

**2022 IEEE Eighth International
Conference on Big Data
Computing Service and
Applications
(BigDataService 2022)**

**Newark, California, USA
15 – 18 August 2022**



**IEEE Catalog Number: CFP22A91-POD
ISBN: 978-1-6654-5891-7**

**Copyright © 2022 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP22A91-POD
ISBN (Print-On-Demand):	978-1-6654-5891-7
ISBN (Online):	978-1-6654-5890-0

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2022 8th IEEE International Conference on Big Data Computing Service and Machine Learning Applications (BigDataService) **BigDataService 2022**

Table of Contents

Welcome from the IEEE BDS 2022 Workshop Chair	x
Message from the IEEE BDS 2022 Technical Program Chairs	xi
IEEE BDS 2022 Organisers	xii

BigDataService Papers

HealthLies: Dataset and Machine Learning Models for Detecting Fake Health News	1
<i>Garima Chaphekar (San Jose State University, USA) and Jorjeta Jetcheva (San Jose State University, USA)</i>	
A Domain Generalization Approach for Out-of-Distribution 12-lead ECG Classification with Convolutional Neural Networks	9
<i>Aristotelis Ballas (Harokopio University of Athens, Greece) and Christos Diou (Harokopio University of Athens, Greece)</i>	
Allocating Resource Capacities for an Offload-enabled Mobile Edge Cloud System	14
<i>Zhiyan Chen (University of Warwick, United Kingdom) and Ligang He (University of Warwick, United Kingdom)</i>	
A Novel Method for Enhancing the Accuracy of Box Detection under Noise Effect of Tags and Complex Arrangement of Pile with Cycle-GAN and Mask-RCNN	22
<i>Thong Nguyen, Seong-Je Kim, Hyung-gyu Kim, Jooyeop Han, and Jonghun Yoon</i>	
Reversible Data Hiding Scheme Based on Image Partitioning and Histogram Shifting	27
<i>Jie Hao (Hohai University, China), Ping Ping (Hohai University, China), Xia Peng (Hunan International Economics University, China), and Zeyu Gao (San Jose State University USA)</i>	
An Expansion on Prioritized Experience Replay with Round Robin Scheduling	35
<i>Angel Solis (University of Colorado, Colorado Springs, USA)</i>	

A Big Data Approach for Fuel Oil Consumption Estimation in the maritime industry	39
<i>Dimitrios Kaklis (Harokopio University of Athens, Danaos Research Center, NCSR Demokritos - Greece), Pavlos Eirinakis (University of Piraeus), George Giannakopoulos (NCSR Demokritos, Greece), Constantine Spyropoulos (NCSR Demokritos, Greece), Takis Varelas (Danaos Research Center, Greece), and Iraklis Varlamis (Harokopio University of Athens, Greece)</i>	
Assessing the Impact of Movie Plot Summaries on Box Office Sales	48
<i>Andreea Pocol (University of Waterloo, Canada) and Lesley Istead (Carleton University, Canada)</i>	
SMARTREC - A SMART Conversational Recommendation System using Semantic Knowledge Graphs	53
<i>Sudha Vijayakumar (TigerGraph, Inc.) and Jorjeta G. Jetcheva (San Jose State University, USA)</i>	
A Parallel Community Detection Algorithm Based on Spanning Trees	61
<i>Petros Potikas (National Technical University of Athens, Greece), Dora Souliou (National Technical University of Athens, Greece), Aris Pagourtzis (National Technical University of Athens, Greece), and Katerina Potika (San Jose State University, USA)</i>	
An Artificial Intelligence Outlook for Colorectal Cancer Screening	66
<i>Panagiotis Katrakazas (Zelus P.C. Athens, Greece), Aristotelis Ballas (Harokopio University of Athens, Greece), Marco Anisetti (Università degli Studi di Milano, Italy), and Ilias Spais (Zelus P.C. Athens, Greece)</i>	
Real-Time Detection of Objects on Roads for Autonomous Vehicles Using Deep Learning	73
<i>Max Ng (San José State University, USA), Darshit Jagetiya (San José State University, USA), Xing Gao (San José State University, USA), Haoran Shi (San José State University, USA), Jerry Gao (San José State University, USA), and Jia Liu (Jilin Engineering Normal University, China)</i>	
Spike2Signal: Classifying Coronavirus Spike Sequences with Deep Learning	81
<i>Sarwan Ali (Georgia State University, USA), Taslim Murad (Georgia State University, USA), Prakash Chourasia (Georgia State University, USA), and Murray Patterson (Georgia State University, Atlanta, USA)</i>	
On Building Real Time Intelligent Agricultural Commodity Trading Models	89
<i>Jiayu Zhou (San Jose State University, USA), Jing Ye (San Jose State University, USA), Yue Ouyang (San Jose State University, USA), Manyuan Tong (San Jose State University, USA), Xuedong Pan (San Jose State University, USA), and Jerry Gao (San Jose State University, USA)</i>	
Active Learning for Node Classification using a Convex Optimization approach	96
<i>Deepesh Agarwal (Kansas State University Manhattan, USA) and Balasubramian Natarajan (Kansas State University Manhattan, USA)</i>	
SoBigDemicSys: A Social Media based Monitoring System for Emerging Pandemics with Big Data.	103
<i>Tung Khanh Tran (FPT Software Company Limited, Vietnam), Xuan-Son Vu (Umea University, Sweden), and Lili Jiang (Umea University, Sweden)</i>	

MAP: Design, Development, Deployment, and Maintenance of Industrie 4.0 AI Applications	108
<i>Aldo Dagnino (GBS IS Innovation and Emerging Services, ABB Inc., USA), Maciej Kolomycki (GBS IS Innovation and Emerging Services, ABB Inc., USA), and Alok Kucheria (GBS IS Innovation and Emerging Services, ABB Inc., USA)</i>	
Small Data Deep Learning for Lung Cancer Detection in CT	114
<i>Kenji Suzuki (Biomedical Artificial Intelligence Research Unit (BMAI), Institute of Innovative Research Tokyo Institute of Technology, Japan)</i>	
Image-Based Mapping of Cells onto a Cell Cycle Continuum for Integration of Single Cell Sequencing and Live Cell Imaging	119
<i>Andrew Schultz (Moffitt Cancer Center, USA), Saeed Alahmari (Najran University, Najran, Saudi Arabia), Joeseeph Johnson (Moffitt Cancer Center, USA), Thomas Veith (Moffitt Cancer Center, USA), Zaid Siddiqui (UPMC, USA), Issam El Naqa (Moffitt Cancer Center, USA), Laura Heiser (OHSU, USA), and Noemi Andor (Moffitt Cancer Center, USA)</i>	
Distribution Grid Topology Estimation: A New Approach-Based on Bayesian Network Models	124
<i>Ahmed Mabrouk (CSAI Lab CRIGEN, ENGIE Group, France) and Ram Rajagopal (Stanford University, USA)</i>	

BigDataService Workshop Papers

A Non-Bottleneck Residual Approach for QR Code Printed Source Identification	132
<i>Min-Jen Tsai (Institute of Information Management, National Yang Ming Chiao Tung University, Taiwan, R.O.C) and Te-Ming Chen (Institute of Information Management, National Yang Ming Chiao Tung University, Taiwan, R.O.C)</i>	
CACRM: Cross-Attention Based Image-Text CrossModal Retrieval	137
<i>Huimin Yu (Hohai University, China), Jinlin Suo (Beijing Goldenwater Information Technology Co., Ltd, China), Yingchi Mao (Hohai University, China; Key Laboratory of Water Big Data Technology of Ministry of Water Resources (Hohai University), China), Jiahong Sui (Hohai University, China), and Yong Qian (Hohai University, China)</i>	
A Federated Learning Backdoor Attack Defense Method Based on Dual Attention Mechanism	143
<i>Yan Jin (College of Computer and Information, Hohai University Nanjing, China), Yingchi Mao (College of Computer and Information, Hohai University Nanjing, China), Hua Nie (College of Computer and Information, Hohai University Nanjing, China), Zijian Tu (College of Computer and Information, Hohai University Nanjing, China), and Jianxin Huang (College of Computer and Information, Hohai University Nanjing, China)</i>	
IEMAM: Image-Feature Extraction Based on Multi-scale Attention Mechanism	149
<i>Zheyuan Wang (College of Computer and Information Hohai University, China), Yingchi Mao (College of Computer and Information Hohai University, China), Shuai Zhang (College of Computer and Information Hohai University, China), Yong Qian (College of Computer and Information Hohai University, China), Zeyu Zhang (College of Computer and Information Hohai University, China), and Zhihao Chen (College of Computer and Information Hohai University, China)</i>	

Comparing Human and Algorithmic Anomaly Detection for HVAC Systems Applications	155
<i>Reuben Borrison (ABB AG, Germany), Marc Syndicus (RWTH Aachen University, Germany), Andre Orth (KIT Karlsruhe Institute of Technology, Germany), Romana Markovic (KIT Karlsruhe Institute of Technology, Germany), Marcel Dix (ABB AG, Germany), Antonio Liguori (RWTH Aachen University, Germany), Matthias Berning (ABB AG, Germany), Andreas Wagner (KIT Karlsruhe Institute of Technology, Germany), and Christoph van Treeck (RWTH Aachen University, Germany)</i>	
Convolutional Kernel-Based Transformation and Clustering of Similar Industrial Alarm Floods	161
<i>Gianluca Manca (Institute of Automation Technology, Helmut-Schmidt-University, Germany), Marcel Dix (Industrial Data Analytics, ABB Corporate Research Center, Germany), and Alexander Fay (Institute of Automation Technology, Helmut-Schmidt-University, Germany)</i>	
Using Siamese Neural Networks on Threshold Maps of Infrared Images to Detect Equipment Faults	167
<i>Ralf Gitzel (DECRC, ABB AG, Germany), Holger Kaul (DECRC, ABB AG, Germany), and Marcel Dix (DECRC, ABB AG, Germany)</i>	

BigDataService Poster Papers

Streamflow Prediction Using a Hybrid Methodology Based on Convolutional Neural Network and Long Short-Term Memory	173
<i>Juan Ramirez Rochac (University of the District of Columbia Washington, D.C., USA), Nian Zhang (University of the District of Columbia Washington, D.C., USA), Tolessa Deksissa (Water Resources Research Institute University of the District of Columbia Washington, D.C., USA), and Wagdy Mahmoud (University of the District of Columbia Washington, D.C., USA)</i>	
Fine-Tuning Language Models to Mitigate Gender Bias in Sentence Encoders	175
<i>Tommaso Dolci (Politecnico di Milano – DEIB Milan, Italy)</i>	
Vision: Delphi Study to Identify Criteria for the Systematic Assessment of Data Protection Risks in the Context of Big Data Analytics	177
<i>Georgios Georgiadis (Ghent University, Belgium) and Geert Poels (Ghent University, Belgium)</i>	
Network Flow Detection of Semantic Relationship between Flow and Byte	179
<i>Yiqing Luo (Beijing University of Posts and Telecommunications, China), Mingshu He (Beijing University of Posts and Telecommunications, China), Xiaojuan Wang (Beijing University of Posts and Telecommunications, China), and Lei Jin (Beijing University of Posts and Telecommunications, China)</i>	
An Energy-Efficient Ensemble-Based Computational Social System for Fake News Detection in MANET Messaging	181
<i>Amit Neil Ramkissoon (The University of the West Indies at St Augustine, Trinidad & Tobago) and Wayne Goodridge (The University of the West Indies at St Augustine, Trinidad & Tobago)</i>	

Enhanced Algorithmic Job Matching Based on a Comprehensive Candidate Profile using Machine Learning 183
Vishnu S. Pendyala (San Jose State University, USA), Nishtha Atrey (IBM, USA), Tina Aggarwal (Amazon, USA), and Saumya Goyal (Playstation, USA)

Author Index 185