Lecture Notes in Electrical Engineering 948

Rajeev Agrawal
Chandramani Kishore Singh
Ayush Goyal
Dinesh Kumar Singh *Editors*

Modern Electronics Devices and Communication Systems

Select Proceedings of MEDCOM 2021

Springer

Lecture Notes in Electrical Engineering

Volume 948

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering—quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** This series is indexed by EI Compendex and Scopus databases. ****

Rajeev Agrawal · Chandramani Kishore Singh ·
Ayush Goyal · Dinesh Kumar Singh
Editors

Modern Electronics Devices and Communication Systems

Select Proceedings of MEDCOM 2021

Springer

Editors
Rajeev Agrawal
G. L. Bajaj Institute of Technology
and Management
Greater Noida, India

Ayush Goyal
Department of Electrical Engineering
and Computer Science
Texas A&M University—Kingsville
Kingsville, USA

Chandramani Kishore Singh
Indian Institute of Science Bangalore
Bengaluru, India

Dinesh Kumar Singh
G. L. Bajaj Institute of Technology
and Management
Greater Noida, India

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-19-6382-7 ISBN 978-981-19-6383-4 (eBook)
https://doi.org/10.1007/978-981-19-6383-4

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

MEDCOM 2021 Luminaries

Chief Patron

Dr. Ram Kishore Agarwal, Hon'ble Chairman, G. L. Bajaj Institute of Technology and Management, Greater Noida, U.P., India

Patron

Mr. Pankaj Agarwal, Hon'ble Vice-Chairman, G. L. Bajaj Institute of Technology and Management, Greater Noida, U.P., India

General Chair

Dr. Rajeev Agrawal, Director, G. L. Bajaj Institute of Technology and Management, Greater Noida, U.P., India

Technical Program Committee Chair

Dr. Subodh Wariya, Institute of Engineering and Technology, Lucknow, U.P., India

Conference Chair

Dr. Satyendra Sharma, HOD, ECE, G. L. Bajaj Institute of Technology and Management, Greater Noida, U.P., India

Organizing Secretary

Dr. Dinesh Kumar Singh, ECE, G. L. Bajaj Institute of Technology and Management, Greater Noida, U.P., India

Co-organizing Secretary

Dr. Amrita Rai, ECE, G. L. Bajaj Institute of Technology and Management, Greater Noida, U.P., India

Conveners

Dr. Piyush Yadav, ECE, G. L. Bajaj Institute of Technology and Management, Greater Noida, U.P., India
Dr. Shilpa Choudhary, ECE, G. L. Bajaj Institute of Technology and Management, Greater Noida, U.P., India

Preface

Innovation and research in a broad area of modern electronics and communication engineering needs to enhance its capabilities for recent challenges. Recent advances in different domains like sensor technologies, electronic devices, wireless communications, signal processing, computer vision medical image, data processing and globalization of digital society across various social areas have intensified the growth of sustainable electronics and communication systems. To effectively maintain a sustainable environment, it is constantly needed to meet the quality of service requirements of the emerging smart ICT applications. In this perspective, the book presents select proceedings of the International Conference on Modern Electronics Devices and Communication Systems (MedCom 2021) that brings together the state-of-the-art research works to propose novel architectures, models, and algorithms for solving the emerging challenges in almost all the areas of modern electronics devices and communication systems.

The book aims to bring teachers, researchers, scientists, engineers and scholars together, where they will exchange and share their experiences, new ideas and research results about all aspects of mentioned areas. This book will also present a strong fostering integration between modern electronics devices, communication and computational technologies. The global dependence in almost every sphere of human activity on computational analysis in modern era cannot be over emphasized. These trends are bound to increase with more and more new applications in the field of finance, marketing, health care, education, etc., which are coming under the gambit of computers.

Greater Noida, India — Rajeev Agrawal
Bengaluru, India — Chandramani Kishore Singh
Kingsville, USA — Ayush Goyal
Greater Noida, India — Dinesh Kumar Singh

Contents

About the Editors

Rajeev Agrawal is currently working as a professor and director at G. L. Bajaj Institute of Technology and Management, Greater Noida, India. Has an illustrated experience of more than 25 years in teaching and research, holds a B.E. degree in Electronics Engineering and M.Tech. degree in System Engineering. He received Ph.D. in the area of Wireless Communication Channels from the School of Computer and System Sciences, JNU, New Delhi. He was visiting professor at Kennesaw State University, Georgia, the USA under a joint research project in the area of Remote Patient Monitoring and Medical Imaging in the years 2010 and 2011. His research areas include planning and performance analysis of wireless networks and medical image analysis for automated diagnosis. He has more than 56 publications in international journals and proceedings. He has been awarded by various state and national agencies for his contribution to research and academics. He is also serving as a reviewer for several reputed international journals.

Chandramani Kishore Singh is currently an assistant professor with the department of electronic systems engineering, Indian Institute of Science, Bengaluru. He received M.E. and Ph.D. degrees in electrical communication engineering from the Indian Institute of Science, Bengaluru, India, in 2005 and 2012, respectively. He worked as a wireless communications engineer in ESQUBE Communications Solutions Pvt. Ltd., Bengaluru, from 2005 to 2006. He was a post-doctoral researcher with TREC, a joint research team between INRIA Rocquencourt and ENS Paris from 2012 to 2013, and with the Coordinated Science Laboratory, the University of Illinois at Urbana Champaign, USA, from 2013 to 2014. His research interests include communication networks, data centers, and smart grids. He has published more than 30 papers in respected international journals and conferences. Dr. Chandramani Kishore Singh has also received the Microsoft Research India Rising Star Award in 2011.

Ayush Goyal is an Assistant Professor in the Department of Electrical Engineering and Computer Science at Texas A&M University, Kingsville, Texas. He completed his B.S. in electrical engineering from Boise State University (BSU), Boise, USA,

and Ph.D. in Computer Science from Oxford University, Oxford, UK. He was a post-doctoral research fellow in the department of computer science, Tulane University, New Orleans, US. His areas of research interest are neurological disease prediction, brain MRI image analysis, and bioinformatics for connected healthcare. He has more than 40 publications in international journals and proceedings. He was awarded the Clarendon scholarship to pursue Ph.D. at Oxford University.

Dinesh Kumar Singh is an Associate Professor in the Department of Electronics and Communication Engineering, G. L. Bajaj Institute of Technology and Management. He completed his B.E. from Kumaon Engineering College, Almora, and M.E. from Rajeev Gandhi Proudyogiki Vishwavidyalaya, M.P. He got his Ph.D. degree from the Indian Institute of Technology Dhanbad. He has published several research papers in international journals and conferences. Dr. Singh's teaching and research areas are microwave engineering, antenna and wave propagation, and electromagnetic field theory. Dr. Singh is presently supervising 2 Ph.D. students.

A Review of State of Art Techniques for 3D Human Activity Recognition System

Bhavana Sharma and Jeebananda Panda

Abstract Recognizing human activities through video sequences and images is still a challenge due to background jumble, partial occlusion, changes in scale, viewpoint, lighting and appearance. A human activity classification technique has been comprehensively reviewed by the researchers. We have categorized human activity methodologies with object detection and feature extraction along with their subcategorization, advantages and restrictions. Moreover, we provide a comprehensive analysis of the existing, publicly available human activity datasets with applications and examine the prerequisites for an ideal human activity recognition dataset. At last, we present some open issues on human activity recognition and characteristics of future research directions.

Keywords Human activity recognition · Object detection · Feature extraction · Object classification · HAR datasets

1 Introduction

Population of elders is increasing with a rapid rate in most of the western countries and hence the challenges [1]. If we convert this in the form of percentage, by 2050, it will be 30% in Europe and China which is maximum globally and then 20.2% in United States of America (USA). This fact was established after a survey of WHO, that on an average, 28–35% of elderly people meet with an accident, because of falling, annually. According to WHO report, 37 million fall accidents are reported every year out of which 64.6 thousand people lose their life because of these accidents [2, 3]. In today's modern world of nuclear families, elderly people are living alone at their homes and they are more prone to meeting with such accidents while staying

B. Sharma (✉) · J. Panda
Department of Electronics and Communication Engineering, Delhi Technological University, New Delhi, India
e-mail: bhavanasharma.iec@gmail.com

J. Panda
e-mail: jpanda@dce.ac.in

R. Agrawal et al. (eds.), *Modern Electronics Devices and Communication Systems*, Lecture Notes in Electrical Engineering 948,
https://doi.org/10.1007/978-981-19-6383-4_1

at home. Hence, the study of fall detection, to improvise the science of detection, is more crucial and important [4–6]. In human activity recognition, to solve the existing and upcoming challenges toward activity recognition, researchers are using different techniques to beat the challenges of analysis. Based on the defined categories in different areas, methods and approaches may differ from one to another. Most commonly used cases of activity recognition is in medical and surveillance, where we talk about the devices and systems which are beneficial to humankind, in improving the life by mitigating the threats. By having a direct impact on saving lives, researchers are very keen to work in the field of video surveillance.

Human activity recognition through a comprehensive survey covers human activity recognition, i.e., 2D and 3D HAR based on RGB, depth and skeleton-based methodologies. The literature is updated with the application of recent advances field of human action recognition in Sect. 2. A structured arrangement of 2D and 3D object detection techniques has been discussed in Tables 1 and 2 highlighting different feature extraction techniques. Organization of our survey is as follows. Section 2 provides a panoramic summary of the related state-of-the-art survey works in the area of abnormal human action recognition followed by paper count analysis per year. It will help the reader to get an overview of key contributions of previous surveys done. Section 3 provides that human action recognition system is closely discussed with methodologies for detection, extractions and classification techniques evolved. Sections 4 and 5 outline recently introduced publicly available datasets used for activity recognition with challenges and applications. Finally, peculiar observations and possible directions are highlighted that need to further explore for research in the field of HAR.

2 Literature Survey

Shian-Ru Ke presented trends of HAR in video signals, and article explains the three different areas in activity recognition using core technology, human activity recognition systems and applications. This article throws light on application areas like surveillance, healthcare and entertainment industry where major focus is on surveillance in healthcare including its challenges [7]. Pau Climent-Pérez's article is based on HBA for ambient assisted living using AI. This study beautifully covers the estimation based on pose and gaze for movement identification. Later, it represents the latest work showcasing latest data tools and new datasets are described here [8]. Paul explained the techniques which are used in identification of human objects in surveillance video data with a benchmark datasets including directions for further research in living human identification and detection [9]. Fei Han represented an extensive survey of space time representations of human based on 3D skeletal data on categorization and analysis including modality, feature engineering, structure and transition including representation encoding [10]. Tej Singh explained key specifications of vision-based human activity recognition datasets which are discussed along with the algorithms according to the datasets best performance. Resolutions, actions/actors,

Table 1 Comparison of object detection methods

Techniques		Accuracy	Computational time	Advantages	Disadvantages
Background subtraction	Mixture of Gaussian model	Moderate	Moderate	Better response with Simple implementation and multi-modal scenarios	Not suitable for dynamic background and need to defined parameters
	Non-parametric background model	Moderate to high	Low to moderate	With significant post-processing, performs better in moving background	In occlusion, cannot performed
	Temporal differencing	High	Low to moderate	With sudden illumination changes, gives well performance in indoor environment	
	Warping background	High	Moderate to high	With high dynamic background, it is good in outdoor environment	Cannot work with occlusion
	Hierarchical background model	High	Low to moderate	Block-based and pixel-based approaches both are used and faster than pixel-based approach	Not good quality
Optical flow		Moderate	High	Good with dynamic camera and crowd detection	Highly computation intensive
Spatio-temporal filter		Moderate to high	Low to moderate	Perform good with low-resolution scenarios	More noise

frame rate, background and application domain are discussed in the paper [11]. Allah Bux described the image segmentation techniques and reviewed including challenges and future scope of research [12]. Athanasios Lentzas focused on the ABHAR for senior citizens. Analysis is done based on the taxonomy [13]. Michalis Vrigkas provided a comprehensive analysis of available datasets and examine the

Table 2 Comparison between different feature extraction methods

Techniques	Accuracy	Computational time	Advantages	Disadvantages
Shape-based method	Moderate	Low	With appropriate templates a simple pattern-matching approach is used	Not able to determine internal movements and in dynamic situations cannot performed
Motion-based method	Moderate	High	There is no need to predefined pattern templates	Cannot identify a non-moving human
Texture-based method	High	High	Good quality	More computation time

requirements for ideal datasets [14]. Tej Singh focused on the issue of benchmark datasets. Here article provide the comprehensive review to address this issue of benchmark datasets, action recognition-related RGB-D video datasets with 27 single-view datasets, 10 multi-view datasets, are provided [15], and various human activity recognition handcrafted and deep approaches are explained with 2D and 3D RGB and RGB_D dataset in this paper [16].

3 Methodology

See Fig. 1.

3.1 Input Image/Video

In the general process of recording the procedure of RGB cameras, a differentiation in the activity was created to analyze the sequence of actions, but at the same time, we have the challenges related to background clarity and lightning effect of the images which leads to the complexity while working on a design of the solution [17, 18]. Afterward, there was a regular improvement in the research methods to improve the factors like capturing of the depth of action in an optimal cost and real-time with the help of infrared radiation which provided a relief to the challenges related to lightning effects.

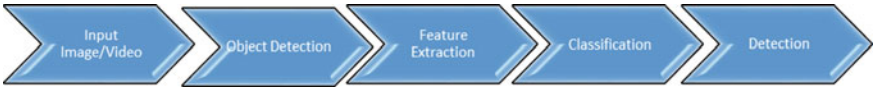

Fig. 1 Process of human activity detection

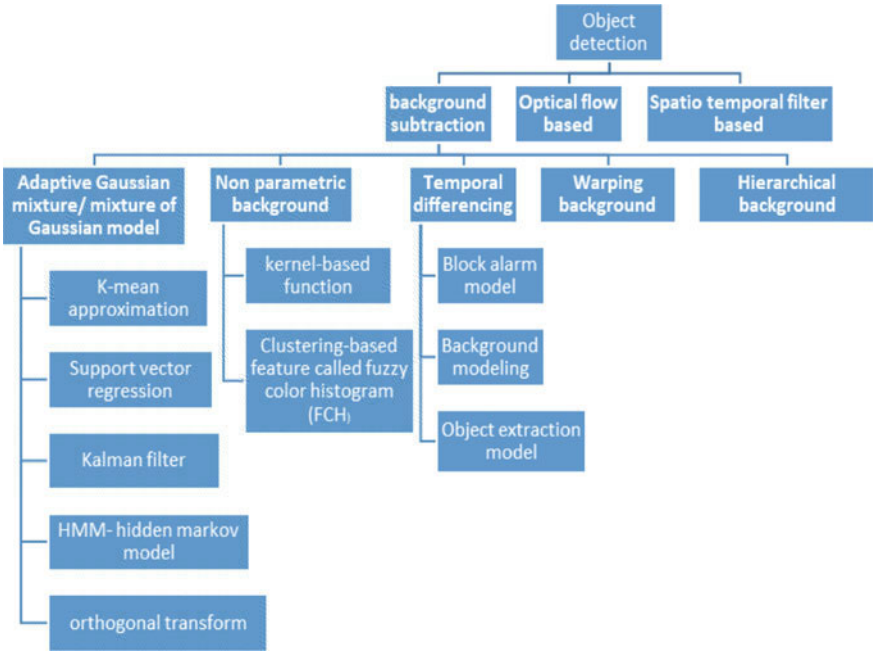

Fig. 2 Three types of human object detection

3.2 Object Detection

Background subtraction—In this techniques, a comparison of the moving object has been done based on the difference between current frames with the background frame. This comparison is done either pixel by pixel or block to block [19–21].

Optical flow based—Here in this technique, detection of moving object w.r.t. time based on the characteristic of flow vector has been used. There are challenges also related to lightening effect, motion sensitivity or noise which leads to high computational time [22].

Spatio-temporal filter based—This method is used to have reduced the computational requirement and noise by using data volume spanned by the moving object in a video signal [23]. This method is called 3D spatio-temporal because it works with spatial as well as time (Fig. 2).

3.3 Feature Extraction

This is a technique for the reduction of data dimension in which transforms lower dimension data into a modified featured space. Researcher selects a subset of features from the superset which will meet the forecasting requirements of target labels

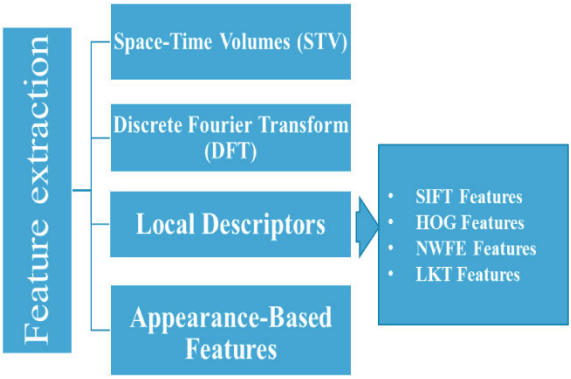

Fig. 3 Types of feature extraction

correctly diminishing the complexity of computation of different algorithm of learnings and predictions by subtracting the cost of remaining features left in the list [24]. Out of all methods, principal component analysis (PCA) gives the most reliable results in reduction of dimensions and extraction of attributes in case of linear structures [25, 26]. On the other hand, for nonlinear structure, linear discriminant analysis (LDA) technique is used to mitigate the challenges of PCA [27]. Linear discriminant analysis is used to separate the features with the aim of establishment of a linear transformation to attain the biggest class discrimination. The traditional LDA is used to find out a standard discriminant subspace (Fig. 3).

3.4 Object Classification

Shape-based method—The shape data of moving regions such as points, boxes and blobs is described foremost and after that deemed as a standard pattern recognition. While using the aforementioned approach, the large number of possible impressions of the body creates chaos between a moving human and other moving objects [28, 29]. An enormous challenge with this method is that it cannot apprehend the internal motion of the object in the contour area.

Motion-based method—In this method, we can overcome with the confusion between a moving human and other moving objects by using object motion characteristics with patterns analysis means to identify people in other moving objects, it uses the periodic property of captured images [30–32].

Texture-based method—Texture-based methods use intensity patterns for nearby pixels. This technique counts the gradient directions of local area of image and does calculations on a dense grid of evenly spaced cells. For better accuracy, it uses overlapping local contrast normalization.

4 Datasets

There are some important datasets.

KTH human motion dataset—This dataset contains six human actions performed by 25 subjects in four different situations. Running, jogging, walking, boxing, hand waving or clapping are performed in more than 2000 sequences. The backgrounds are homogeneous and uncluttered. Video files are classified by operation, to eliminate unnecessary operations easily [13].

Weizmann human action dataset—It uses static front-side cameras to record individual human movements from 10 subjects in different environments. Approximately 340 MB of video sequences are available. The actions performed include bending, walking, running, hand waving and different types of jumping.

HOHA—(Hollywood human actions)—HOHA dataset contains video sequences from 32 movies with annotations for eight action types: AnswerPhone, GetOutCar, HandShake, HugPerson, Kiss, SitDown, SitUp and Stand.

INRIA Xmas motion acquisition sequences—The video images of 390 $\times$ 291 pixel which is recorded from five different angles are included in these sequences. 11 actors perform 13 actions: check watch, cross arms, scratch head, sit down, get up, turn around, walk, wave, punch, kick, point, pick up, throw overhead and throw from bottom to top.

TUM kitchen dataset—This dataset aims ADLs in a kitchen scene with a low level of action. Multiple subjects perform tablet setting in different ways; transporting items one by one; and other behaviors are natural, grabbing multiple objects at once.

5 Challenges in HAR Dataset

In this section, we discuss the various current challenges in the dataset.

Background and environment conditions—In videos if there is moving object or background, it is very difficult to recognize human activity. There are so many types of background in a video signal like slow and fast, dynamic and static, airy and rainy, and crowded. Same recognition activity in environment conditions which contains various issues like rain, waves, trees and water is affected.

Similarity and Difference of actions—There are many actions that looks same in the videos like running, jogging, walking, etc. The same type of procedure affects classification accuracy. Similarity between classes of actions in datasets provides a fundamental challenge.

Occlusion—Occlusion occurs when an object is hiding the another object. The occlusion can be classified into two parts one is self-Occlusion and another one is partial occlusion. Occlusion is a greater challenge in computer vision applications such as human posture, object tracking and video monitoring.

View variations—In human identification system, any action recorded inside the video is the most crucial characteristic. Multiple views have larger facts

comparetively single view which leads to fair analysis of captured perspective in dataset.

6 Conclusion and Future Work

The literature survey encloses a wide area around HAR covering different methodologies and techniques of identification, detection and limitations along with its pros and cons. It also throws light on dataset benchmark and its quality which leads to the variation in results. Numerous HAR dataset challenges discussed.

In future, HAR systems need to address specific issues related to the quality of dataset and connect it to the real-life application development. In future, researcher will need to work on the challenges relating to noise, input quality data and various process-related challenges. Some meaningful datasets to represent abnormal actions in different scenarios are still a problem. Working on deep architectures from primary CNN to RCNN, RNN, auto-encoders can be extended to enhance the parameters of recognition systems.

References

1. Merrouche F, Baha, N (2016) Depth camera based fall detection using human shape and movement. In: IEEE international Conference on signal and image processing
2. Ma X, Wang H, Xue B, Zhou M, Ji B, Li Y (2014) Depth-based human fall detection via shape features and improved extreme learning machine. IEEE J Biomed Health Inf
3. Bian Z-P, Chau L-P, Magnenat-Thalmann N (2014) Fall detection based on body part tracking using a depth camera. IEEE J Biomed Health Inf
4. Lentzas A, Vrakas D (2019) Non-intrusive human activity recognition and abnormal behavior detection on elderly people: a review. Springer Nature B.V
5. Pham C, Nguyen-Thai S, Tran-Quang H, Tran S, Vu H, Tran T-H, Le T-L (2020) SensCapsNet: deep neural network for non-obtrusive sensing based human activity recognition. IEEE Access
6. Popoola OP, Wang K (2012) Video-based abnormal human behavior recognition—a review. IEEE Trans Syst Man Cybern C: Appl Rev
7. Ke S-R, Thuc HLU, Lee Y-J, Hwang J-N, Yoo J-H, Choi K-H (2013) A review on video-based human activity. Recognition 2:88–131. https://doi.org/10.3390/computers2020088
8. Chaaraoui AA, Climent-Pérez P, Flórez-Revuelta F (2012) A review on vision techniques applied to human behaviour analysis for ambient-assisted living. Elsevier
9. Paul M, Haque SME, Chakraborty S (2013) Human detection in surveillance videos and it applications—a review. EURASIP J Adv Signal Process
10. Han F, Reily B, Hoff W, Zhang H (2016) Space-time representation of people based on 3D skeletal data: a review. Elsevier
11. Dhiman C, Vishwakarma DK (2019) A review of state-of-the-art techniques for abnormal human activity recognition. In: Engineering Applications of Artificial Intelligence Elsevier, pp 21–45
12. Dhiman C, Vishwakarma DK (2020) View-invariant deep architecture for human action recognition using two-stream motion and shape temporal dynamics. IEEE Trans Image Process
13. Singh T, Vishwakarma DK (2019) Human activity recognition in video benchmarks: a survey. Springer Nature Singapore

14. Jankowski S, Szymański Z, Mazurek P, Wagner J (2015) Neural network classifier for fall detection improved by Gram-Schmidt variable selection. In: The 8th IEEE international conference on intelligent data acquisition and advanced computing systems
15. Brun L, Percannella G, Saggese A, Vento M IAPR Fellow (2017) Action recognition by using kernels on aclets sequences. Elsevier
16. Jing C, Wei P, Sun H, Zheng N (2019) Spatiotemporal neural networks for action recognition based on joint Loss. Springer-Verlag, London Ltd., part of Springer Nature
17. Thien Huynh- Cam-Hao Hua, Nguyen Anh Tu , Taeho Hur , Jaehun Bang , Dohyeong Kim , Muhammad Bilal Amin , Byeong Ho Kang , Hyonwoo Seung , Soo-Yong Shin , Eun-Soo Kim , Sungyoung Lee (2018) "Hierarchical topic modeling with pose-transition feature for action recognition using 3D skeleton data", Elsevier
18. Sarakon S, Tamee K (2020) An individual model for human activity recognition using transfer deep learning. In: Joint international conference on digital arts
19. Cai X, Zhou W, Wu L, Luo J, Li H (2016) Effective active skeleton representation for low latency human action recognition. IEEE Trans Multimedia 18(2)
20. Suto J, Oniga S, Lung C, Orha I (2018) Comparison of offline and real-time human activity recognition results using machine learning techniques. Springer
21. Dhiman C, Vishwakarma DK (2019) A robust framework for abnormal human action recognition using R-transform and Zernike moments in depth videos. IEEE Sens J
22. Ladjailia A, Bouchrika I, Merouani HF, Harrati N, Mahfouf Z (2019) Human activity recognition via optical flow: decomposing activities into basic actions. Springer-Verlag London Ltd., part of Springer Nature
23. Ji X, Cheng J, Feng W, Tao D (2017) Skeleton embedded motion body partition for human action recognition using depth sequences. Elsevier
24. Vishwakarma DK, Rawat P, Kapoor R (2015) Human activity recognition using Gabor wavelet transform and Ridgelet transform. In: 3rd international conference on recent trends in computing—ICRTC
25. Lahiri D, Dhiman C, Vishwakarma DK (2017) Abnormal human action recognition using average energy images. In: Conference on information and communication technology
26. Abdull Sukor AS, Zakaria A, Abdul Rahim N (2018) Activity recognition using accelerometer sensor and machine learning classifiers. In: 2018 IEEE 14th international colloquium on signal processing & its applications (CSPA 2018), Penang, Malaysia, 9–10 March [2018]
27. Tao D, Jin L, Yuan Y, Xue Y (2016) Ensemble manifold rank preserving for acceleration-based human activity recognition. IEEE Trans Neur Netw Learn Syst
28. Akagündüz E, Aslan M, Şengür A, Wang H, İnce MC (2015) Silhouette orientation volumes for efficient fall detection in depth videos. IEEE J Biomed Health Inf
29. Mazurek P, Morawski RZ (2015) Application of Naïve Bayes classifier in fall detection systems based on infrared depth sensors. In: The 8th IEEE international conference on intelligent data acquisition and advanced computing systems
30. Wagner J, Morawski RZ (2015) Applicability of mel-cepstrum in a fall detection system based on infrared depth sensors. In: The 8th IEEE international conference on intelligent data acquisition and advanced computing systems
31. Jankowski S, Szymański Z, Dziomin U, Mazurek P, Wagner J (2015) Deep learning classifier for fall detection based on IR distance sensor data. In: The 8th IEEE international conference on intelligent data acquisition and advanced computing systems
32. Zhang H, Parker LE (2011) 4-dimensional local spatio-temporal features for human activity recognition. In: IEEE international conference on intelligent robots and systems, San Francisco

Comparative Analysis for Email Spam Detection Using Machine Learning Algorithms

Gayatri Gattani, Shamla Mantri, and Seema Nayak

Abstract The prominence of the use of communication over the Internet is increasing progressively. Being economical, faster, and easy user interface, the number of email users is increasing tremendously. These led to the gradually increasing activity of spam. Spam emails are unrequested and unimportant emails in bulk. Due to this, there arise major Internet and email security issues that also include a problem of electronic storing space and waste of time. Thus, the identification of spam emails is very necessary. In this paper, four supervised machine learning algorithms, which are Naïve Bayes, support vector machine (SVM), logistic regression, and random forest classifier, are proposed for spam and ham emails classification. Experiments using these four algorithms are performed on prepared feature sets on two different datasets to select the best model with the highest accuracy and less overfitting or underfitting for spam detection. To automate the workflow of building the model and its evaluation, a machine learning pipeline is used in this project. Experimental results show that the overall accuracy of the random forest classifier model is the highest and also has less complexity.

Keywords Emails · Spam email · Spam email identification · Machine learning · Naïve Bayes · Support vector machine · Logistic regression · Random forest classifier

G. Gattani (✉) · S. Mantri
School of Computer Engineering and Technology, Dr. Vishwanath Karad MIT WPU, Pune, India
e-mail: gayatrigattani2001@gmail.com

S. Mantri
e-mail: shamla.mantri@mitcoe.edu.in

S. Nayak
Department of Electronics and Communication Engineering, IIMT College of Engineering, Greater Noida, India

R. Agrawal et al. (eds.), *Modern Electronics Devices and Communication Systems*,
Lecture Notes in Electrical Engineering 948,
https://doi.org/10.1007/978-981-19-6383-4_2

1 Introduction

In recent decades, the use of technology and the Internet has reached to peak. Being fast, cheap, and accessible, the extension of the use of email has increased tremendously. This resulted in a dramatic increase in spam emails [1]. These emails are junk emails that are almost identical and sent to multiple recipients randomly [2]. The changing way of communication by Internet on a very large scale has led to the expansion of new communication services, such as email [3]. According to a recent study, over 4 billion of the population use email. Due to its simplicity and accessibility, the mark of people using email is increasing day by day. It is extremely fast and cost-effective. With the escalation in the broadening of emails, there is also a rise in spam emails, and the unnecessary and undesirable bulk mails sent to several users haphazardly. Spam mails not only cause the problem of electronic storing space but also are the carrier of malware and hoard the network bandwidth, space, and computational power [4]. A study estimates that approximate measure of spam emails is 85%.

While the number of spam emails increases, the certainty of a user not reading a non-spam email increases. Due to the loss of network bandwidth and time consumed by users to demarcate between normal and spam [5], various spam filtering techniques have been introduced. These techniques can be categorized based on the use and non-use of machine learning algorithms. The use of ML algorithms provides an automated approach where the model trains itself based on features extracted from the dataset. As easy to implement and short training time, Naïve Bayes is a popular spam filter [6]. The main objective is to collate the accuracy of four major classification systems that include SVM, random forest classifier, Naïve Bayes, and logistic regression and select the best model for spam detection.

2 Literature Survey

Spam: Unnecessary emails sent by unknown people randomly in bulk are spam mails. These spam mails are vulnerable to major user security and also cause the problem for electronic storing space. The following are the major spam categories (Table 1).

Table 1 Frequency of major scam categories and danger level caused by them

Categories	Frequency of receiving	Danger
Ads	High	Moderate
Chain letter	Low	High
Email spoofing	Low	High
Hoaxes	Moderate	Moderate
Money scams	Moderate	High

Table 2 Some previously used techniques in spam filtering and their accuracy [9–12]

Year	Author	Classification technique	Dataset	Highest accuracy
2008	Bo Yu, Zong-ben Xu	Naive Bayes, NN, SVM, RVM	SpamAssassin & Babble Text	SVM: 95.2% and 96.0%
2011	W. A. Awad & S. M. ELseuofi	Naïve Bayes, KNN, ANN, SVM, artificial immune	SpamAssassin spam corpus	NB: 99.46%
2013	Sumant Sharma & Amit Arora	ML techniques provided by WEKA tool	SPAMBASE	94.28%
2014	Andronicus A. Akinyelu & Aderemi O. Adewumi	Classification of phishing email Using random forest ML technique	2000 phishing and ham emails set	99.7%
2017	A. S. Yuksel, S. F.Cankya, & I. S. Uncu	Cloud-based approach combining predictive analysis and ML techniques (SVM and decision tree)	SpamAssassin	SVM: 97.6%
2018	Deepika Mallampati	Naïve Bayes, J48, MLP	–	MLP: 99.3%
2021	Manoj Sethi, Sumesha Chandra, Vinayak Chaudhary & Yash	NB-multinomial, logistic regression, SVM, NN	SpamAssassin	NN: 99.02%

Spam classification: Email systems without spam classification techniques are highly open to risks. The dangers open to email systems without spam filtering are spyware, phishing, ransomware [7]. Thus, the classification of such messages can be seen as another defense mechanism against such dangers. In the previous years, various techniques of spam identification have been developed. Domain name server blacklist (DNSBL) and white list, high-volume spammers (HVSs) and low-volume spammers (LVSs) classification, machine learning-based Web spam classification, support vector machine classifier model, TruSMS systems, cloud-based approach [3], and ML algorithms like Naïve Bayes, random forest classifier, neural networks [8] are some of the classification techniques developed by researchers earlier (Table 2).

2.1 Existing Approaches

Global email users are increasing day by day. In 2024, it is set to grow up to 4.48 billion [13]. As the use of email increases, spam increases too. This causes a decrement in productivity since manually spam filtering is time-consuming, and also the electronic

storing space is reduced. Spam also increases the cyber threat to users through various phishing and malware attacks. Not only this, it has been discovered that on yearly basis, spam is accountable for over 77% of whole global email traffic [13].

In today, two common approaches, namely knowledge engineering and machine learning, are used for spam filtering. A collection of rules in knowledge engineering are used to identify mails as ham or spam. This method can lead to large time wastage and also does not guarantee the results as there is a continued need for an update in the specified set of rules. Thus, it is mainly used by naïve users [14].

Machine learning is completely based on the datasets. It just needs the training datasets, and the algorithm used itself learns the classification rules from the set of training samples from datasets. Thus, machine learning is proved to be more effective than knowledge engineering [14]. Examples of machine algorithms used for the classification of spam include Naïve Bayes, support vector machine, artificial immune systems, neural networks, logistic regression, deep learning, and many more.

The best possible outcome for any algorithm can be checked using various evaluation techniques in machine learning. This evaluation technique also helps in recognizing the overfitting and underfitting of the model. Cross-validation score, F1-score, confusion matrix, precision, recall, accuracy, regression metrics, and mean squared error can be used for evaluating the model. The three major metrics to weigh up a classification model are accuracy, precision, and recall [15].

Three major methods that are reliable for present spam detection systems are linguistic-based (used in places like a search engine), behavior-based (user-dependent since the need of change in rules from time to time), and graph-based (detect abnormal forms in data showing the behavior of spammers) [16].

3 Proposed Method

3.1 Proposed Algorithm and Workflow

Two datasets are used for this experiment to select the best algorithm with the highest accuracy. Dataset 1 is taken from Kaggle SMS Spam Collection [17]. This dataset contains 5574 messages tagged ham or spam. Dataset 2 is taken from the collection of emails from_Apache SpamAssassin's public datasets_ available on Kaggle as spam or not spam dataset [18]. There are 2500 non-spam and 500 spam emails in this dataset. The experiment is performed using four simple machine learning classification algorithms that are logistic regression, support vector machine (SVM), random forest classifier, and logistic regression on a prepared feature set of two datasets.

Through evaluation using confusion matrix, evaluation metrics, k cross-validation score, and accuracy, the perfect model with the highest accuracy and reduced underfitting or overfitting is selected. Selection of parameter k in k cross-validation score and splitting ratio of datasets play an efficient contribution in assessing the accuracy

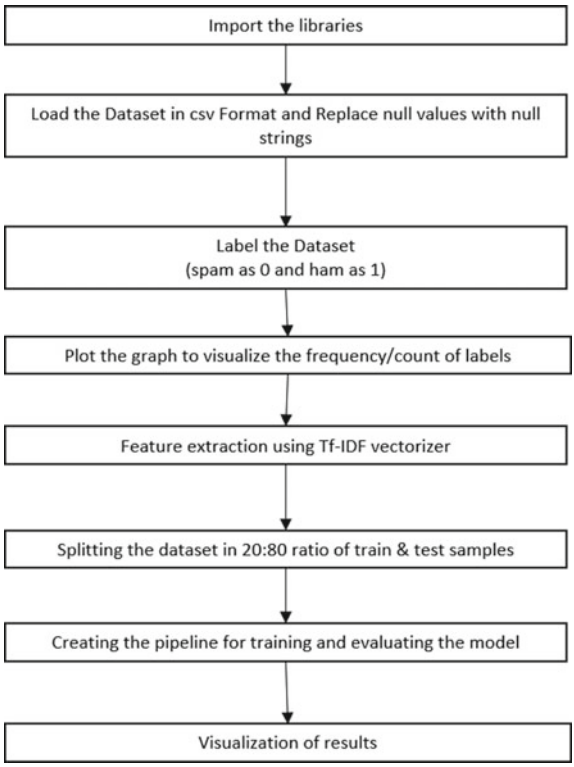

Fig. 1 Flowchart of model

of the model. The accuracy and overfitting/underfitting results are visualized using a heat map (Fig. 1).

Data Preprocessing

Both the datasets are taken from Kaggle [17, 18]. Dataset 1 comprises 5574 messages tagged according to being ham or spam. Here, we need to label the spam messages as 0 and ham messages as 1 for further simplicity. Dataset 2 comprises a collection of 3000 emails taken from "_Apache SpamAssassin's public datasets_". There are 2500 non-spam emails and 500 spam emails in this dataset. Here, the dataset initially contains the labeled data, that is, spam mails as 1 and ham emails as 0. All the null values in both datasets are converted to null strings for the normalization of plain text (Fig. 2).

Feature Extraction

The feature set will be prepared using term frequency-inverse document frequency (TF-IDF) vectorizer by transforming the feature text into feature vectors and converting it to lowercase. Parameter min_df is set to 1 that means to ignore the terms that appear in less than one document. The terms that appear irregular, min_df is used to remove them [19]. The next parameter, stopwords, is set to English to return

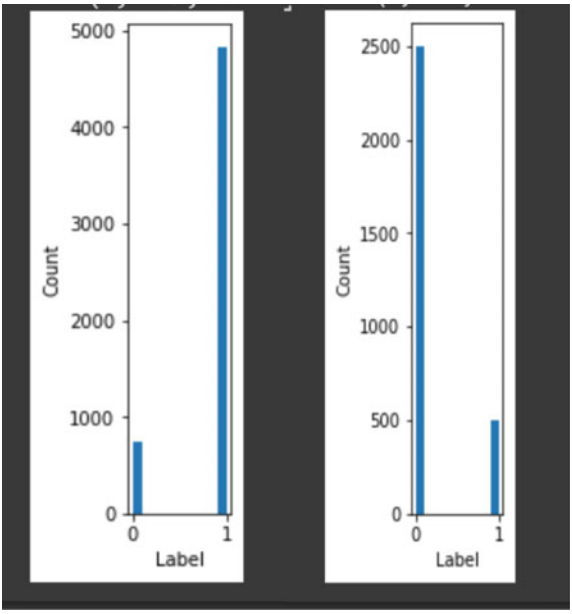

Fig. 2 Visualization of labeled dataset 1 (left) and dataset 2 (right)

the relevant stop list. The parameter lowercase is set to true to convert all characters to lowercase.

Pipeline

To automate the workflow of producing a machine learning model and evaluation of spam detection using different algorithms, a pipeline is created. The different algorithms used in this experiment are as follows:

Logistic Regression

A supervised machine learning algorithm is used for solving classification problems. It is a simple yet very effective algorithm for binary classification. The basis of this algorithm is the logistic function (sigmoid function), which takes any real-valued number and maps it in the value between 0 and 1 [20].

$$\text{Logistic Function: } y = 1/(1 + e^{-x})$$
$$\text{i.e., } 1 + e^{-x} = 1/y$$
$$e^{-x} = (1-y)/y$$
$$e^{x} = y/(1-y)$$
$$x = \log(y/(1-y))$$