decades, efforts have been underway to improve the security levels of passenger and cargo transport (Abadie and Gardeazabal, 2008; Baum, 2016). With the increasing risk of air transport-related terrorism, especially after the September 11 attacks in 2001 (START, 2021), the detection of prohibited objects, including explosives, via X-ray scans of personal baggage has become ever more important (Baum, 2016; Novakoff, 1993; Singh and Singh, 2003). Formerly, object detection via X-ray imaging depended heavily on human vision; therefore, it had high fallibility and limitations in discrimination (Gallen and Morrison, 2015; Sterchi and Schwaninger, 2015). Nowadays, cabin baggage inspection systems have advanced using computer vision technology. Furthermore, with the developments in AI, a sophisticated baggage inspection system has been introduced to improve the detection performed by security personnel. With this new technology, the accuracy and speed of baggage inspection have been gradually enhanced (Hättenschwiler et al., 2018; Eom, 2020).

However, despite the recent advancements, some existing studies on cabin baggage inspection have revealed different problems in automated object detection via X-ray scans. This technology is either limited to supporting security personnel in the field (Gallen and

Morrison, 2015; Sterchi and Schwaninger, 2015; Hättenschwiler et al., 2018) or is slow as it most commonly performs two object detection processes at a time (Morera et al., 2020; Tong et al., 2020). Furthermore, this technology often fails to detect relatively small hazardous objects because it has been constructed with a focus on recognizing large objects linked to explosions and killings (Adarsh et al., 2020; M. Liu et al., 2020).

Lee and Cho (2020) used the airport baggage X-ray data disclosed and provided by the Korea National Information Society Agency, which is the same dataset used in this study, and reported on the detection of prohibited objects in baggage based on images taken by X-ray scanners. The Xception algorithm, a lightening model in which input data are received to reduce the number of channels through a 1 × 1 convolution product and made to pass through a 3 × 3 convolution product for establishing an individual output channel, was used. Subsequently, a detection and categorization model for 12 prohibited objects was developed, which demonstrated a high performance based on its F-1 score. Then, an experiment was conducted to detect and categorize large hazardous items and single items from a single image. The model exhibited a generally good performance when the latest exponentially developed image detection algorithm was applied. However, only a fraction of the numerous item types that must be detected in the field were used; thus, there were limitations in applying the model to an actual environment in which objects of two or more classes must be detected. However, a dataset of X-ray images comprising six prohibited items was used, and performance comparison tests were conducted against models such as YOLOv2, R-CNN, and region-based fully CNN (R-FCN) to determine the best performing model (Akcay et al., 2018). This study focused on large objects, such as laptops and cameras; therefore, it failed to evaluate the detection performance with small objects and related situations. Similarly, using YOLOv3 solely in the detection of large, harmful, prohibited objects, such as razor blades, knives, and guns, among others has been studied.

2.3 Research Gap and Opportunities

Research in the fields of medical imaging, tax delinquency, and security detection has experienced significant advances in terms of development and application of AI technology. For example, in the field of radiology, ML algorithms are commonly used; in tax delinquency research, the focus is increasingly shifting towards deep learning algorithms. Although the usage of traditional ML methods remains the norm in tabular data-based research, the implementation of AI in object detection in the field of security underlines the active improvement and integration of deep learning algorithms (Zheng et al., 2021). The literature review presented in this dissertation indicates that there is an inconsistency in the level of research and development speed across various AI domains, which provides new opportunities for novel AI applications. In medical imaging and tax delinquency, researchers can focus on finding solutions to specific problems by exploiting AI technologies. In the field of security, object detection techniques should be studied more thoroughly; moreover, the efficiency of existing deep learning algorithms should be improved. In addition, the research should focus on exploring how AI can be employed to deal with both tabular and image-based datasets, as well as creating more generalizable solutions for various problems.

The research gap and opportunities identified in this dissertation lie in using advanced AI technologies in tabular data-based and object detection-based research. The current landscape of AI applications indicates to discrepancies between domains in terms of the level of research and development speed, with some fields lagging in terms of their implementation of deep learning algorithms. For example, radiology relies on shallow ML algorithms, whereas in the financial sector, deep learning algorithms are being implemented. Therefore, there is a clear need for increased investment in radiology and tabular data-based research in general. This can lead to improved AI technologies, which can then be applied to problems in these domains. Moreover, object detection within the security field is experiencing active improvements, which suggests that there may be less opportunity for novel solutions and research in this field.

The current AI research landscape provides numerous opportunities for innovation, from both established and upcoming domains. Tabular data-based research still has ample opportunities for improvement, whereas object detection-based research has developed various existing deep-learning models that could be used for integration with tabular data. Therefore, the researchers need to identify the most appropriate opportunities to improve AI technologies, regarding both the present and future studies. Finally, research can also focus on creating novel AI algorithms and solutions. For instance, developing more accurate AI models for medical diagnosis, predicting tax delinquency with deep learning models, and improving object detection models for security purposes exhibit potential research opportunities. The ability to design and implement sophisticated AI models could lead to significant advancements in AI technology while providing insights and practical solutions to various important problems. In the following chapters, the main characteristics and results of research on radiology, tax delinquency, and security detection study cases are identified.

Chapter 3

AI Research on Tabular Data I (Study 1)[①]

3.1 Research on Prediction of Regional Lung Function

3.1.1 Research Objective

In this study, CT images are used to quantitatively characterize local lung function for detecting pulmonary diseases. As confirmed in previous studies (Shin et al., 2020; Li et al., 2021; Chae et al., 2021), disease likelihood can be determined through CT-based quantitative measurements, such as relative regional air volume changes in the lung. However, owing to the inherent variability of the lung, there is a need to establish baseline variation levels for regional air volume even in healthy lungs. Therefore, this study aims to model regional lung function distribution based on inspiratory and expiratory CT images to quantitatively evaluate local lung regions. Specifically, the aim of this study is to develop a standardized model by predicting the

[①] This study is based on the following publication by **Kim, E.**, Lee, Y.H., Choi, J., Yoo, B., Chae, K. J., and Lee, C. H. (2023). Machine Learning-based Prediction of Relative Regional Air Volume Change from Healthy Human Lung CTs. *KSII Transactions on Internet and Information Systems*.

RRAVC, a measure associated with diseases such as COPD, and validating effective latent variables.

3.1.2 Key Features of Research

This study explores the potential of using ML algorithms for predicting the RRAVC in a healthy human lung from CT scans. Specifically, the research focuses on applications in the chronic obstructive pulmonary disease (COPD) field. The performance of six ML algorithms was compared: multivariate linear regression, ridge regression, lasso regression, ElasticNet, extreme gradient boosting (XGBoost), and light gradient boosting machine (LightGBM), to explore the quantitative characteristics of CT scans. In addition, the performances of three ML algorithms, namely, XGBoost, LightGBM, and multi-layer perceptron (MLP), were compared by additionally processing and utilizing lung relative coordinate information.

This study proposes a novel approach that uses relative rather than absolute coordinates to increase the accuracy of these predictions, and calculates subject relative proportion coordinates to reduce inter-subject variability in results. The presented approach is validated by performing a comparative analysis, which indicated a significant performance improvement with relative proportion coordinates over conventional absolute coordinates. Therefore, the key features of this research includes the application of ML for RRAVC prediction, comparative assessment of different algorithms in terms of their effectiveness, and novel use of relative coordinate systems for reducing inter-subjective variability. In addition, this research offers some practical suggestions for improving CT scan acquisition methods and relative coordinate systems, to optimize predictions of RRAVC for detecting COPD.

3.2 Research Design

3.2.1 Data Collection and Preprocessing

The dataset for this study was obtained from Jeonbuk National University Hospital, and consent for processing the expiratory CT scan data was obtained in the original study. Two CT images acquired from a patient were used: one at full inspiration and another at full expiration. From these CT images, an image processing called "segmentation" identifies separate regions of airways, lungs, lobes, and blood vessels using the VIDA vision software (Coralville, IA, USA), which is a tool used to analyze airways and segment lung CT images. After the segmentation, an image registration process is performed to anatomically align corresponding lung regions in both CT images. This process maps the expiratory CT image onto the inspiratory CT image within the identical coordinate system. From this, not only local lung regions can be compared from the two images but also various variables, including air ventilation, can be computed. The RRAVC, can be computed at small lung units, to quantitatively characterize the three-dimensional map of regional air ventilation distribution attributes, and it is calculated by applying Equation (3.1).

$$\text{RRAVC} = \frac{(\ v_{air}^{insp} - v_{air}^{exp})/v_{air}^{TLC}}{(v_{air}^{insp} - v_{air}^{exp})/TLC} \tag{3.1}$$

A total of 292 subjects were used for data collection, each with approximately 60,100 data points (rows). Accordingly, a total of 8,772,704 non-zero RRAVC values were used. The 37 columns for each row included coordinates for the data point, its changes in coordinates, length, diameter, lobe region, air and tissue volumes, Horsfield ordering, Jacobian matrix (J), anisotropic deformation index (ADI), slab-rob index (SRI), reference index of bronchi, displacement, normalized displacement, and angle. Additionally, a total of 5 data scaling changes were made (i.e., non-scaling, standard, robust, minmax, and maxabs). The entire data collection process started by submitting a formal proposal to the Institutional Review Board of Jeonbuk National University Hospital and detailing the intention of the study. After approval was granted and participants consented, full inspiratory CT and full expiratory CT images were taken. Then, the collected images were organized into a format readable by the experimental setup, as detailed above. Finally, the experiments were conducted and results were analyzed against the performance metrics, providing valuable insights into results of the study. The data were divided into training and testing sets, with 80% of the total number of subjects being used for training and 20% for testing. The performance metrics for the experiments were the R2 score, mean squared error (MSE), and mean absolute error (MAE). To accurately analyze the data, an Nvidia Quadro RTX 5000 GPU, Intel Xeon Gold 5220R CPU, and 187 GB of RAM were used to process the results.

1) Relative Coordinates

Relative coordinates indicate how far each data point coordinate is from the minimum position. The formula is stated in Equation (3.2) as follows.

$$(x_{rpc_i}, y_{rpc_i}, z_{rpc_i}) = (x_i - \min(x), y_i - \min(y), z_i - \min(z)) \tag{3.2}$$

For example, suppose a simple lung modeling of subject S1 is represented in two-dimensional coordinates, namely, x and y coordinates. Assume that the lower-left data point (the data point with the smallest x and y coordinates) is (5, 10) and upper-right data point (the data point with the largest x and y coordinates) is (95, 100). In this case, min(x) = 5, min(y) = 10, max(x) = 95, and max(y) = 100. Assume to have a data point A (50, 55) above that lung. In this case, the absolute coordinate of point A is (50, 55) but its relative coordinates are (50-5, 55-10) = (45, 45). The absolute coordinate for data point A is not above the y=x line. However, it is above the y=x line for the relative coordinates system.

2) Relative Proportional Coordinates

The relative proportional coordinates indicate how far each coordinate is located between the minimum and maximum values. For the abovementioned example with subject S1, the relative coordinate value of point A is (45, 45). Moreover, the top right coordinate of (95, 100) is also converted to (95-5, 100-10) = (90, 90) in the relative coordinates. Therefore, observe that point A is located at the center of the lung modeling for that subject.

Suppose there is another subject, namely, S2, who is taller than S1. This means that coordinate values of the upper-right data point of S2 is larger than that of the previous subject S1. Assume that the lower-left corner coordinate and point A are kept as those used in the previous example with subject S1, and only the coordinates of the upper-right corner are changed to (125, 130). In this case, the relative coordinates of S2 are (45, 45), which is the same as that of S1. However, observe that point A is not located at the center of the lungs of S2 like it was in S1. If the upper-right coordinate of S2 is transformed to relative coordinates, the result is (125-5, 130-10) = (120, 120). Therefore, point A of subject S2 is only located approximately at 1/3 of the lower left corner of the overall coordinate system. To solve this problem, a relative proportional coordinate system that normalizes each relative coordinate, as stated in Equation (3.3), is proposed.

$$\left(x_{rppc_i}, y_{rppc_i}, z_{rppc_i}\right) = \left(\frac{x_i - min(x)}{max(x) - min(x)}, \frac{y_i - min(y)}{max(y) - min(y)}, \frac{z_i - min(z)}{max(z) - min(z)} \right) \quad (3.3)$$

Using the above relative proportional coordinates, point A in S1 is transformed to $\left(\frac{50-5}{95-5}, \frac{55-10}{100-10} \right) = (0.5, 0.5)$ and point A in S2 to $\left(\frac{50-5}{125-5}, \frac{55-10}{130-10} \right) = (0.375, 0.375)$ respectively. Observe that the abovementioned problem is solved.

3) Potential Relative Coordinates

This method does not compute the relative coordinates, relative proportion coordinates, or other new features of each data point. However, the max and min values of each coordinate for a subject are added as new features to every data point. The max and min values of coordinates are different for each subject. Therefore, experiments were conducted to assess whether these new feature columns can potentially represent the relative coordinates.

3.2.2 Methodology

Various ML techniques for pulmonary nodule automatic detection have been applied in the respiratory field (C. Liu et al., 2020; Mei et al., 2021). Similarly, ML techniques are used in this study to predict the RRAVC. The predictive target variable, RRAVC, is a continuous value. Extensive research has been conducted on regression algorithms for predicting continuous values prior to the recent surge in popularity of machine learning applications. The simplest multivariate linear regression model among various linear regression models is presented in Equation (3.4).

$$\underset{\alpha \in R,\, \beta \in R^p}{\operatorname{argmin}} \sum_{i=1}^{N} \left(y_i - \alpha - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 \tag{3.4}$$

There are more advanced linear regression models, for example, the ridge and the LASSO regression models, which are well-known regularization techniques. Those techniques can reduce the weight of independent variables with low explanatory power by constraining the size of the coefficients when calculating the regression coefficients. Ridge regression includes L2-norm regularization and LASSO includes L1-norm regularization. Ridge regression can be expressed as presented in Equation (3.5), and LASSO regression is expressed as presented in Equation (3.6).

$$\underset{\alpha \in R,\, \beta \in R^p}{\operatorname{argmin}} \sum_{i=1}^{N} \left(y_i - \alpha - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{p} \left| \beta_j \right|^2 \tag{3.5}$$

$$\underset{\alpha \in R,\, \beta \in R^p}{\operatorname{argmin}} \sum_{i=1}^{N} \left(y_i - \alpha - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{p} \left| \beta_j \right| \tag{3.6}$$

ElasticNet combines ridge regression and LASSO regression methods; accordingly, it includes both the L1- and L2-norm (Zou and Hastie, 2005). The formula is stated in Equation (3.7) as follows.

$$\underset{\alpha \in R,\ \beta \in R^p}{\operatorname{argmin}} \sum_{i=1}^{N} \left(y_i - \alpha - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \lambda_1 \sum_{j=1}^{p} \left\| \beta_j \right\|_1 + \lambda_2 \sum_{j=1}^{p} \left\| \beta_j \right\|_2^2 \tag{3.7}$$

Support vector regression is a method that learns to fit as much data as possible within a margin (tube). In this study, both linear and nonlinear kernel functions were used (Zhang and O'Donnell, 2020). Two well-known tree-based algorithms were used as well. XGBoost (Chen and Guestrin, 2016) and LightGBM (Ke et al., 2017) are well-known models that use multiple decision trees, which are known for their good performance in various fields. XGBoost employs a level-wise tree growth strategy, while LightGBM adopts a leaf-wise tree growth approach. MLP is an ANN architecture where perceptrons are stacked using two or more hidden layers. In this study, the Adam optimizer was used as the MLP setting with a dropout probability of 0.2. Rectified linear unit (ReLU) is used as the nonlinear activation function. In addition to the input and output layers, 3 hidden layers are used. The structural information on the number of nodes per layer is shown in Table 3.1. The input layer requires 46 or 52 nodes depending on which coordinate is used. Experiments with more diverse MLP structures are conducted in the follow-up.

Table 3.1 MLP Architecture for RRAVC Prediction.

Layers	# of nodes
Input layer	46 or 52
Hidden layer 1	30
Hidden layer 2	20
Hidden layer 3	10
Output layer	1

In the MLP framework, the decision on the number of layers (i.e., how many hidden layers to stack) is a crucial process of hyper-parameter tuning. Therefore, several experiments are conducted to determine the optimal model while changing the number of layers. Intuitively, as the number of hidden layer changes, the number of nodes in each hidden layer also changes.

After segmentation and image registration using the CT images, features for each data point of the lung were obtained. Moreover, the x, y, and z coordinates of each lung were obtained, which were absolute coordinates. However, these absolute coordinates may not reflect the patient's characteristics properly. Because humans have different heights and different chest thicknesses, the x, y, and z coordinate values differ even if the specific data point is located exactly at the center coordinates of each human. To address this problem, the performance of the proposed model was attempted to be maximized by introducing the following three relative coordinate systems: 1) relative coordinates, 2) relative proportional coordinates, and 3) potential relative coordinates.

3.2.3 Limitations and Considerations

When considering the limitations and considerations of the experimental design used in this research, it is important to recognize that to achieve an optimal model, several experiments must be conducted to determine the best model. The most important reason for this is appropriately choosing the number of layers for the problem at hand. Additionally, with an increase in the number of layers, the number of nodes must also change. In addition, when using the potential relative coordinates method, the relative coordinates, relative proportion coordinates, or other new features of each data point are not calculated. Instead, the max and min values of each coordinate of the subject are used as new features for every data point. Because the max and min values of coordinates are unique for each subject, this could potentially be an effective representation of the relative coordinates.

Regarding the segmentation and image registration using the CT images, features are produced for each data point of the lung. To reflect patient characteristics more accurately, these absolute coordinates must be complemented with the aforementioned three relative coordinate systems. However, there remain some issues to be considered when designing the experiments. One issue is the bias in the data. For example, if the data is skewed in favor of certain parameters, the results of the experiment may be impacted and skewed as well. Additionally, it is important to properly consider the training, validation, and testing datasets to avoid potential problems.

Finally, while utilizing certain experimental designs may be successful, this does not necessarily guarantee that these results can be reliably transferred to a real-world situation. It is always important to consider the external validity of the experiments before the results can be deemed reliable. While designing an experimental study to obtain the optimal model, certain limitations and considerations should be remembered to ensure the accuracy and reliability of the results. It is also important to be aware of the various considerations involved in working with different relative coordinate systems. Considering all these aspects will help construct the most effective experiments possible.

3.3 Experimental Results

3.3.1 RRAVC Prediction Using Machine Learning

A total of five scaling changes were considered in the experiments (i.e., non-scaling, standard, robust, minmax, and maxabs) to robustly determine which model performs best. Because each scaling has different characteristics, the scaling changes were implemented to compare the general performance. Based on these experiments the best performing model can be determined, which can be used as a baseline model for future indicators of performance evaluation. Grid experiments are conducted on a total of 40 settings (5 scaling changes and 8 models mentioned above). Excluding non-scaling, the four scaling formulas are presented below: standard formula in Equation (3.8), robust in Equation (3.9), minmax in Equation (3.10), and maxabs in Equation (3.11).

$$x_{standard} = \frac{x - \mu}{\sigma} \tag{3.8}$$

$$x_{robust} = \frac{x - x_{median}}{x_{Q_3} - x_{Q_1}} \tag{3.9}$$

$$x_{minmax} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{3.10}$$

$$x_{maxabs} = \frac{x}{|max(x)|} \tag{3.11}$$

Table 3.2 reports the results of comparing the R2 scores of each model for different scaling techniques. MSE of all models are listed in Table 3.3. MAE of all models are listed in Table 3.4. MLP models are additionally tested, whose results were not included in Table 3.2 but still reported latter. Among the 8 models, the trained linear support vector regressor model does not converge. Therefore, it is excluded from the experimental results of this research. Nonlinear support vector regressor takes more than 24 hours to train, which is not considered as realistic. Therefore, nonlinear support vector regressor is also not included in the experimental results.

Table 3.2 R2 scores of RRAVC prediction.

	none	standard	minmax	robust	maxabs
Linear	0.0622803	0.0622803	0.0622803	0.0622803	0.0622803
Ridge	0.0622806	0.0622808	0.0623275	0.0622812	0.0623058
LASSO	0.0298870	0.0003844	-0.0002219	0.0004708	-0.0002219
ElasticNet	0.0431111	0.0004880	-0.0002219	0.0003826	-0.0002219
XGBoost	**0.3683279**	**0.3775327**	**0.3207365**	**0.3608702**	**0.3369287**
LightGBM	0.2975713	0.3328849	0.3100740	0.3340164	0.2975726

Table 3.3 Mean squared error of RRAVC prediction.

	none	standard	minmax	robust	maxabs
Linear	1.0818174	1.1550367	0.0001345	6.5694631	0.0003202
Ridge	1.0818170	1.1550361	0.0001345	6.5694563	0.0003202
LASSO	1.1191885	1.2312770	0.0001434	7.0024867	0.0003415
ElasticNet	1.1039323	1.2311495	0.0001434	7.0031050	0.0003415
XGBoost	**0.7287400**	**0.7667244**	**0.0000974**	**4.4776060**	**0.0002264**
LightGBM	0.8103700	0.8217194	0.0000989	4.6657379	0.0002398

Table 3.4 Mean absolute error of RRAVC prediction.

	none	standard	minmax	robust	maxal
Linear	0.329931465	0.340913857	0.00367840	0.81303959	0.00567
Ridge	0.329931405	0.340913626	0.00367823	0.81303862	0.00567
LASSO	0.309842234	0.346390702	0.00373153	0.82971405	0.00575
ElasticNet	0.307328835	0.347636997	0.00373153	0.83158253	0.00575
XGBoost	**0.230503337**	**0.235101718**	**0.00270653**	**0.57226700**	**0.00416**
LightGBM	0.253868842	0.258941210	0.00285972	0.62120382	0.00436

These tables report that XGBoost has the best performance in all scaling applications. When the parameters of the trained XGBoost model used to predict the RRAVC are examined, J, the local volume expansion ratio, plays the most important role among 37 columns. This may reflect that J as a geometrical deformation index is important for accurate prediction of air ventilation distribution, even though it is not explicitly used in the formula for calculating the RRAVC. J was also used in a deep learning pattern cluster-based detection of regional lung features in a COPD population (Cai et al., 2015).

3.3.2 Effect of Relative Coordinates for Standardization

Figure 3.1 shows the experimental results for XGBoost related to comparing the absolute coordinate and three types of relative coordinates proposed in this research. Similarly, the experimental results for LightGBM and MLP are shown in **Figure 3.2** and **Figure 3.3**, respectively. In all cases of XGBoost, LightGBM, and MLP, the relative proportional coordinates record the highest adjusted R2 score. This coordinate system exhibits better performance than the conventional absolute coordinate system. XGBoost, LightGBM, and MLP also exhibit the lowest MSE when relative proportional coordinates are used among a total of four coordinates. The relative coordinates of XGBoost and LightGBM show similar or slightly worse performance compared to that of the conventional absolute coordinate. From this result, it can be deduced that in what proportion between the min and max is more important than how far from the min.

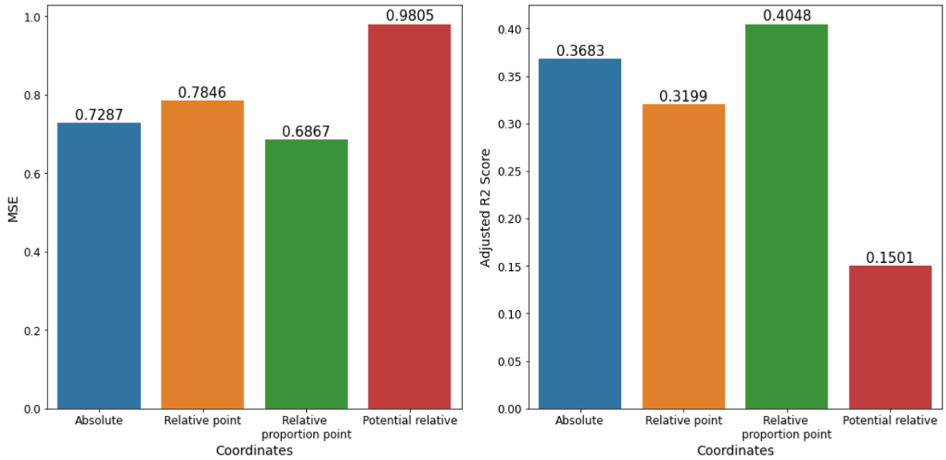

Figure 3.1 XGBoost performance according to different coordinates; MSE (left) and adjusted R2 score (right).

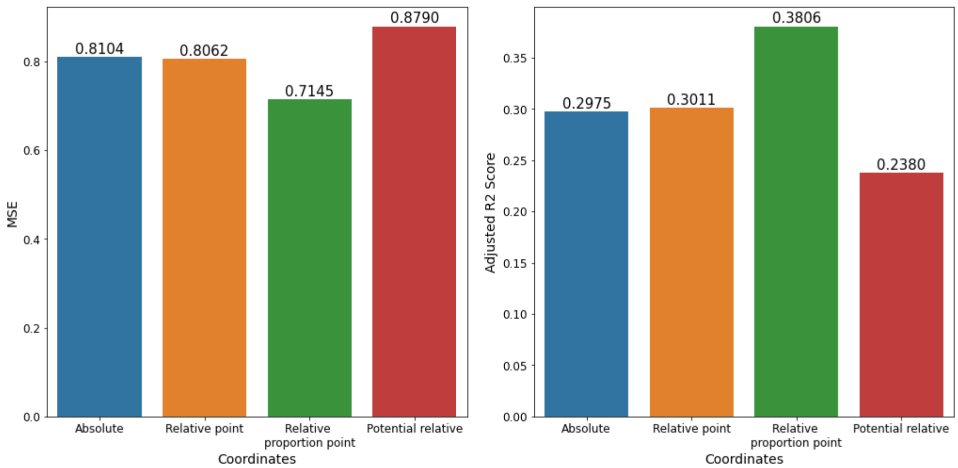

Figure 3.2 LightGBM performance according to different coordinates; MSE (left) and adjusted R2 score (right).

In XGBoost and LightGBM, relative coordinates perform worse than absolute coordinates. However, for MLP, other interesting results are obtained. In MLP the relative proportional coordinates show the best performance, while the relative coordinates show the second-best performance with a relatively small performance gap. This is probably the effect of deep-stacking the neural network. MLP also outperforms other regression algorithms, including XGBoost and LightGBM. MLP with absolute coordinate performs even better than XGBoost or LightGBM with relative proportional coordinates. Therefore, MLP exhibits the best performance among the various ML models experimented with.

In case of potential relative coordinates, the performance remains subpar in XGBoost, LightGBM, and MLP. The MSE increases and adjusted R2 score decreases compared to those of the absolute coordinates. Accordingly, the method of adding the max and min values of each coordinate as a feature is found to be inefficient.

In conclusion, among the conventional absolute coordinates and three relative coordinates, adopting the relative proportional coordinates results in the best performance. Therefore, when experimenting with MLPs of various structures, the relative proportional coordinates are used to determine the best performing model.

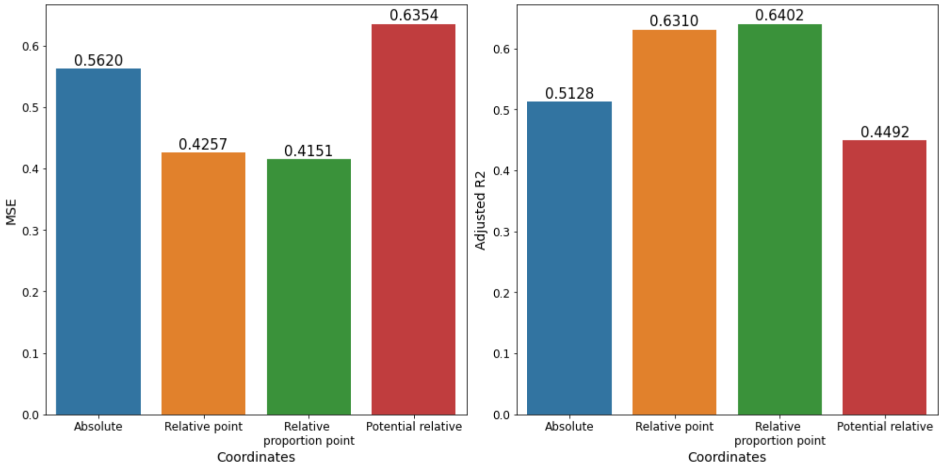

Figure 3.3 MLP performance according to different coordinates; MSE (left) and adjusted R2 score (right).

3.3.3 Performance According to the Number of MLP Layers

MLP has relatively numerous hyper-parameters compared to other ML algorithms, including, dropout probability, learning rate, optimizer type, number of hidden layers, number of nodes in each hidden layer, and number of epochs. In this study, to determine the optimal MLP model, various experiments are conducted by changing the number of hidden layers and number of nodes in each hidden layer. The number of hidden layers is changed from 1 to 6, while the number of nodes in each hidden layer is set arbitrarily based on the number of hidden layers. To ensure the reproducibility of the experimental results, the numbers of nodes in each layer are listed in **Table 3.5**. The following settings are adopted for the hyperparameters: dropout probability = 0.2, learning rate = 10^{-4}, Adam optimizer, and number of epochs = 10. The number of input layers is 46, and the prediction model used is regression; thus, the number of output layers is 1.

Table 3.5 Various MLP architectures.

Number of hidden layer	Nodes in each layer
1	{ 46, 23, 1 }
2	{ 46, 23, 12, 1 }
3	{ 46, 30, 20, 10, 1 }
4	{ 46, 32, 20, 13, 6, 1 }
5	{ 46, 36, 26, 18, 10, 4, 1 }
6	{ 46, 32, 20, 14, 10, 7, 4, 1 }

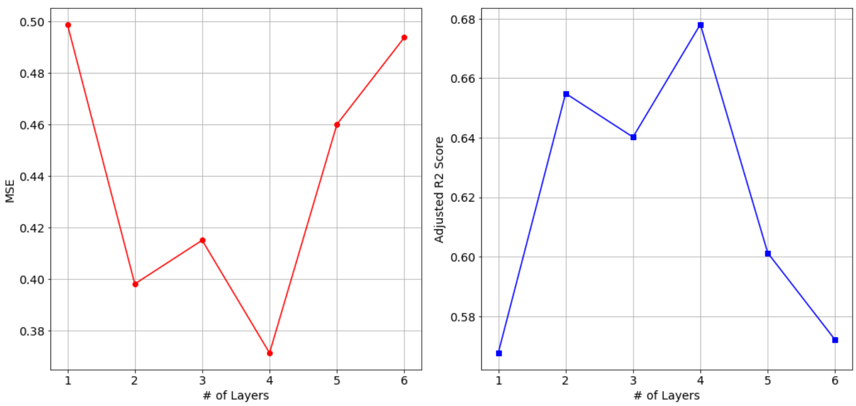

Figure 3.4 MLP performance according to the different number of hidden layers; MSE (left) and adjusted R2 score (right).

Figure 3.4 shows the MSE and R2 scores used for the evaluation. Observe that the best performance is achieved when the number of hidden layers is 4, where the MSE is lowest and R2 score is the highest. Adopting 2 and 3 hidden layers follow in terms of performance. As the number of hidden layers increases, the MSE increases as well. Therefore, it may be effective to use MLP models with 2–4 hidden layers for RRAVC prediction.

3.4 Discussion and Implications

The findings of the research concerning the application of AI for tabular data analysis suggest that a careful consideration of scaling techniques is important when determining which ML model to use. The XGBoost model achieved the highest R2 score when predicting the RRAVC. Among the eight models tested, the linear and nonlinear support vector regressor models were not included owing to not converging and taking a long time to train, respectively. One important finding was that the parameter, J, which is the local volume expansion ratio, was deemed to be the most important among the 37 columns for the XGBoost model when predicting the RRAVC. J was used in a deep learning pattern cluster-based detection of regional lung features in a COPD population. Parameter J is a geometrical deformation index that is used to predict air ventilation distribution. Therefore, it is one of the more important parameters in the XGBoost model.

The findings indicate that parameter importance needs to be explored more thoroughly when ML models are being used, as these parameters can make a major difference in the accuracy of the model results. For example, when predicting the RRAVC, J is the most important parameter. Therefore, the impact of J should be further explored and examined in future research to further understand the role that this parameter plays when predicting the RRAVC. Additionally, it would be beneficial to examine the parameter roles when implementing different ML and deep learning models. Accordingly, these results highlight the importance of scaling techniques when determining which ML model to use, as well as clarifying the importance of parameters and understanding their roles when predicting outcomes. More exploration into the parameter roles in different ML models is needed to obtain a better understanding of the role parameters play and their relevance for certain models. Understanding this information is essential to obtain better and more accurate results from ML models.

The experiment also revealed some important insights into the application of AI for tabular data analysis. For example, the highest adjusted R2 score was obtained using relative proportional coordinates compared to using absolute coordinates for all three algorithms, namely, XGBoost, LightGBM, and MLP. This suggests that when using XGBoost and LightGBM algorithms for quantitative lung function studies, determining "in what proportion between the min and max" is more important than "how far from the min." For MLP, the best performance was achieved with relative proportional coordinates, while the relative coordinates achieved the second-best performance with a performance minimal gap. This can be attributed to the deep stacking of the neural network. Compared to XGBoost and LightGBM, MLP achieved the best performance, even with absolute coordinates. This indicates that MLP may be the most suitable model when handling tabular and image data.

On the other hand, it is possible that some data points could be overlooked during the experiment owing to the nature of tabular and image data. Therefore, it is important to conduct additional tests to confirm the reliability of the findings and assess whether there is a noticeable difference between the performance of relative and absolute coordinates in different scenarios. Furthermore, it would be useful to explore the potential benefits of using multiple models simultaneously to leverage the strengths of each model and produce better overall performance. Accordingly, the experiment provided some valuable insights into the use of AI for tabular and image data analysis. Using relative proportional coordinates outperformed the other regression algorithms, underlining its importance when applying XGBoost and LightGBM. Moreover, MLP achieved the best performance even when using absolute coordinates, suggesting it could be the most suitable model for this type of data. To strengthen the validity of these findings, further experiments should be conducted to compare different scenarios and explore the benefits of using multiple models simultaneously.

When evaluating the MLP model, the performance of the model varied depending on the number of layers. Experiments indicated that four hidden layers were the best configuration for this model in terms of performance. Not only did it produce the lowest MSE and a high R2 score but adopting two and three hidden layers yielded moderate and adequate performance results, respectively. Interestingly, the performance decreased when more than four layers were used. This suggests that when predicting the RRAVC with an MLP model, one should opt for a configuration of two to four hidden layers. Furthermore, note that manipulating various hyperparameters, including the dropout probability, learning rate, optimizer type, and several nodes in each hidden layer, had a direct influence on the model performance. To obtain the best results, these factors must be optimized.

The insight gathered from these application studies is invaluable, particularly in the field of MLP optimization. It indicates the number of layers and nodes that need to be adopted to maximize the RRAVC prediction performance. In addition, understanding the influence of the various hyperparameters helps create the most efficient model for a given purpose. Generally, AI has become an increasingly important part of modern-day technology and these application studies underline only certain uses for it. Understanding how to optimize MLP models can significantly improve the prediction process for the RRAVC, yielding more reliable results. Such studies can provide insights that can have a significant impact on the success of AI various fields.

3.5 Conclusion and Future Research

In this study, the RRAVC, which is a quantitative CT imaging biomarker useful for distinguishing COPD patients from normal subjects, was predicted using ML rather than a defined formula. Among the various regression models, XGBoost yielded the largest R2 score and smallest MAE and MSE values. Accordingly, three relative coordinate methodologies were proposed and experimentally confirmed to yield a better performance than conventional naive coordinates. In particular, among the proposed three coordinates, relative proportional coordinates, which reflect the relative position of each point in the subject's lung, yielded the best performance. Moreover, the RRAVC was predicted by using MLP, which is a basic deep learning model. The best MLP model architecture for RRAVC prediction was determined by varying the number of MLP layers and structure of MLP models. Subsequently, the performance improvements were showcased by comparing the results to those obtained with conventional ML models. The XGBoost or MLP model tested in this study can be used as a baseline for RRAVC prediction using ML in the future.

In conclusion, this research provides important insights into the application of AI for tabular data analysis. Specifically, it demonstrated that parameter importance needs to be explored more thoroughly when determining which ML model to use because this can significantly impact the accuracy of the model results. Furthermore, the experiment demonstrated that for XGBoost and LightGBM algorithms, "in what proportion between the min and max" is more important than "how far from the min." MLP, on the other hand, achieved the best performance even when using absolute coordinates, suggesting it could be the most suitable model for this type of data.

The results of this experiment present numerous avenues for future research. For instance, it would be beneficial to further investigate the parameter roles when implementing different ML and deep learning models. Moreover, it would be valuable to compare different scenarios and explore the benefits of using multiple models simultaneously to produce better overall performance. Additionally, it would be useful to explore the potential impact of data points that may have been overlooked during the experiment owing to the nature of tabular data. Finally, in the future, more sophisticated ML models and deep learning models will be implemented to predict other functional variables, such as J and functional small airway disease (fSAD). The lung bronchi can also be viewed as a graph structure; therefore, several graph neural networks will be attempted to be applied these structures. Furthermore, in this study, data extracted from two images were used, namely, inspiratory and expiratory CT images. However, attempts will be made to predict the RRAVC by using only an inspiratory or expiratory CT image for improving efficiency. Ultimately, this experiment has provided an interesting overview of the application of AI for tabular data and has highlighted some key considerations and potential avenues for further research.

Chapter 4

AI Research on Tabular Data II (Study 2)[②]

4.1 Research on Tax Deliquency

4.1.1 Research Summary

This study focused on deep learning-based delinquent taxpayer prediction models for local taxation analyses in South Korea. To assess the viability of using AI models in predicting local tax arrears, popular ML and deep learning algorithms, including convolutional neural networks (CNNs), long short-term memory (LSTM), and sequence-to-sequence (seq2seq) algorithms, were employed. To evaluate the predictive power of these algorithms, credit information was combined with the algorithmic outputs. Specifically, the dataset comprised loan and delinquency information, credit card information, and public arrears. This wider feature set would prove advantageous when making accurate predictions. By better predicting local tax

---

[②] This study is extended based on the following publication by **Kim, E.** and Yoo, B. (2018). Credit information and public big data analysis: Development of prediction model for the possibility of recovering tax arrears and improvement of tax arrears information system. *Korea Society of Management Information System International Conference (KMIS International Conference).*

arrears, a waste of administrative resources including manpower could be avoided.

As with most AI-based models, a central part of the selection criteria adopted in the research focused on selecting and optimizing the data. This involved selecting a large dataset with clear outcomes and trends, as well as clean input variables that accurately represented the features to be analyzed. As local taxation can be affected by various inputs, it was essential to select a dataset that could capture these subtle changes. The selected dataset needed to be valid and consistent across countries and states, with recent data reflecting more current trends. Given the present research aim, it was essential to determine the viability of each algorithm. Performance metrics, such as accuracy and precision, were utilized to evaluate the models used for predicting tax arrears. The research found that the CNN framework demonstrated the highest accuracy and LSTM also showed excellent results. Ultimately, for improved performance, it was determined that a combination of several algorithms and application of an ensemble technique would be needed in future local tax evaluation and delinquent prediction.

4.1.2 Key Features of Research

The research assesses deep learning techniques for predicting individual local tax arrears. As a result, it is now possible to gain greater insights into a person's creditworthiness. Applying such a model can help reduce administrative waste, such as manpower, for pursuing tax defaults. The research leverages four different deep learning algorithms, namely, CNN, LSTM, residual neural network (ResNet), and seq2seq. Several ML algorithms are used as well, namely, random forest and SVM. For each model, different sets of input data are used, such as loan and delinquency data, credit card records, and public arrears data. This combination of inputs provides a greater range of features compared to traditional models that rely only on statistical approaches.

The CNN framework is used to process information contained in two or more related datasets. For example, a person's credit score and one's payment history on credit cards can be processed together, providing greater insights than individual pieces of information. On the other hand, LSTMs focus on longer-term temporal dependencies and can analyze larger sets of data than other deep learning algorithms. Finally, the seq2seq algorithm helps to identify patterns in data by automatically "mapping" the relationships between related elements. To evaluate the effectiveness of these algorithms, experiments were conducted in which accuracy metrics, such as the F1 score and AUC values, were recorded. It was found that CNN had the highest accuracy, while LSTM also yielded good results. These models could, therefore, be effectively used to accurately predict local tax arrears in South Korea. The research proposed standard models for predicting individual local tax arrears using deep learning algorithms. By analyzing various input data types reported in Table

4.1, such as loan and delinquency information, credit card information, and public arrears records, these models achieved satisfactory accuracy in experiments, proving the effectiveness of deep learning in this field. As such, these algorithms can help avoid the wasteful expenditure of resources on tax defaults in the Republic of Korea.

4.2 Research Design

4.2.1 Data Collection and Preprocessing

The data collection and preprocessing conducted in this study required a significant amount of time and resources. To begin, a stratified random sampling technique was used to select the individuals in the Republic of Korea for three fiscal tax years (from January 2015 to December 2017). The sampling result showed that the total proportion of local tax defaulters in Korea was calculated as 12.05%, which was found to be very close to the sample result of 12.07%, proving the reliability and validity of the dataset. This was an important step to ensure a fair representation of local taxpayers and defaulters, which corresponded to a total of 354,130 anonymized individuals excluding corporations. After the selection process, 192 variables were extracted from the chosen sample to be further studied. A brief description of the variables is provided in Table 4.1. The variables included the target variable of whether or not the individual defaulted on the local tax, which was determined after 30 days had passed since the date of the tax notice, with the payment status remaining as not received. This metric could easily be checked with existing automated admin systems.

The preprocessing stage of the research consisted of transforming and analyzing the data, as well as cleaning and standardizing the information to a suitable format. Specifically, the study preprocessed the data by selecting a range of local taxes, implementing summary functions to group certain values, creating additional calculated variables from the extracted data, deleting rows with any missing data, and renaming the remaining variables for easier understanding. Finally, the dataset was split into two sets: the training (first two-fiscal years) and test datasets (last one-fiscal year), which allowed the study to further evaluate the performance of the model and increase its overall accuracy. The entire research was conducted using an Nvidia RTX 3080 GPU, 4 Intel i7-10700 CPU, and 128 GB of RAM.

Table 4.1 Utilized variables: The credit information of local tax arrears including the public information. The target variable is local tax arrears status, which indicates normal tax payers or arrears.

Credit information	Loan and overdue information	Number of institutional lenders, Number of loans, Loan amount, Loan balance, Loan expiry date, Loan status, Number of overdue cases, overdue status, Number of default cancellations, Cancellation of arrears, Days after delinquency cancellation, Credit score
	Credit card information	Amount of credit card usage, Number of credit cards, Debit card usage amount, Credit card usage amount, Credit limit amount, Lump-sum usage amount, Installment usage amount, Installment usage rate, Short-term card loan amount, Number of short-term card loans, Long-term card loan amount, Number of long-term card loan, Number of using card companies, Loan expiry date, Loan balance, Number of overdue cases, Overdue status, Number of default cancellations, Cancellation of arrears, Days after delinquency cancellation
Public arrears information	Arrears information	Days after local taxation, Local tax amount, Local tax types, **Local tax arrears status (TARGET variable),** Amount of local tax arrears, Number of local tax arrears, Status of local tax urge by the government, Days after local tax arrears
	Public information	Missing person information, Incompetent person information, Quasi-incompetence person information, Personal bankruptcy / rehabilitation person information, Disclaimer / exemption person information, Transaction suspension information, death information, Suspicion information for multiple resident numbers, Repayment ability information, Annual income information, Mobile communication information, Utility bill payment information, Employment status, Period of employment, Four

The de-identified dataset was provided solely for research purposes by the National Information Society Agency in South Korea.

4.2.2 Methodology

1) Prediction Models

Despite the numerous and diverse predictive ML models for classification, in this study, such models were divided into three categories and applied accordingly. First, the seq2seq algorithm was used, which is commonly used in translations. Seq2seq implements LSTM from the family of recurrent neural networks (RNNs). Moreover, a CNN that utilizes the local information of adjacent data and ResNet, which is an extended CNN version combining the output of each layer, was used. Finally, the delinquent taxpayer prediction was modeled using a random forest and SVM, which are the most popular ML methods other than deep neural networks. A brief overview of the six models used for default prediction is presented as follows.

① LSTM

LSTM is an extension of the RNN algorithm, and it can remember older information more effectively than an RNN owing to its long-term memory (Hochreiter and Schmidhuber, 1997). Because RNNs by nature learn using a short-term memory, they suffer from a long-term dependency problem in which the information that comes out during the early phase vanishes during the later phase. The cell state of the LSTM solves this problem. The cell state, a key aspect of LSTM, has several elements called “gates” that allow eliminating unnecessary information or adding and storing new information. It also acts as a gradient highway to prevent a vanishing gradient, allowing information to spread farther. LSTM consists of four stages. As the first step, LSTM uses a forget gate to determine how much of the previous information has been forgotten.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{4.1}$$

The forget gate ($\mathrm{f_t}$) takes the previous hidden state ($\mathrm{h_{t-1}}$) and current input data ($\mathrm{x_t}$) as inputs and converts them into values between zero and one as they pass through the sigmoid function. If the output of the forget gate has a value of zero, the past memory completely disappears. Conversely, if the output of the forget gate has a value of one, the past memory is completely preserved.

The second step is the input gate ($\mathrm{i_t}$), which determines how much of the current data will be reflected. As with the forget gate, it takes $\mathrm{h_{t-1}}$ and $\mathrm{x_t}$ as inputs, which are converted into values between zero and one using the sigmoid function. The new information, $\tilde{C}_t$, was also obtained using $\mathrm{h_{t-1}}$ and $\mathrm{x_t}$:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i \tag{4.2}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{4.3}$$

In other words, the new information ($\tilde{C}_t$) to be stored in the cell state and how much new information (i_t) has been stored were computed through this step. During the third step, the cell state is updated. After determining the amount of past information to forget using $\mathrm{f_t}$, and how much current information is to be reflected using $\mathrm{i_t}$, their sum becomes the input to the next cell state as follows:

$$C_t = f_t \circ C_{t-1} \ + i_t \circ \tilde{C}_t \tag{4.4}$$

The final step is to determine the value of $\mathrm{o_t}$ to be sent to the output. Accordingly, $\mathrm{o_t}$ is multiplied by $\tanh(\mathrm{C_t})$ to determine $\mathrm{h_t}$, which will be entered as the input of the next state.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o \tag{4.5}$$

$$h_t = o_t \circ \tanh(C_t) \tag{4.6}$$

② Seq2seq

Seq2seq (Sutskever et al., 2014), which is widely used in machine translation and document summaries, receives a sequence of words as input and then outputs another sequence of words. Seq2seq consists of an encoder and decoder, in which an RNN is primarily used. In the case of machine translation, the encoder generates the context of one sentence as a vector, which is passed to the decoder. For example, in an English-Korean translation, the encoder encodes one English sentence sequence into a context vector (compression). Then, the decoder decodes the encoded context vector to generate the corresponding Korean sentence (reproduction). However, because the seq2seq model represents the context of an entire sentence as a single fixed vector, it struggles with handling particularly long sentences. The longer the sentence is, the more information that disappears regarding the words that appeared first. Bahdanau et al. (2015) introduced an attention technique to address this problem. The seq2seq model with attention solves the aforementioned problem by setting an attention score, which is a weight for each word that is fed as the input to the model. For all words of both the encoder and decoder, an attention score is calculated through a specific operation (e.g., dot product), which is reflected in the output of the decoder. This seq2seq model performs well in language-related domains and is also known to yield a good classification performance. Therefore, this model was selected as a candidate.

③ CNN

A CNN is a model that mimics the human optic nervous system and since the advent of AlexNet has been a commonly used framework in the deep-learning era (Krizhevsky et al., 2017). A CNN consists of a convolution layer and pooling layer and achieves a good performance when the locality of the data is important within the spatial structure. In particular, it shows an excellent performance in image processing, where the spatial structure and locality are extremely important. Because the convolution layer only connects to the input data or the output of the previous layer belonging to the received field, only the part belonging to the received field reflects the information. The pooling layer extracts only the necessary information from the window. Typical examples are max pooling, which extracts only the maximum value from the window, and average pooling, which extracts the average value of the window. If more convolution and pooling layers are stacked, information regarding a larger area in an image or video can be obtained.

④ ResNet

Deep learning techniques, including CNN, perform well in various fields; however, when the layer is too deep, the performance deteriorates owing to a vanishing or exploding gradient problem. In most ANN models, the weights are updated using a backpropagation method, which involves computing the derivatives of the loss function for each weight. The deeper the neural network is, the more derivatives are multiplied. If a small derivative is multiplied several times, it approaches zero, which indicates a gradient loss. Conversely, when a large derivative is multiplied several times, the value becomes extremely large, which is called a gradient explosion. In general, ANNs are predicted to perform better as the layers deepen. However, in a research conducted by He et al. (2016), the authors showed that a plain 20-layer network has fewer training and test errors than a plain 50-layer network for an image classification problem, which they stated was not caused by overfitting. This problem is called degradation and is caused by the loss of gradient. To solve this problem, ResNet, which uses the skip connection method where input x is added to the output value after several layers, has been proposed. Assume that when the input value of a neural network is x, the value created through several neural-network layers is F(x). In neural networks, the goal is to find a function H(x) that maps an input x to a target y, typically with H(x) = F(x). However, ResNet, which uses a skip connection, defines H(x) = F(x) + x and optimizes it. Using

ResNet, the authors built deeper neural networks that achieved a better performance.

⑤ Random Forest

A single decision tree is extremely susceptible to overfitting; accordingly, a random forest using multiple decision trees is applied as an ensemble approach to solve the overfitting problem. A random forest uses a process called bagging, which applies selected subsets of data or features, rather than using all data and features, to build each tree (Ham et al., 2005). When a decision tree is modeled, all features are considered and those with the highest information gain are selected; subsequently, the data are split based on this gain. However, with a random forest, multiple trees are created, where each tree considers only a fraction of the total features. Here, the term "random" means that when creating each decision tree, the features to be used are randomly selected. After creating multiple trees, voting is conducted on the results of the classification of each tree, and the final result is selected by a majority vote. A random forest is excellent at preventing overfitting and improves the prediction accuracy. In addition, for classification tasks, using random forests, relatively important features can be ranked.

⑥ SVM

An SVM is a model that defines a decision boundary, which is a classification criterion (Noble, 2006; Hearst et al., 1998). Classification is applied on new data by determining which side of the boundary it falls on. This decision boundary is a line in two dimensions, plane in three dimensions, and hyperplane in higher dimensions. Support vectors are the closest data points to the decision boundary. The objective of a Support Vector Machine (SVM) is to identify a decision boundary that maximizes the separation distance between the support vectors. The margin is the distance between the decision boundary and support vectors. In other words, with an SVM, the optimal decision boundary maximizes the margin. If the given data are linearly separable, the decision boundary can be found directly through an optimization, which minimizes the objective function L.

$$L = \min \frac{1}{2}\|W\|^2, \quad s.t \quad y_i(w \cdot x_i + b) \geq 1 \tag{4.7}$$

Here, y_i are the labels and x_i are the input data. Therefore, SVM can be expressed as follows for linearly separable data:

$$f(x) = w^t x + b = \sum_{i=1}^{N} y_i \alpha_i x_i x + b \tag{4.8}$$

However, the above formula does not allow any errors. If the data are not divided by a straight boundary line and linear separation is impossible, errors can be accommodated using the slack variable ξ, as shown below. This means determining the hyperplane while allowing some errors to occur. Here, C is a hyperparameter that determines the strength of the error penalty.

$$\min \frac{1}{2} \|W\|^2 + C \sum_{i=1}^{N} \xi_i \ , \tag{4.9}$$
$$s.t \quad y_i(w \cdot x_i + b) \geq 1$$

The above formula is converted into a dual problem using a Lagrangian as follows.

$$\min \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j (x_i \cdot x_j) - \sum_{i=1}^{N} \alpha_i \ , \tag{4.10}$$
$$s.t. \sum_{i=1}^{N} y_i \alpha_i = 0, a \leq \alpha_i \leq C$$

In addition, if the data are not linearly separable, a kernel trick is used to map the data into a high-dimensional space and then determine the decision boundaries in that hyperplane. Instead of using the original D-dimensional independent variable vector x, the M-dimensional vector $\phi(x)$ can be used, which is transformed by the basis function as the independent variable ($\phi(\cdot): R^D \rightarrow R^M$). In this case, the original vector x is transformed as follows:

$$x = (x_1, x_2, \dots x_D) \rightarrow \phi(x) = (\ \phi_1(x), \phi_2(x), \dots \phi_M(x)\ ) \tag{4.11}$$

For this expression, x is changed into $\phi(x)$ with a basis function transformation, and because all basis functions are used only in the form of $\phi(x_i)^T\phi(x_j)$, they can be represented as $K(x_I, x_j)$. The resulting formula is expressed as follows:

$$\min \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j K(x_i \cdot x_j) - \sum_{i=1}^{N} \alpha_i, \tag{4.12}$$
$$s.t. \sum_{i=1}^{N} y_i \alpha_i = 0, a \leq \alpha_i \leq C$$

Kernel, $K(x_i, x_j)$ can be viewed as a criterion to measure the similarity between two data samples, and the linear kernel and radial basis function (RBF) kernel are often used in practice.

2) Performance Metric and Model Architecture

In this study, an experimental design was employed to determine the accuracy of various deep-learning architectures. The performance metric was evaluated by calculating the number of true positives (TP) and true negatives (TN) with respect to the total number of true positives, false positives, true negatives, and false negatives (TP + FP + TN + FN). For the experiment, an LSTM network was implemented with five layers (512 x 256 x 128 x 32 x 2). The seq2seq model consisted of four layers (256 x 128 x 32 x 2), while the CNN consisted of two convolution layers and one pooling layer (conv → conv → pooling) followed by five fully connected layers (2024 x 512 x 128 x 64 x 2). Finally, a ResNet model was built with four layers (ResNetLayer → Pooling) followed by three fully connected layers (512 x 128 x 2). These architectures were carefully chosen to provide optimal performance while reducing the need for any extensive parameter tuning and adjustments. To determine the performance of each model, the TP and TN scores were computed and compared with the FP, TN, and FN scores, which provided insight into the accuracy of each architecture.

Various performance metrics, such as accuracy, were collected, stored, and then analyzed. An experimental protocol was applied with multiple test sets to produce more reliable and representative results. Random samples from the two groups of people, namely, local tax defaulters and non-defaulters (i.e., normal), were selected and fed into the models for performance evaluation. The statistical model that was created helped assess the local tax defaulters and non-defaulters. All results and data collected were cross-examined and confirmed to obtain trustworthy and significant results. Furthermore, parameters and thresholds were optimized through several test sets, such as precision and recall, to avoid overfitting or bias.

This experimental research provided a detailed investigation of the research problem and an improved understanding of the dataset. A performance metric was selected, which allowed to make predictions on the target classes of local tax defaulters and non-defaulters. Appropriate models were chosen with well-designed architectures, providing enhanced and more reliable results. Lastly, extensive tests were conducted to obtain meaningful and useful conclusions from the results.

4.2.3 Limitations and Considerations

This study had various limitations and aspects that must be considered when interpreting its findings. The experimental design was used to assess the accuracy of various predictive models for the task of distinguishing between local tax defaulters and non-defaulters (i.e., normal). Accuracy was measured where the weighted average of the group was determined based on the number of people in each group and TP, FP, TN, and FN indicated actual results against the prediction made by the model.

When designing the architecture of the predictive models, efforts were made to keep the models relatively simple and efficient considering the load on the operation systems. The LSTM had five layers (512 x 256 x 128 x 32 x 2), seq2seq had four layers (256 x 128 x 32 x 2), CNN was designed with two convolution layers and one pooling layer followed by five fully connected layers, and ResNet was designed with four ResNetLayer→Pooling layers followed by three fully connected layers. Accordingly, the model layers were not very deep and model complexities were not high. Therefore, further refinements may be necessary to improve the models' performance and usability for practical application.

Another limitation of this research is the number of people studied. In general, a larger sample size can provide more reliable results than a small one. A larger population also increases confidence in the representativeness of the sample data, allowing for more detailed inferences. Using only accuracy as a metric to assess the models can also be considered a limitation, as it fails to measure other parameters, such as computational speed, sensitivity, or false-negative rate. Other measures, such as the area under the curve and sensitivity analysis, might be more suitable for the task of distinguishing between local tax defaulters and non-defaulters.

In summary, this research is subject to several limitations that must be considered when interpreting its results. Although an experimental design was used and various predictive models were designed, their architectures may not have been optimal and sample size may have been small; moreover, accuracy was the only measure used. Thus, the results should be taken as a starting point for further investigation and development.

4.3 Experimental Results

The results, compared the performance of various models for local tax arrears prediction, are shown in **Table 4.2**. The most accurate model was the CNN, achieving slightly higher accuracy than the seq2seq with attention. Other models performed similarly, achieving 82 – 83% accuracy. This indicates that these models, while achieving satisfactory performance, could be improved further. Given the aim of this research was to develop relatively shallow models for administrative operations, the relatively low performance of the ResNet model could be attributed to its lack of depth in comparison to that of the other models used in this study. Therefore, future research could explore more sophisticated architectures with deeper stacking to boost accuracy, along with other potential improvements to architecture and hyperparameter settings.

Table 4.2 Various model prediction accuracies for local tax default prediction.

Model	Accuracy (%)
LSTM	83.39
Seq2Seq with attention	84.43
CNN	**87.64**
ResNet	83.72
Random Forest	82.86
SVM	82.23

The overall accuracy results for the models considered in this research demonstrate the usefulness of AI algorithms in administrative operations and for predicting arrears for local taxes. Optimizing models for accuracy and relying on more sophisticated models such as CNNs and seq2seq with attention can ensure greater accuracy, and ultimately, performance and productivity can be improved. Accordingly, this research has highlighted the importance of deep learning models and their potential for achieving greater accuracy and predictive success in local tax arrears prediction.

The experimental results, as displayed in Figure 4.1, indicate that while there were no significant differences between the models, LSTM performed relatively well. This can be observed when looking at the receiver operating characteristic (ROC) curve. Generally, the area under the curve (AUC) represents the accuracy of the models in correctly distinguishing between positive and negative results; therefore, it is usually used for checking the variance and stability of model. Notice that the LSTM model had a slightly higher AUC than that of the other models, underlining its superiority in experimental performance. Despite the LSTM model demonstrating an overall better performance than the other models, note that the standard deviation of its results was also the highest. This suggests that, although LSTM had superior results on average, its accuracy may vary drastically in certain situations. This is further reinforced by the other models' AUCs remaining consistent and not showing any extreme outliers.

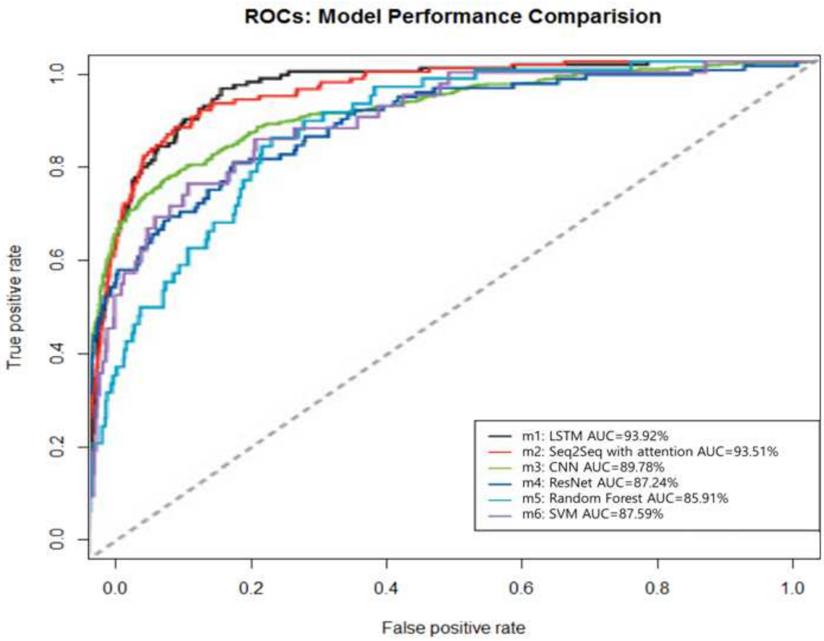

Figure 4.1 ROC curves of various models for predicting local tax default status.

Moreover, a major difference between the LSTM model and other models is its temporal component. Because LSTM stores memory from past inputs, it can pick up subtle trends from previous iterations and yield more accurate results. This has proven to be quite useful in understanding complicated trends, such as speech patterns or changes in stocks. Accordingly, the experiment yielded results that, while showing no major difference between the models, showcased LSTM's potential superiority in terms of performance. Its ability to utilize data from past inputs and draw more accurate conclusions, although making its results vary extremely in certain situations, could still provide it with a definitive advantage over the other models. Therefore, more research should be done to truly gauge the LSTM model's capability and gain further insight into its behavior.

4.4 Discussion and Implications

The findings of this research illustrate the potential of AI in aiding local tax arrears prediction. Using CNN and seq2seq with attention, a slightly higher accuracy than that obtained with other models tested can be achieved. Despite other models providing satisfactory accuracy, namely, between 82–83%, their accuracies could be improved further by increasing the depth of the architectures or optimizing hyperparameters. Note that the shallow model of ResNet achieved the same accuracy as the other models tested. This demonstrates that a model's depth has a direct effect on the accuracy achieved and may warrant further exploration and experimentation of different architectures in the future. Such experiments should strive for improved accuracy, to ensure that AI can truly be a powerful tool in optimizing performance and productivity in this context.

This study has highlighted the potential of AI applications and associated technologies in both administrative operations and prediction tasks. While it has been established that AI can effectively be used for predicting arrears for local taxes and other related tasks, there is room for improving its accuracy and efficiency. Thus, more studies should investigate such enhancements to leverage the most out of AI and attain better organization performance. In terms of improving accuracy, there are various advanced architectures that could be utilized to maximize the potential of AI and obtain more accurate predictions. Examples of such architectures include combination or ensemble of different models. By applying the appropriate architecture for a particular task, the results are likely to be more accurate as AI algorithms would learn patterns better. Additionally, other more innovative and specialized AI techniques such as reinforcement learning can also be employed in such tasks to improve accuracy.

In addition to accuracy, other important aspects to consider in terms of exploring the potential applications of AI are efficiency and scalability. To maximize the benefits of AI, it is important that the processes and operations involving it are efficient and scalable. Such aspects could be achieved through various methods, such as transferring more work to the server side and optimizing algorithm configurations. Furthermore, more efficient computing devices can also be leveraged, such as using powerful graphical processing units (GPUs) and better storage systems for data. Moreover, this study has shed light on the capability of AI in predicting arrears for local taxes and other administrative tasks, which can be further enhanced with additional research and developments in terms of both accuracy and efficiency. Consequently, AI-based solutions could bring substantial value to businesses and other organizations, which makes the continued exploration and exploitation of AI in various scenarios a promising research direction.

Although the accuracy has already been identified, the reason for introducing AUC is that while it is important to predict delinquents, it is also increasingly important to catch normal tax payers as normal, so the corresponding indicator is used to check the stability of the model. The result of the experiment indicates that while there were no drastic differences between the various models, the LSTM model exhibited superior performance with its slightly higher AUC value. Moreover, the standard deviation of its results was the highest, which implies that the accuracy of the LSTM model is more varied compared to those of the other models. A significant distinction between the LSTM model and other models is its temporal aspect. This attribute of the LSTM model is crucial because of its capacity to record information from prior inputs, thereby enabling it to predict subtle trends that could potentially yield more accurate results. This feature was certainly utilized during the experiment, which further underlines its value in distinguishing patterns from large and complicated datasets. For example, LSTM models have demonstrated remarkable performance in recognizing and translating speech, as well as in making predictions about the variations in stock markets. These results highlight the potential of the LSTM model over the other models. It is also clear that, although it exhibited superior performance during the experiment, its accuracy still varied drastically at times, which should not be ignored. Therefore, further research should be conducted to study the behaviors and abilities of the LSTM model more thoroughly. To sum up, while there were no drastic differences between the models during the experiment, the results are nevertheless indicative of the superior performance of the LSTM model compared to the other models. Its temporal attribute enables it to make predictions based on prior inputs, which was clearly demonstrated in the experiment. However, its results tend to fluctuate, thus necessitating further study and analysis of its behaviors and capabilities.

4.5 Conclusion and Future Research

This research concludes that AI can be an effective tool for predicting local tax arrears, with the CNN model attaining the highest accuracy among the tested models. Despite achieving an accuracy of 87.64%, further research can be conducted to improve this accuracy. Moreover, studies can be conducted to analyze and explore latent variables indicating the presence of subtle patterns in a dataset. Additionally, given the rising popularity of AI, future studies should focus on further improving the accuracy and efficiency of AI-based solutions. To facilitate further exploration of AI applications in tabular data analysis, future research should also aim to discover and test additional innovative uses for AI and analyze the advantages and challenges associated with implementing such solutions. For example, more advanced algorithms and architectures should be tested, as well as deep learning models, to optimize the accuracy of AI models. In addition, further investigations should focus on making AI-based solutions more cost-effective for organizations, so that their adoption is accelerated in the industry. Generally, the experiment conducted in this research has demonstrated that there is potential in applying AI for predicting local tax arrears. To take advantage of the possibilities created by using AI, future studies should focus on studying its capabilities and ethical considerations, optimizing the accuracy of AI models, and exploring potential uses of AI to provide solutions in the domain. With further research, AI applications can continue to make an impact on the industry and ultimately lead to improvements in efficiency and accuracy of prediction of local tax arrears.

Chapter 5

AI Research on Image Data (Study 3)[③]

5.1 Research on Security Detection

5.1.1 Research Objective

You look only once (YOLO) algorithm is relatively ineffective at detecting small objects owing to its high detection speed (Adarsh et al., 2020). Therefore, studies are being conducted in the aeronautical field to enhance the detection of small objects that are difficult to identify with the naked eye. M. Liu et al. (2020) developed a model that detects automobiles and people from a video of the ground filmed by an unmanned airplane. Generative adversarial networks (GANs) have been utilized for data augmentation to develop a network that produces high-resolution videos from low-resolution

③ This study is based on the following publication by Kim, E., Lee, J., Jo, H., Na, K., Moon, E., Gweon, G., Yoo B., and Kyung, Y. (2022). SHOMY: Detection of Small Hazardous Objects using the You Only Look Once Algorithm. *KSII Transactions on Internet and Information Systems*.

satellite-filmed footage (Rabbi et al., 2020). In this application, object detection and model learning were performed simultaneously, leading to enhanced detection of small objects. However, despite these efforts, prior research on the detection of small objects have failed to maintain performance or provide enhancements for datasets that include objects of different sizes. Owing to the difficulties encountered in analyzing large datasets and securing proper data labeling, studies on X-ray object detection have failed to achieve the same performance for detecting small, medium, and large objects. Therefore, studies on the topic remain scarce.

Accordingly, this study aims to improve the detection performance regarding small hazardous objects and develop a model that can effectively detect even information protection objects, such as USB, which are related to the possibility of information leakage. Therefore, in this study, a model to improve the performance regarding small-object detection without degrading the performance related to detecting large or medium-sized objects is proposed.

5.1.2 Key Features of Research

The research explores an important topic related to safety, namely, the detection of small hazardous objects in airport cargo. Research on object detection, and especially the detection of small objects, has been lacking. Therefore, the goal of this paper was to develop a new model to improve the detection performance related to small objects. The proposed model is called the "small hazardous object detection enhanced and reconstructed model," and is improved based on the recently launched object detection model you look only once version 5 (YOLOv5). Through different experiments based on YOLOv5, the detection performance of the proposed model was improved by 0.3 in terms of mean average precision (mAP) index and by 1.1 in terms of mAP (.5:.95). One of the main advantages of the proposed model is its ability to detect small objects of different types in overlapping environments where objects of different sizes are densely packed. This study also states that the proposed model does not require data preprocessing for immediate industrial application, which helps reduce the cost of implementing this technology.

Another advantage of this research is that it offers the possibility for rapid and effective application in industries related to airport cargo and passenger safety without performance degradation. This could prove especially helpful in terrorism prevention in the near future. Additionally, it is suggested that further experiments and improved preprocessing methods can further enhance the model's performance in this field. Generally, the study has made a significant contribution to the detection of small hazardous objects in airport cargo. Through the development of the small hazardous object detection enhanced and reconstructed model based on the YOLOv5 algorithm (SHOMY), the study has created a technology that offers an improvement in performance, which is also easily implementable in the industry.

5.2 Research Design

In this study, the characteristics of real-time detection of target objects through aeronautical X-ray scans were considered; accordingly, slow two-stage detectors were deemed unfit for baggage search and detection. Subsequently, model development and research on YOLO (Adarsh et al., 2020; M. Liu et al., 2020), the fastest network among one-stage detectors, were performed. The research modified YOLOv5, the latest YOLO model, and developed SHOMY as a new model with enhanced performance.

To overcome the limitations of existing research, object detection with 38 objects of different sizes applicable to actual baggage search was conducted using video data. Furthermore, YOLOv5 was applied as a one-stage detector. Although it is common to manipulate data to enhance detection performance on small objects, as done in existing studies (Kisantal et al., 2019; Ozge Unel et al., 2019), data preprocessing would require considerable additional calculation costs and may have an adverse effect on learning. Therefore, the study chose to proceed with only model tuning instead of learning data alteration or manipulation, such as increasing the data volume and resolution of the input data, oversampling, copy-pasting small objects, and tiling. This ensured that the object detection performance was maintained in actual systems.

5.2.1 Data Collection and Preprocessing

Image data related to prohibited objects obtained by a scanner from Rapiscan and provided by the National Information Society Agency, South Korea was utilized in this study. To build a system that detects target objects using X-ray during airport security searches at airports, ports, train stations, private companies, and public offices, a sizable dataset is required.

The dataset contained 38 discernable target objects. Based on Microsoft common objects in context (MS COCO), small, medium, and large objects were defined according to their surface areas, as outlined in Table 5.1. They were categorized into 5 small, 26 medium, and 7 large objects. Among the small objects, USB flash drives were considered sensitive storage media linked closely to the leakage of confidential information, whereas bullets, nail clippers, batteries, and lighters were considered objects that must be detected for flight safety (Kim, 2019; Andriyanov et al., 2021; Wei and Liu, 2020).

Table 5.1 Definition of different object sizes (small, medium, and large) and classification of target objects based on MS COCO standard.

Size	MS COCO Definition (Pixel) Min.	Max.	Classification of X-ray Target Objects (38 EA)
small	1 × 1	32 × 32	USB, bullet, nail clippers, battery, lighter (5 EA)
medium	32 × 32	96 × 96	throwing knife, match, electronic cigarettes, electronic cigarettes (liquid), awl, thinner, SSD, screwdriver, Zippo oil, liquid, aerosol, knife, portable gas, supplementary battery, smart phone, HDD, alcohol, scissors, spanner, handcuffs, gun parts, solid fuel, pliers, chisel, gun, firecracker (26 EA)
large	96 × 96	∞ × ∞	hammer, tablet PC, laptop, saw, axe, bat, metal pipe (7 EA)

Four types of data were provided according to their provision methods, as reported in Table 5.2. Because multiple target items for detection and general items are included together in actual images obtained in the field, “Multiple & Categories” and “Multiple & Others” were selected as data types. The former is a type of dataset that includes multiple target items for detection, other hazardous objects, and other general objects. The total number of distinct items for detection was 38, with the dataset containing a total of 54,949 data points. Of these, 41,211 (75%) were used for training, whereas the remaining 13,738 (25%) were used for validation. Data were randomly extracted and produced from each item at an equal rate of 25%.

Table 5.2 Types of X-ray baggage datasets.

Type	Details
Single Default	1 target object only
Single & Others	1 target object + other non-targeted objects
Multiple & Categories	multiple target objects + other target objects
Multiple & Others	multiple target objects + other non-targeted objects

The experiments were conducted using an Nvidia RTX 3080 GPU, Intel i7-10700 2 CPU, and 64 GB of RAM.

5.2.2 Methodology

1) YOLOv5 Model

YOLO realizes a one-stage detector model by defining object detection problems as regression problems. Available YOLO versions range from v1 to v5; among these, YOLOv5 was used in this study. Its object detection structure comprises a backbone layer for extracting traits, neck layer for aiding the detection of objects in different scales, and head layer for detecting objects.

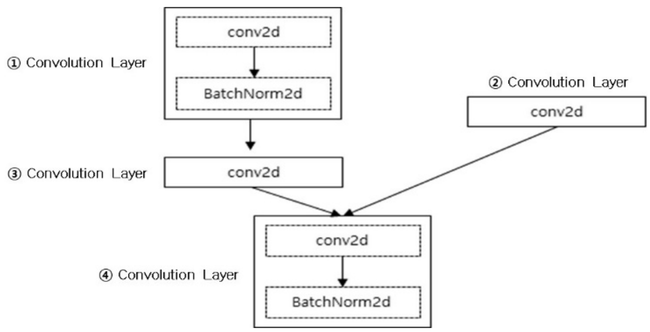

Figure 5.1 Structure of cross-stage partial network (CSPNet).

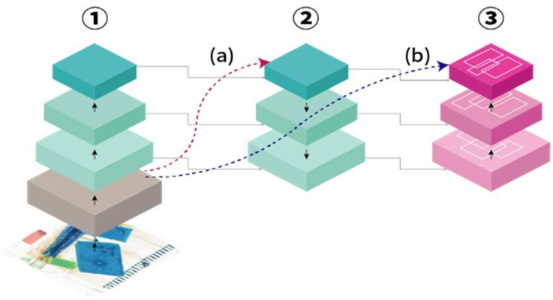

Figure 5.2 Architecture of path aggregation network (PANet).

The first stage, referred to as a backbone layer–cross-stage partial network, is for extracting image traits. In this stage, a cross-stage partial network (CSPNet), which is a model with a reduced number of calculations developed for being used in low-function computer environments or real-time image detection, is used. While a general CNN model requires large number of calculations with duplicate gradient problems, CSPNet integrates the feature map provided at the beginning and at the end to innovatively reduce the number of calculations (Wang et al., 2020).

The architecture of the CSPNet utilized by YOLOv5 is presented in **Figure 5.1.** The output value in the early base layer is divided into convolutions ① and ②. Subsequently, the value gained through convolution layers ① and ③ and that gained through convolution layer ② are merged. The value also passes through convolution layer ④, and consequently, the output value of the base layer is connected directly to the final convolution layer, serving as the gradient shortcut.

The second stage, referred to as a neck layer–path aggregation network, responds to different object scales. By using a path aggregation network (PANet) (Liu et al., 2018) as its backbone for modeling, the feature pyramid network (FPN) resolves the problem of the information from the first layer not being reflected properly in the final prediction.

In Figure 5.2, layer ① denotes the FPN backbone and comprises considerably large networks, such as ResNet-50. Therefore, low-level feature information must go through numerous layers to be conveyed to (a) a higher level, wherein loss of information is inevitable. PANet is designed to fully convey the information even if the information is passed through a shortcut (b) and through multiple convolution layers with the addition of layer ③.

The last stage, or the head layer, is used for predicting the possible locations of objects and utilizes the complete intersection over union (CIoU) loss proposed by Zheng et al. (2021). CIoU has a high learning speed compared to IoU, generalized IoU (GIoU), and distance IoU (DIoU), which are often used as object detection loss rates and are useful when detecting small objects.

$$\text{CIoU Loss} = S(B, B^{gt}) + D(B, B^{gt}) + V(B, B^{gt}) \tag{5.1}$$

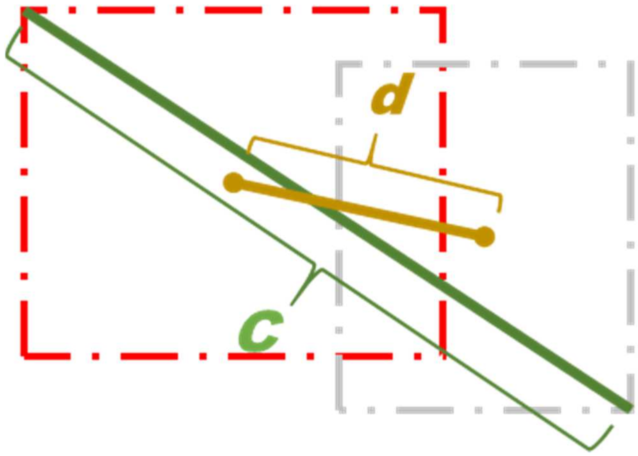

Figure 5.3 Definition of D in CIoU.

As presented in Equation (5.1), CIoU comprises the surface area (S), distance (D), and aspect ratio (V) of two boxes (B: box coordinates predicted by the model, B^{gt}: coordinates of the actual box (ground truth)).

$$S = 1 - IoU \tag{5.2}$$

where S is the loss of the overlapping area of the two boxes.

$$D = \frac{\rho^2 \ (p, p^{gt})}{c^2} \tag{5.3}$$

The distance, *D,* of the box is calculated based on the diagonal distance (c) and distance of the central points (d), as shown in Figure

5.3. In Equation (5.3), c is the diagonal distance and ρ is the distance between central points. Therefore, if the two objects are close, D approaches zero and loss decreases.

$$V = \frac{4}{\pi^2}(\arctan\frac{w^{gt}}{h^{gt}} - \arctan\frac{w}{h})^2 \tag{5.4}$$

Finally, V calculates the difference in the proportions of the widths, presented in Equation (5.2), and heights (h) of the two boxes, as stated in Equation (5.4), to determine whether the two boxes have similar forms. Ultimately, CIoU induces learning by simultaneously maximizing the overlapping area between the two boxes, minimizing the distance between them, and maintaining a similar form.

$$\begin{aligned}
&\lambda_{coord}\sum_{i=0}^{S^2}\sum_{j=0}^{B} 1_{ij}^{obj}[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] \\
&+\lambda_{coord}\sum_{i=0}^{S^2}\sum_{j=0}^{B} 1_{ij}^{obj}[(\sqrt{w_i} - \sqrt{\widehat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] \\
&+\sum_{i=0}^{S^2}\sum_{j=0}^{B} 1_{ij}^{obj}(C_i - \hat{C}_i)^2 \\
&+\lambda_{noobj}\sum_{i=0}^{S^2}\sum_{j=0}^{B} 1_{ij}^{noobj}(C_i - \hat{C}_i)^2 \\
&+\sum_{i=0}^{S^2} 1_{i}^{obj}\sum_{c\in classes}(p_i(c) - \hat{p}_i(c))^2
\end{aligned} \tag{5.5}$$

YOLO defines the loss by separating cases wherein there are objects on the grid cell and wherein there are no objects on the grid cell. If the elements of Equation (5.5) were to be labeled (a)–(e) from the top, (a) would denote the mean squared error (MSE) for the central point coordinates (x, y) of the object when there is an object in the grid cell. Element (b) would denote the MSE for the height (h) and width (w) of the object when there is an object in the grid cell. Here, square root is used to decrease the scale difference between large and small objects. Elements (c) and (d) would denote the confidence scores for when there is an object in the grid cell and there is no object in the grid cell, respectively. This is defined as *P(object)* × *CIoU*. Finally, (e) would denote the loss value of the conditional probability for the class when there is an object in the grid cell.

2) SHOMY: Performance Enhancement Model for Small-Object Detection

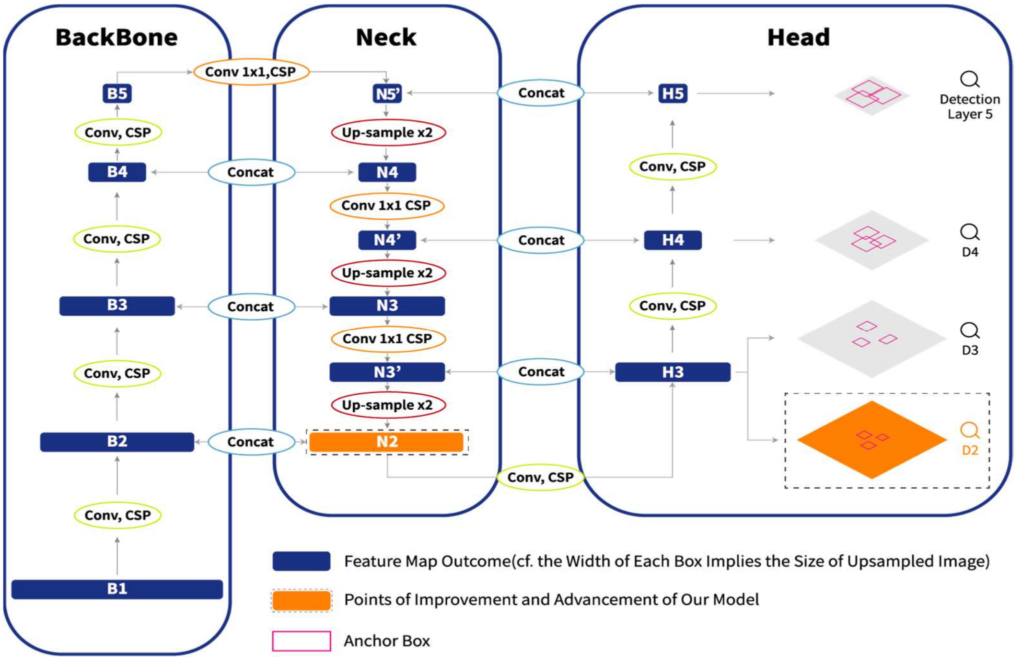

Figure 5.4 Architecture of SHOMY.

As reported in prior studies, X-ray baggage detection approaches differ between large and small objects. In addition, while the detection performance regarding large objects is now adequate to a certain extent, the detection rates for small objects, such as USB flash drives, bullets, and lighters, remain relatively low. Scaled-YOLOv4 (Wang et al., 2020) adds detection layers to the existing framework improve the probability of acquiring enhanced results compared with those of the basic model for increasing the detection rates for such small objects. The backbone network of YOLOv5 detects the location and spatial information of objects, whereas the neck network obtains the semantic information. Based on the characteristics of Scaled-YOLOv4 and YOLOv5, this study added a neck layer, namely, N2, which produces a massive feature map, as shown in **Figure 5.4.** This further expands the last feature map (upsample) and secures as much semantic information as possible.

The added N2 is connected to B2 of the same scale in the backbone to help the network detect the traits and spatial information of small objects. Meanwhile, the head is the final stage of object detection. As discussed, a stage dedicated to detecting small objects was added (Wang et al., 2020).

As shown in **Figure 5.4,** the detection layer D2, which utilizes an anchor box for small objects, was added to H3 of the head. The sizes of the boxes (5 × 7, 8 × 15, and 17 × 12) are half of those of three anchor boxes combined (10 × 13, 16 × 30, and 33 × 23). Consequently, regression learning is performed with a smaller anchor box in a space expanded further to detect small objects, such as USB flash drives (10 × 17). By adding N2 to the neck, a large feature map is produced; whereas by adding D2 to the produced feature map and utilizing a smaller anchor box, the detection of small objects is improved.

5.2.3 Measurement Indices

The basic indices for measuring object detection performance are precision and recall. The former refers to the proportion of actual true values among the experimental values identified to be true, whereas the latter refers to the proportion of values identified to be true among the actual true values. They are defined as follows:

$$\begin{aligned} \text{Precision} &= \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \\ \text{Recall} &= \frac{\text{True Positive}}{\text{True Positive} + \text{False } Neg\text{ative}} \end{aligned} \tag{29}$$

Precision and recall are complementary. When recall is enhanced in object detection, precision is bound to decrease, and it is important to allow the model to learn such that the values of both indices are high. Therefore, the mean average precision (mAP) index, an average index for detection that considers both precision and recall, is generally used. Generally, the mAP is measured based on an intersection over union (IoU) of 0.5, which refers to the overlapping ratio of the predicted area of the box where an object is placed to the actual area of the box (ground-truth of bounding box), as shown in Figure 5.5. In other words, the IoU is the area of overlap divided by the area of union. When the IoU is 0.5 or higher, mAP categorizes the predicted value as true (Vuola et al., 2019).

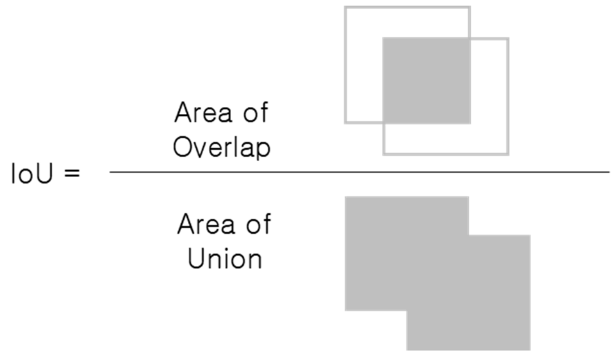

Figure 5.5 Calculation of IoU.

Another index measured with mAP is mAP (.5:.95), which is the average value measured based on increasing 0.5 to 0.95 at intervals of 0.05 and used as a more rigorous object detection performance index compared to mAP (Vuola et al., 2019). This study focused on enhancing the performance regarding the detection of small objects in terms of the two indices.

5.2.4 Limitations and Considerations

The research for advanced detection of harmful objects in airport cargo for passenger safety against terrorism is constrained by various limitations and considerations. This research is primarily focused on detecting larger objects. The current state of research concerning the detection of small objects has been comparatively lagging. This underlines the need for developing more sophisticated models to detect these objects more accurately, as well as the main motivation behind the creation of the small hazardous object detection enhanced and reconstructed model based on the YOLOv5 algorithm. While the developed model offers to improve object detection performance, note that object detection is an ever-evolving technology and existing models require consistent refinement. The potential of the proposed model will depend on various technical and material factors. Different environments can produce varying results, as conditions and accuracy are affected by low-light settings. When different object types of various sizes are densely packed and overlapping in crowded environments, where it will be particularly difficult to identify small objects in comparison to larger objects. Therefore, additional experiments will be needed to ensure the accuracy and effectiveness of the model in such complex situations.

In addition, one should consider the importance of data pre-processing. It is possible to greatly reduce errors during the object detection process by removing errors contained in the collected data beforehand. Unfortunately, at the current state of research in pre-processing technology, data pre-processing is a major challenge and requirement for high-performance object detection. Additionally, depending on user specific applications and requirements for the preprocessing techniques, additional expertise and specific data may need to be acquired. While the new model exhibits significant potential to improve detection accuracy, one should consider all the aforementioned issues before using this model for industrial purposes. By examining and fully understanding these limitations and considerations, further development and refinement can be achieved, potentially leading to even better models for hazardous object detection.

5.3 Experimental Results

This section presents the experimental results of performance enhancement evaluations regarding small objects detection in airport baggage using X-ray videos. In particular, the results achieved by YOLOv5 and SHOMY are compared and analyzed based on the definition of object size and classification of target objects that are described in **Table 5.1**.

The default YOLOv5 model achieved mAP = 98.7 and mAP (.5:.95) = 87.8, as reported in **Table 5.3**. The detection performance on small objects was lower compared with those on medium and large objects. In particular, the drop in small-object detection performance is notably large in terms of mAP (.5:.95) with rigorous IoU settings.

In contrast, unlike in the default YOLOv5 model, which performs learning based on an anchor box with a pre-designated size, an auto-anchor box uses the K-means algorithm to automatically produce an optimal anchor box within the box coordinate distribution of the dataset class. However, although the model managed to learn how to automatically produce 3–5 auto-anchor boxes, the performance was similar to or slightly lower than that of the default model. The underlying causes may need to be analyzed in the future; however, it is assumed that optimal groupings were not produced from the 38-object box coordinates.

Table 5.3 Experimental results of default YOLOv5 and SHOMY with auto-anchor settings.

	mAP%				mAP(.5:.95)%			
	All	**S**	M	L	**All**	**S**	M	L
Default YOLOv5	**98.7**	**96.4**	99.2	98.8	**87.8**	**76.8**	88.6	92.6
/w Auto Anchor 3	**98.7**	**96.2**	99.2	98.8	**87.7**	**76.7**	88.7	91.5
/w Auto Anchor 4	**98.6**	**96.0**	99.2	98.8	**87.4**	**76.0**	88.3	92.2
/w Auto Anchor 5	**98.6**	**95.8**	99.2	98.8	**87.4**	**75.6**	88.5	91.8
SHOMY (N2+D2)	**99.0**	**98.3**	99.2	98.7	**88.9**	**81.6**	89.5	92.1
/w Auto Anchor 3	**98.7**	**96.7**	99.1	98.8	**88.1**	**77.8**	89.3	90.9
/w Auto Anchor 4	**98.7**	**96.7**	99.1	98.9	**88.3**	**77.6**	89.3	92.3
/w Auto Anchor 5	**98.7**	**96.8**	99.2	98.8	**88.5**	**78.0**	86.6	92.7

(S: Small, M: Medium, L: Large; : YOLOv5, : SHOMY).

For each small item, the default YOLOv5 model demonstrated lower performance for smaller sizes, as reported in **Table 5.4**.

Table 5.4 Results of small-object detection for default YOLOv5 vs. SHOMY with different performance indices.

	USB drive (10 × 17)	Bullet (15 × 25)	Nail Clippers (18 × 34)	Battery (20 × 40)	Lighter (21 × 43)
YOLOv5 mAP%	92.0	96.3	99.5	95.6	98.5
YOLOv5 mAP(.5:.95)%	64.6	74.0	85.1	75.1	85.1
SHOMY mAP% (vs. YOLOv5 Default)	96.7 (▲4.7)	97.4 (▲1.1)	99.7 (▲0.1)	98.7 (▲3.1)	99.1 (▲0.6)
SHOMY mAP(.5:.95)% (vs. YOLOv5 Default)	72.7 (▲8.1)	79.3 (▲5.3)	87.1 (▲2.0)	81.3 (▲6.2)	87.5 (▲2.4)

(: YOLOv5, : SHOMY).

Table 5.5 Experimental results of the SHOMY model.

	mAP%				mAP(.5:.95)%			
	All	**S**	M	L	**All**	**S**	M	L
SHOMY (N2+D2)	**99.0**	**98.3**	99.2	98.7	**88.9**	**81.6**	89.5	92.1
SHOMY (N2)	**98.8**	**96.9**	99.2	98.8	**88.1**	**77.3**	88.9	92.8
Default YOLOv5	**98.7**	**96.4**	99.2	98.8	**87.8**	**76.8**	88.6	92.6

(S: Small, M: Medium, L: Large).

SHOMY achieved an average overall performance of mAP = 99.0 and mAP (.5:.95) = 88.9 (see **Table 5.5**), which is an enhancement of 0.3 and 1.1, respectively, compared with those of YOLOv5. In particular, enhancements of 1.9 and 4.8 were observed for small objects. Simultaneously, the detection performances of medium-sized and large objects were maintained without significant reductions or, in some cases, even increased slightly. Even models without the detection layer D2 exhibited slight increases in performance compared with that of YOLOv5 but remained weak compared to that of SHOMY. This result verifies that adding the detection layer D2, which is dedicated to the detection of small objects, greatly enhanced the overall performance. Additionally, in learning, there were weight losses for each of the detection layers D2, D3, D4, and D5, as shown in Figure 5.4. A weight of 4 was allotted to D2, whereas those of 1, 0.3, and 0.1 were allotted to layers D3, D4, and D5, respectively, which are dedicated to detecting medium-sized and large objects. The purpose of attributing large losses to errors in the detection of small objects is to ensure that the models have strong learning abilities in detecting small objects.

The USB drive, the smallest object analyzed in this study, had a size of 10 × 17. When the basic anchor box value was used, the size of the box was set higher than that of the object, eventually leading to a decreased learning ability. Therefore, it was necessary to optimize the anchor boxes with respect to the object sizes of the used data. As a result, new anchor boxes with dimensions of 5 × 7, 8 × 15, and 17 × 12, or half the size of the smallest anchor box used in the default YOLOv5 model were used.

As reported in Table 5.3, the performance of the newly developed SHOMY(N2+D2) model with the new anchor setting value was superior to that of the SHOMY model with auto-anchor boxes. The decline in detection performance on small objects was especially significant when auto-anchor settings were used. This result underlines that when learning a dataset with small objects, it will be more effective to manually set anchor boxes according to the object characteristics. As reported in Table 5.4, SHOMY exhibited an enhanced performance for small-object detection compared with that of the default YOLOv5 model. Notably, the mAP and mAP(.5:.95) indices for the USB drive, the smallest object analyzed in this study, increased by 4.7 and 8.1 to 96.7 and 72.7, respectively. Figure 5.6 depicts an example of an inference by SHOMY, demonstrating its high performance in detecting objects of different sizes in an X-ray baggage image featuring various objects. In this study, SHOMY was demonstrated to be capable of detecting small objects, such as USB drives, lighters, and bullets.

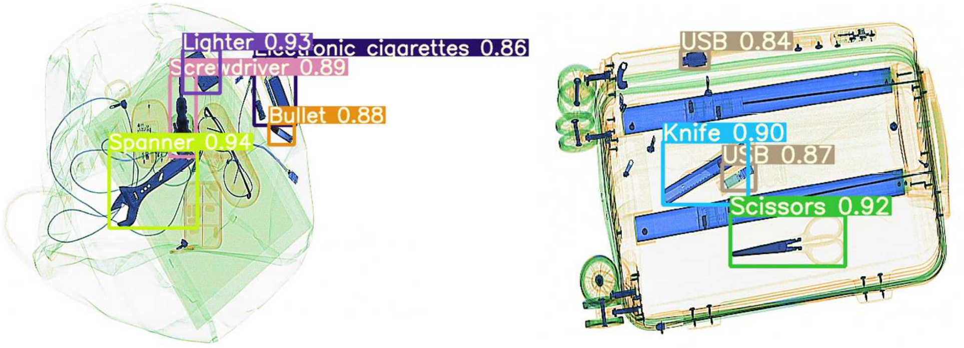

Figure 5.6 Example of inference images by SHOMY model.

5.4 Discussion and Implications

The object detection research field is currently experiencing important challenges because existing research has focused exclusively on the detection of medium- and large-sized objects, resulting in low detection performance for small objects. However, the hazardous potential of certain small baggage items cannot be ignored, as highlighted by incidents such as smartphone explosions and continuous industrial espionage cases using ultrasmall storage media, which have been reported in prior research. Therefore, it was aimed to enhance the detection performance for small objects while maintaining the existing detection performance levels for medium-sized and large objects.

This study developed the SHOMY model based on YOLOv5 for enhancing the detection performance for small objects contained in airport baggage captured by X-ray videos. A comparative analysis was performed between the performances of two models with different auto-anchors to verify the superior small-object detection performance of the newly developed model. To improve the detection of small objects, a neck layer, namely, N2, was added, which further expands the last feature map by one stage (2×). Then, this feature map was added to the early feature map of the backbone to ensure the effectiveness of spatial and semantic information. A new detection layer, namely, D2, was also added to the N2-produced feature map, and an extremely small anchor box was created such that learning would specialize on small objects.

Consequently, small-object detection performance was enhanced compared with that achieved in previous studies and YOLOv5 model, while maintaining a suitable detection performance of medium-sized and large objects. This has great academic significance in that this has not been achieved in the object detection field. In addition, prior studies used synthetic target data and oversampling. Thus, when the models were implemented on actual systems, the performance degraded compared with the reported performance. Therefore, in this study, the model was reconstructed only through tuning and enhancing the architecture without data manipulation. Consequently, an improved performance was achieved, which is of great significance.

It is believed that SHOMY will contribute to the research dedicated to overcoming the limitations of one-stage object detectors. It is also expected that air transport and security industries, which require detection of various small objects in environments packed with various-sized objects, can benefit from the development of an embedded system or application of an experimental method utilizing the results of this study.

5.5 Conclusion and Future Research

This research was conducted to enhance the detection of small objects. Accordingly, through the development of the SHOMY model, which was based on YOLOv5, the detection performance was successfully improved. Not only did the proposed model provide more effective detection performance on small objects compared with existing methods but also it maintained a suitable detection performance for medium-sized and large objects. Thus, this study contributes to the research in the object detection field, as this model can detect objects of various sizes with one model. Furthermore, the proposed method also maintains a good performance with minimal data manipulation. Although this research provides a potential solution to the issue of poor detection performance on small objects, future research can further improve the model's detection accuracy by addressing other related challenges, such as processing speed, spatial invariance, and robustness to extreme environmental changes. It would also be beneficial to improve real-time performance by implementing existing accelerators and hardware. Additionally, research into improving the detection accuracy regarding small objects with deep learning methods, such as capsule networks and attention networks, should be investigated. By addressing these challenges and focusing on small objects contained in airport baggage captured using X-ray videos, ultimately, improvements can be observed in overall detection performance and efficient identification of objects.

The necessity to detect small hazardous items in baggage checks at transport hubs, such as airports, ports, and train stations, is directly linked to the safety of passengers and expected to become more critical in the future. Events of terrorism are increasing, methods involved are continuously becoming more complex, and types of items used are becoming more diverse, leading to more hazards. In addition to the small items analyzed in this study, many other small and ultrasmall items that may be used for terrorism must be rapidly identified for detection and researched continuously. In addition to the model enhancement methods proposed herein, continuous efforts to overcome the limitations and achieve improved results, such as research on bold layer composition, diversification of feature map information utilization, and improvement of detection efficiency, are extremely necessary.

Chapter 6

Conclusion

6.1 Summary of Each Study

6.1.1 Necessity and Justification for Each Study

1) Study 1

The utilization of artificial intelligence (AI) techniques in medical research has been steadily increasing, including in the field of respiratory studies. Noteworthy examples encompass investigations into early prediction of asthma, classification of chronic obstructive diseases, and the categorization of lung cancer cases. Lung Function-related recent investigations have underscored the significance of the relative regional air mass change value (RRAVC) in local lung respiration. While prior studies have successfully identified potential diseases through quantitative values derived from computed tomography (CT), the inherent characteristics of the lungs give rise to regional disparities even in healthy individuals. Thus, it became imperative to establish a baseline for assessing changes in air volume within specific areas.

In essence, this study employed AI techniques to develop a standard lung prediction model based on the RRAVC value, aiming to discern the normal state of the lungs in a COPD patient group, such as those with asthma. The research objective was to predict the RRAVC value, a direct indicator of chronic obstructive pulmonary disease (COPD), based on quantitative lung data extracted from CT images. By verifying the validity of latent variables, the study sought to develop a standardized model.

2) Study 2

Since 2016, there has been a growing emphasis on data-driven scientific administration, with notable advancements made in administrative experiments such as big data standard analysis models. These developments have ultimately led to the implementation of data-based administrative laws in 2020. In the context of local taxes, it is worth noting that the accumulated national local tax arrears amount to approximately 3.4 trillion won through the fiscal year of 2022. Unlike national taxes, which are governed by consistent policies established by the central government, local taxes are collected based on policies determined by local governments and serve as their primary financial resource. However, given the recent surge in local tax arrears, the importance and necessity of scientific administrative operations cannot be overstated.

Consequently, this study aimed to enhance the management of local taxes, which constitute the principal financial mechanism for local governments, by employing AI techniques to predict the early increase in delinquent taxpayers. To achieve this, the study developed a national standard model for identifying local tax delinquents by incorporating deep learning algorithms such as Convolutional Neural Networks (CNN). This departure from the machine learning algorithms utilized in previous studies aimed to verify the applicability of deep learning in the context of local tax administration.

3) Study 3

Following the events of September 11, there has been a concerted effort to enhance safety measures, including those pertaining to the security of passengers. In response, X-ray scans and physical searches became mandatory procedures at airports. While advancements in AI technology have facilitated some degree of automation in X-ray scanning, the ultimate decision-making still relies on human operators. The integration of AI technology aims to mitigate the potential for human error.

However, the investigation revealed that existing object detection systems frequently fall short in identifying small yet hazardous items. These systems predominantly prioritize the detection of larger objects that may pose a more immediate risk to passenger safety, such as fire equipment and weapons like firearms and knives. Consequently, the primary objective of this study was to address this concern by developing strategies to enhance the detection of small dangerous goods in light of this prevailing problem.

6.1.2 Summary of Each Study Results

The results of three separate studies demonstrate the superiority of certain models. These findings reflect and characterize the context and nuances of the three different subfields. In the first study, the experimental results of RRAVC predictions using ML showed that XGBoost had the best performance; however, MLP showed the best performance when using the relative coordinate information proposed in the study. After the parameters of the trained XGBoost model were examined, J, a local volume expansion ratio, was found to play the most important role in prediction. XGBoost and MLP models will be available as standard models for modeling regional lung function distribution. In addition, by using MLP to confirm the possibility of applying deep learning in the future, it was shown that future research developments are possible. The second study evaluated the performance of different models for local tax arrears prediction in terms of accuracy and found that the CNN framework was the most accurate, slightly better than the seq2seq with attention. The ResNet model had a relatively low performance, which was likely due to its lack of depth compared to other models. For reference, LSTM performed excellently when AUC was selected as the model evaluation metric in addition to accuracy. It is worth trying to propose suitable ensemble models based on deep learning in the future or improve well-known deep learning-based algorithms.

The third study compared the performance of the default YOLOv5 model and that of the proposed SHOMY model regarding image-based object detection. The results showed that the default YOLOv5 model significantly underperformed with smaller items, in terms of accuracy, precision, and recall. In contrast, the SHOMY model used an enhanced YOLOv5 structure with extra layers to drastically improve image-based object detection performance. This included addressing feature representation challenges to characterize objects by abstracting finer details and efficiently scaling conventional layers to maintain accuracy with larger datasets. Therefore, the SHOMY model is far superior to the default YOLOv5 model in terms of its accuracy and efficiency when it comes to the detection of small objects, including medium to large objects.

6.2 Comparison of AI Application Studies

6.2.1 Comparing and Structuring of AI Techniques

In the three studies, a wide range of ML and deep learning models, as well as a newly proposed model, was employed to accurately predict outcomes, reduce variability, and improve detection accuracy. Although all studies focus on improving AI performance, every study is focused on very different contexts to solve each and separate problem using intelligent information: the first with respiratory functional research, second with tax arrears prediction, and third with airport cargo safety. Therefore, although the AI techniques are employed here commonly, the outcomes, data, problems, and attributes studied in each case vary significantly. In addition, the first and second studies are characterized by a confirmative study approach that utilizes well-known algorithms, but the third research is characterized by an explorative study in developing and proposing new algorithms. Table 6.1 summarizes the comparative analysis of the three studies.

Table 6.1 Comparative analysis of study 1, 2, and 3.

Study	Commonality	Research Approach	Data	Problem	Method / Algorithm	Subset of AI Application
Study 1 – Chapter 3	Research to Solve Separate Research Field's Problem using Intelligent Information with Supervised Learning Approach	Confirmative Study: Utilize Existing Well-known Algorithms	Tabular Data: Extracted from CT Images	Prediction: Regression Problem – Prediction of RRAVC	Linear Ridge LASSO ElasticNet XGBoost LightGBM MLP	Application of Machine Learning Algorithms
Study 2 – Chapter 4			Tabular Data: Financial and Public Records	Prediction: Classification Problem – Normal Tax Payers / Arrears	LSTM Seq2Seq CNN ResNet Random Forest SVM	Application of Deep Learning Algorithms
Study 3 – Chapter 5		Explorative Study: Develop and Propose a New Algorithm	Image Data: X-ray Images	Detection: Objects Recognition	YOLOv5 SHOMY	Application of Deep Learning Algorithm and Propose a New Method

Table 6.1 was prepared by referring to the framework for AI techniques proposed by Dwivedi (2021) and Regona (2022) to confirm the characteristics of each study. To confirm the characteristics of each study, 1) data type, 2) problem to be solved, 3) methodology and algorithm, and 4) AI technology subset and characteristics should be confirmed. The first study was conducted using tabular data extracted from CT images, which can be classified as a regression problem predicting the calculated continuous value. In addition, the methodology used a total of eight ML algorithms, among which SVM was eliminated. Therefore, the final seven algorithms (linear regression, ridge regression, LASSO regression, ElasticNet, XGBoost, LightGBM, and MLP) were selected and corresponding results were reported.

The second study focused on solving the problem of discriminating between delinquents and normal payers through tabular data containing delinquent information, credit information, and public information. For this purpose, deep learning algorithms (i.e., LSTM, seq2seq, CNN, and ResNet) including some ML algorithms (i.e., random forest and SVM) were used, and their performances were confirmed based on the accuracy and AUC metrics. The third study focused on solving the problem of detecting hazardous objects using air transport X-ray image datasets and used and improved the deep learning-based well-known YOLOv5 to propose the SHOMY model. In comparison, all three studies are in the field of AI application research. However, their characteristics are divided into the use of ML algorithms, deep learning algorithms, improvement of deep learning algorithms, and related proposals. When the framework is conceptually derived based on the research development level for each field (i.e., radiology, tax arrears, and security detection) according to Chapter 2, along with the classification of AI subsets (Regona, 2022), the conceptual research framework for this dissertation can be proposed as shown in **Figure 6.1**.

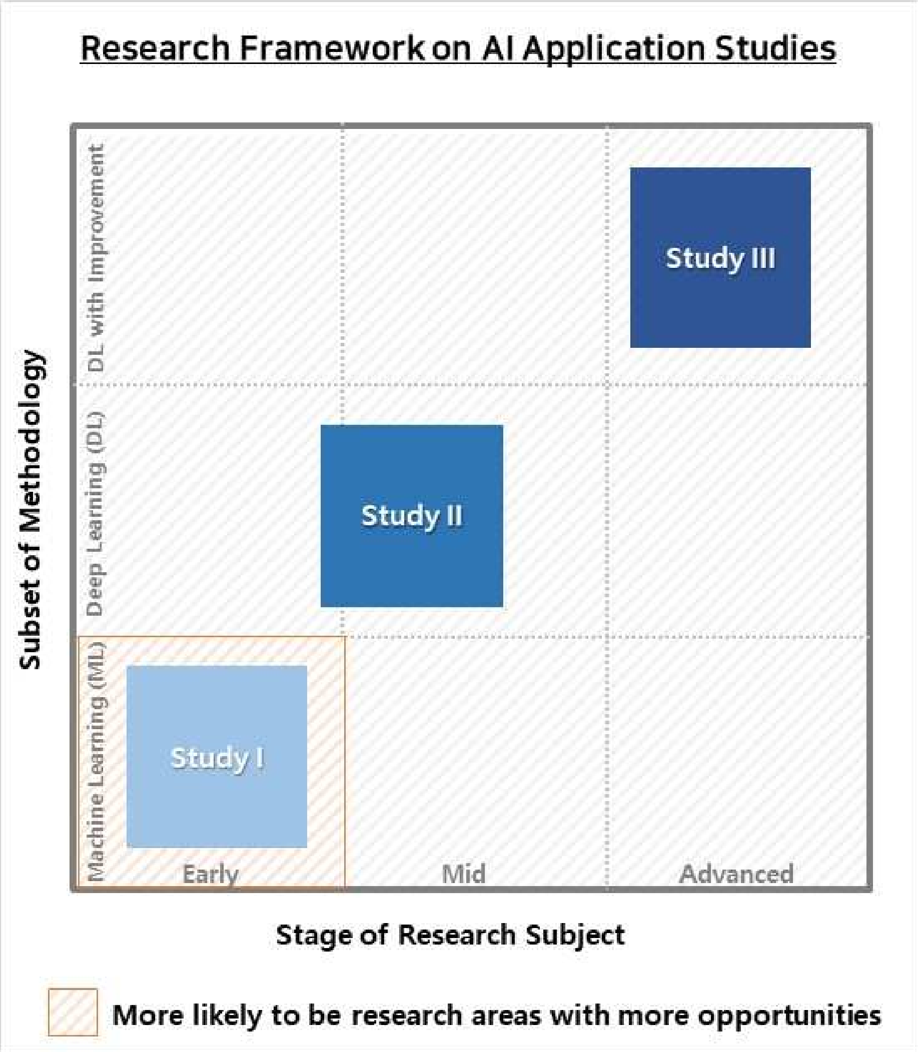

Figure 6.1 Research framework of three AI application studies.

The x-axis represents the level of research progress in each separate domain based on what was confirmed in Chapter 2 as the level of research stage. The y-axis denotes the subset of methodology differences in AI applications studies, which shows the types and categories of AI applications (i.e., ML, deep learning, and deep learning with improvement) of the three studies covered in this dissertation as a result of the comparison. Additionally, as discussed in Chapter 2, the opportunities for each study differ. In other words, study 1 is still at the level of introducing relatively basic and shallow ML algorithms; accordingly, there will be more research opportunities if the subject is in its early stage. However, even in such cases, having numerous opportunities does not necessarily imply that it is easy to convince the validity of applying AI techniques in the existing domain.

Relatively, study 2 is at the level of introducing deep learning algorithms, and the subject is assessed to be at the mid-term level. Comparatively, if the existing research already provides various ensemble techniques based on deep learning with improvement, such as study 3, and research topic is also highly advanced, the opportunity seems to be relatively low. Ultimately, this framework can be applied as a classification method for other studies in other fields as well.

6.3 Contributions and Implications

6.3.1 Each Study

In study 1, how ML models can be used to accurately predict the RRAVC in respiratory research was investigated. Experiments confirmed MLP and XGBoost as standard models among ML models. Additionally, the use of a subject's relative proportion coordinates was proposed to minimize intersubjective variability. This study contributes to the medical field by providing new insights into the role of ML in the prediction of RRAVCs. In addition, the development of the subject's relative proportion coordinates shows promise for decreasing intersubjective variability and can result in more accurate predictive models. This study also demonstrates how different coordinate systems can be applied to reduce intersubjective variability and suggests that relative proportion coordinates produce the best results. The implications of this study are important for both radiology and medical research because an efficient and precise method is presented to efficiently detect respiratory illness or function in a short period. Moreover, this study can lead to a better understanding of latent traits of diseases and provide new insights into diseases that have previously been undiagnosed or unaddressed because the value of the *J*, which is the ratio of local lung volume expansion among the lung-related key variables and is the most important variable in predicting the RRAVC, was confirmed. In addition, the standard model identified in this study can be used as a baseline model for further research.

In study 2, the problem of predicting individual local tax arrears was addressed by using popular ML and deep learning algorithms. Deep learning algorithms, such as LSTM, seq2seq, CNN, and ResNet, and shallow ML algorithms, such as random forest and SVM, were employed. Unlike previous studies on tax delinquents that remain at the level of ML algorithms, this study paves the way for using deep learning algorithms. In addition, existing financial sector prediction studies mainly used a single dataset type but this study developed a prediction model using convergence data obtained from heterogeneous datasets (i.e., financial information, credit rating information, public information, etc.). It was confirmed that CNN performed the best in terms of accuracy and concluded that LSTM attained the highest performance when AUC was used as model evaluation metric. By predicting local tax arrears better, the study contributed to the reduction of administrative resources. This research reveals that the CNN framework offers the highest accuracy, while the LSTM network also performs somewhat well. It is expected that the results of this study will be used as a basic model when conducting follow-up studies regarding the application of explainable models in the future. The implications of this research are wide-ranging and include better detection of fraudulent credit cases and efficient government policies to combat them. It could also help minimize the waste of administrative resources by reducing the amount of manpower needed.

In study 3, the aim was to develop a new model called SHOMY to better detect small objects in airport cargo. The proposed model was based on the YOLOv5 algorithm, whose performance was tested experimentally. The performance was improved in terms of the mean average precision (mAP) index and mAP (0.5:0.95) compared to those of the YOLOv5 model. The findings of the study show promise for better detection of small objects in dense areas, increasing the safety of air cargo with respect to potential threats. The implications of this research could lead to improved security at airports as the ability to accurately detect hazardous small objects can reduce the risk of terrorism. This research also presents an innovative approach for immediate industrial application without any performance degradation.

6.3.2 Entire Dissertation

This dissertation offers a comprehensive literature review that critically examines, compares, and distinguishes three AI application studies in the domains of interest, namely radiology, tax arrears, and security detection. The aim is to provide an academically rigorous exploration of the current state of research and development within these domains. Due to the expansive and multifaceted nature of AI applications research (i.e. broad academic gray area), this literature review endeavors to explore various areas within the AI domain that are currently under investigation. Moreover, it aims to emphasize disparities in the levels of research undertaken in these areas, thus contributing to a comprehensive understanding of the academic landscape. As a result, it is confirmed that this can provide new opportunities for AI application research. In particular, the research framework provided in this dissertation has significance in the managerial aspects of both information systems and engineering, and it would serve as a theoretical basis for establishing AI application research in the future.

In detail, it confirms that there are clearly more opportunities to apply advanced AI techniques in tabular data study than in unstructured data research, such as image and text data. This too is achieved by identifying research gaps. It also suggests that there is a clear need for increased AI graft research in radiology and tax delinquency. Moreover, object detection within the field of security is already highly advanced, suggesting that new solutions and research opportunities may diminish. These contents are all derived and explained as the final research framework presented by the dissertation, and clearly reconfirmed through three empirical studies conducted in practice.

6.4 Limitations and Future Research Directions

One significant constraint of this dissertation is its limited scope, as only three particular fields (i.e., radiology, taxation, and security detection) are validated and examined. Ideally, a more comprehensive literature review and meta-analysis should be conducted across multiple fields to determine more practical and sophisticated implications. In addition, towing to the literature review being limited to image and tabular datasets, additional research on more comprehensive unstructured data, such as text, audio, and video, are required. Through this, the research frame derived from the academic comparison and distinction process of AI application research provided by this dissertation would be able to serve as the proper theoretical basis.

Moreover, it should be acknowledged that this study may not comprehensively encompass all the rapidly evolving advancements in AI research, as ongoing research is imperative to identify potential research opportunities rather than solely proposing the introduction of novel technologies. Therefore, in conjunction with the sustainable development of AI techniques, continuous efforts should be made to recognize and address the limitations of prior studies through persistent follow-up, evaluation, and deliberation. Through these academicization processes of AI, research in the realm of AI applications would be able to establish a more specialized scholarly foundation. Furthermore, conducting studies to extract the managerial aspects of technologies, including their social and economic impact, implications for businesses and industries, and policy considerations related to the diverse techniques and methods employed in AI application research, would contribute significantly to the establishment of a more dependable and stable AI research environment while fostering an academic culture.

Bibliography

Abadie, A. and Gardeazabal, J. (2008). Terrorism and the world economy. *European economic review*, vol. 52, no. 1, pp. 1-27.

Abedin, M. Z., Hassan, M. K., Khan, I., and Julio, I. F. (2022). Feature Transformation for Corporate Tax Default Prediction: Application of Machine Learning Approaches. *Asia-Pacific Journal of Operational Research*, vol. 39, no. 4, 2140017.

Abedin, M. Z., Moon, M. H., Hassan, M. K., and Hajek, P. (2021). Deep learning-based exchange rate prediction during the COVID-19 pandemic. *Annals of Operations Research*, pp. 1-52.

Adarsh, P., Rathi, P., and Kumar, M. (2020, March). YOLO v3-Tiny: Object Detection and Recognition using one stage improved model. In *2020 6th international conference on advanced computing and communication systems (ICACCS),* IEEE, pp. 687-694.

Ahmad, T. and Chen, H. (2020). A review on machine learning forecasting growth trends and their real-time applications in different energy systems. *Sustainable Cities and Society*, vol. 54, 102010.

Akcay, S., Kundegorski, M. E., Willcocks, C. G., and Breckon, T. P. (2018). Using deep convolutional neural network architectures for object classification and detection within x-ray baggage security imagery. *IEEE transactions on information forensics and security*, vol. 13, no. 9, pp. 2203-2215.

Al-Dhabi, Y. and Zhang, S. (2021, August). Deepfake Video Detection by Combining Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN). In *2021 IEEE International Conference on Computer Science, Artificial Intelligence and Electronic Engineering (CSAIEE),* IEEE, pp. 236-241.

Ali, S. E. A., Rizvi, S. S. H., Fong-Woon, L., Rao, F. A., and Jan, A. A. (2021). Predicting delinquency on Mortgage loans: an exhaustive parametric comparison of machine learning techniques. *International Journal of Industrial Engineering and Management*, vol. 12, no. 1, pp. 1-13.

Amelon, R., Cao, K., Ding, K., Christensen, G. E., Reinhardt, J. M., and Raghavan, M. L. (2011). Three-dimensional characterization of regional lung deformation. *Journal of biomechanics*, vol. 44, no. 13, pp. 2489-2495.

Andriyanov, N. A., Volkov, A. K., Volkov, A. K., and Gladkikh, A. A. (2021, February). Research of recognition accuracy of dangerous and safe x-ray baggage images using neural network transfer learning. In *IOP Conference Series: Materials Science and Engineering,* IOP Publishing, vol. 1061, no. 1, pp. 012002.

Arik, S. Ö. and Pfister, T. (2021, May). Tabnet: Attentive interpretable tabular learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 8, pp. 6679-6687.

Bahdanau, D., Cho, K., and Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.

Bahmei, B., Birmingham, E., and Arzanpour, S. (2022). CNN-RNN and data augmentation using deep convolutional generative adversarial network for environmental sound classification. *IEEE Signal Processing Letters*, vol. 29, pp. 682-686.

Bai, Q., Li, S., Yang, J., Song, Q., Li, Z., and Zhang, X. (2020). Object detection recognition and robot grasping based on machine learning: A survey. *IEEE access*, vol. 8, pp. 181855-181879.

Baum, P. (2016). *Violence in the skies: a history of aircraft hijacking and bombing*. Summersdale Publishers LTD-ROW.

Bhatore, S., Mohan, L., and Reddy, Y. R. (2020). Machine learning techniques for credit risk evaluation: a systematic literature review. *Journal of Banking and Financial Technology*, vol. 4, pp. 111-138.

Bibi, R., Saeed, Y., Zeb, A., Ghazal, T. M., Rahman, T., Said, R. A., Abbas, S., Ahmad, M., and Khan, M. A. (2021). Edge AI-based automated detection and classification of road anomalies in VANET using deep learning. *Computational intelligence and neuroscience*, vol. 2021, pp. 1-16.

Bodduluri, S., Newell Jr, J. D., Hoffman, E. A., and Reinhardt, J. M. (2013). Registration-based lung mechanical analysis of chronic obstructive pulmonary disease (COPD) using a supervised machine learning framework. *Academic radiology*, vol. 20, no. 5, pp. 527-536.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., and Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, vol. 33, pp. 1877-1901.

Cai, Z., Xu, D., Zhang, Q., Zhang, J., Ngai, S. M., and Shao, J. (2015). Classification of lung cancer using ensemble-based feature selection and machine learning methods. *Molecular BioSystems*, vol. 11, no. 3, pp. 791-800.

Cannarile, A., Dentamaro, V., Galantucci, S., Iannacone, A., Impedovo, D., and Pirlo, G. (2022). Comparing deep learning and shallow learning techniques for api calls malware prediction: A study. *Applied Sciences*, vol. 12, no. 3, 1645.

Castaldi, P. J., Boueiz, A., Yun, J., Estepar, R. S. J., Ross, J. C., Washko, G., Cho, M. H., Hersh, C. P., Kinney, G. L., Young, K. A., Regan, E. A., Lynch, D. A., Criner, G. J., Dy, J. G., Rennard, S. I., Casaburi, R., Make, B. J., Crapo, J., Silverman, E. K., Hokanson, J. E., and COPDGene Investigators. (2020). Machine learning characterization of COPD subtypes: insights from the COPDGene study. *Chest*, vol. 157, no. 5, pp. 1147-1157.

Chae, K. J., Choi, J., Jin, G. Y., Hoffman, E. A., Laroia, A. T., Park, M., and Lee, C. H. (2020). Relative regional air volume change maps at the Acinar scale reflect variable ventilation in low lung attenuation of COPD patients. *Academic Radiology*, vol. 27, no. 11, pp. 1540-1548.

Chamberlin, J., Kocher, M. R., Waltz, J., Snoddy, M., Stringer, N. F., Stephenson, J., Sahbaee, P., Sharma, P., Rapaka, S., Schoepf, U. J., Abadia, A. F., Sperl, J., Hoelzer, P., Mercer, M., Somayaji, N., Aquino, G., and Burt, J. R. (2021). Automated detection of lung nodules and coronary artery calcium using artificial intelligence on low-dose CT scans for lung cancer screening: accuracy and prognostic value. *BMC medicine*, vol. 19, no. 1, pp. 1-14.

Chen, T. and Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 785-794.

Cheng, B., Misra, I., Schwing, A. G., Kirillov, A., and Girdhar, R. (2022). Masked-attention mask transformer for universal image segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,* pp. 1290-1299.

Choi, J., Hoffman, E. A., Lin, C. -L., Milhem, M. M., Tessier, J., and Newell Jr, J. D. (2017). Quantitative computed tomography determined regional lung mechanics in normal nonsmokers, normal smokers and metastatic sarcoma subjects. *PLoS One*, vol. 12, no. 7, e0179812.

Choi, S., Hoffman, E. A., Wenzel, S. E., Tawhai, M. H., Yin, Y., Castro, M., and Lin, C. -L. (2013). Registration-based assessment of regional lung function via volumetric CT images of normal subjects vs. severe asthmatics. *Journal of applied physiology*, vol. 115, no. 5, pp. 730-742.

Choudhury, A. and Asan, O. (2020). Role of artificial intelligence in patient safety outcomes: systematic literature review. *JMIR medical informatics*, vol. 8, no. 7, e18599.

Chui, K. T., Liu, R. W., Zhao, M., and De Pablos, P. O. (2020). Predicting students' performance with school and family tutoring using generative adversarial network-based deep support vector machine. *IEEE Access*, vol. 8, pp. 86745-86752.

Dabiri, H., Farhangi, V., Moradi, M. J., Zadehmohamad, M., and Karakouzian, M. (2022). Applications of Decision Tree and Random Forest as Tree-Based Machine Learning Techniques for Analyzing the Ultimate Strain of Spliced and Non-Spliced Reinforcement Bars. *Applied Sciences*, vol. 12, no. 10, 4851.

DeLancey, E. R., Simms, J. F., Mahdianpari, M., Brisco, B., Mahoney, C., and Kariyeva, J. (2019). Comparing deep learning and shallow learning for large-scale wetland classification in Alberta, Canada. *Remote Sensing*, vol. 12, no. 1, 2.

Dong, D., Tang, Z., Wang, S., Hui, H., Gong, L., Lu, Y., Xue, Z., Liao, H., Chen, F., Yang, F., Jin, R., Wang, K., Liu, Z., Wei, J., Mu, W., Zhang, H., Jiang, J., Tian, J., and Li, H. (2020). The role of imaging in the detection and management of COVID-19: a review. *IEEE reviews in biomedical engineering*, vol. 14, pp. 16-29.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duanf, Y., Dwivedig, R., Edwardsh, J., Eirugi, A., Galanosj, V., Ilavarasank, P. V., Janssenl, M., Jonesm, P., Kark, A. K., Kizginb, H., Kronemannm, B., Lalf, B., Lucinin, B., Medagliao, R., Meunier-FitzHughp, K. L., Meunier-FitzHughp, L. C. L., Misraq, S., Mogajir, E., Sharmas, S. K., Singhs, J. B., Raghavant, V., Ramanu, R., Ranab, N. P., Samothrakisv, S., Spencerw, J., Tamilmanib, K., Tubadjix, A., Waltony, P., and Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management,* vol. 57, 101994.

Eom, H. S. (2020). Softonnet Inc. Introduces Artificial Intelligence X-Ray Security Search Automatic Reading System at Incheon Airport. *Boan News*. Retrieved from: https://www.boannews.com/media/view.asp?idx=93264.

Fang, H., Zhang, D., Shu, Y., and Guo, G. (2020). Deep learning for sequential recommendation: Algorithms, influential factors, and evaluations. *ACM Transactions on Information Systems (TOIS)*, vol. 39, no. 1, pp. 1-42.

Finkelstein, J. and Jeong, I. C. (2017). Machine learning approaches to personalize early prediction of asthma exacerbations. *Annals of the New York Academy of Sciences*, vol. 1387, no. 1, pp. 153-165.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, vol. 29, no. 5, pp. 1189-1232.

Gillen, D. and Morrison, W. G. (2015). Aviation security: Costing, pricing, finance and performance. *Journal of Air Transport Management*, vol. 48, pp. 1-12.

Global Initiative for Chronic Obstructive Lung Disease. (2020). *Global Strategy for Prevention, Diagnosis and Management of Chronic Obstructive Pulmonary Disease*. Retrieved from: https://goldcopd.org/gold-reports/

Grinsztajn, L., Oyallon, E., and Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on tabular data?. *arXiv preprint arXiv:2207.08815*.

Gupta, A., Anpalagan, A., Guan, L., and Khwaja, A. S. (2021). Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues. *Array*, vol. 10, 100057.

Haghighi, B., Choi, S., Choi, J., Hoffman, E. A., Comellas, A. P., Newell Jr, J. D., Lee, C. H., Barr, R. G., Bleecker, E., Cooper, C. B., Couper, D., Han, M. L., Hansel, N. N., Kanner, R. E., Kazerooni, E. A., Kleerup, E. A. C., Martinez, F. J., O'Neal, W., Paine III, R., Rennard, S. I., Smith, B. M., Woodruff, P. G., and Lin, C. -L. (2019). Imaging-based clusters in former smokers of the COPD cohort associate with clinical characteristics: the SubPopulations and intermediate outcome measures in COPD study (SPIROMICS). *Respiratory research*, vol. 20, no. 1, pp. 1-14.

Ham, J., Chen, Y., Crawford, M. M., and Ghosh, J. (2005). Investigation of the random forest framework for classification of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 492-501.

Haq, I., Mazhar, T., Malik, M. A., Kamal, M. M., Ullah, I., Kim, T., Hamdi, M., and Hamam, H. (2022). Lung Nodules Localization and Report Analysis from Computerized Tomography (CT) Scan Using a Novel Machine Learning Approach. *Applied Sciences*, vol. 12, no. 24, 12614.

Hättenschwiler, N., Sterchi, Y., Mendes, M., and Schwaninger, A. (2018). Automation in airport security X-ray screening of cabin baggage: Examining benefits and possible implementations of automated explosives detection. *Applied ergonomics*, vol. 72, pp. 58-68.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition,* pp. 770-778.

He, Y., Luo, Z., Zhou, Y., Xue, R., Li, J., Hu, H., Yan, S., Chen, Z., Wang, J., and Lou, M. (2022). U-net models based on computed tomography perfusion predict tissue outcome in patients with different reperfusion patterns. *Translational Stroke Research*, vol. 13, no. 5, pp. 707-715.

Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., and Scholkopf, B. (1998). Support vector machines. *IEEE Intelligent Systems and their applications*, vol. 13, no. 4, pp. 18-28.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, vol. 9, no. 8, pp. 1735-1780.

Howard, R., Rattray, M., Prosperi, M., and Custovic, A. (2015). Distinguishing asthma phenotypes using machine learning approaches. *Current allergy and asthma reports*, vol. 15, pp. 1-10.

Höglund, H. (2017). Tax payment default prediction using genetic algorithm-based variable selection. *Expert Systems with Applications*, vol. 88, pp. 368-375.

Huang, Y. P. and Yen, M. F. (2019). A new perspective of performance comparison among machine learning algorithms for financial distress prediction. *Applied Soft Computing*, vol. 83, 105663.

Kadir, T. and Gleeson, F. (2018). Lung cancer prediction using machine learning and advanced imaging techniques. *Translational lung cancer research*, vol. 7, no. 3, pp. 304-312.

Kadra, A., Lindauer, M., Hutter, F., and Grabocka, J. (2021). Well-tuned simple nets excel on tabular datasets. *Advances in neural information processing systems*, vol. 34, pp. 23928-23941.

Karthick, A. and Gopalsamy, S. (2022, November). Artificial Intelligence: Trends and Challenges. In *2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC)*, IEEE pp. 7-12.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T. Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, vol. 30, pp. 3149-3157.

Kim, S. M. (2019). 8 Types of Aviation Security Equipment. *Boan News*. Retrieved from: https://www.boannews.com/media/view.asp?idx=80922.

Kisantal, M., Wojna, Z., Murawski, J., Naruniec, J., and Cho, K. (2019). Augmentation for small object detection. arXiv preprint arXiv:1902.07296.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, vol. 60, no. 6, pp. 84-90.

Kwekha-Rashid, A. S., Abduljabbar, H. N., and Alhayani, B. (2023). Coronavirus disease (COVID-19) cases analysis using machine-learning applications. *Applied Nanoscience*, vol. 13, pp. 2013–2025.

Lai, L. (2020, August). Loan default prediction with machine learning techniques. In *2020 International Conference on Computer Communication and Network Security (CCNS),* IEEE, pp. 5-9.

Laucho-Contreras, M. E. and Cohen-Todd, M. (2020). Early diagnosis of COPD: myth or a true perspective. *European Respiratory Review*, vol. 29, no. 158.

Lee, J. N. and Cho, H. C. (2020). Development of Artificial Intelligence System for Dangerous Object Recognition in X-ray Baggage Images. *Journal of Transactions of the Korean Institute of Electrical Engineers*, vol. 69, no. 7, pp. 1067-1072.

Lessmann, S., Baesens, B., Seow, H. V., and Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, vol. 247, no. 1, pp. 124-136.

Li, F., Choi, J., Zou, C., Newell Jr, J. D., Comellas, A. P., Lee, C. H., Ko, H., Barr, R. G., Bleecker, E. R., Cooper, C. B., Abtin, F., Barjaktarevic, I., Couper, D., Han, M., Hansel, N. N., Kanner, R. E., Paine III, R., Kazerooni, E. A., Martinez, F. J., O'Neal, W., Rennard, S. I., Smith, B. M., Woodruff, P. G., Hoffman E. A., and Lin, C. -L. (2021). Latent traits of lung tissue patterns in former smokers derived by dual channel deep learning in computed tomography images. *Scientific reports*, vol. 11, no. 1, 4916.

Li, Y. (2019, August). Credit risk prediction based on machine learning methods. In *2019 14th International Conference on Computer Science and Education (ICCSE),* IEEE, pp. 1011-1013.

Liu, C., Hu, S. C., Wang, C., Lafata, K., and Yin, F. F. (2020). Automatic detection of pulmonary nodules on CT images with YOLOv3: development and evaluation using simulated and patient data. *Quantitative Imaging in Medicine and Surgery*, vol. 10, no. 10, 1917.

Liu, M., Wang, X., Zhou, A., Fu, X., Ma, Y., and Piao, C. (2020). Uav-yolo: Small object detection on unmanned aerial vehicle perspective. *Sensors*, vol. 20, no. 8, 2238.

Liu, S., Qi, L., Qin, H., Shi, J., and Jia, J. (2018). Path aggregation network for instance segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 8759-8768.

Loey, M., Manogaran, G., Taha, M. H. N., and Khalifa, N. E. M. (2021). A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. *Measurement*, vol. 167, 108288.

Lu, Y. (2019). Artificial intelligence: a survey on evolution, models, applications and future trends. *Journal of Management Analytics*, vol. 6, no. 1, pp. 1-29.

Masita, K. L., Hasan, A. N., and Shongwe, T. (2020, August). Deep learning in object detection: A review. In *2020 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD),* IEEE, pp. 1-11.

Meddage, D. P. P., Ekanayake, I. U., Herath, S., Gobirahavan, R., Muttil, N., and Rathnayake, U. (2022). Predicting Bulk Average Velocity with Rigid Vegetation in Open Channels Using Tree-Based Machine Learning: A Novel Approach Using Explainable Artificial Intelligence. *Sensors*, vol. 22, no. 12, 4398.

Mei, S., Jiang, H., and Ma, L. (2021, October). YOLO-lung: a practical detector based on imporved YOLOv4 for pulmonary nodule detection. In *2021 14th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI),* IEEE, pp. 1-6.

Morera, Á., Sánchez, Á., Moreno, A. B., Sappa, Á. D., and Vélez, J. F. (2020). SSD vs. YOLO for detection of outdoor urban advertising panels under multiple variabilities. Sensors, vol. 20, no. 16, 4587.

Mun, S. K., Wong, K. H., Lo, S. C. B., Li, Y., and Bayarsaikhan, S. (2021). Artificial intelligence for the future radiology diagnostic service. *Frontiers in molecular biosciences*, vol. 7, 614258.

Munoko, I., Brown-Liburd, H. L., and Vasarhelyi, M. (2020). The ethical implications of using artificial intelligence in auditing. *Journal of Business Ethics*, vol. 167, pp. 209-234.

National Consortium for the Study of Terrorism and Responses to Terrorism (START). (2021). *The Global Terrorism Database*. University of Maryland. Retrieved from: https://start.umd.edu/gtd

Noble, W. S. (2006). What is a support vector machine?. *Nature biotechnology*, vol. 24, no. 12, pp. 1565-1567.

Novakoff, A. K. (1993). FAA bulk technology overview for explosives detection. *Applications of signal and image processing in explosives detection systems*, vol. 1824, pp. 2-12.

OECD. (2018, November 22). *Main causes of mortality among women and men in EU countries.* Retrieved from: https://doi.org/10.1787/health_glance_eur-2018-graph26-en

Ozge Unel, F., Ozkalayci, B. O., and Cigla, C. (2019). The power of tiling for small object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*.

Papanastasopoulos, Z., Samala, R. K., Chan, H. P., Hadjiiski, L., Paramagul, C., Helvie, M. A., and Neal, C. H. (2020, March). Explainable AI for medical imaging: deep-learning CNN ensemble for classification of estrogen receptor status from breast MRI. In *Medical imaging 2020: Computer-aided diagnosis,* SPIE, vol. 11314, pp. 228-235.

Paschen, U., Pitt, C., and Kietzmann, J. (2020). Artificial intelligence: Building blocks and an innovation typology. *Business Horizons*, vol. 63, no. 2, pp. 147-155.

Patel, D., Hall, G. L., Broadhurst, D., Smith, A., Schultz, A., and Foong, R. E. (2022). Does machine learning have a role in the prediction of asthma in children?. *Paediatric Respiratory Reviews*, vol. 41, pp. 51-60.

Paul, M., Maglaras, L., Ferrag, M. A., and AlMomani, I. (2023). Digitization of healthcare sector: A study on privacy and security concerns. *ICT Express*.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., and Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems,* vol. 31.

Rabbi, J., Ray, N., Schubert, M., Chowdhury, S., and Chao, D. (2020). Small-object detection in remote sensing images with end-to-end edge-enhanced GAN and object detector network. *Remote Sensing*, vol. 12, no. 9, 1432.

Raschka, S., Patterson, J., and Nolet, C. (2020). Machine learning in python: Main developments and technology trends in data science, machine learning, and artificial intelligence. *Information*, vol. 11, no. 4, 193.

Regona, M., Yigitcanlar, T., Xia, B., and Li, R. Y. M. (2022). Opportunities and adoption challenges of AI in the construction industry: a PRISMA review. *Journal of open innovation: technology, market, and complexity*, vol. 8, no. 1, 45.

Rueda, D. F., Caviedes, J. C., and Muñoz, W. Y. C. (2022). A Hybrid Intrusion Detection Approach Based on Deep Learning Techniques. In *Computer Networks, Big Data and IoT: Proceedings of ICCBI 2021,* Springer Nature, pp. 863-878.

Saleh, A., Sheaves, M., and Rahimi Azghadi, M. (2022). Computer vision and deep learning for fish classification in underwater habitats: A survey. *Fish and Fisheries*, vol. 23, no. 4, pp. 977-999.

Sanil, N., Rakesh, V., Mallapur, R., and Ahmed, M. R. (2020, January). Deep learning techniques for obstacle detection and avoidance in driverless cars. In *2020 International Conference on Artificial Intelligence and Signal Processing (AISP),* IEEE, pp. 1-4.

Santos, D., Cardoso-Fernandes, J., Lima, A., Müller, A., Brönner, M., and Teodoro, A. C. (2022). Spectral analysis to improve inputs to random forest and other boosted ensemble tree-based algorithms for detecting NYF pegmatites in Tysfjord, Norway. *Remote Sensing*, vol. 14, no. 15, 3532.

Shafaat, O., Bernstock, J. D., Shafaat, A., Yedavalli, V. S., Elsayed, G., Gupta, S., Sotoudeh, E., Sair, H. I., Yousem, D. M., and Sotoudeh, H. (2021). Leveraging artificial intelligence in ischemic stroke imaging. *Journal of Neuroradiology*. vol. 49, no. 4, pp. 343-351.

Shenfield, A., Day, D., and Ayesh, A. (2018). Intelligent intrusion detection systems using artificial neural networks. *ICT Express*, vol. 4, no. 2, pp. 95-99.

Shin, K. M., Choi, J., Chae, K. J., Jin, G. Y., Eskandari, A., Hoffman, E. A., Hall, C., Castro, M., and Lee, C. H. (2020). Quantitative CT-based image registration metrics provide different ventilation and lung motion patterns in prone and supine positions in healthy subjects. *Respiratory Research*, vol. 21, no. 1, pp. 1-9.

Shwartz-Ziv, R. and Armon, A. (2022). Tabular data: Deep learning is not all you need. Information Fusion, vol. 81, pp. 84-90.

Singh, S. and Singh, M. (2003). Explosives detection systems (EDS) for aviation security. *Signal processing*, vol. 83, no. 1, pp. 31-55.

Song, S., Liu, J., Liu, Y., Feng, G., Han, H., Yao, Y., and Du, M. (2020). Intelligent object recognition of urban water bodies based on deep learning for multi-source and multi-temporal high spatial resolution remote sensing imagery. *Sensors*, vol. 20, no. 2, 397.

Stahl, B. C., Antoniou, J., Ryan, M., Macnish, K., and Jiya, T. (2022). Organisational responses to the ethical issues of artificial intelligence. *AI and SOCIETY*, vol. 37, no. 1, pp. 23-37.

Sterchi, Y. and Schwaninger, A. (2015, September). A first simulation on optimizing EDS for cabin baggage screening regarding throughput. In *2015 International Carnahan conference on security technology (ICCST),* IEEE, pp. 55-60.

Stiglic, G., Kocbek, P., Fijacko, N., Zitnik, M., Verbert, K., and Cilar, L. (2020). Interpretability of machine learning-based prediction models in healthcare. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 10, no. 5, e1379.

Sutskever, I., Vinyals, O., and Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in neural information processing systems*, vol. 27.

Tong, K., Wu, Y., and Zhou, F. (2020). Recent advances in small object detection based on deep learning: A review. *Image and Vision Computing*, vol. 97, 103910.

Voter, A. F., Meram, E., Garrett, J. W., and John-Paul, J. Y. (2021). Diagnostic accuracy and failure mode analysis of a deep learning algorithm for the detection of intracranial hemorrhage. *Journal of the American College of Radiology*, vol. 18, no. 8, pp. 1143-1152.

Vuola, A. O., Akram, S. U., and Kannala, J. (2019, April). Mask-RCNN and U-net ensembled for nuclei segmentation. In *2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019),* IEEE, pp. 208-212.

Wang, C. Y., Liao, H. Y. M., Wu, Y. H., Chen, P. Y., Hsieh, J. W., and Yeh, I. H. (2020). CSPNet: A new backbone that can enhance learning capability of CNN. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops,* pp. 390-391.

Wei, W., Ke, Q., Nowak, J., Korytkowski, M., Scherer, R., and Woźniak, M. (2020). Accurate and fast URL phishing detector: a convolutional neural network approach. *Computer Networks*, vol. 178, 107275.

Wei, Y. and Liu, X. (2020). Dangerous goods detection based on transfer learning in X-ray images. *Neural Computing and Applications*, vol. 32, pp. 8711-8724.

World Health Organization. (2020, December 9). *The top 10 causes of death*. Retrieved from: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

Zhang, C. and Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. *Journal of Industrial Information Integration*, vol. 23, 100224.

Zhang, F. and O'Donnell, L. J. (2020). Support vector regression. In *Machine learning* (pp. 123-140). Academic Press.

Zhang, W., Li, H., Li, Y., Liu, H., Chen, Y., and Ding, X. (2021). Application of deep learning algorithms in geotechnical engineering: a short critical review. *Artificial Intelligence Review*, pp. 1-41.

Zhang, Y., Qin, J., Park, D. S., Han, W., Chiu, C. C., Pang, R., Le., Q. V., and Wu, Y. (2020). Pushing the limits of semi-supervised learning for automatic speech recognition. arXiv preprint arXiv:2010.10504.

Zheng, Z., Wang, P., Ren, D., Liu, W., Ye, R., Hu, Q., and Zuo, W. (2021). Enhancing geometric factors in model learning and inference for object detection and instance segmentation. *IEEE Transactions on Cybernetics*, vol. 52, no. 8, pp. 8574-8586.

Zhou, L., Pan, S., Wang, J., and Vasilakos, A. V. (2017). Machine learning on big data: Opportunities and challenges. *Neurocomputing*, vol. 237, pp. 350-361.

Zou, C., Li, F., Choi, J., Haghighi, B., Choi, S., Rajaraman, P. K., Comellas, A. P., Newell Jr, J. D., Lee, C. H., Barr, R. G., Bleecker, E., Cooper, C. B., Couper, D., Han, M., Hansel, N. N., Kanner, R. E., Kazerooni, E. A., Kleerup, E. C., Martinez, F. J., O'Neal, W., Paine III, R., Rennard, S. I., Smith, B. M., Woodruff, P. G., Hoffman, E. A., and Lin, C. -L. (2021). Longitudinal imaging-based clusters in former smokers of the copd cohort associate with clinical characteristics: The subpopulations and intermediate outcome measures in copd study (spiromics). *International journal of chronic obstructive pulmonary disease*, vol. 16, pp. 1477-1496.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, vol. 67, no. 2, pp. 301-320.

국문초록

최근 인공지능(AI)의 응용과 접목은 다양한 산업 및 연구 분야에서 지속적으로 확대 전개되고 있습니다. 여러 산업 및 연구 분야중에서 특히 농업, 제조업, 교육, 금융, 행정, 의료, 보안 등은 다수의 학자들에 의해 AI 기술을 적극적으로 도입하고 있는 분야임이 확인되었습니다. 본 학위 논문은 이러한 추세를 확인하고 상기 분야 중에서도 의료, 금융, 보안 산업의 각 하위 분야인 영상 의학, 세금 징수, 보안 탐지 분야에서 AI 응용 연구를 제시합니다. 본 학위 논문에서 제시하는 세 가지 인공지능 응용 연구는 각각 정형 데이터와 이미지 데이터를 기반으로 수행되었으며, 정형 데이터 기반 연구로는 1) 만성 폐쇄성 폐 질환(COPD) 관련 폐 기능 예측 연구와 2) 지방세 체납자 예측 연구를 소개하며, 이미지 데이터 기반 연구로는 3) 공항 수하물 엑스레이 영상에 대한 물체 검출 연구를 제시합니다. 본 학위 논문의 목적은 이 세 가지 연구 각각의 특정 분야뿐만 아니라 정형 데이터와 비정형 이미지 데이터 연구에서 현재 수준의 인공지능 응용 연구에 대한 문헌 검토를 수행하여, AI 응용 연구에서 연구 격차와 기회를 식별하는 것입니다.

궁극적으로, 본 학위 논문은 AI 응용 연구를 위한 방법론에 대한 통찰을 제공하고, 넓은 범위의 문헌 검토 수행과 제시된 세 가지 연구 결과를 통해 식별된 AI 응용 연구를 비교 분석합니다. 그 결과 각 분야에 걸쳐 수행된 AI 응용 연구의 불일치 정도 및

수준 차이를 확인 및 실증하였고, 신규 연구 주제일수록 더 큰 연구 기회가 존재하는 것을 식별하였습니다. 또한, AI 응용 기법 수준을 각각 머신러닝, 딥러닝, 딥러닝의 개선으로 식별함에 따라 머신러닝을 도입해보는 연구 초기 단계일수록 더 많은 연구 기회가 존재하는 것을 확인하였습니다. 이러한 학문적 구분 과정을 통해 도출 및 제시하는 연구 프레임워크는 향후 인공지능 응용 연구를 학문적으로 확립하기 위한 이론적 근거로 활용될 수 있을 것입니다. AI 기술이 다양한 분야에서 점점 더 많이 사용되고 있고 앞으로 더 중요한 역할을 할 것으로 예상됨에 따라 이러한 연구 결과는 AI 응용 분야의 연구자 및 업계 전문가들에게 중요한 사실을 제공합니다. 나아가 본 학위 논문은 이러한 연구 결과를 바탕으로 연구의 한계점을 확인함으로써 AI 응용 분야의 추후 연구 방향을 제시하고, 다양한 분야에서 활용할 수 있는 AI 기술 개발 및 학문적 체계화를 위한 지속적인 노력이 필요함을 강조합니다.

주요어: 인공지능 응용 연구, 정보 시스템 공학 및 경영, 정형 및 이미지 데이터, 영상 의학, 세금 징수, 보안 탐지

학번: 2019-31312

Acknowledgements

As the momentous culmination of my doctoral journey draws near, I find myself overwhelmed with a profound sense of gratitude, an emotion that words alone struggle to capture. It is with great humility and heartfelt appreciation that I wish to acknowledge the unwavering support and guidance bestowed upon me by those who have played an instrumental role in shaping my path.

Foremost among these guiding lights is Professor Byungjoon Yoo, to whom my indebtedness knows no bounds. With astute discernment, he illuminated the depths of my weaknesses, propelling me towards my own unique expression of excellence. The benevolence of his time and counsel has proven invaluable, but it is his profound wisdom that shall forever be etched upon the fabric of my soul. His resolute belief that perseverance and endurance pave the way to triumph resonates deep within me, igniting a fervent passion for knowledge that has become the cornerstone of my being.

I am also compelled to express my sincerest appreciation to the members of my dissertation committee, whose profound insights and perspectives have indelibly shaped my growth as a researcher. Each one has been a vital catalyst in my intellectual development, propelling me forward on this arduous quest for knowledge. To Professor Gahgene Gweon, I owe an immeasurable debt of gratitude for her steady support and guidance from the very inception of my doctoral odyssey. Profound appreciation is also extended to Professor Yeunwoong Kyung, whose dedicated encouragement nurtured the seeds of my passion, enabling them to blossom into a flourishing dedication to the pursuit of research. Professor Jiwoong Choi imparted upon me the foundational principles of application research, broadening the horizons of my knowledge, while Professor Jong Woo Choi lent his expertise, assisting me in organizing my research endeavors with greater efficiency.

Furthermore, I would be remiss not to extend my heartfelt gratitude to the esteemed individuals who have selflessly aided me along this arduous journey. To Professor Wookjae Heo of Purdue University, Professor Chang-Hyun Lee of Seoul National University Hospital, Professor Kum Ju Chae of Chonbuk National University Hospital, Professor Yun Seog Lee of Seoul National University, and Senior Manager Minjae Kim of the National Information Society Agency, I offer my deepest appreciation. Their invaluable advice and sagacious guidance proved to be beacons of light, particularly during moments of adversity that inevitably beset the course of research.

Lastly, I reserve a special place in my heart for the researchers who have been unwavering pillars of support throughout this expedition. Dr. Hyunji Lee of the National Institute of Green Technology, YongHyun Lee and Jungwook Rhim of Seoul National University, and Professor Jaewhan Lee of Hallym University your counsel and encouragement have been sources of solace and fortitude both in my academic pursuits and personal life. Moreover, I extend my profound gratitude to Director Dr. Hyochan Lee and Senior Researcher Kwangtek Na of Hanwha Group, Director Hyukjae Lee of Korea Credit Bureau, Director Dr. Jong Mun Yoon of Credit Finance Association of Korea, Professor Seyoung Park of University of Nottingham, and Commissioner Dr. Jae Eun You of Office for Government Policy Coordination for their steadfast consideration and support.

Of course, I cannot embark upon this voyage of gratitude without acknowledging the indomitable love and support of my cherished family, including my dear parents and sister, beloved wife, and the precious son soon to be born. And, God, your presence in my life is an immeasurable blessing, and I humbly express my gratitude to the divine forces that have bestowed upon me these countless blessings, which have made this moment possible.

With profound sincerity and a heart brimming with appreciation, I offer my deepest thanks to all those who have touched my life and have been instrumental in my journey. Your unwavering belief in me has kindled a flame that shall forever illuminate my path. Thank you, from the depths of my soul.

Summer 2023,
Eunchan Kim