Information Systems Engineering and Management 11

Mohammed Serrhini
Kamal Ghoumid *Editors*

# Advances in Smart Medical, IoT & Artificial Intelligence

## Proceedings of ICSMAI'2024, Volume 1

Springer

# Information Systems Engineering and Management 11

The book series “Information Systems Engineering and Management” (ISEM) publishes innovative and original works in the various areas of planning, development, implementation, and management of information systems and technologies by enterprises, citizens, and society for the improvement of the socio-economic environment.

The series is multidisciplinary, focusing on technological, organizational, and social domains of information systems engineering and management. Manuscripts published in this book series focus on relevant problems and research in the planning, analysis, design, implementation, exploration, and management of all types of information systems and technologies. The series contains monographs, lecture notes, edited volumes, pedagogical and technical books as well as proceedings volumes.

Some topics/keywords to be considered in the ISEM book series are, but not limited to: Information Systems Planning; Information Systems Development; Exploration of Information Systems; Management of Information Systems; Blockchain Technology; Cloud Computing; Artificial Intelligence (AI) and Machine Learning; Big Data Analytics; Multimedia Systems; Computer Networks, Mobility and Pervasive Systems; IT Security, Ethics and Privacy; Cybersecurity; Digital Platforms and Services; Requirements Engineering; Software Engineering; Process and Knowledge Engineering; Security and Privacy Engineering, Autonomous Robotics; Human-Computer Interaction; Marketing and Information; Tourism and Information; Finance and Value; Decisions and Risk; Innovation and Projects; Strategy and People.

Indexed by Google Scholar. All books published in the series are submitted for consideration in the Web of Science.

For book or proceedings proposals please contact Alvaro Rocha (amrrocha@gmail.com).

Mohammed Serrhini · Kamal Ghoumid
Editors

# Advances in Smart Medical, IoT & Artificial Intelligence

## Proceedings of ICSMAI'2024, Volume 1

Springer

*Editors*
Mohammed Serrhini
Department of Computer Sciences, Faculty of Sciences
University Mohammed Premier Oujda Morocco
Oujda, Morocco

Kamal Ghoumid
University Mohammed Premier Oujda Morocco
Oujda, Morocco

ISSN 3004-958X ISSN 3004-9598 (electronic)
Information Systems Engineering and Management
ISBN 978-3-031-66849-4 ISBN 978-3-031-66850-0 (eBook)
https://doi.org/10.1007/978-3-031-66850-0



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

# Preface

This book includes a selection of articles from the First International Conference on Smart Medical, IoT & Artificial Intelligence – ICSMAI'24, held in Saidia, Morocco, from April 18 to 20, 2024.

This first edition aims to bring together students, researchers, and experts in the field of Smart Medical, IoT, and artificial intelligence to share their latest research findings, exchange ideas, and discuss challenges and opportunities in the field.

The Program Committee of ICSMAI'24 was composed of a multidisciplinary group of experts and those who are intimately concerned with, artificial intelligence, Internet of Things, Smart Medical, Information & communication technologies and security. They have had the responsibility for evaluating, in a 'blind-review' process, the papers received for each of the main themes proposed for the conference.

The main topics covered are:

Smart Healthcare/Smart Technologies/Smart Industry; AI, Machine Learning and Deep Learning; Parallel/concurrent/distributed algorithms and programming; Neuromorphic Systems; Distributed database, embedded and operating systems; Cloud/Fog/Edge Computing; Distributed ledgers and blockchain technologies; Internet of Things - IoT, 5G, URLLC; Robotics, Electrical and Electronics Engineering; Mobile, wireless, ad-hoc and sensor networks; Low-Power Wide-Area Networks; Virtual and augmented reality; Graph and Image Processing; Static and dynamic analysis and testing; Collaborative intelligent systems; Information/Network Security and privacy; Web of Things and Semantic Interoperability; Game Theory, mechanisms/hardware design; Computer Vision; Ethics and Cybercrime; Cryptocurrencies, Biometric, Cryptography, Authentication and Access Control; Fuzzy/Agents/Multi-agent Systems; Natural Language Processing; Data Analysis and Big Data; High Performance Computing; Scientific Calculation, Environment and Renewable Energy; Numerical modelling; AI-Optimized Medical Supply Chain in Smart Transportation; Emerging Technologies in Smart Medical Supply Transportation.

The book is aimed at all those dealing with Smart Medical, Internet of Things & artificial intelligence issues, including practitioners, researchers, and teachers as well as undergraduate, graduate, master's and doctorate students.

ICSMAI'24 received contributions from 14 countries around the world. The papers accepted for presentation and discussion at the conference are published by Springer (this book) and will be submitted for consideration in the Web of Science, Google Scholar, among others. Extended versions of selected best papers will be published in relevant journals, including WoS and Scopus indexed journals.

We acknowledge all those who contributed to the staging of ICSMAI'24 (authors, committees, and sponsors); their involvement and support was very much appreciated.

April 2024 Mohammed Serrhini

# Organization

## General Chairs

| | |
|---|---|
| Mohammed Serrhini | FSO, Med 1st University, Oujda, Morocco |
| Kamal Ghoumid | ENSAO, Med 1st University, Oujda, Morocco |

## General Co-chairs

| | |
|---|---|
| Mohcine Kodad | ESTO, UMP, Oujda, Morocco |
| F. Picaud | Franche-Comté Univ. Besançon, France |
| R. Yahiaoui | Franche-Comté Univ. Besançon, France |

## Organizing Committee

| | |
|---|---|
| Aissa Kerkour El Miad | UMP, Oujda, Morocco |
| Salah Daoudi | UMP, Oujda, Morocco |
| Rachid Malek | ENSAO, UMP, Morocco |
| E.-M. AR-Reyouchi | SNRT, Morocco |
| Driss Bria | FSO, UMP, Oujda, Morocco |
| Hamza Barboucha | ENSAO, Med. UMP, Oujda, Morocco |
| Omar Moussaoui | ESTO, UMP, Oujda, Morocco. |
| Hicham Bouali | ESTO, UMP, Oujda, Morocco |
| Rachid Benkaddour | ESTO, UMP, Oujda, Morocco |
| Imad Kissami | Mohamed VI Polytechnic University, Morocco |
| El Miloud Amamou | EHEI, Oujda, Morocco |
| Ouafae El Melhaoui | EHEI, Oujda, Morocco |
| Khalid Salhi | EHEI, Oujda, Morocco |
| Sara Said | EHEI, Oujda, Morocco |
| Driss Bahia | FSO, UMP, Oujda, Morocco |
| Khalil Kassmi | FSO, UMP, Oujda, Morocco |
| Alae Azouzoute | FSO, UMP, OUJDA, Morocco |
| Mohammed Benhaddou | ENA, Oujda, Morocco |
| Abderrahim Essadek | FSO, UMP, Oujda, Morocco |
| Sanae Mazouz | FSO, UMP, Oujda, Morocco |
| Moussa El Ayachi | FSO, UMP, Oujda, Morocco |
| El Houcine Sebbar | FMPO, UMP, Oujda, Morocco |

| | |
|---|---|
| Dounia El Moujtahide | FMPO, UMP, Oujda, Morocco |
| Abderrazak Saddari | FMPO, UMP, Oujda, Morocco |
| Saadiâ Chabel | EST, Guelmim, Morocco |

## Program Chairs

| | |
|---|---|
| Kamal Ghoumid | ENSAO, Med 1st University, Oujda, Morocco |
| Mohammed Serrhini | FSO, Med 1st University, Oujda, Morocco |
| Mohcine Kodad | ESTO, Med 1st University, Oujda, Morocco |

## Technical Program Committee

| | |
|---|---|
| I. Lajoie | Franche-Comté Univ. Besançon, France |
| O. Barakat | Franche-Comté Univ. Besançon, France |
| F. Auber | Franche-Comté Univ. Besançon, France |
| A. Mostefaoui | Montbebilart Univ. Franche-Comté, France |
| E. Yusifli | ATISOL Paris, France |
| M. Abboud-Abid | ENP, Oran, Algeria |
| K. Cetinkaya | Karabuk University, Turkey |
| S. Yevseiev | Kharkiv Polytechnic Institute, Ukraine |
| L. Bekirova | ASOIU Baku, Azerbaijan |
| A. G. Armada | Universidad Carlos III de Madrid, Spain |
| H. Bourouina | QPPC, Normal School Boussada, Algeria |
| E.-M. AR-Reyouchi | SNRT, Morocco |
| N. Derguini | ENS de Kouba, Algeria |
| G. Varshney | JSS Academy of Technical Education Noida, UP, India |
| A. Khare | Sharda University, Greater Noida, India |
| S. Varshney | Sharda University, Greater Noida, India |
| P. K. Rajput | UPES, Dehradun, India |
| F. Abdi | FSDM, Fez, Morocco |
| M. Kadic | Franche-Comte Univ, FEMTO-ST, France |
| K. Kassmi | FSO, UMP, Oujda, Morocco |
| F. Kardous | ENSTA, Tunis, Tunisia |
| A. Subasi | Univ. Jeddah, Saudi Arabia |
| S. Mian Qaisar | LINEACT, CESI Eng School, France |
| S. Zhgoon | Univ. Russia, Moscow, Russia |
| K. Salhi | EHEI, Oujda, Morocco |
| S. Said | EHEI, Oujda, Morocco |
| C. Lepers | INT Paris, France |

| | |
|---|---|
| A. Ahaitouf | Univ. Sidi Med ben Abdellah, Fez, Morocco |
| A. Boukerch | Ottawa Univ., Canada |
| F. Belkacem | L.I.S.I.C., FEI, USTHB, Algeria |
| L. Ferhat-Hamida | L.I.S.I.C, FEI, USTHB, Algeria |
| A. Telman | Baku Institute, Academician, Azerbaijan |
| F. Sadikoghlu | Univ. Nicosia, Cyprus |
| M. Farhat | UWI, Mona, FST, Jamaica |
| F. Benkhaldoun | Sorbonne Paris Nord University, France |
| M. Azizi | ESTO, UMP, Oujda, Morocco |
| O. Moussaoui | ESTO, UMP, Oujda, Morocco |
| A. Talbi | Centrale Lille, France |
| M. Kodad | ESTO, UMP, Oujda, Morocco |
| A. Kerkour Elmiad | FS, Oujda, Morocco |
| B. El Mahdawi | Engineering Technical College Baghdad, Iraq |
| C. Drocourt | UPJV, Amiens, France |
| D. Durand | UPJV, Amiens, France |
| I. Driouch | ENSAH, Al Hoceima, Morocco |
| A. El Moussaouy | CRMEF, Oujda, Morocco |
| M. El Kodadi | CRMEF, Oujda, Morocco |
| R. Touzani | FSO, UMP, Oujda, Morocco |
| M. Emharref | ENSAO, UMP, Oujda, Morocco |
| M. Saber | ENSAO, UMP, Oujda, Morocco |
| A. Benlghazi | ENSAO, UMP, Oujda, Morocco |
| M. L. Elhafyani | ENSAO, UMP, Oujda, Morocco |
| W. Mrabti | ENCGO, UMP, Oujda, Morocco |
| K. Laroussi | FSO, UMP, Oujda, Morocco |
| M. A. Madani | ENSAO, Oujda, Morocco |
| A. Kerkri | ENSAO, Oujda, Morocco |
| E. A. Moussa | UMP, Oujda, Morocco |
| M. Bellafkih | INPT, Rabat, Morocco |
| N. Kerzazi | ENSIAS, Rabat, Morocco |
| M. A. Kasmi | FS, Oujda, Morocco |
| N. Ouerdi | FS, Oujda, Morocco |
| I. Hamzaoui | Univ. Valenciennes, France |
| M. Umar Khan | COMSATS University, Islamabad, Pakistan |
| A. Mokhtari | Univ. Grenoble, France |
| M. A. Koulali | ENSA, Oujda, Morocco |
| Y. Jabri | ENSA, Oujda, Morocco |
| H. Bouali | ESTO, UMP, Oujda, Morocco |
| R. Benkaddour | ESTO, UMP, Oujda, Morocco |
| M. Jeyar | ESTO, UMP, Oujda, Morocco |
| M. Abboud-Abid | ENP, Oran, Algérie |

| | |
|---|---|
| I. Kissami | Mohamed VI Polytechnic University, Morocco |
| M. Fayad | Franche-Comté Univ. Besançon, France |
| N. Kharmoum | CNRST, Rabat, Morocco |
| Y. Khlifi | ENSAO, Med. UMP, Oujda, Morocco |
| H. Barboucha | ENSAO, Med. UMP, Oujda, Morocco |
| M. Benhaddou | ENA, Oujda, Morocco |
| A. Benbrike | FS, Oujda, Morocco |
| R. Chikh | ENSA, Oujda, Morocco |
| S. Daoudi | FS, Oujda, Morocco |
| A. Abbadi | ENSA, Oujda, Morocco |
| R. Korikache | ESTO, UMP, Oujda, Morocco |
| J. Zidouni | ENSA, Oujda, Morocco |
| H. Ouedfell | ENSA, Oujda, Morocco |
| M. El Boukhari | ESTO, UMP, Oujda, Morocco |
| M. Grari | ESTO, UMP, Oujda, Morocco |
| S. Chabel | EST, Guelmim, Morocco |
| D. Bahia | FS, Oujda, Morocco |
| O. Moussaoui | ESTO, UMP, Oujda, Morocco |
| K. Haboubi | ENSA, Al Hoceima, Morocco |
| F. Dimane | ENSA, Al Hoceima, Morocco |
| A. Essadek | FS, Oujda, Morocco |
| T. Hajji | ENSEM, Meknes, Morocco |
| N. Kharmoum | CNRST, Rabat, Morocco |
| M. Boutaounte | ENCG, Dakhla, Morocco |
| M. S. Elyoubi | FS, Kenitra, Morocco |
| M. Blej | CRMEF, Oujda, Morocco |
| M. Tahrichi | ESTO, UMP, Oujda, Morocco |
| A. Serghini | ESTO, UMP, Oujda, Morocco |
| K. Reklaoui | ENSA, Tetouan, Morocco |
| A. Nasser | ESTO, UMP, Oujda, Morocco |
| I. Bouabdallah | CRMEF Oujda, Morocco |
| F. Falyouni | FSO Oujda, Morocco |
| M. Merzougui | UMP Oujda, Morocco |
| S. Mazouz | FSO, UMP, Oujda, Morocco |
| B. Al-Mahdawi | TCE, MT University, Baghdad, Iraq |
| A. Al-Azzawi | TCE, MT University, Baghdad, Iraq |
| A. Abdulazeez | CS, Al-Mustansiriyah University, Baghdad, Iraq |
| A. Moussa | FSO, UMP, Oujda, Morocco |
| A. Ghadbane | FSO, UMP, Oujda, Morocco |
| A. Benlghazi | ENSAO, Oujda, Morocco |
| Z. Bougroune | ENSAO, Oujda, Morocco |
| K. Azghiou | ENSAO, Oujda, Morocco |

| | |
|---|---|
| E. H. Sebbar | FMPO, UMP, Oujda, Morocco |
| D. El Moujtahide | FMPO, UMP, Oujda, Morocco |
| A. Saddari | FMPO, UMP, Oujda, Morocco |

# Contents

## Computer Vision, NLP, LLMs and Data Analysis



## Mobile, Internet of Things, 5G and Embedded System

# About the Editors

**Mohammed Serrhini** is a professor with decades of academic and research experience in the Department of computer science at Faculty of Sciences, Mohamed First University, Oujda, Morocco. Mohammed Serrhini finished his engineering degree in computer science in 1996 from Polytechnic Institute of Applied Mathematics and Computer Science at Tula State University and PhD in 2012, at the Faculty of Sciences of the University Mohamed Premier of Oujda. He leads innovative research in artificial intelligence (AI) and cyber-security. His research interests are in the areas of computer vision, image processing, brain–computer interface in education, IA security; his works have appeared in several prestigious journals. He has served as General Chair of more than 15 international conferences and has delivered a number of keynote talks on AI and cyber-security. Serrhini receives many awards such from UNESCO in 2022 among others.

**Kamal Ghoumid** received his PhD degree from the Institut TELECOM, TELECOM Sud-Paris, Evry, France, and Institute FEMTO-ST of the Franche-Comté University (Besançon, France), in 2008. He previously graduated as an engineer in electronics and telecommunications at CNAM-Paris (France), received his Masters in communication systems from Paris-Est University (France) and specialized Masters in technics of radio communications. He has worked as a post-doctoral researcher at Jean Lamour Institute of Henri Poincaré University (Nancy, France) from 2008 to 2009, and the Institut FEMTO-ST of the Franche-Comté University, Besançon as well. He is also an Assistant Professor in the National School of Applied Sciences (ENSAO), Mohammed Premier University of Oujda (Morocco), where he is the head of the research team on Signals, Systems and Information Processing. His research interests are mainly in signal processing and integrated optic components in the fields of telecommunications, wireless and networks Communications, radio over fiber and communications systems.

# AI, Machine Learning and Deep Learning

# Forest Fire Surveillance Through Deep Learning Segmentation and Drone Technology

Mimoun Yandouzi[1,4](✉), Sokaina Boukricha[1,4], Mounir Grari[2,4], Mohammed Berrahal[3,4], Omar Moussaoui[2,4], Mostafa Azizi[2,4], Kamal Ghoumid[1,4], and Aissa Kerkour Elmiad[3,4]

[1] LSI, ENSAO, Mohammed First University, Oujda, Morocco
m.yandouzi@ump.ac.ma
[2] MATSI, ESTO, Mohammed First University, Oujda, Morocco
[3] LMC, PFS, Cadi Ayyad University, Safi, Morocco
[4] LARI, FSO, Mohammed First University, Oujda, Morocco

**Abstract.** Forests are essential to our planet's well-being, playing a vital role in climate regulation, biodiversity preservation, and soil protection, thus serving as a cornerstone of our global ecosystem. The threat posed by forest fires highlights the critical need for early detection systems, which are indispensable tools in safeguarding ecosystems, livelihoods, and communities from devastating destruction. In combating forest fires, a range of techniques is employed for efficient early detection. Notably, the combination of drones with artificial intelligence, particularly deep learning, holds significant promise in this regard. Image segmentation emerges as a versatile method, involving the partitioning of images into multiple segments to simplify representation, and it leverages deep learning for fire detection, continuous monitoring of high-risk areas, and precise damage assessment. This study provides a comprehensive examination of recent advancements in semantic segmentation based on deep learning, with a specific focus on Mask R-CNN (Mask Region Convolutional Neural Network) and YOLO (You Only Look Once) v5, v7, and v8 variants. The emphasis is placed on their relevance in forest fire monitoring, utilizing drones equipped with high-resolution cameras.

**Keywords:** Forest fires · Deep Learning · Segmentation · UAV (Drone) · Mask R-CNN · YOLO

## 1 Introduction

Forests play a crucial role in maintaining the ecological balance of our planet, providing oxygen, sheltering precious biodiversity, and regulating the climate. However, these ecosystems are constantly threatened by wildfires, which can cause irreparable damage in a short amount of time. Early detection of forest fires is crucial to limit their spread and reduce the ecological and economic losses associated with them [1].

Unfortunately, the issue of wildfires has become increasingly pressing over time, due in part to climate change and deforestation. Rapid detection has therefore become


M. Serrhini and K. Ghoumid (Eds.): ICSMAI 2024, ISEM 11, pp. 3–12, 2024.
https://doi.org/10.1007/978-3-031-66850-0_1

an absolute priority. This is where technological advancements, such as the use of drones equipped with sophisticated sensors, come into play [2]. Paired with computer vision techniques, particularly those based on Deep Learning, these drones can provide real-time monitoring and early detection of forest fires over vast expanses, which was previously difficult to achieve with traditional methods [3].

Several techniques are employed for forest fire detection, including analysis of infrared images, satellite monitoring, and the use of sensor networks deployed on the ground. However, the use of drones offers significant advantages, including higher spatial resolution and the ability to cover hard-to-reach areas. By combining these capabilities with Deep Learning algorithms, drones can identify fire signals with increased accuracy, thereby enabling swift intervention by firefighting teams [4].

Regarding image segmentation, a key technique in fire detection, Deep Learning offers promising possibilities. By using Convolutional Neural Networks (CNNs), it is possible to efficiently segment images to isolate fire and smoke areas, even in reduced visibility conditions. This capability allows for continuous monitoring of forest fires, thus facilitating decision-making for authorities and improving emergency management. In summary, the integration of Deep Learning into forest fire detection and supervision paves the way for more effective strategies in preventing and combating these natural disasters [5].

In this study, we broaden our research scope after exploring image classification and object detection in two previous works [4, 6]. We now focus on semantic segmentation, delving into a comparison of the most prevalent Deep Learning models for this task. Our aim is to assess their potential for wildfire detection and supervision using drones.

## 2 Backgrounds

### 2.1 Deep Learning

Deep learning is a subset of artificial intelligence focused on learning multi-level data representations. In computer vision, it enables machines to understand visual content in images or videos, mimicking human visual perception. Thanks to deep learning, computer vision has made tremendous strides in recent years, particularly in object recognition, pattern detection, and image segmentation [3]. This advancement has revolutionized image processing capabilities for various applications, including forest fire monitoring. With this combination, it's now possible to quickly classify images as containing or lacking forest fires, accurately detect objects such as fire or smoke in forest environments, and efficiently segment areas affected by fires for better natural disaster management.

### 2.2 Image Segmentation

Segmentation, a fundamental technique in computer vision, divides images into meaningful regions for analysis, offering fine granularity by identifying objects, boundaries, and structures. It finds diverse applications, from medical imaging for diagnosing conditions to autonomous driving for obstacle detection, augmented reality, video surveillance, facial recognition, and robotics [7]. Moreover, segmentation proves invaluable for forest fire detection and monitoring, providing insights into fire size, affected areas, and differentiating fire types [8].

Segmentation techniques for forest fire monitoring include instant, semantic, and Panoptic segmentation. Instant segmentation enables real-time monitoring for swift response to fire outbreaks, ideal for early detection using drone imagery. Semantic segmentation identifies fire-affected areas, aiding decision-making for response strategies. Panoptic segmentation provides comprehensive image representation, crucial for understanding fire dynamics. The choice of technique depends on specific criteria; for our research focusing on fire extent, semantic segmentation is preferred (Fig. 1).

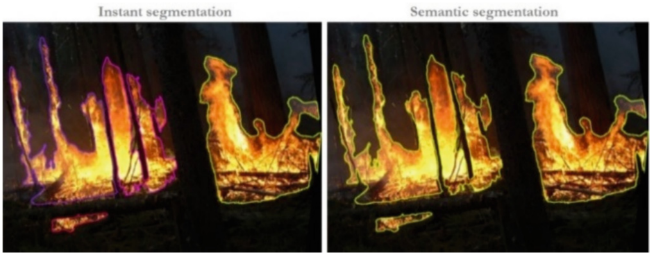


**Fig. 1.** Examples of Instant Segmentation and Semantic Segmentation.

### 2.3 Drones for Forest Fire Surveillance

A drone, also known as an Unmanned Aerial Vehicle (UAV), is an aircraft operated without a human pilot onboard. Drones have gained significant attention for their versatility and applications across various industries, including forest fire detection and monitoring. Utilizing drones for these purposes offers several advantages, such as their ability to access remote or hazardous areas, providing real-time aerial footage for enhanced situational awareness, and covering large geographical areas quickly and efficiently. However, there are constraints associated with drone usage, notably limited battery life and onboard resources [6]. These limitations necessitate careful planning of flight paths, payload management, and resource optimization to maximize the effectiveness of drone operations in forest fire management scenarios.

## 3 Related Work

In their research, Bulatov et al. [9] utilized instance segmentation to monitor and analyze the spatial and temporal distribution of deadwood, a significant factor linked to forest fire occurrences. Their study highlights the considerable promise of achieving precise instance segmentation within the RGB and elevation domains, even with constraints on training data availability. Employing a high-performing Mask R-CNN model, the team effectively mapped standing and fallen deadwood instances across German forests. The outcomes of their investigation were particularly noteworthy, demonstrating an impressive overall accuracy rate of 92.4%.

Tran et al. [10] utilized images from a forest fire in Andong, South Korea, in April 2020. They employed a two-stage deep learning approach at the patch level. The first network, based on UNet++, operated at patch-level 1. Its output predictions served as input for a second network, using UNet architecture, to refine results based on positional information. Their method's performance was evaluated against state-of-the-art algorithms, including a comparative analysis of loss functions. However, the approach requires training dual patch-level models in various conditions for optimal performance and transitioning to an online platform for practicality and reduced processing time.

Zhao et al. [11] proposed an innovative saliency detection algorithm geared towards swiftly pinpointing and segmenting core fire areas in aerial images. Their approach addresses the issue of feature loss caused by direct resizing, rendering it well-suited for tasks like data augmentation and the creation of the 'UAV_Fire' dataset, which serves as a standard repository of fire images captured by drones. Additionally, they introduced 'Fire_Net,' a 15-layer deep convolutional neural network (DCNN) architecture, designed to function as both a self-learning fire feature extractor and classifier. Through rigorous evaluation of various architectures and critical parameters (e.g., dropout ratio, batch size) of the DCNN model concerning its validation accuracy, their proposed design outperformed previous methods, achieving an impressive overall accuracy of 98%. Furthermore, 'Fire_Net' exhibited a remarkable average processing speed of 41.5 ms per image, enabling real-time wildfire monitoring and inspection.

## 4 Proposed Method

Our proposed method (Fig. 2) consists of the following six steps: Data Collection, Preprocessing, Deep Learning Model, Object Detection, In-Box Segmentation, and Output Data and Evaluation.

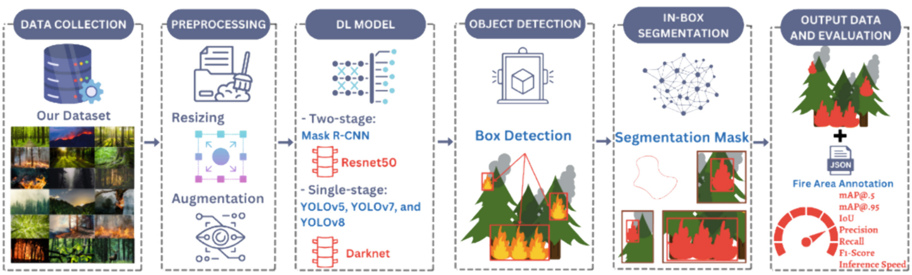


**Fig. 2.** Approach to Experimental Research Methodology.

### 4.1 Data Collection

The primary phase of our proposed method involves assembling a dataset of images depicting instances of forest fires. This dataset comprises a total of 4236 labeled images. It includes photographs taken from ground-level cameras as well as aerial drones, offering both intricate real-time portrayals from ground cameras and broader perspectives from drones to monitor the fire's scope. We also incorporated publicly accessible datasets to enhance the dataset's authenticity. Our rigorous approach ensures a comprehensive range of images, encompassing diverse fire types, intensities, and environmental settings. The labeling process was streamlined using the Darwin training data platform, known for its effectiveness in creating AI solutions [12]. Team members efficiently distributed the dataset to optimize workflow and resolve tasks swiftly.

### 4.2 Preprocessing

Our approach encompasses preprocessing and augmenting the gathered datasets. Preprocessing entails resizing images to dimensions of $640 \times 640$ and converting them into JPG format, crucial for effective image segmentation. Augmentation introduces distortions such as rotations and flips to diversify the dataset and improve system resilience. The combination of resizing and augmenting data enhances the likelihood of successful segmentation outcomes [13].

### 4.3 Deep Learning Models

Our examination of deep learning models for segmentation encompasses two key classes: YOLO and Mask R-CNN.

YOLO, particularly versions v5, v7, and v8, has significantly advanced computer vision, initially devised for object detection but now integrating image segmentation capabilities [14]. This integration marks a crucial milestone, seamlessly combining object detection with segmentation, enabling unified extraction of object information and precise contour delineation. These advancements offer immense potential, from bolstering automated surveillance to enabling advanced environmental analysis.

Similarly, Mask R-CNN, introduced in 2017 by He et al., represents a historic breakthrough by merging object detection with semantic segmentation. Unlike YOLO, Mask R-CNN operates with a two-stage architecture, evolving from Faster R-CNN. It introduces a segmentation branch to precisely identify and delineate detected objects, revolutionizing applications such as multi-instance object detection and contextual understanding [15]. As a backbone, we will utilize ResNet50, a convolutional neural network renowned for its depth and efficiency, which has been validated as an effective choice in prior studies [16].

### 4.4 Object Detection

The first step in segmentation is object detection, a pivotal stage in identifying and delineating specific entities within an image, such as fires. Object detection focuses on locating and classifying these objects of interest, enabling the system to proficiently recognize and outline them amidst intricate visual backgrounds. This capability serves as the cornerstone for subsequent segmentation tasks, facilitating deeper analysis and comprehension of the image content.

### 4.5 In-Box Segmentation

The second step in segmentation is the In-Box segmentation, which involves refining the initial object detection by precisely delineating the boundaries of detected objects within their respective bounding boxes. This process enhances the accuracy and granularity of segmentation, enabling a more detailed understanding of the objects' spatial extent and characteristics within the image. Through advanced algorithms and techniques, such as pixel-wise classification and boundary refinement, the In-Box segmentation further refines the segmentation process, laying the groundwork for comprehensive analysis and interpretation of the image content.

### 4.6 Output Data and Evaluation

The final stage of the process involves "Output Data and Evaluation", where the segmented results undergo comprehensive processing and evaluation using a range of metrics. Central to this evaluation are key metrics such as the F1-Score, mAP@0.5, and mAP@0.95, which serve to gauge the accuracy and efficacy of the segmentation models.

The F1-Score stands as a pivotal metric, striking a balance between precision and recall, thereby offering a holistic assessment of the model's performance [17]. It is calculated using the formula:

$$F1score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{1}$$

Complementing this, mAP@0.5 and mAP@0.95 metrics quantify the mean Average Precision across varying Intersection over Union (IoU) thresholds, providing valuable insights into the model's capacity for precise object localization [18]. Mean Average Precision (mAP) is calculated as:

$$mAP = \frac{1}{N} \sum_{i=1}^{N} APi \tag{2}$$

where $APi$ is the Average Precision of class i and N is the number of classes.

To deepen the evaluation, it's imperative to consider the inference speed, evaluating the model's efficiency in real-time applications. This consideration ensures that the model can deliver timely and responsive performance in practical scenarios, which is crucial for its real-world utility.

By subjecting the segmentation models to rigorous evaluation utilizing these metrics, their capabilities can be thoroughly scrutinized, thereby facilitating further refinement and optimization aimed at enhancing their performance [18]. Additionally, the Intersection over Union (IoU) is calculated as:

$$IoU = \frac{(Area\ of\ \ Overlap)}{(Area\ of\ \ Union)} \tag{3}$$

# 5 Results and Discussions

## 5.1 Hardware Characteristics

The footage capturing the simulated fires at the university campus was obtained through a combination of a DJI Mavic Air drone and the mobile phone cameras of team members. For model training, TensorFlow v2.13.0 [19] was utilized, which is an open-source software library for data analysis and machine learning. The training took place on a high-performance computing (HPC) system equipped with powerful hardware, including 2× Intel Gold 6148 (2.4 GHz/20 cores) CPUs and 2× NVIDIA Tesla V100 graphics cards, each boasting 32GB of RAM. These hardware specifications were essential to provide the necessary computational power for effective model training.

## 5.2 Experimental Results

In Table 1 and Fig. 3, we present the results obtained for various metrics, including accuracy, precision, recall, Intersection over Union, mAP at confidence thresholds of 0.5 and 0.95, as well as the inference time for each model. To maximize effectiveness, we trained all of the models for seventy epochs. The initial observation to highlight is that all the models under examination demonstrated the capability to be trained for forest fire segmentation. However, there was variability in the number of epochs required for convergence. Notably, the YOLO models exhibited swifter training and achieved convergence in fewer than ten epochs.

While segmentation speed is crucial for real-time drone processing, our focus here is on the mAP@0.95 metric. We aim to utilize segmentation not just for fire detection but also for assessing post-fire damage and aiding reforestation efforts. Drones must operate at high altitudes to survey large areas effectively, requiring segmentation with near-perfect precision to accurately depict burnt areas and tree counts.

The Mask R-CNN model excels in terms of image segmentation performance. It achieves an mAP@0.5 of 97.93% and an mAP@0.95 of 97.53% for object detection "Fire", along with an mAP@0.5 of 97.91% and an mAP@0.95 of 96.42% for fire segmentation, accompanied by an impressive F1 score of 99.01%. Mask R-CNN stands out as the top-performing model among all the models examined. However, it also demands the highest computational resources, requiring approximately 0.028 s per frame for object detection and around 0.0046 s per frame for the segmentation operation. This translates to an overall processing time of 0.0346 s per frame for the entire image analysis pipeline.

All variations of the YOLO model demonstrated slightly inferior performance compared to the Mask R-CNN model, yet offered significantly improved inference times. These findings remain compelling, particularly those attributed to the YOLOv8 model, which achieved an F1 score of approximately 98.3%. The model attained a mAP@0.95 of 96.28% and 89.61% for object detection and segmentation, respectively. YOLOv8's performance is noteworthy, especially given its swift overall inference time of 3.8 ms per image, rendering it an optimal choice for real-time detection tasks.

In most of the models analyzed, the object detection phase consistently shows longer processing times compared to the segmentation stage, which aims to delineate the fire's outline within the specified area. Remarkably, segmentation time remains relatively stable across various models. Significant improvements in both performance and speed within the object detection phase, especially with the YOLOv8 model, have been noted. Figure 4 presents several instances of forest fire image segmentation performed by the YOLOv8 model.

**Table 1.** Achieved results for the implemented models.

| DEEP LEARNING MODEL | STAGE | IoU | mAP@0.5 | mAP@0.95 | Inference | Precision | Recall | F1-Score |
|---|---|---|---|---|---|---|---|---|
| | | % | % | % | s/Image | % | % | % |
| MASK R-CNN | BOXING | 91.02 | 97.93 | 97.53 | ~0.0280 | 99.13 | 98.89 | 99.01 |
| | SEGMENTATION | 92.14 | 97.91 | 96.42 | ~0.0046 | | | |
| YOLOv5 | BOXING | 88.94 | 95.43 | 92.04 | ~0.0051 | 98.18 | 97.95 | 98.06 |
| | SEGMENTATION | 89.02 | 95.25 | 84.20 | ~0.0039 | | | |
| YOLOv7 | BOXING | 89.17 | 95.10 | 95.96 | ~0.0027 | 98.35 | 98.17 | 98.26 |
| | SEGMENTATION | 89.14 | 94.90 | 85.00 | ~0.0037 | | | |
| YOLOv8 | BOXING | 89.36 | 96.50 | 96.28 | ~0.0011 | 98.47 | 98.22 | 98.34 |
| | SEGMENTATION | 90.54 | 96.37 | 89.61 | ~0.0027 | | | |

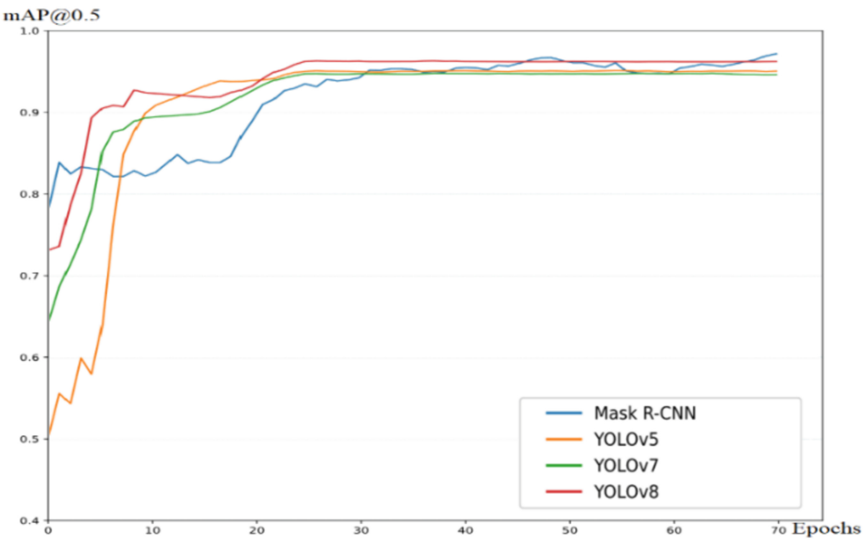


**Fig. 3.** Achieved mAP@0.5 over epochs on validation set for the implemented models.


**Fig. 4.** Segmentation of forest fires: Illustrative examples with YOLOv8.

## 6 Conclusion

The utilization of drones in conjunction with deep learning techniques proves to be a cost-effective method for accurately and swiftly detecting forest fires in real-time. Image segmentation models based on deep learning play a critical role not only in fire detection but also in conducting thorough environmental analyses. This sophisticated technological approach aids in ecosystem monitoring, thereby contributing to the preservation of nature and facilitating the assessment of damage caused by wildfires. In this study, recent advancements in deep learning-based semantic segmentation, with a specific focus on models such as Mask R-CNN and YOLO versions 5, 7, and 8, are examined. The main focus lies in assessing their applicability for monitoring forest fires using drones equipped with RGB cameras. All the models demonstrate promising performance across various metrics, establishing themselves as valuable tools for the semantic segmentation of forest fire images. While Mask R-CNN showcases exceptional image segmentation performance, achieving an impressive F1 score exceeding 99%, along with an mAP@0.5 of nearly 97.5% for the "Fire" object detection step and 98% for segmentation within the bounding boxes, YOLO models exhibit commendable performance along with outstanding inference speeds, making them ideal choices for real-time detection tasks. Notably, YOLOv8 demonstrates an impressive overall inference time of just 3.8 ms per image.

Moving forward, we plan to develop a comprehensive framework for the detection and monitoring of forest fires. This framework will capitalize on the collaboration between IoT sensor networks, drone networks, and state-of-the-art deep learning algorithms.

## References


1. Grari, M., Boukabous, M., et al.: Forest fire detection and monitoring through environment sound spectrum using deep learning. J. Theor. Appl. Inf. Technol. **15**, 19 (2023)
2. Grari, M., et al.: IoT-based approach for wildfire monitoring and detection, pp. 205–213 (2023). https://doi.org/10.1007/978-3-031-35251-5_19
3. Yandouzi, M., Grari, M., et al.: Review on forest fires detection and prediction using deep learning and drones. J. Theor. Appl. Inf. Technol. **100**, 4565–4576 (2022)
4. Yandouzi, M., et al.: A lightweight deep learning model for forest fires detection and monitoring, pp. 697–705 (2023). https://doi.org/10.1007/978-981-19-6223-3_71
5. Grari, M., Yandouzi, M., et al.: Using IoT and ML for forest fire detection, monitoring, and prediction: a literature review. J. Theor. Appl. Inf. Technol. **100**, 5445–5461 (2022)
6. Yandouzi, M., et al.: Investigation of combining deep learning object recognition with drones for forest fire detection and monitoring. IJACSA) Int. J. Adv. Comput. Sci. Appl. **14** (2023). https://doi.org/10.14569/IJACSA.2023.0140342
7. Mo, Y., Wu, Y., et al.: Review the state-of-the-art technologies of semantic segmentation based on deep learning. Neurocomputing **493**, 626–646 (2022). https://doi.org/10.1016/J.NEUCOM.2022.01.005
8. Kherraki, A., et al.: Traffic scene semantic segmentation by using several deep convolutional neural networks. In: Proceedings of the 2021 3rd IEEE Middle East North Africa Communications Conference, pp. 1–6 (2021). https://doi.org/10.1109/MENACOMM50742.2021.9678270
9. Bulatov, D., Leidinger, F.: Instance segmentation of deadwood objects in combined optical and elevation data using convolutional neural networks **11863**, 299–308 (2021). https://doi.org/10.1117/12.2599837
10. Tran, D.Q., et al.: Damage-map estimation using UAV images and deep learning algorithms for disaster management system. Rem. Sens. **12**, 4169 (2020). https://doi.org/10.3390/RS12244169
11. Zhao, Y., et al.: Saliency detection and deep learning-based wildfire identification in UAV imagery. Sensors **18**, 712 (2018). https://doi.org/10.3390/S18030712
12. Darwin. https://darwin.v7labs.com/datasets. Accessed 5 Sep 2023
13. Berrahal, M., et al.: Investigating the effectiveness of deep learning approaches for deep fake detection. Bull. Electr. Eng. Inform. **12**, 3853–3860 (2023). https://doi.org/10.11591/EEI.V12I6.6221
14. Islam, A., et al.: Enhanced leafy vegetable analysis: image classification and disease instance segmentation using deep learning techniques. https://doi.org/10.2139/SSRN.4470131
15. Bharati, P., Pramanik, A.: Deep learning techniques—R-CNN to mask R-CNN: a survey. Adv. Intell. Syst. Comput. **999**, 657–668 (2020). https://doi.org/10.1007/978-981-13-9042-5_56/COVER
16. Yandouzi, M., et al.: Forest fires detection using deep transfer learning. Int. J. Adv. Comput. Sci. Appl. **13**, 268–275 (2022). https://doi.org/10.14569/IJACSA.2022.0130832
17. Cakir, F., et al.: Deep metric learning to rank. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp 1861–1870 (2019)
18. Average Precision—CloudFactory Computer Vision Wiki. https://wiki.cloudfactory.com/docs/mp-wiki/metrics/average-precision. Accessed 5 Jan 2024
19. API Documentation—TensorFlow v2.13.0. https://www.tensorflow.org/api_docs. Accessed 5 Sep 2023

# Transfer Learning for Efficiency in Elderly Fall Detection with Limited Data Samples

Moustafa Fayad[1](✉), Mohammed Amine Merzoug[2], Ahmed Mostefaoui[2], Kamal Ghoumid[3], Isabelle Lajoie[1], and Réda Yahiaoui[1]

[1] University of Franche-Comté, SINERGIES, 25000 Besançon, France
{moustafa.fayad,isabelle.lajoie,reda.yahiaoui}@univ-fcomte.fr
[2] DISC Department, University of Franche-Comté, FEMTO-ST Institute, 90000 Belfort, France
{mohammedamine.merzoug,ahmed.mostefaoui}@univ-fcomte.fr
[3] Mohammed First University, ENSAO, 60050 Oujda, Morocco
k.ghoumid@ump.ac.ma

**Abstract.** In light of significant demographic shifts worldwide, elderly fall detection is an ongoing, vital area of research. Deep learning, known for its effectiveness in healthcare applications, is challenged by limited accessibility to substantial datasets, especially in the case of fall detection. Moreover, training deep learning models is both time-consuming and costly. To address these issues, in this paper, we implemented a sample size technique called N×Subsampling and utilized transfer learning with MobileNetV2. Our study leveraged the public URFD database, and the obtained experimental results demonstrated a notable achievement: an accuracy range of 94.74% to 98.94%, using only a 15% training subset consisting of 732 images of activities of daily living and 369 images of fall scenarios.

**Keywords:** Fall detection · Transfer learning · Sample size

## 1 Introduction

In recent decades, the elderly population has continued to increase significantly. According to the World Health Organization (WHO), the number of seniors aged 60 and over was 1 billion in 2019 [20]. This statistic is expected to rise to 1.4 billion by 2024 and reach 2.1 billion by 2050. In France, the National Institute of Statistics and Economic Studies (INSEE) estimates that the primary growth in the French population concerns elderly individuals [11]. This growth explicitly affects seniors aged 75 and over; by 2040, the proportion of seniors is projected to exceed a quarter of the total population.

Aging is generally characterized by a gradual and irreversible decline in the physical functions of the human body [10]. The frailty resulting from this universal natural phenomenon exposes seniors to a higher risk of domestic accidents, including falls [5]. In other words, the combination of aging and frailty increases the risk of falling. Indeed, falls constitute a major public health problem and a significant source of injury-related deaths for the elderly [19, 27]. Globally, it is estimated that one in three individuals aged over 65 experiences at least one fall per year [8, 18]. In France, this issue has dramatic


M. Serrhini and K. Ghoumid (Eds.): ICSMAI 2024, ISEM 11, pp. 13–20, 2024.
https://doi.org/10.1007/978-3-031-66850-0_2

human and economic consequences [1, 24]: i) nearly one-third of individuals over 65 and half of those over 80 experience falls each year, ii) these falls result in approximately 9,000 deaths per year in individuals over 65, iii) falls have a significant cost for society, estimated at 2 billion euros per year, with 1.5 billion attributed to health insurance.

To address this demographic challenge, scientific researchers are committed to aligning with societal needs through the development of innovative solutions. These solutions rely on technological advances to ensure a safe aging process. Various types of artificial intelligence have proven to be valuable tools for developing effective fall detection approaches [12, 25]. These models outperform traditional methods based on predefined thresholds [21], as they minimize detection errors and adapt to various contexts and situations, thus providing better precision, adaptability, and generality.

Developing fall detection approaches based on deep learning presents numerous challenges for researchers. One of the major challenges lies in the need for large datasets for training and the associated high costs [2]. However, in specific domains, the number of publicly accessible samples is very limited [4]. In the context of fall detection, datasets witness a shortage of voluminous data and a limitation of real-world samples [5, 13]. Additionally, creating such a high-quality database involves high costs and considerable time investment, exposing volunteers to an increased risk of injury. Hence, transfer learning is a promising solution for overcoming obstacles related to the lack of voluminous data [26].

The originality of this study lies in evaluating the sample size and applying transfer learning to develop an efficient fall detection system. Our approach involves several key steps. First, the N$\times$Subsampling technique [3] is employed to partition the training dataset. Then, these subsamples are used in the fine-tuning training of the model based on MobileNetV2 [23], with modifications made to the classification layers. By combining transfer learning and determining the sample size, our approach demonstrates high performance in fall detection, even with small fractions of the dataset, especially in the context of limited, specific fall data.

This paper is organized into five sections. Section 2 presents previous studies related to our research on transfer learning for human fall detection. Section 3 provides the background context for our study. Section 4 details the experimentation conducted and then presents and analyzes the obtained results. Finally, Sect. 5 concludes the paper with future work.

## 2 Related Work

Researchers dedicate efforts to all phases of the fall detection process, including relevant feature selection, data preprocessing, and other steps, aiming to optimize system performance and ensure high-quality detection [6, 7]. The fall detection sector is confronted with a significant challenge due to the limited availability of training data. To address this gap, the scientific community has opted for transfer learning. Several studies have demonstrated the effectiveness of transfer learning in the context of fall detection.

McCall et al. [17] conducted a study focused on fall detection and prediction, leveraging transfer learning and the transformer model. Initially, the latter was trained on the extensive MPOSE dataset, consisting of 15429 samples of 20 actions performed by

100 subjects. These data samples comprised sequences of 2D poses involving walking, jogging, running, and kicking. This initial pre-training phase allows the model to acquire valuable representations of features related to human poses and actions. For the fine-tuning process, the weights of the integration and transformer encoder layers are frozen, leaving only the two last adjustable MLP layers. The output network architecture was also adapted by modifying the number of outputs from 20 to 2 (fall/non-fall or high/low fall risk). For this purpose, the CAUCAFall dataset was used, with a distribution of 70% for training, 10% for validation, and 30% for testing. The experimental results of the small transformer with transfer learning revealed the best performance, reaching 97.95% $\pm$ 0.87% for fall detection.

Lobanova et al. [15] applied the AlexNet model, enhanced by the use of transfer learning, to address the issue of fall detection. They modified the last fully connected layer, replacing it with a new one with an output size of two (Fall/No Fall). The weights of this layer were adjusted during training. The fine-tuning process and validation were conducted using the Le2i dataset, comprising 191 videos representing daily activities and simulated falls by 9 volunteers. The experimental results demonstrated an accuracy of 96% for fall detection from images.

Sadreazami et al. [22] introduced an approach based on a deep neural net-work, using VGG16, for fall detection by leveraging radar data and transfer learning. Radar data were transformed into spectrograms using the squared magnitude of the short-term Fourier transform. These data, collected from radars, include 121 fall scenarios and 85 non-fall activities, simulated by five subjects in a cluttered room. The experimental results revealed that the model, resulting from fine-tuning the last convolutional block in conjunction with the classification part (including maximum pooling layers, spatial global average pooling, and the output layer), achieved a significant accuracy of 95.64% with a 3-block cross-validation.

Yhdego et al. [28] proposed a fall detection method using transfer learning and Support Vector Machines (SVM). They converted 3D acceleration data and amplitude into RGB images through continuous wavelet transform to make them compatible with the AlexNet architecture. The last three layers of AlexNet were then fine-tuned for classification. Utilizing the public URFD database with 80% for training and 20% for testing, their approach achieved an accuracy of 96.43%.

Lobanova et al. [16] presented a fall detection method using bioradar data and transfer learning. They evaluated different combinations of bioradars to determine the optimal configuration. For this purpose, they created a database simulated by 8 volunteers, comprising 400 records for each bioradar (200 fall records and 200 non-fall records). Their approach involves pre-trained convolutional neural networks (CNNs) like AlexNet, which process bioradar signals converted into scalograms. During the fine-tuning step, the network's last layer is adjusted to produce a two-value output corresponding to falls and non-falls. The experimental results using the leave-one-out method and hold-out validation revealed an accuracy rate of 92% for two bioradars, 98% for three bioradars, and 99% for four bioradars.

Research in the literature has demonstrated the effectiveness of transfer learning in achieving highly accurate results better suited to real-world needs in fall detection. However, to the best of our knowledge, no study has explored the impact of sample size