Proceedings Of

3rd IEEE International Conference on AI in Cybersecurity (ICAIC)

07-09 February 2024

University of Houston, Houston, USA

Proceeding Editors:

Hardik Gohel, Bishwajeet Pandey

About ICAIC'2024

3rd International Conference on AI in Cybersecurity (ICAIC) to be held in Hybrid mode on 07-09 February 2024 (ONLINE for those participants, who will not reach USA, offline for those participants, who will reach USA. ICAIC'2024 intended to attract innovative technical and scientific work in the field of computer science and electronics engineering. The response to the conference was overwhelming and we are proud to state that we have received really good quality contributions and we are sure as either online or offline participant you will share the same sentiment. All accepted papers will be submitted to IEEE Explore and hopefully these papers will be available online by mid of 2024.

As a chair and on behalf of the organizing committee, we are extremely happy to host you in the USA. And as a participant, you shall be able to visit the USA from different parts of the world to share and you will contribute in the areas of your expertise. We hope to provide a good hybrid platform to the participants of ICAIC'2024 where not only they meet and share their vision, ideas but also fertilize their thoughts in the ever-growing area of computer science and electronics engineering technologies. We are also confident that our keynote speakers will be able to enrich your knowledge during the conference and we wish you a very safe stay at your home country if you are not coming to the USA and happy Journey if you are coming to the USA. It is the 29th conference hosted by Gyancity Research Consultancy in association with partner university across the globes, next three conferences in 2024 are following:

4th International Conference On Business, Management, Emerging technologies, and Social Science 2024 (BMESS), 25-26 April 2024 https://bmess.gyancity.com/
& 7th International Multi-Topic Conference on Engineering and Science (IMCES), 25-26 April 2024 https://imces.tech
Bath Spa University, Academic Center Ras Al-Khaimah-UAE

9th International Conference on Green Computing and Engineering Technologies (ICGCET)
27 - 28 September 2024, Zanzibar, https://icgcet.org/

Best wishes.

Dr Bishwajeet Pandey, Gyancity Research Consultancy Pvt Ltd, India

Dr Hardik Gohel, University of Houston-Victoria, Texas, USA

Mobile/Whatsapp: +91-742-864-0820, +1-361-570-4219

Email: dr.pandey@ieee.org, gohelh@uhv.edu

ICAIC-2024 Schedule

07 February 2024

10:00AM-10:15AM (USA-Texas Time)
Welcome Speech: General Chair Prof Hardik Gohel, University of Houston-Victoria, USA

10:15AM-10:30AM (USA-Texas Time)
Inaugural Speech: Rector, University of Houston, USA

10:30AM-10:11.00 AM (USA-Texas Time)
First Keynote by Dr Leila Rzayeva, Astana IT University, Astana, Kazakhstan

Physical Session 1 @ University of Houston, USA
11:00-13:00 PM (USA-Texas Time)
Session Chair: Dr Suresh Kumar Peddoju, University of Houston, USA
Paper Id: 322, 1315, 1976, 2439, 4287

13.00-13:30 (USA-Texas Time)
Second Keynote by Dr Ciro Rodriguez, National University of San Marcos, Lima, Peru

13.30-14.30: LUNCH

Physical Session 2 @ University of Houston, USA
14:30-16:30 PM (USA-Texas Time)
Session Chair: Dr Bhagwan Das, SP Jain School of Global Management, Sydney, Australia
Paper Id: 7125, 7498, 7600, 8520, 8623

8 February 2024

9:00-12:30 (USA-Texas Time)

Google Meet Session Chaired by **Dr Bishwajeet Pandey, Astana IT University, Astana, Kazakhstan**

Google Meet Link: https://meet.google.com/sdn-rdny-xba
Paper Id: 944, 2827, 3393, 3553, 4126, 5086, 5238, 5990, 6791, 6858, 7169, 7383, 7616, 7743, 7859, 9618

9 February 2024

INDUSTRY VISIT

OFFLINE SESSION 11:00-13:00 (USA-Texas Time) 7 February 2024
SESSION CHAIR: Dr Suresh Kumar Peddoju, University of Houston, USA

Paper Id	Paper Title	Author Name	Presenter Name
322	Link-based Anomaly Detection with Sysmon and Graph Neural Networks	Charlie Grimshaw, Brian Lachine, Taylor Perkins and Emilie Coote	Brian Lachine
944	A Novel Deep Learning Method for Segmenting the Left Ventricle in Cardiac Cine MRI	Wenhui Chu, Aobo Jin, Hardik A. Gohel	Wenhui Chu
1315	Deep Reinforcement Learning-based Malicious URL Detection with Feature Selection	Antonio Maci, Nicola Tamma and Antonio Coscia	Antonio Maci
1976	AI-Based Cybersecurity Policies and Procedures	Shadi Jawhar, Zeina Bitar and Jeremy Miller	Shadi Jawhar
2439	AI-Driven Customized Cyber Security Training and Awareness	Shadi Jawhar, Jeremy Miller and Zeina Bitar	Jeremy Miller
4287	YSAF: Yolo with Spatial Attention and FFT to Detect Face Spoofing Attacks	Rathinaraja Jeyaraj, Barathi Subramanian, Karnam Yogesh, Aobo Jin and Hardik A Gohel	Rathinaraja Jeyaraj,
6791	Toward robust systems against sensor-based adversarial examples based on the criticalities of sensors	Ade Kurniawan, Yuichi Ohsita, Masayuki Murata	Ade Kurniawan

OFFLINE SESSION 14:30-16:30 (USA-Texas Time) 7 February 2024
SESSION CHAIR: Dr Bhagwan Das, SP Jain School of Global Management, Sydney, Australia

Paper Id	Paper Title	Author Name	Presenter Name
7125	Secure federated learning applied to medical imaging with fully homomorphic encryption	Xavier Lessage, Leandro Collier, Charles-Henry Bertrand Van Ouytse, Axel Legay, Saïd Mahmoudi and Philippe Massonet	Leandro Collier
7600	Video key concept extraction using Convolution Neural Network	Tanvir H Sardar, Ruhul Amin Hazarika, Bishwajeet Pandey, Guru Prasad M S, Sk Mahmudul Hassan, Radhakrishna Dodmane, Hardik Gohel	Bishwajeet Pandey
8520	Prescriptive Analytics-based Robust Decision-Making Model for Cyber Disaster Risk Reduction	Joseph Ponnoly, John Puthenveetil and Patricia D'Urso	Dr Joseph Ponnoly
8623	Leveraging Advanced Visual Recognition Network Classifier for Pneumonia Prediction	Maulin Raval, Jin Aobo and Hardik Gohel	Maulin Raval
9618	Simulations and Advancements in MRI-Guided Power-Driven Ferric Tools for Wireless Therapeutic Interventions	Wenhui Chu, Aobo Jin, Hardik A. Gohel	Wenhui Chu

LISTENER

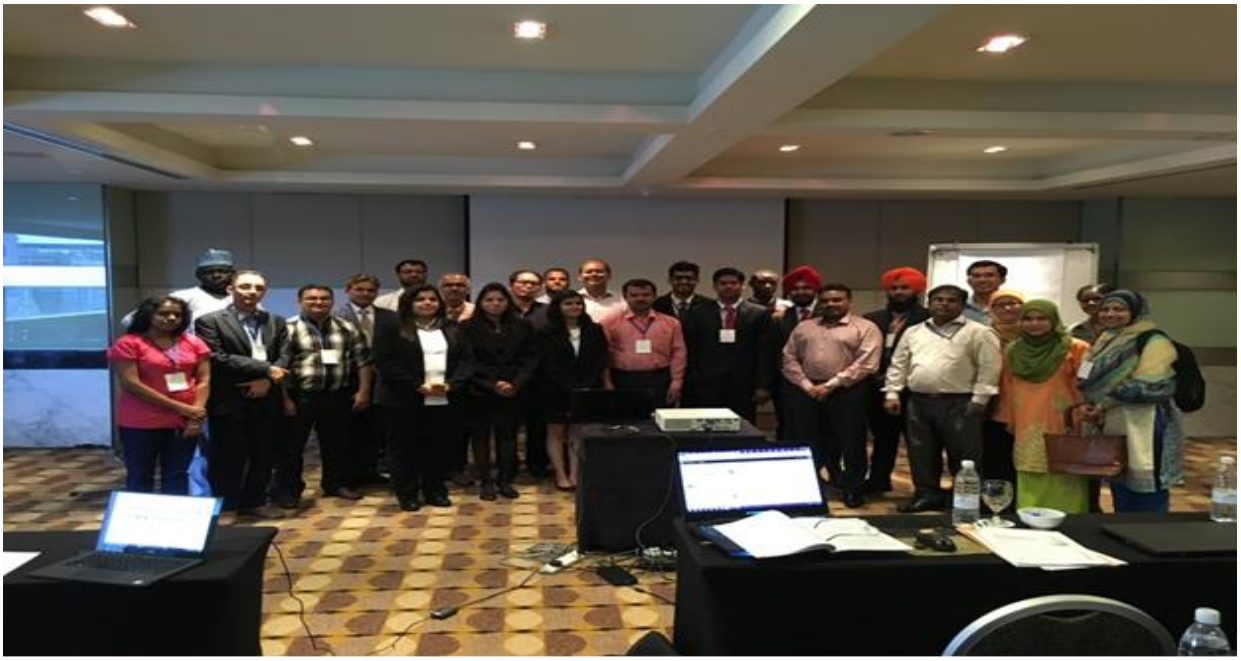
- Minho Park
- Sarah Nagro
- Humphrey Malisa
- Nicola Tamma
- Kate Highnam
- Saeed A. S. Shurrab
- Joshua Kiihne
- Anthony Adoyele
- Paige Reynolds

ONLINE SESSION 09:00-12:30 (USA-Texas Time) 8 February 2024
Google Meet Link: https://meet.google.com/sdn-rdny-xba
SESSION CHAIR: Dr Bishwajeet Pandey, Astana IT University, Kazakhstan

Paper Id	Paper Title	Presenter Name
72	Mobile Application Security Risk Score: A sensitive user input-based approach	**Trishla Shah**
2827	A Secure Open-Source Intelligence Framework For Cyberbullying Investigation	Sylvia worlali Azumah
3393	Improving Network Intrusion Detection Performance An Empirical Evaluation Using Extreme Gradient Boosting (XGBoost) with Recursive Feature Elimination	Gerard Shu Fuhnwi
3553	The Application of the Fifth Discipline Strategies in the Learning City Concept	Chipo Mutongi
4126	zkFDL: An efficient and privacy-preserving decentralized federated learning with zero knowledge proof	Mojtaba Ahmadi
4438	Risk-Aware Mobile App Security Testing: Safeguarding Sensitive User Inputs	**Trishla Shah**
5086	DataAgent: Evaluating Large Language Models' Ability to Answer Zero-Shot, Natural Language Queries	Manit Mishra, Dakshesh Sidnerlikar
5238	Leveraging Weak Supervision and BiGRU Neural Networks for Sentiment Analysis on Label-Free News Headlines	Shahin Alipour
5990	Identifying Race and Gender Bias in Stable Diffusion AI Image Generation	Aadi Chauhan
6858	Enhanced Network Intrusion Detection System Using PCGSO-Optimized BI-GRU Model in AI-Driven Cybersecurity	Balasubramanian Prabhu kavin
7169	Federated Leaming Based Smart Horticulture and Smart Storage of Fruits Using E-Nose, and Blockchain: A Proposed Model	Akniyet Nurzhaubayev
7383	A Holistic Review on Detection of Malicious Browser Extensions and Links using Deep Learning	Rama Abirami K
7616	CANAL - Cyber Activity News Alerting Language Model : Empirical Approach vs. Expensive LLMs	Chinmay Gondhalekar
7743	Sentiment Analysis of Financial News Data using TF-IDF and Machine Learning Algorithms	Gideon Popoola
7859	Robotics in Healthcare: The African Perspective	Dr Chipo Mutongi

ICGCET'2015: 1st International Conference of Gyancity at Dubai, UAE

RTCSE'16: 2nd International Conference of Gyancity at Kuala Lumpur, Malaysia



ICGCET'2016: 3rd International Conference of Gyancity at Aalborg University, Esbjerg, Denmark

Institut i Esbjerg samler forskere fra hele verden

DEL Af Edmund Jacobsen 15. august 2016 kl. 05:31

40 forskere og studerende fra hele verden samles på Institut for Energiteknik, Aalborg Universitet Esbjerg, i tre dage i denne uge, når der afvikles en international konference, der handler om at gøre computerteknologi mere grøn.

D.M. Akbar Hussain, lektor ved Institut for Energiteknik på Aalborg Universitet Esbjerg, har sammen med en kollega fra Indien arrangeret konferencen International Conference on Green Computing and Engineering Technologies.

Det er planen, at disse konferencer skal afvikles i Esbjerg hvert andet år – ganske enkelt fordi Institut for Energiteknik i Esbjerg er internationalt anerkendt.

RTCSE'17: 4th International Conference of Gyancity at Kuala Lumpur, Malaysia

IMCES'17: 5th International Conference of Gyancity at Kuala Lumpur, Malaysia

ICGCET'2018: 6th International Conference of Gyancity at Limerick, Ireland

RTCSE'2018: 7th International Conference of Gyancity at Bangkok, Thailand

ICGCET'18: 8th International Conference of Gyancity at Aalborg University, Esbjerg, Denmark

RTCSE'2019: 9th International Conference of Gyancity at Univeristy of Hawaii, USA

IMCES'2019:10th International Conference of Gyancity at Port Louis, Mauritius

ICGCET'2019: 11th International Conference of Gyancity at Casablanca, Morocco

RTCSE'2020: 12th International Conference of Gyancity at University of Hawaii, USA

IMCES'2020: 13th International Conference by Gyancity at Jakarta, Indonesia

ICGCET'2020: 14th Conference by Gyancity at St Petersburg, Russia

Jammu, September 18: Dr. Amit Kant Pandit, Faculty, SoECE, SMVDU chaired an online session in 6th International Conference on Green Computing and Engineering Technologies (ICGCET®).

The international conference is scheduled from 16th-18th September 2020 at Herzen State Pedagogical University, St Petersburg, Russia. The traditional face-to-face meeting was replaced by the online meeting due to a pandemic situation. The first online session was conducted through CISCO WebEx app.

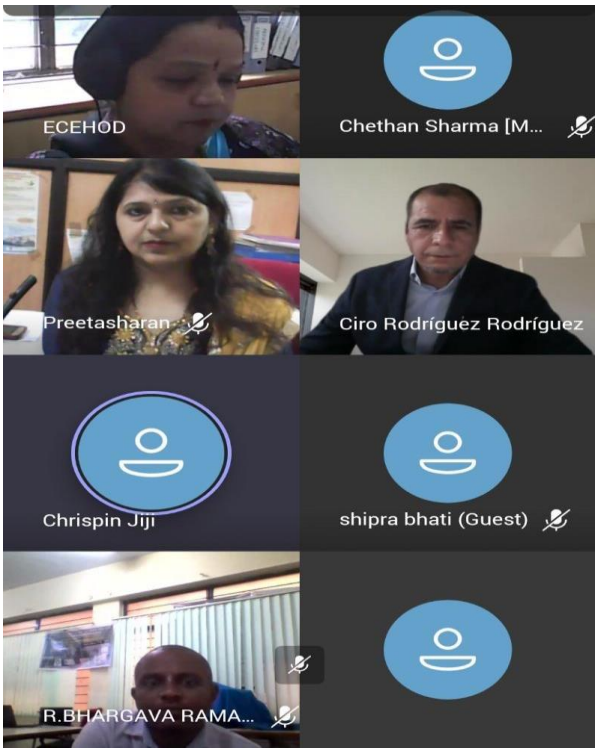
Dr. Pandit along with co-chair Dr. Bishwajeet Pandey, Birla Institute of Applied Sciences, Bhimtal Uttarakhand, and associated with Gyancity Research consultancy conducted the first session and an introductory talk.

The attendees across the world presented their work through an online meeting and recorded video presentations. The presentation and other videos are uploaded for public viewing on YouTube channel for wider academic sharing.

The convener of the conference Prof. Jason Levy, University of Hawaii, USA. Prof. Geetam S Tomar, Director Birla Institute of Applied Sciences, Bhimtal, India, congratulated on the successful organizing of the session.

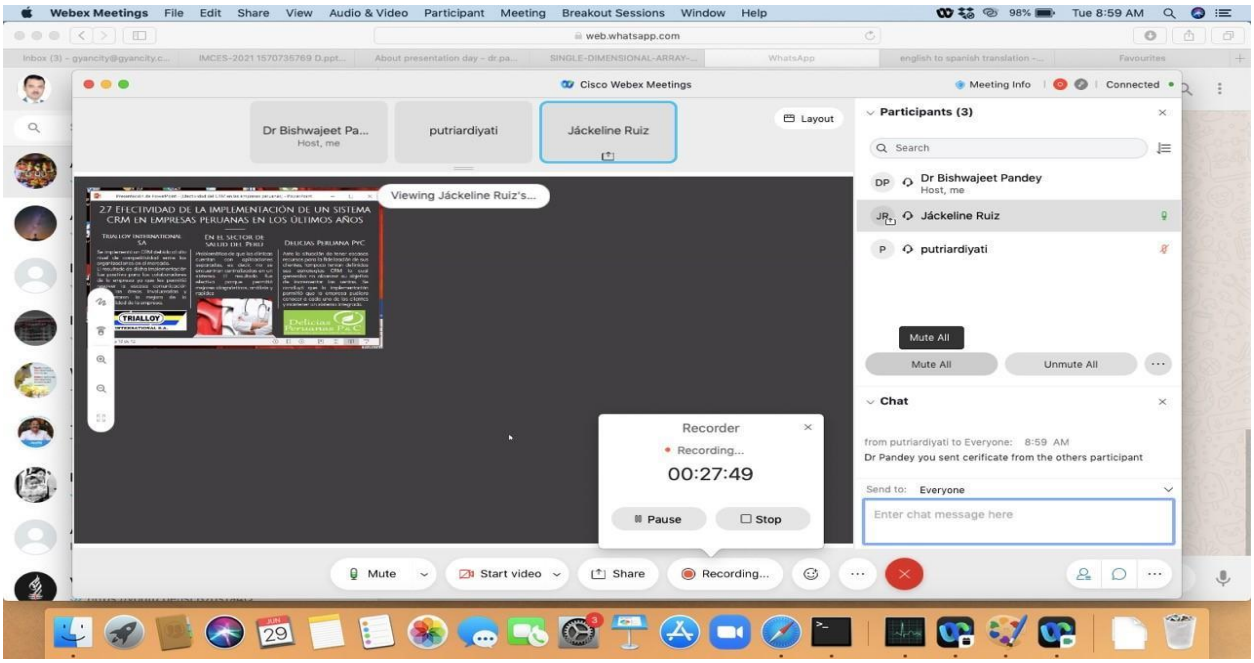
Dr. Amit Kant Pandit thanked coordinators for arranging such academic meetings in difficult times.

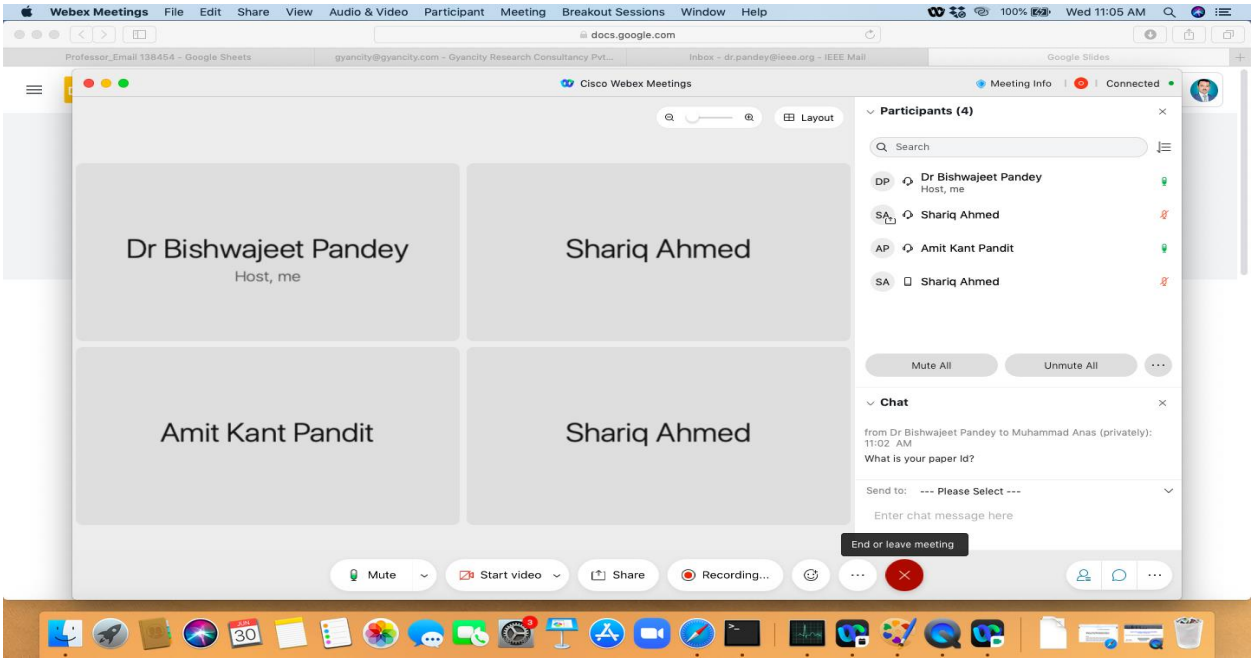
RTCSE'2021: 15th International Conference of Gyancity at University of Hawaii, USA



BMESS'2021: 16th Virtual Conference by Gyancity

IMCES'2021: 17th International Conference by Gyancity at Yarsi University, Indonesia





ICGCET'2021: 18th International Conference by Gyancity at National University of Federico Villareal, Lima, Peru

Evento se dará el 22 y 23 de septiembre. Foto: difusión

La República
larepublica_pe
ediciondigital@glr.pe

16 Set 2021 | 12:40 h
Actualizado el 16 de Setiembre 2021 | 12:40 h

Este 22 y 23 de septiembre se realizará la 7ª Conferencia Internacional sobre Tecnologías de Ingeniería y Computación Ecológicas 2021 (ICGCET-2021) y la 13ª Conferencia Internacional en Inteligencia Computacional y Redes de Comunicación 2021 (CICN 2021), eventos que tendrán como sede a la Universidad Villareal (UNFV).

Juan Alfaro, rector de la UNFV, será el encargado de inaugurar los referidos certámenes, el miércoles 22 a las 10.00 a. m. Previamente, Akbar Hussain, de la Universidad Aalborg de Dinamarca, será el encargado de brindar las palabras de bienvenida.

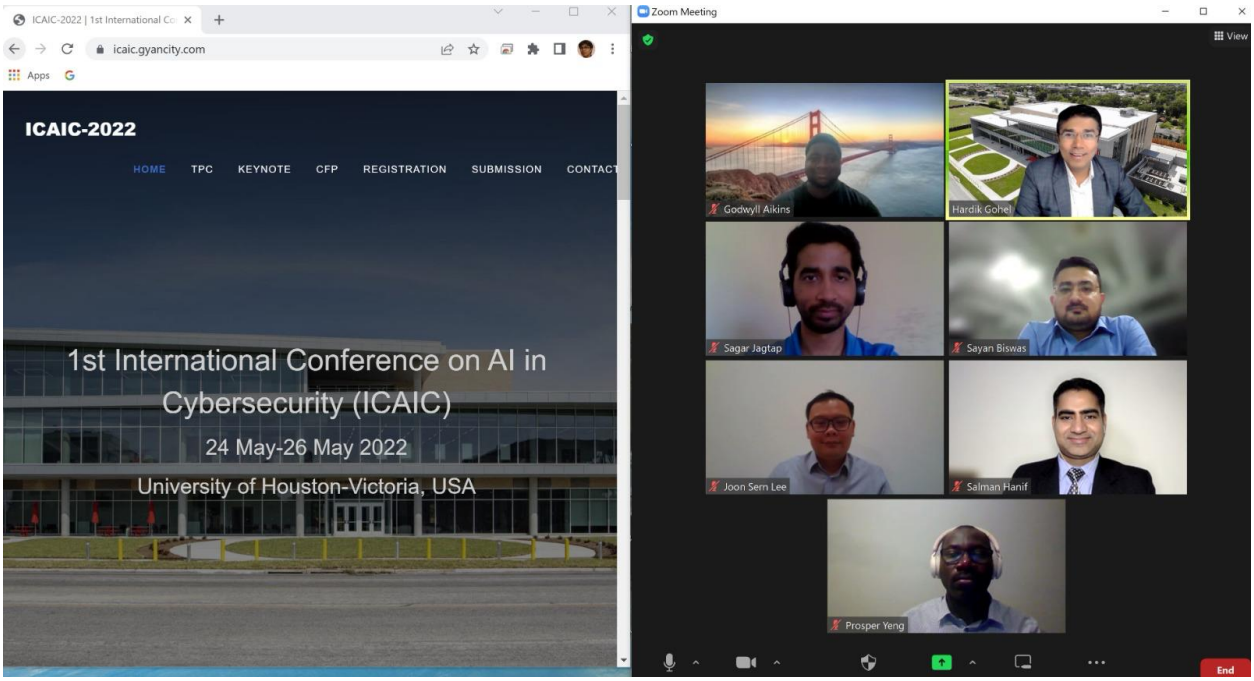
La ICGCET-2021 presentará las investigaciones de diferentes áreas de la ciencia y la tecnología, y proporcionará una plataforma para que investigadores y científicos de todo el mundo intercambien y compartan sus experiencias y resultados de investigación.

ICGCET'2021: 18th International Conference by Gyancity at National University of Federico Villareal, Lima, Peru

RTCSE'2022: 19th International Conference of Gyancity at University of Hawaii USA

BMESS'2022: 20th International Conference by Gyancity at Bath Spa University UAE

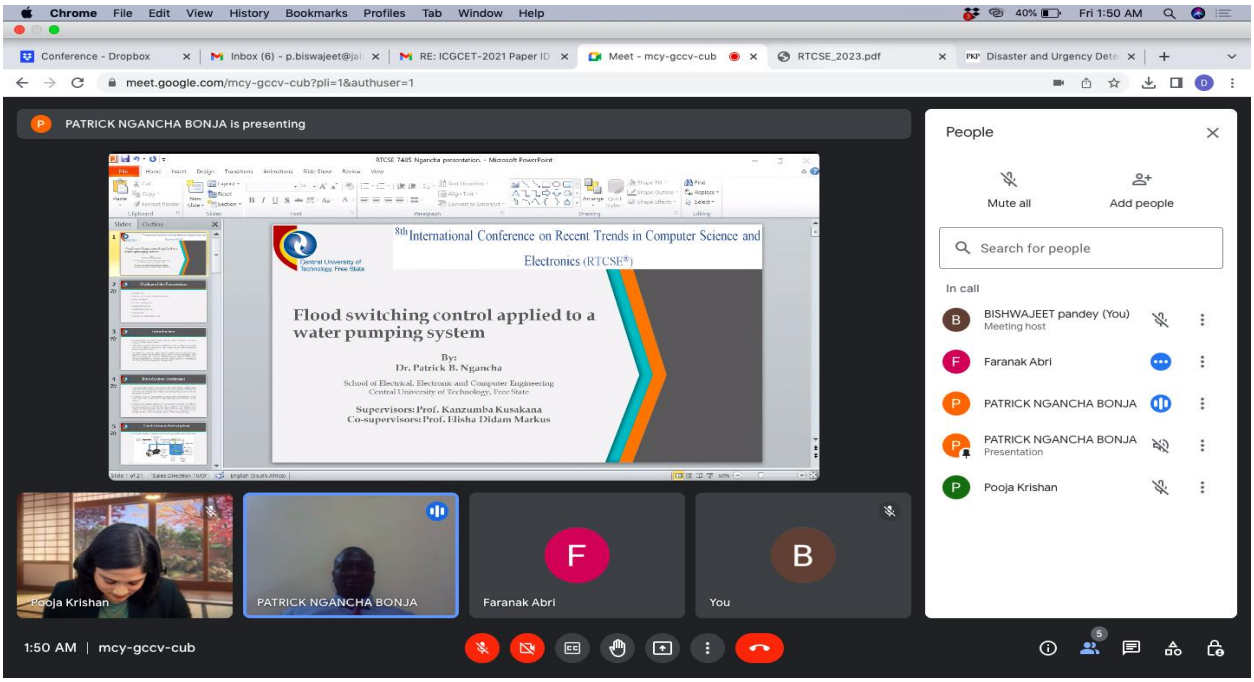
ICAIC'2022: 21st International Conference by Gyancity at University of Houston-Victoria, USA



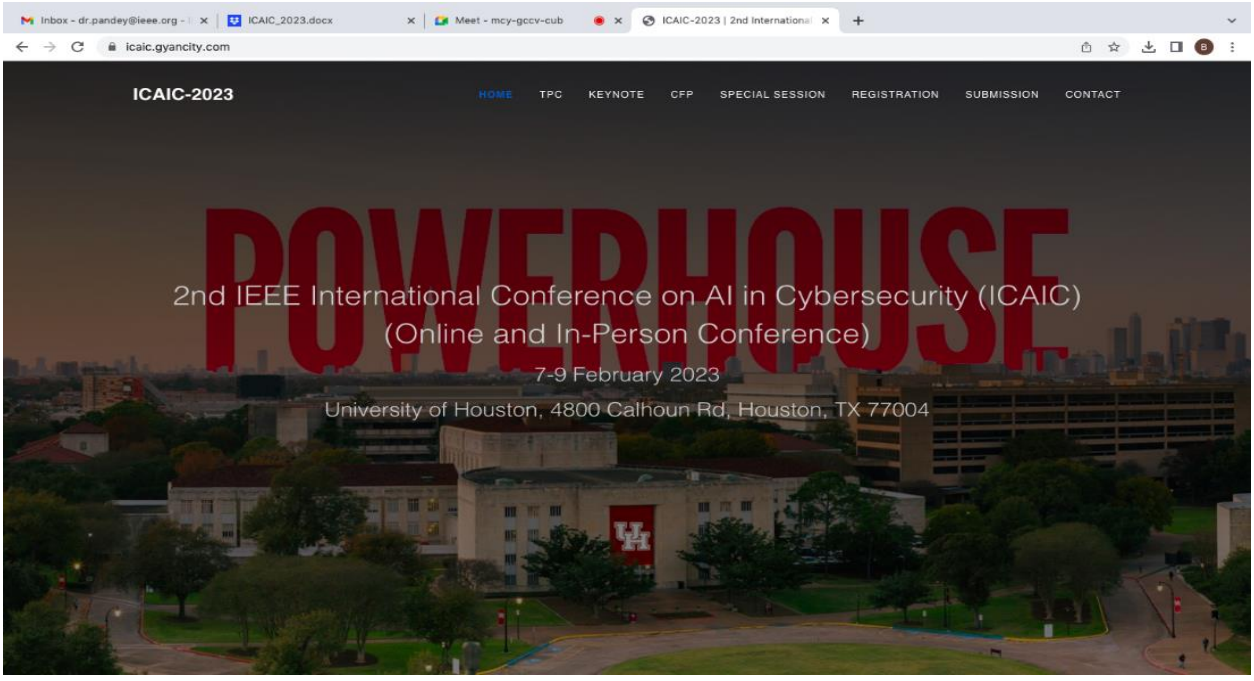
IMCES'2022: 22nd International Conference by Gyancity at Aalborg University, Esbjerg, Denmark

ICGCET'2022: 23rd International Conference by Gyancity at Mauritius

RTCSE'2023: 24th International Conference of Gyancity at University of Hawaii USA



ICAIC'2023 GROUP PHOTO: 25th International Conference of Gyancity at University of Houston-Victoria, USA



BMESS'2023 GROUP PHOTO: 26th International Conference of Gyancity at Bath Spa University, UAE

IMCES'2023: 27th International Conference by Gyancity at Yarsi University, Jakarta, Indonesia

ICGCET'2024: 28th International Conference by Gyancity at Cape Town South Africa

Abstract of Paper Accepted in ICAIC'2024

322

Link-based Anomaly Detection with Sysmon and Graph Neural Networks

Charlie Grimshaw[1], Brian Lachine[2] , Taylor Perkins[3], Emilie Coote[4]
[1]Director Naval Platform Systems, Department of National Defence, Ottawa, Canada
charles.grimshaw@forces.gc.ca
[2]Electrical and Computer Engineering, Royal Military College Kingston, Canada
brian.lachine@rmc.ca
[3,4]Cyber & Strategic Risk Deloitte Canada Toronto, Canada
taperkins@deloitte.ca, ecoote@deloitte.ca

ABSTRACT

Anomaly detection is a challenge well-suited to machine learning and in the context of information security, the benefits of unsupervised solutions show significant promise. Recent attention to Graph Neural Networks (GNNs) has provided an innovative approach to learn from attributed graphs. Using a GNN encoder-decoder architecture, anomalous edges between nodes can be detected during the reconstruction phase. The aim of this research is to determine whether an unsupervised GNN model can detect anomalous network connections in a static, attributed network. Network logs were collected from four corporate networks and one artificial network using endpoint monitoring tools. A GNN-based anomaly detection system was designed and employed to score and rank anomalous connections between hosts. The model was validated against four realistic experimental scenarios against the four large corporate networks and the smaller artificial network environment. Although quantitative metrics were affected by factors including the scale of the network, qualitative assessments indicated that anomalies from all scenarios were detected. The false positives across each scenario indicate that this model in its current form is useful as an initial triage, though would require further improvement to become a performant detector. This research serves as a promising step for advancing this methodology in detecting anomalous network connections. Future work to improve results includes narrowing the scope of detection to specific threat types and a further focus on feature engineering and selection.

Index Terms—Anomaly Detection, Graph Neural Networks, Unsupervised Learning, Link Prediction

Abstract of Paper Accepted in ICAIC'2024

944

A Novel Deep Learning Method for Segmenting the Left Ventricle in Cardiac Cine MRI

Wenhui Chu†, Aobo Jin†, Hardik A. Gohel †
†Dept. of Computer Science, University of Houston-Victoria, Victoria, USA.
ChuW1@uhv.edu, Jina@uhv.edu, gohelh@uhv.edu

ABSTRACT

This research aims to develop a novel deep learning network, GBU-Net, utilizing a group-batch-normalized UNet framework, specifically designed for the precise semantic segmentation of the left ventricle in short-axis cine MRI scans. The methodology includes a down-sampling pathway for feature extraction and an up-sampling pathway for detail restoration, enhanced for medical imaging. Key modifications include techniques for better contextual understanding crucial in cardiac MRI segmentation. The dataset consists of 805 left ventricular MRI scans from 45 patients, with comparative analysis using established metrics such as the dice coefficient and mean perpendicular distance. GBU-Net significantly improves the accuracy of left ventricle segmentation in cine MRI scans. Its innovative design outperforms existing methods in tests, surpassing standard metrics like the dice coefficient and mean perpendicular distance. The approach is unique in its ability to capture contextual information, often missed in traditional CNN-based segmentation. An ensemble of the GBU-Net attains a 97% dice score on the SunnyBrook testing dataset. GBU-Net offers enhanced precision and contextual understanding in left ventricle segmentation for surgical robotics and medical analysis

Index Terms—segmentation; MRI; CNN; left ventricle

Abstract of Paper Accepted in ICAIC'2024

1315

Deep Reinforcement Learning-based Malicious URL Detection with Feature Selection

Antonio Maci, Nicola Tamma, Antonio Coscia
Cybersecurity Laboratory BV TECH S.p.A.
Milan, Italy
0000-0002-6526-554X, 0009-0009-2161-7345, 0000-0002-7263-4999

ABSTRACT

Data theft through web applications that emulate legitimate platforms constitutes a major network security issue. Countermeasures using artificial intelligence (AI)-based systems are often applied because they can effectively detect malicious websites, which are extremely outnumbered by legitimate ones. In this domain, deep reinforcement learning (DRL) emerges as an attractive field for the development of network intrusion detection models, even in the case of highly skewed class distributions. However, DRL requires training time that increases with data complexity. This paper combines a DRL-based classifier with state-of-the-art feature selection techniques to speed up training while retaining or even improving classification performance. Our experiments used the Mendeley dataset and five different statistical and correlation-based feature-ranking strategies. The results indicated that the selection technique based on the calculation of the Gini index reduces the number of columns in the dataset by 27%, saving more than 10% of training time and significantly improving classification scores compared with the case without selection strategies.

Index Terms—Artificial Intelligence in Cybersecurity, Deep Reinforcement Learning, Feature Selection, Malicious URL, Web Phishing

Abstract of Paper Accepted in ICAIC'2024

1976

AI-Based Cybersecurity Policies and Procedures

Shadi Jawhar, Jeremy Miller, Zeina Bitar
Cyber Security Department Lionfish Cyber Security Fort Wayne, Indiana
Shadi@lionfishcybersecurity.com, Jeremy@lionfishcybersecurity.com,
zeinajb@hotmail.com

ABSTRACT

The use of artificial intelligence (AI) in cyber security [1] has proven to be very effective as it helps security professionals better understand, examine, and evaluate possible risks and mitigate them. It also provides guidelines to implement solutions to protect assets and safeguard the technology used. As cyber threats continue to evolve in complexity and scope, and as international standards continuously get updated, the need to generate new policies or update existing ones efficiently and easily has increased [1][2]. The use of (AI) in developing cybersecurity policies and procedures can be key in assuring the correctness and effectiveness of these policies as this is one of the needs for both private organizations and governmental agencies. This study sheds light on the power of AI-driven mechanisms in enhancing digital defense procedures by providing a deep implementation of how AI can aid in generating policies quickly and to the needed level.

Index Terms—Cyber Security Policies and Procedures, AI-driven, National Cyber Security Frameworks, Security Compliance.

Abstract of Paper Accepted in ICAIC'2024

2439

AI-Driven Customized Cyber Security Training and Awareness

Shadi Jawhar, Jeremy Miller, Zeina Bitar
Cyber Security Department Lionfish Cyber Security Fort Wayne, Indiana
Shadi@lionfishcybersecurity.com, Jeremy@lionfishcybersecurity.com,
zeinajb@hotmail.com

ABSTRACT

Artificial intelligence (AI) has been successfully used in cyber security for enhancing comprehending, investigating, and evaluating cyber threats. It can effectively anticipate cyber risks in a more efficient way. AI also helps in putting in place strategies to safeguard assets and data. Due to their complexity and constant development, it has been difficult to comprehend cybersecurity controls and adopt the corresponding cyber training and security policies and plans. Given that both cyber academics and cyber practitioners need to have a deep comprehension of cybersecurity rules, artificial intelligence (AI) in cybersecurity can be a crucial tool in both education and awareness. By offering an in-depth demonstration of how AI may help in cybersecurity education and awareness and in creating policies fast and to the needed level, this study focuses on the efficiency of AI-driven mechanisms in strengthening the entire cyber security education life cycle.

Index Terms—Cyber Security Awareness, Cyber Security Education, AI-driven, Cyber Security Compliance.

Abstract of Paper Accepted in ICAIC'2024

2827

A Secure Open-Source Intelligence Framework For Cyberbullying Investigation

Sylvia Worlali Azumah, Victor Adewopo, Zag Elsayed, Nelly Elsayed, Murat Ozer
School of Information Technology University of Cincinnati, Cincinnati, USA
azumahsw@mail.uc.edu, adewopva@mail.uc.edu, elsayezs@ucmail.uc.edu, elsayeny@ucmail.uc.edu, ozermm@ucmail.uc.edu

ABSTRACT

Cyberbullying has become a pervasive issue based on the rise of cell phones and internet usage affecting individuals worldwide. This paper proposes an open-source intelligence pipeline using data from Twitter to track keywords relevant to cyberbullying in social media to build dashboards for law enforcement agents. We discuss the prevalence of cyberbullying on social media, factors that compel individuals to indulge in cyberbullying, and the legal implications of cyberbullying in different countries also highlight the lack of direction, resources, training, and support that law enforcement officers face in investigating cyberbullying cases. The proposed interventions for cyberbullying involve collective efforts from various stakeholders, including parents, law enforcement, social media platforms, educational institutions, educators, and researchers. Our research provides a framework for cyberbullying and provides a comprehensive view of the digital landscape for investigators to track and identify cyberbullies, their tactics, and patterns. An OSINT dashboard with real-time monitoring empowers law enforcement to swiftly take action, protect victims, and make significant strides toward creating a safer online environment.

Index Terms— Cyberbullying, Open Source Intelligence, Investigation, Data Visualization

Abstract of Paper Accepted in ICAIC'2024

3393	**Improving Network Intrusion Detection Performance**

Gerard Shu Fuhnwi, Matthew Revelle, Clemente Izurieta
Gianforte School of Computing Montana State University Bozeman, MT, USA
gerard.shufuhnwi@student.montana.edu, matthew.revelle@montana.edu, clemente.izurieta@montana.edu

ABSTRACT

In cybersecurity, Network Intrusion Detection Systems (NIDS) are essential for identifying and preventing malicious activity within computer networks. Machine learning algorithms have been widely applied to NIDS due to their ability to identify complex patterns and anomalies in network traffic. Improvements in the performance of an IDS can be measured by increasing the Matthew Correlation Coefficient (MCC), the reduction of False Alarm Rates (FARs), and the maintenance of up-to-date signatures of the latest attacks to maintain confidentiality, integrity, and availability of services. Integrating machine learning with feature selection for IDSs can help eliminate less important features until the optimal subset of features is achieved, thus improving the NIDS. In this research, we propose an approach for NIDS using XGBoost, a popular gradient boosting algorithm, with Recursive Feature Elimination (RFE) feature selection. We used the NSLKDD dataset, a benchmark dataset for evaluating NIDS, for training and testing. Our empirical results show that XGBoost with RFE outperforms other popular machine learning algorithms for NIDS on this dataset, achieving the highest MCC for detecting NSL-KDD dataset attacks of type DoS, Probe, U2R, and R2L and very high classification time.

Index Terms— Network Intrusion Detection; feature selection; RFE; XGBoost; recursive feature elimination; NSL-KDD.

Abstract of Paper Accepted in ICAIC'2024

3553	**The Application of the Fifth Discipline Strategies in the Learning City Concept**

Dr Chipo Mutongi, [2]Dr Billy Rigava, [3]Joyce Goredema and [4]Rudo Manyere
[1]Midlands State University, [2]UZ, [1,3,4]City of Harare
mutongic@gmail.com, mutongic@staff.msu.ac.zw

ABSTRACT

The stone age did not end because there were no more stones, it ended because of continuous improvement, innovation, creativity and learning. Local government has always been around since time immemorial. Even in the Stone Age period there was some form of local government, leaning and continuous improvement. In this DVUCADD environment, an environment characterized by dynamic, volatile uncertainty, ambiguous, diversity and disruptive, cities should be in a position to employ Peter Senge's fifth discipline in order to survive and be in a position to learn faster. The Local government in Africa and Zimbabwe in particular has the role of proving a range of vital learning city services delivery for residents and organisations in defined areas. Among them are well known functions such as social services like primary education, libraries, vocational training and recreational facilities. Local government administration has a great role to play in bringing citizenry's lifelong learning, engagement and participation. This then brings in economic and social development. One of the important aspects that ever happened in our life, is when Peter Senge came up with the fifth disciplines that any organisation can apply in order to achieve a learning organisation. These disciplines are personal mastery, mental shared vision, team learning and systems thinking. The City of Harare is used as a case study in the application of Peter Senge's fifth discipline to foster the learning city concept.

Index Terms— fifth discipline, strategies, learning city, lifelong learning, smart city

Abstract of Paper Accepted in ICAIC'2024

4126

zkFDL: An efficient and privacy-preserving decentralized federated learning with zero knowledge proof

Mojtaba Ahmadi
Shahid Beheshti University Tehran, Iran
mojtaba27ahmadi@gmail.com
Reza Nourmohammadi
The University Of British Columbia Kelowna, Canada
reza.nourmohammadi@ubc.ca

ABSTRACT

Federated leaning (FL) has been frequently used in various field of studies and businesses. Traditional centralized FL systems suffer from serious issues. To address these concerns, decentralized federated learning (DFL) systems have been introduced in recent years in which with the help of blockchains, try to achieve more integrity and efficiency. On the other hand, privacypreserving is an uncovered part of these systems. To address this, and also scaling the blockchain-based computations, we propose a zero knowledge proof (ZKP) based aggregator (zkDFL) that allows clients to share their large-scale model parameters with a trusted centralized server without revealing their individual data to other clients. We utilize blockchain technology to manage the aggregation algorithm via smart contracts. The server performs a ZKP algorithm to prove to the clients that the aggregation is done according to the accepted algorithm. The server can also prove that all inputs of clients have been used. We evaluate our measure through a public dataset about wearable internet of things. As demonstrated by numerical evaluations, zkDFL introduces verifiability of correctness of aggregation process and enhances the privacy protection and scalability of DFL systems, while the gas cost has declined significantly.

Index Terms—federated learning, blockchain, zero knowledge proof, aggregator algorithm, scaling systems

Abstract of Paper Accepted in ICAIC'2024

4287

YSAF: Yolo with Spatial Attention and FFT to Detect Face Spoofing Attacks

Rathinaraja Jeyaraj, Karnam Yogesh, Aobo Jin, Hardik A Gohel
University of Houston Victoria, USA. jrathinaraja@gmail.com
Barathi Subramanian
Kyungpook National University Daegu, South Korea. barathi.sn93@gmail.com

ABSTRACT

Besides biometrics, face authentication is quite popular on smart devices like smartphones and other electronic gadgets to verify and authenticate individuals. In the face authentication method, there is a chance of spoofing attacks, in which a static image or recorded video can be substituted for a real person's face to breach security and gain access. To solve this problem, smart devices use additional hardware like a dual camera or an infrared sensor, which adds extra cost, weight, and incompatibility to different gadgets. Alternatively, software-based methods may be confused with a video of the user to gain the access. To overcome these problems, in this paper, we present a framework, YSAF, that combines Yolo v8 object detection, spatial attention, and fast Fourier transform (FFT) to restrict facialbased spoofing attacks without additional hardware. In YSAF, spatial attention is first used to focus on relevant features and reduce noise in the input image. Next, frequency analysis through FFT is applied to embed information in the collected images to help the classification model differentiate live faces from static ones. As a final step, Yolo detects whether the object present in the collected images is real or fake (spoof). The YSAF is trained using real images collected by volunteers from different sources and pre-processed with spatial attention and FFT before training with Yolo. The results show that the YSAF accurately blocks spoofing attacks with still images/videos in real-time.

Index Terms—Face authentication, FFT, Spatial attention, Spoofing detection, Yolo

Abstract of Paper Accepted in ICAIC'2024

5086

DataAgent: Evaluating Large Language Models' Ability to Answer Zero-Shot, Natural Language Queries

Manit Mishra
Irvington High School Fremont, United States mshmanit@gmail.com
Abderrahman Braham
Pioneer High School Sousse, Tunisia br.abderrahman.contact@gmail.com
Charles Marsom
Davis Senior High School Davis, USA, charleshenrymarsom@gmail.com
Bryan Chung
The Loomis Chaffee School Windsor, United States bryan chung@loomis.org
Gavin Griffin, Chatanya Sarin
Bellarmine College Preparatory Sunnyvale, USA
gavgriffin563@gmail.com
Dakshesh Sidnerlikar
Rutgers University New Brunswick, USA
dakshesh.sid@gmail.com, chatanya.sarin@gmail.com
Arjun Rajaram
University of Maryland, College Park Frisco, United States
arajara1@terpmail.umd.edu

ABSTRACT

Conventional processes for analyzing datasets and extracting meaningful information are often time-consuming and laborious. Previous work has identified manual, repetitive coding and data collection as major obstacles that hinder data scientists from undertaking more nuanced labor and high-level projects. To combat this, we evaluated OpenAI's GPT-3.5 as a "Language Data Scientist" (LDS) that can extrapolate key findings, including correlations and basic information, from a given dataset. The model was tested on a diverse set of benchmark datasets to evaluate its performance across multiple standards, including data science code-generation based tasks involving libraries such as NumPy, Pandas, Scikit-Learn, and TensorFlow, and was broadly successful in correctly answering a given data science query related to the benchmark dataset. The LDS used various novel prompt engineering techniques to effectively answer a given question, including Chain-ofThought reinforcement and SayCan prompt engineering. Our findings demonstrate great potential for leveraging Large Language Models for low-level, zero-shot data analysis.

Index Terms—GPT, data science, natural language processing, large language model, data processing, RefleXion, Chain-of-Thought, SayCan, action plan generation, zero-shot prompting, plain language

Abstract of Paper Accepted in ICAIC'2024

5238

Leveraging Weak Supervision and BiGRU Neural Networks for Sentiment Analysis on Label-Free News Headlines

Ahmadali Jamali, Shahin Alipour , Audrey Rah
Islamic Azad University (IAU), Science and Research Branch - Department of Computer science, Iran
University of Houston - Department of Biomedical Engineering, USA
University of Houston - Department of Electrical and Computer Engineering, USA
ahmadali.jamali@srbiau.ac.ir, salipour@uh.edu, arahimi@cougarnet.uh.edu

ABSTRACT

Auto-labeling of text is a useful and necessary technique for creating large and high-quality training data sets for machine learning models. Label-free sentiment classification is a challenging semi-supervised task in the natural language processing domain. This study leveraged the weak supervision framework to generate weak labels in three categories for millions of news headlines from Australian Broadcasting Corporation (ABC). A Bidirectional Gate Recurrent Unit (BiGRU) was then trained with neural network dense layers to achieve a validation accuracy of 96.76% with 99.99% accuracy. The performance of this method was also compared with traditional and deep learning natural language processing techniques.

Index Terms— Label-free, Sentiment classification, Headline News, Weak-Supervision, BiGRU Neural Network

Abstract of Paper Accepted in ICAIC'2024

5990

Identifying Race and Gender Bias in Stable Diffusion AI Image Generation

Aadi Chauhan[1], Taran Anand[1], Tanisha Jauhari[2]
Arjav Shah[1], Rudransh Singh[1], Arjun Rajaram[3], Rithvik Vanga[4]
[1]Bellarmine College Preparatory
[2]Foothill High School
[3]University of Maryland
[4]University of Michigan
emailaadichauhan@gmail.com, tarananand24@gmail.com,
tanisha.jauhari@gmail.com, arjavshah21@gmail.com,
rudransh.singh18@gmail.com, arjun.rajaram2404@gmail.com,
rithvikvanga@gmail.com

ABSTRACT

In this study, we set out to measure race and gender bias prevalent in text-to-image (TTI) AI image generation, focusing on the popular model Stable Diffusion from Stability AI. Previous investigations into the biases of word embedding models—which serve as the basis for image generation models—have demonstrated that models tend to overstate the relationship between semantic values and gender, ethnicity, or race. These biases are not limited to straightforward stereotypes; more deeply rooted biases may manifest as microaggressions or imposed opinions on policies, such as paid paternity leave decisions. In this analysis, we use image captioning software OpenFlamingo and Stable Diffusion to identify and classify bias within text-to-image models. Utilizing data from the Bureau of Labor Statistics, we engineered 50 prompts for profession and 50 prompts for actions in the interest of coaxing out shallow to systemic biases in the model. Prompts included generating images for "CEO", "nurse", "secretary", "playing basketball", and "doing homework". After generating 20 images for each prompt, we document the model's results. We find that biases do exist within the model across a variety of prompts. For example, 95% of the images generated for "playing basketball" were African American men. We then analyze our results through categorizing our prompts into a series of income and education levels corresponding to data from the Bureau of Labor Statistics. Ultimately, we find that racial and gender biases are present yet not drastic.

Index Terms— generated bias, neural network, stable diffusion, text-to-image, race classification, gender classification, openflamingo

Abstract of Paper Accepted in ICAIC'2024

6791

Toward robust systems against sensor-based adversarial examples based on the criticalities of sensors

Ade Kurniawan[1], Yuichi Ohsita[2], Masayuki Murata[1]
[1]Department of Information Networking Graduate School of Information Science and Technology, Osaka University Osaka, Japan
[2]Cybermedia Center Osaka University Osaka, Japan
k-ade@ist.osaka-u.ac.jp, yuichi.ohsita.cmc@osaka-u.ac.jp, murata@ist.osaka-u.ac.jp

ABSTRACT

In multi-sensor systems, certain sensors could have vulnerabilities that may be exploited to produce AEs. However, it is difficult to protect all sensor devices, because the risk of the existence of vulnerable sensor devices increases as the number of sensor devices increases. Therefore, we need a method to protect ML models even if a part of the sensors are compromised by the attacker. One approach is to detect the sensors used by the attacks and remove the detected sensors. However, such reactive defense method has limitations. If some critical sensors that are necessary to distinguish required states are compromised by the attacker, we cannot obtain the suitable output. In this paper, we discuss a strategy to make the system robust against AEs proactively. A system with enough redundancy can work after removing the features from the sensors used in the AEs. That is, we need a metric to check if the system has enough redundancy. In this paper, we define groups of sensors that might be compromised by the same attacker, and we propose a metric called criticality that indicates how important each group of sensors are for classification between two classes. Based on the criticality, we can make the system robust against sensor-based AEs by interactively adding sensors so as to decrease the criticality of any groups of sensors for the classes that must be distinguished.

Index Terms— Adversarial examples, sensors

Abstract of Paper Accepted in ICAIC'2024

6858

Enhanced Network Intrusion Detection System Using PCGSO-Optimized BI-GRU Model in AIDriven Cybersecurity

Priyan Malarvizhi Kumar, Kavya Vedantham,
Department of Data Science, University of North Texas, USA.
Priyan.Malarvizhikumar@unt.edu, kavyavedantham@my.unt.edu,
Jeeva Selvaraj
Department of ISE, Jain University-Global Campus, Bangalore, India.
sassyjeeva@gmail.com
Balasubramanian Prabhu kavin
Department of Data Science and Business Systems, SRM Institute of Science and Technology, Tamil Nadu
ceaserkavin@gmail.com

ABSTRACT

The detection of complex attacks by Network Intrusion Detection Systems (NIDS) is hindered by evasion strategies including encrypted traffic and polymorphic malware. Attackers frequently take advantage of holes in NIDS algorithms, emphasising the never-ending cat-and-mouse game between cybersecurity defences and dynamic attack tactics. In the context of cybersecurity, this study offers a sophisticated method for supporting Network Intrusion Detection Systems (NIDS). The tactic includes a thorough preprocessing stage that include functions for normalisation and standardisation in order to recover the accuracy and consistency of the input data. The Perceptive Craving Game Search Optimisation (PCGSO) algorithm is then used for feature selection, maximising the effectiveness of the NIDS. Bidirectional Gated Recurrent Unit (BI-GRU) representations are used in the classification phase because of their ability to identify sequential dependencies in network traffic data. A second PCGSO programme is used to carry out hyperparameter tuning, which guarantees the best possible model performance. The ISCXIDS2012, a popular benchmark dataset in the field, has been selected as the dataset for evaluation. The suggested approach demonstrates how PCGSO may be used to improve feature selection and hyperparameter tweaking, leading to an NIDS that is more accurate and resilient to cyberattacks. Performance evaluations and experimental findings show that the suggested technique outperforms other current models with 99% accuracy

Index Terms—Network Intrusion Detection Systems, Cybersecurity, Gated Recurrent Unit, Game Search Optimization, Normalization Function.

Abstract of Paper Accepted in ICAIC'2024

7125

Secure federated learning applied to medical imaging with fully homomorphic encryption

Xavier Lessage, Leandro Collier, Philippe Massonet
Applied Research Centre
CETIC Charleroi, Belgium
xavier.lessage@cetic.be, leandro.collier@cetic.be, Philippe.massonet@cetic.be
Charles-Henry Bertrand Van Ouytsel, Axel Legay
Faculty of Engineering
UCLouvain Louvain-la-Neuve, Belgium
charles-henry.bertrand@uclouvain.be, axel.legay@uclouvain.be
Saıd Mahmoudi
Faculty of Engineering UMONS Mons, Belgium
said.mahmoudi@umons.ac.be

ABSTRACT

This study explores the convergence of Federated Learning (FL) and Fully Homomorphic Encryption (FHE) through an innovative approach applied to a confidential dataset composed of mammograms from Belgian medical records. Our goal is to clarify the feasibility and challenges associated with integrating FHE into the context of Federated Learning, with a particular focus on evaluating the memory constraints inherent in FHE when using sensitive medical data. The results highlight notable limitations in terms of memory usage, underscoring the need for ongoing research to optimize FHE in real-world applications. Despite these challenges, our research demonstrates that FHE maintains comparable performance in terms of Receiver Operating Characteristic (ROC) curves, affirming the robustness of our approach in secure machine learning applications, especially in sectors where data confidentiality, such as medical data management, is imperative. The conclusions not only shed light on the technical limitations of FHE but also emphasize its potential for practical applications. By combining Federated Learning with FHE, our model preserves data confidentiality while ensuring the security of exchanges between participants and the central server

Index Terms—breast cancer, masses and microcalcifications detection, federated learning, homomorphic encryption, convolutional neural networks.

Abstract of Paper Accepted in ICAIC'2024

7169

Federated Learning Based Smart Horticulture and Smart Storage of Fruits Using E-Nose, and Blockchain: A Proposed Model

Shakhmaran Seilov, Akniyet Nurzhaubayev, Dias Abildinov, Assem Konyrkhanova
Faculty of Information Technology
Eurasian National University Astana, Kazakhstan
seilov@mail.ru, nurzhaubaev.akniet@gmail.com, abildinov.ds@yandex.kz, erkeshank@mail.ru
Bishwajeet Pandey
Dept of CSE UCSI University Kuala Lumpur, Malaysia
dr.pandey@ieee.org
Bibinur Zhursinbek
AI Department
Kazakh Infocommunications Academy Astana, Kazakhstan zhursinbek99@gmail.com

ABSTRACT

The main objective of this project is to increase the productivity of farmers producing fruits and vegetables in Kazakhstan. We are planning to use technology during production at orchards and also using technology during storage. At the production stage, we shall capture images of fruit flowers, growing fruits, and a ripe fruit. Then we shall apply federated learning to train our model with healthy fruits and flowers and then we shall be able to predict any ongoing pest infections with either fruits or flower. At the storage phase, we shall use e-nose to check the current status of apple and save it from any possible degradation. We shall also use blockchain to store data related to fruits at both stages of production and storage to create an e-passport that will give access to data related to production and storage of fruits. At the same time, we shall also use various width clustering algorithms to detect intrusion in our sensor based IoT networks.

Index Terms—IoT, Blockchain, Federated Learning, EPassport, Intrusion Detection, E-Nose, Width Clustering Algorithm

Abstract of Paper Accepted in ICAIC'2024

7383

A Holistic Review on Detection of Malicious Browser Extensions and Links using Deep Learning

Rama Abirami K[1,2], Tiago Zonta[2], Mithileysh Sathiyanarayanan[3]
[1]Curtin University, Malaysia
[2]University of Western Santa Catarina, Brazil
[3]MIT Square London, UK
rama.abirami@curtin.edu.my, tiago.zonta@unoesc.edu.br,
s.mithileysh@gmail.com

ABSTRACT

The growth of the Internet has aroused people's attention toward network security. A secure network environment is fundamental for the expeditious and impeccable development of the Internet. The majority of internet-based tasks can be completed with the help of a web browser. Although many web applications add browser extensions to improve their functionality, some of these extensions are malicious and can access sensitive data without the user's knowledge. Browser extensions with malicious intent present a growing security concern and have quickly become one of the most prevalent methods used to compromise Internet security. This is largely due to their widespread usage and the extensive privileges they possess. After being installed, these malicious extensions are executed and make an attempt to compromise the victim's browser. This makes them particularly elusive and challenging to combat. It is crucial to promptly develop an effective strategy to address the threats posed by these extensions. A comprehensive review of the research on browser extension vulnerabilities is presented in this paper. The role of malicious links in web browser extensions are examined for several attacks. Detection of malicious browser extension on various aspects are represented namely Intrusion malicious web browser extensions detection using Intrusion detection, Machine learning based detection methods and Deep learning based techniques to mitigate malicious web browser extensions are examined. This study investigates the critical function of malicious detection in protecting web browsers, looking at the changing threats and risk-reduction tactics. A robust cybersecurity frameworks can be created that not only respond to known threats but also anticipate and thwart the strategies of future cyber adversaries by realizing the significance of proactive detection. Thus this survey provides a detailed comparison of various solutions for malicious browser extension.

Index Terms— Cyber-attacks, Machine learning, Malicious Browser extension, and Malicious Uniform Resource Locator (URL).

Abstract of Paper Accepted in ICAIC'2024

7498

Navigating Data Privacy and Analytics: The Role of Large Language Models in Masking conversational data in data platforms

Mandar Khoje
Independent Researcher, Dublin, California, USA
Email: mandar.khoje@gmail.com

ABSTRACT

In the rapidly evolving landscape of data analytics, safeguarding conversational data privacy presents a pivotal challenge, especially with third-party enterprises commonly offering analytic services. This paper delves into the innovative application of Large Language Models (LLMs) for real-time masking of sensitive information in conversational data. The focus is on balancing privacy protection and data utility for analytics within a multi-stakeholder framework. The significance of data privacy is underscored across sectors, with specific attention to challenges in industries like healthcare, particularly when analytics involve external entities. A comprehensive literature review reveals limitations in existing data masking techniques and explores the role of LLMs in diverse contexts, extending beyond direct healthcare applications. The proposed methodology utilizes LLMs for real-time entity recognition and replacement, effectively masking sensitive information while adhering to privacy regulations. This approach is particularly pertinent for third-party analytics providers dealing with conversational data from various sources. Hypothetical case studies, including healthcare scenarios, showcase the practical application and efficacy of the method in real-world settings with external data analytics providers. The dual assessment evaluates the method's efficiency in preserving privacy and maintaining data utility for analytical purposes. Experimental results using synthetically generated [1] healthcare conversational data sets further illustrate the effectiveness of the approach in typical third-party analytics service scenarios. The discussion highlights broader implications, addressing challenges and limitations [2] across industries, and emphasizes ethical considerations in handling sensitive data by external entities. In conclusion, the paper summarizes the significant strides achievable with LLMs for data masking, with implications for diverse sectors and analytics providers. Future research directions, especially fine-tuning LLMs for enhanced performance in varied analytic scenarios, are suggested. This study sets the stage for a harmonious coexistence of customer data protection and utility in the intricate ecosystem of data analytics services, facilitated by the advanced capabilities of LLM technology.

Index Terms—Data Analytics, Data Platform, AI, LLMs, Data Masking, Healthcare

Abstract of Paper Accepted in ICAIC'2024

7600

Video key concept extraction using Convolution Neural Network

Tanvir H Sardar, Ruhul Amin Hazarika, Bishwajeet Pandey, Guru Prasad M S, Sk Mahmudul Hassan, Radhakrishna Dodmane, Hardik Gohel
School of Technology (GST), GITAM University, Bengaluru, India
Manipal Institute of Technology Bengaluru, India
Astana IT University, Astana, Kazakhstan
Graphic Era (Deemed to be University), Dehradun, India
School of Computer Science and Engineering, VIT-AP, India
NMAM Institute of Technology Nitte, India
University of Houston, USA
tsardar@gitam.edu, rahazarika@gmail.com, bk.pandey@astanait.edu.kz, guru0927@gmail.com, hassan.m@vitap.ac.in, rkdodmane@gmail.com
gohelh@uhv.edu

ABSTRACT

This work aims to develop an automated video summarising methodology and timestamping that uses natural language processing (NLP) tools to extract significant video information. Methods: The methodology comprises extracting the audio from the video, splitting it into chunks by the size of the pauses, and transcribing the audio using Google's speech recognition. The transcribed text is tokenised to create a summary, sentence and word frequencies are calculated, and the most relevant sentences are selected. The summary quality is assessed using ROUGE criteria, and the most important keywords are extracted from the transcript using RAKE. Our proposed method successfully extracts key points from video lectures and creates text summaries. Timestamping these key points provides valuable context and facilitates navigation within the lecture. Our method combines video-to-text conversion and text summarisation with timestamping key concepts, offering a novel approach to video lecture analysis. Existing video analysis methods focus on keyword extraction or summarisation, while our method offers a more comprehensive approach. Our timestamped key points provide a unique feature compared to other methods. Our method enhances existing video reports by (i) providing concise summaries of key concepts and (ii) enabling quick access to specific information through timestamps. (iii) Facilitating information retrieval through a searchable index. Further research directions: (i) Improve the accuracy of the multi-stage processing pipeline. (ii) Develop techniques to handle diverse accents and pronunciations. (iii) Explore applications of the proposed method to other video genres and types. This approach is practical in giving accurate video summaries, saving viewers time and effort when comprehending the main concepts presented in a video.

Index Terms—CNN, transcribe, tokenise, ROUGE, RAKE

Abstract of Paper Accepted in ICAIC'2024

7616

CANAL - Cyber Activity News Alerting Language Model : Empirical Approach vs. Expensive LLMs

Urjitkumar Patel, Fang-Chun Yeh, Chinmay Gondhalekar
Ratings Data Science
S&P Global
New York, USA
urjitkumar.patel@spglobal.com, jessie.yeh@spglobal.com,
chinmay.gondhalekar@spglobal.com

ABSTRACT

Abstract—In today's digital landscape, where cyber attacks have become the norm, the detection of cyber attacks and threats is critically imperative across diverse domains. Our research presents a new empirical framework for cyber threat modeling, adept at parsing and categorizing cyber-related information from news articles, enhancing real-time vigilance for market stakeholders. At the core of this framework is a fine-tuned BERT model, which we call CANAL - Cyber Activity News Alerting Language Model, tailored for cyber categorization using a novel silver labeling approach powered by Random Forest. We benchmark CANAL against larger, costlier LLMs, including GPT-4, LLaMA, and Zephyr, highlighting their zero to few-shot learning in cyber news classification. CANAL demonstrates superior performance by outperforming all other LLM counterparts in both accuracy and cost-effectiveness. Furthermore, we introduce the Cyber Signal Discovery module, a strategic component designed to efficiently detect emerging cyber signals from news articles. Collectively, CANAL and Cyber Signal Discovery module equip our framework to provide a robust and cost-effective solution for businesses that require agile responses to cyber intelligence.

Index Terms—Large Language Models (LLM), BERT, Natural Language Processing (NLP), Machine Learning, Generative AI (Gen AI), Cyber Risk Modeling, Cyber Signal Discovery, Cyber News Alerts, Empirical Cost Analysis

Abstract of Paper Accepted in ICAIC'2024

7743

Sentiment Analysis of Financial News Data using TF-IDF and Machine Learning Algorithms

Gideon Popoola[1], Khadijat-Kuburat Abdullah[2], Gerard Shu Fuhnwi[1], Janet Agbaje[1]
[1]Gianforte School of Computing Montana State University Bozeman, Montana, 59715, USA
[2]Olabisi Onabanjo University Ago-Iwoye, Ogun state, Nigeria
gideon.popoola@student.montana.edu, abdullah.adebisi@oouagoiwoye.edu.ng, gerard.shufuhnwi@student.montana.edu, jagbaje@mtech.edu

ABSTRACT

Blogs, online forums, comment sections, and social networking sites like Facebook, Twitter (now known as X), and Instagram can all be called social media. The growing use of social media has made some unstructured data available, which can benefit us if we clean, structure, and analyze the data. Twitter is a popular microblogging social media platform where people share and express their opinions about any topic. The act of analyzing these opinions of people is called sentimental analysis. Sentimental analysis can be helpful to individuals, businesses, government agencies, etc. In this study, tweets related to financial news were extracted, labeled, and analyzed to capture the opinions of people around the world. This paper proposes a novel machine learning-based approach to analyze social media data for sentiment analysis. The presented approach is divided into three steps. The first stage is preprocessing, where the tweets are refined and filtered. In the second stage, feature extraction was performed using Term Frequency and Inverse Document Frequency (TF-IDF). The third stage involves using the extracted features to make predictions using machine learning algorithms. Three machine learning models were used, namely, random forest classifier (RF), Na¨ıve Bayes (NB), and k-nearest neighbor (KNN). The evaluation results show that both NB and RF perform better than KNN in accuracy, precision, Recall, and F1-score metrics. These results also show an overwhelmingly positive opinion regarding financial news.

Index Terms— Sentiment Analysis, Natural Language Processing, Financial News, Term Frequency-Inverse Document Frequency (TF-IDF), Machine Learning Algorithms.

Abstract of Paper Accepted in ICAIC'2024

7859	**Robotics in Healthcare: The African Perspective**

[1]Dr Chipo Mutongi and [2]Dr Billy Rigava
[1]PhD-Inf and KM, MBA, MSc, BA-Media Studies HND-LIS, Ex Dip-Local Gvt and Admin, Dip-Edu, Dip-LIS, Dip PM-IPMZ, Dip Sal-Admin, Paralegal Training Cert
[2]MMed.FM (Stellenbosch RSA) MBA (UZ), MB.CHB (UZ)
mutongic@gmail.com, mutongic@staff.msu.ac.zw

ABSTRACT

Today one has to run very fast to stay on the same position. We are no longer competing with humans only, we are now also competing with robots as they are involving in learning leading to machine learning. Robots are increasingly being adopted in healthcare to carry out various tasks that enhance patient care. Robots in health care have revolutionized the health ecosystem. There are different types of healthcare robots which include nursing robots, surgical robots, clinical Training, Prescription Dispensing, care robots, Telepresence, Rehabilitation Robots, Health Call Centre Robots, Nursing Robots, Ambulance Robots and Physical Therapy Robots. Healthcare robots are mostly found in the developed countries. This paper seeks to establish robotics in healthcare considering the African perspectives and Zimbabwe in particular. A qualitative study was conducted whereby twenty students at a university were interviewed concerning their views on robots in the African context. It was found out that healthcare robots are still at their conception in Africa and Zimbabwe in particular, there is fear of the unknown, some indicated that robots will affect their indigenous way of life as they are used to interact with each other as human beings and not as robot to human, power challenges, connectivity, lack of awareness challenges, as well as cultural and religious challenges. However, some participants indicated that they greatly welcome the robots as they may cease the health professional shortages in Africa and also they consider them to be more precise and accurate as compared to humans. Some indicated that more privacy will be promoted due to the use of robots hence feeding to the African Ubuntu concept. It was recommended that there is need for immense healthcare robots conscientisation, awareness, training, robots to mimic the African way of living and language.

Index Terms— robots, robotics, healthcare, artificial intelligence, African perspective, health ecosystem

Abstract of Paper Accepted in ICAIC'2024

8520	**Prescriptive Analytics-based Robust Decision Making Model for Cyber Disaster Risk Reduction**

Joseph Ponnoly, John Puthenveetil, Patricia D'Urso
College of Doctoral Studies Grand Canyon University Phoenix, AZ, USA
jponnoly@my.gcu.edu, john.puthenveetil@aya.yale.edu,
pat.durso@my.gcu.edu

ABSTRACT

Decision-making in cyber security attack scenarios involves deep uncertainty and adversarial decision-making. Robust Decision Making (RDM) uses a structured approach to evaluate the performance of various decision strategies under conditions of deep uncertainty to enable adaptive decision-making. Prescriptive analytics enabled by predictive analytics including big data analytics, reinforcement learning, and Monte Carlo simulations, could provide decision-makers with various options to make informed and robust decision-making. Prescriptive analytics based RDM model is proposed for cyber disaster risk reduction. The model extends the proactive cyber defense model based on sensing and sensemaking of early warning signs of cyber disasters.

Index Terms— prescriptive analytics, cyber disaster risk reduction, robust decision-making, proactive cyber defense

Abstract of Paper Accepted in ICAIC'2024

8623

Leveraging Advanced Visual Recognition Network Classifier for Pneumonia Prediction

Maulin Raval, Jin Aobo, Hardik Gohel
Department of Computer Science University of Houston - Victoria Victoria, Texas, USA
ravalm@uhv.edu, jina @uhv.edu, gohelh@uhv.edu

ABSTRACT

Pneumonia prediction using chest X-ray images is a challenging task because of the complex image processing involved. The radiographic features of pneumonia, especially in the earlier stages, easily overlap with other lung conditions, which makes the differentiation even more challenging. Moreover, X-ray image quality varies due to equipment, patient condition, and techniques, particularly in rural areas with undertrained radiologists and medical experts. The use of Artificial Intelligence (AI) models in detecting pneumonia is a novel but crucial research field and rapid advancement in medical imaging technology and neural network models along with the availability of large de-identified public datasets has paved the way for this life-saving biomedical research. In this paper, we propose a unique comprehensive solution for predicting pneumonia using chest X-ray images. We utilize an enhanced VGGNet model tailored for the binary classification task. The modified VGG19 with a binary classifier provides a solid foundation for feature extraction and leverages the pretrained features and deep architecture to differentiate between normal and pneumonia-affected lung images.

Index Terms— deep learning, transfer learning, biomedical classification, pneumonia prediction, x-ray imaging, vision-based, image processing

Abstract of Paper Accepted in ICAIC'2024

9618

Simulations and Advancements in MRI-Guided Power-Driven Ferric Tools for Wireless Therapeutic Interventions

Wenhui Chu[†], Aobo Jin[†], Hardik A. Gohel [†]
[†]Dept. of Computer Science, University of Houston-Victoria, Victoria, USA.
ChuW1@uhv.edu, Jina@uhv.edu, gohelh@uhv.edu

ABSTRACT

Designing a robotic system that functions effectively within the specific environment of a Magnetic Resonance Imaging (MRI) scanner requires solving numerous technical issues, such as maintaining the robot's precision and stability under strong magnetic fields. This research focuses on enhancing MRI's role in medical imaging, especially in its application to guide intravascular interventions using robot-assisted devices. A newly developed computational system is introduced, designed for seamless integration with the MRI scanner, including a computational unit and user interface. This system processes MR images to delineate the vascular network, establishing virtual paths and boundaries within vessels to prevent procedural damage. Key findings reveal the system's capability to create tailored magnetic field gradient patterns for device control, considering the vessel's geometry and safety norms, and adapting to different blood flow characteristics for finer navigation. Additionally, the system's modeling aspect assesses the safety and feasibility of navigating pre-set vascular paths. Conclusively, this system, based on the Qt framework and C/C++, with specialized software modules, represents a major step forward in merging imaging technology with robotic aid, significantly enhancing precision and safety in intravascular procedures.

Index Terms—MRI, medical imaging, robot-assisted devices, magnetic field gradient

Next Conference

4th International Conference On Business, Management, Emerging technologies, and Social Science 2024 (BMESS®-2024)

https://gyancity.com/bmess/

25-26 April 2024

Bath Spa University, Academic Center Ras Al-Khaimah-UAE

7th International Multi-Topic Conference on Engineering and Science (IMCES ®)

https://imces.tech/

25-26 April 2024

Bath Spa University, Academic Center Ras Al-Khaimah-UAE