Lecture Notes in Networks and Systems 1970

Abhishek Bhattacharya
Soumi Dutta
Intan Ermahani A. Jalil
Alvaro Rocha *Editors*

Data Mining and Information Security

Proceedings of ICDMIS 2025, Volume 2

Springer

Lecture Notes in Networks and Systems 1970

The series "Lecture Notes in Networks and Systems" publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the worldwide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, EI Compendex, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

Abhishek Bhattacharya · Soumi Dutta ·
Intan Ermahani A. Jalil · Alvaro Rocha
Editors

Data Mining and Information Security

Proceedings of ICDMIS 2025, Volume 2

Springer

Editors
Abhishek Bhattacharya
Department of Computer Science
and Engineering
Sister Nivedita University
Kolkata, West Bengal, India

Soumi Dutta
Department of Computer Science
and Engineering
Sister Nivedita University
Kolkata, West Bengal, India

Intan Ermahani A. Jalil
Universiti Teknikal Malaysia Melaka (UTeM)
Malacca, Malaysia

Alvaro Rocha
University of Lisbon
Lisbon, Portugal

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-3-032-25951-6 ISBN 978-3-032-25952-3 (eBook)
https://doi.org/10.1007/978-3-032-25952-3

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

Welcome to the 2nd International Conference on Data Mining and Information Security (ICDMIS 2025), held at the University of Puthisastra, Cambodia, and organized by Eminent College of Management and Technology (ECMT), India, and University of Puthisastra, Cambodia, in collaboration with the International Knowledge Research Foundation, during October 7–8, 2025.

As a premier conference in the field, ICDMIS 2025 provides a highly competitive and rigorous forum for the presentation of recent advances in research and applications related to data mining and information security. We are pleased to present the proceedings of the conference as its official scholarly record. The theme of this year, Crossroads of Data Mining, Big Data, and Information Security, reflects areas of growing significance that are rapidly attracting both academic and industrial attention, as data mining increasingly depends on effective information retrieval techniques for collecting and analyzing data from diverse sources.

ICDMIS 2025 is a relatively young conference in the areas of data mining, data science, information retrieval, and information security; however, it has already demonstrated significant growth. As evidence of this progress, ICDMIS 2025 received a record total of 873 paper submissions. The authors of the submitted papers represent 17 countries, while the authors of the accepted papers come from 15 countries.

We hope that this program will further stimulate high-quality research in information security and data mining, while providing practitioners and researchers with advanced techniques, algorithms, and practical tools for real-world deployment. It is our privilege and honor to present these proceedings, which capture the most recent advances and emerging trends in the field of data mining, and we trust that they will serve as a valuable reference for the academic and industrial research communities.

Surajit Goon
Ian Rouse
Conveners

Preface

We are pleased to announce that the Eminent College of Management and Technology, India, in collaboration with the University of Puthisastra, Cambodia, and the International Knowledge Research Foundation, hosted the 2nd International Conference on Data Mining and Information Security (ICDMIS 2025) in hybrid mode during October 7–8, 2025.

The conference attracted a diverse range of engineering practitioners, academicians, scholars and industry delegates, with the reception of abstracts including more than 2500 authors from different parts of the world. The committee of professionals dedicated towards the conference is striving to achieve a high-quality technical program with tracks on data mining, Advance Computing, Information Retrieval, Data Science and Data Analytics and information security. More than 870 full-length papers from 17 different countries have been received, among which the contributions are focused on theoretical, computer simulation-based research, and laboratory-scale experiments. Amongst these manuscripts, only 205 papers (Acceptance Ratio 23.48%) have been included in the Springer proceedings after a thorough two-stage review and editing process. All the manuscripts submitted to the ICDMIS 2025 were peer-reviewed by at least two independent reviewers. The comments from the reviewers were communicated to the authors, who incorporated the suggestions in their revised manuscripts. It is an exalting experience to finally see these elite contributions materialize into five book volumes of proceedings published by Springer. ICDMIS 2025 invited six keynote speakers, who are eminent researchers in the field of computer science and engineering, from different parts of the world. In addition to the keynote sessions on each day of the conference, 12 concurrent technical sessions are held every day to assure the oral presentation of 205 accepted papers. Prominent researchers in the session's subject area have served as keynote speakers and session chair(s) for each of the concurrent sessions. An international conference of such magnitude and release of the ICDMIS 2025 proceedings by Springer has been the remarkable outcome of the untiring efforts of the entire organizing team. Fortunately, ICDMIS 2025 has benefited from contributions and support from all sides since the start of its journey. We extend our gratitude to everyone who sent their best wishes and made any kind of contribution to the success of ICDMIS 2025. The tenacity of the steering, advising, and technical program committee members made the edited proceeding volumes published by Springer possible. The ICDMIS 2025 organizers would like to thank all of the participating authors for their outstanding works and attention. We would like to express our sincere gratitude to the editors, committee members, and production workers whose hard work helped to define the ICDMIS 2025 proceedings. Finally, we express our gratitude to Springer for agreeing to publish the

conference papers of the ICDMIS 2025. The acquisition senior editor, Mr. Aninda Bose, has been a great help throughout the process.

Abhishek Bhattacharya
Alvaro Rocha
Intan Ermahani A. Jalil
Soumi Dutta

About the Book

This book features selected research papers presented at the 2nd International Conference on Data Mining and Information Security (ICDMIS 2025) held at the University of Puthisastra, Cambodia, and organized by Eminent College of Management and Technology (ECMT), India, and University of Puthisastra, Cambodia, in collaboration with the International Knowledge Research Foundation, during October 7–8, 2025.

Data Mining and Information Security are closely related fields that deal with the collection, analysis, and utilization of data and information from various digital sources. These fields are crucial in today's digital age, where vast amounts of data are generated and stored electronically. By combining the strengths of Information Security and information retrieval, organizations can enhance their ability to detect, respond to, and mitigate cyber threats, while also improving their overall decision-making processes.

This book comprises high-quality research work by academicians and industrial experts in the field of Artificial intelligence, data analytics, computing and communication, including full-length papers related to all the areas of Data Mining, Machine Learning, Internet of Things (IoT) and information security etc.

Convolutional Neural Networks for Foliar Disease Diagnosis in Horticulture Crops: A Critical Review

Anshu Khare(✉), Naimur Rahman Kidwai, and Shrish Bajpai

ECE Department, Faculty of Engineering and Information Technology Integral University Lucknow, Lucknow, Uttar Pradesh, India
anshusrivastava09@gmail.com

Abstract. The research aims to examine further how to apply advanced machine learning models and deep learning models, especially convolutional neural network models (CNN), in detecting disease in both tomato and potato plants in advance. It emphasizes how various factors, such as disease-producing microorganisms and environmental components in addition to less varying datasets, especially when using datasets such as PlantVillage, which have shown vast usage in various models in various research worldwide, can particularly influence the accuracy of disease detection models in crops such as potatoes and tomatoes using CNN models talked of in this study. CNN models, especially AlexNet models, ResNet models, and EfficientNet models, which show high accuracy in classifying different kinds of disease in various kinds of crops using 94–99.75% accuracy levels, are discussed in this study as a basis of further research to come up with more efficient models in promoting both machine learning and various uses of CNN models in disease management in promoting food security in many countries worldwide.

Keywords: EfficientNet · CNN · ResNet · deep learning models · machine learning

1 Introduction

Humans depend a lot on agricultural products for their food. But the health of plants can be affected by a range of factors, such as harmful microorganisms like bacteria, fungi, and viruses; changes in soil pH; extreme weather conditions; and variations in moisture and humidity in the air. Human beings rely greatly on agricultural produce for their nourishment. However, plants get their health from diverse factors, one of which being the microbes harmful to them like bacteria, fungi, and viruses. Besides these, plants get damaged big time if the soil pH changes, the weather becomes extreme or the moisture and humidity of the air fluctuate. All these things negatively affect crop development, their structure and functions, consequently, low agricultural output does not only get affected but also farmers who depend on these crops. It is quite a difficult task to discover the presence of a disease in a plant at an early stage of its growth just

A. Bhattacharya et al. (Eds.): ICDMIS 2025, LNNS 1970, pp. 16–24, 2026.
https://doi.org/10.1007/978-3-032-25952-3_2

through manual observation which can lead to low output. Scientists are addressing this issue with the help of deep learning (DL), image processing, and machine learning (ML) techniques that can identify diseases at a very high accuracy level by analysing pictures of plant leaves.

The main aim of image processing [12] is to get the image to be of higher quality so that the image gives more and more information in a clearer manner. This technique can be and is extensively used in various fields including remote sensing, colour correction, pattern recognition, etc. In fact, it is most extensively used in medical and agricultural diagnostics. Image processing can be summarized as a series of consecutive steps that essentially include capturing images, preprocessing them, segmenting, extracting features, and classifying the results. This research work is obviously focused on finding out what diseases are commonly found in tomato and potato plants so that by diagnosing at the earliest and taking immediate action, the loss of yield can be kept at a minimum.

The research specifically tackles the following questions:

1. How can CNN architectures be fine-tuned for detecting diseases in tomato and potato leaves?
2. What performance differences exist among various CNN architectures when it comes to classifying plant diseases?
3. How can we overcome the limitations of current datasets to enhance model generalization in real-world conditions?

The key contributions of this study include:

1. A comprehensive review and performance comparison of state-of-the-art CNN architectures for tomato and potato disease detection.
2. Identification of methodological challenges in existing approaches.
3. Recommendations for future research directions that address current techno-logical and dataset limitations.

2 Related Work

An in-depth literature review of research papers from 2019 to 2024 has been carried out, with a major focus on convolutional neural networks (CNNs), deep learning, and machine learning algorithms for the detection of plant diseases. The first step in scanning the research works was to select the ones based on CNN methods for disease detection in plants like tomatoes and potatoes.

It was possible in [2] to boost the classification accuracy by performing fine-tuning on a pre-trained neural network architecture by adjusting the parameters.

Similarly, the work described in [3] used the Django framework to build a web-based solution to the problem of early detection of diseases in potato leaves.

The majority of the authors refer to popular CNN architectures such as VGG, InceptionNet, ResNet, EfficientNet, and MobileNet [4–6, 20]. In reality, quite a few of these works have been discovered to use the publicly available PlantVillage dataset [7, 16, 20, 15].