Proceedings of the 9th National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics

NCVPRIPG 2024

18th - 20th July

IIST - Thiruvananthapuram

भारतीय अन्तरिक्ष अनुसंधान संगठन
अन्तरिक्ष विभाग
भारत सरकार
अन्तरिक्ष भवन
न्यू बी ई एल रोड, बेंगलूर – 560 094, भारत
दूरभाष : +91-80-2341 5241 / 2217 2333
फैक्स : +91-80-2341 5328

Indian Space Research Organisation
Department of Space
Government of India
Antariksh Bhavan
New BEL Road, Bangalore - 560 094, India
Telphone: +91-80-2341 5241 / 2217 2333
Fax : +91-80-2341 5328
e-mail : chairman@isro.gov.in
secydos@isro.gov.in

डॉ.एस. सोमनाथ / Dr. S. SOMANATH
अध्यक्ष / Chairman

MESSAGE

It is heartening to note that the 9th edition of the ***National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG-2024)*** is being hosted by the Department of Avionics, IIST and organized by the Indian Unit of Pattern Recognition and Artificial Intelligence (IUPRAI).

This esteemed conference, since its inception in 2008, has been a cornerstone for researchers and practitioners dedicated to advancing the frontiers of computer vision, pattern recognition, image processing, and graphics. The themes of this year's conference—future cameras, diffusion model generation, machine intelligence, and more—are particularly timely as they align with the cutting-edge innovations that drive both scientific discovery and technological advancement.

At ISRO, we recognize the transformative potential of these technologies. They play critical roles in enhancing our capabilities and in addressing the challenges faced in space exploration and beyond. The synergy between these fields and our work at ISRO is evident as we continually seek to push the boundaries of what is possible.

I compliment the efforts of all participants, organizers, and contributors for their unwavering commitment to excellence. Your work not only contributes to the body of knowledge but also paves the way for future innovations that can benefit society as a whole.

I look forward to the insightful discussions, groundbreaking presentations, and collaborative spirit that NCVPRIPG-2024 promises. May this conference be a springboard for new ideas, fruitful collaborations, and significant advancements in our fields of interest.

Wishing you all a productive and enriching conference experience.

Dated: July 16, 2024

(डॉ. एस. सोमनाथ / Dr. S. Somanath)

सत्यमेव जयते

भारतीय अंतरिक्ष विज्ञान एवं प्रौद्योगिकी संस्थान
(वि.अ.आयोग अधिनियम 1956 की धारा-3 के अधीन मानित विश्वविद्यालय घोषित)
भारत सरकार, अंतरिक्ष विभाग, वलियमला पोस्ट, तिरुवनंतपुरम 695 547 भारत

INDIAN INSTITUTE OF SPACE SCIENCE AND TECHNOLOGY
(A Deemed to be University u/s 3 of the UGC Act, 1956)
Government of India, Department of Space
Valiamala P. O, Thiruvananthapuram 695 547 India

www.iist.ac.in

MESSAGE

I am delighted to extend a warm welcome to all the delegates, speakers, and participants of the ninth edition of the National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG-2024), organized by the Indian Unit of Pattern Recognition and Artificial Intelligence (IUPRAI) and hosted by the Department of Avionics, IIST. This conference series, initiated in 2008, has grown to become a significant event for researchers and practitioners in the field, providing a platform to share pioneering research, innovative ideas, and valuable insights.

The theme for this year's conference, 'Computer Vision, Pattern Recognition, Image Processing and Graphics,' reflects the rapidly evolving landscape of these domains and their profound impact on technology and society. I am particularly pleased to note the inclusion of plenary talks on future cameras, controlling diffusion model generation to user constraints, and machine intelligence. These topics are at the forefront of research and hold great promise for advancing knowledge and applications in these fields.

The invited research talk sessions, student research symposium, and poster presentations will offer invaluable opportunities for young researchers to showcase their work and engage with experts. I commend the organizers for their efforts in putting together such a comprehensive and engaging program.

I convey my very best wishes to the organizers and all participants.

(Dr B N Suresh)
Chancellor, IIST

सत्यमेव जयते

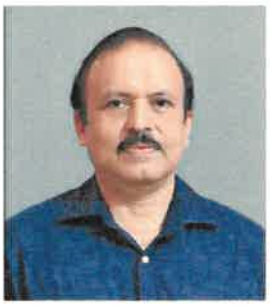
भारतीय अंतरिक्ष विज्ञान एवं प्रौद्योगिकी संस्थान
(वि.अ.आयोग अधिनियम 1956 की धारा-3 के अधीन मानित विश्वविद्यालय घोषित)
भारत सरकार, अंतरिक्ष विभाग, वलियमला पोस्ट, तिरुवनंतपुरम 695 547 भारत

INDIAN INSTITUTE OF SPACE SCIENCE AND TECHNOLOGY
(A Deemed to be University u/s 3 of the UGC Act, 1956)
Government of India, Department of Space
Valiamala P. O, Thiruvananthapuram 695 547 India

www.iist.ac.in

दूरभाष (Tel): +91 471 2568402 फैक्स (Fax): +91 471 2568401 ई-मेल (E-mail): unnikrishnan@iist.ac.in

डॉ. उण्णिकृष्णन नायर एस / Dr. Unnikrishnan Nair S
निदेशक / Director



MESSAGE

Indian Institute of Space Science and Technology (IIST) is proudly hosting the ninth edition of the National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG-2024). Organised by the Department of Avionics, IIST, in association with the Indian Unit of Pattern Recognition and Artificial Intelligence (IUPRAI). This conference serves as a pivotal platform for researchers and practitioners to explore the forefront of advancements in computer vision, pattern recognition, image processing, and graphics.

In this confluence of experts, it is inspiring to witness the dedication and expertise that participants bring to this dynamic field. The theme of this year's conference underscores the transformative potential of these technologies in shaping our future. We are privileged to have distinguished experts delivering plenary talks on topics such as future cameras, controlling diffusion model generation to user constraints, and machine intelligence. These discussions not only highlight current breakthroughs but also set the stage for envisioning the future avenues of research and development.

As we embark on this conference, let us collectively explore how we can leverage our discussions and collaborations to address the pressing challenges and opportunities ahead. Together, we can chart a course towards new discoveries and innovations that will impact society and industry alike. As a national Institute dedicated to nurture and mould scientific and technological manpower in advanced disciplines related to space transportation and exploration, IIST is evolving courses and identifying research domain pertaining to the exciting new areas being presented and deliberated in NCVPRIPG-2024.

I extend my warmest welcome to all participants. I look forward to fruitful discussions, productive exchanges, and new insights that will emerge from NCVPRIPG-2024.

(Unnikrishnan Nair S)
Director, VSSC/IIST

भारत सरकार
अंतरिक्ष विभाग
द्रव नोदन प्रणाली केंद्र
वलियमला पोस्ट
तिरुवनंतपुरम - 695 547, भारत
दूरभाष : 0471 2567257/2567554
फैक्स : 0471 2567242
इमेल : director@lpsc.gov.in

डॉ. वी. नारायणन/ Dr. V. Narayanan
निदेशक/Director

Government of India
Department of Space
Liquid Propulsion Systems Centre
Valiamala P.O.
Thiruvananthapuram - 695 547, India
Telephone : 0471 2567257/2567554
Fax : 0471 2567242
Email : director@lpsc.gov.in

Message

It gives me great pleasure to note that IIST, Trivandrum is organizing a 9th National conference on Computer vision, Pattern recognition, Image processing and Graphics in this beautiful IIST campus. Let me congratulate the organizers for selecting this high technology intensive and most relevant area as the topic for the national conference. I am sure, the deliberations in this conference shall not be limited to space sector alone but shall deal with the overall spectrum of activities in this very latest and important domain.

This seminar shall definitely put forward the space vision for AI/ML which shall go a long way in incorporating this high end technology across our systems. I am sure, it will help us in a big way as we venture into our new initiatives like Gaganyaan continuation programme, Bharathiya Anthariksh station etc. This conference shall provide a birds eye view into high throughput vision based phenotyping, contactless hand biometrics, gesture analysis etc. It shall look into the larger picture of LARGE Language Models (LLMs) and multimodal LLMs for information extraction. Deep Learning models enabling us land cover mapping from high resolution satellite images shall also form a main topic of interest in this conference.

I am sure this National conference shall provide an ideal platform for technical and managerial interaction between related professionals and sharing of expertise between Researchers, Academia and Industries. I am all the more happy that a large number of student community from IIST and elsewhere are participating to derive benefits from this event. At this juncture, let me once again extend my best wishes for NCVPRIPG – 2024.

V Narayanan

भारत सरकार
अंतरिक्ष विभाग
इसरो नोदन कॉम्प्लेक्स (आईपीआरसी)
महेंद्रगिरि पी. ओ., तिरुनेलवेली जिला - 627 133
तमिलनाडु, भारत
दूरभाष : 04637 281900 (प्रचालक)
फैक्स : 04637 281618 (प्रशासन)
04637 232661 (क्रय)
04637 281567 (भण्डार)
04637 281541 (लेखा)

Government of India
Department of Space
ISRO Propulsion Complex (IPRC)
Mahendragiri P.O., Tirunelveli District - 627 133,
Tamil Nadu, India.
Telephone : 04637 281900 (Operator)
Fax : 04637 281618 (Administration)
04637 232661 (Purchase)
04637 281567 (Stores)
04637 281541 (Accounts).

J. ASIR PACKIARAJ,
Distinguished Scientist &
Director

Fax : (04637) 281447
Phone : (04637) 281444

Message

Greetings to the organizers and the participants of the National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG-2024), It gives me immense pleasure to welcome you all to this prestigious conference organized by the Department of Avionics, IIST and the Indian Unit of Pattern Recognition and Artificial Intelligence (IUPRAI). This event serves as a vital platform for exploring the latest advancements and innovations in computer vision, pattern recognition, image processing, and graphics.

At ISRO Propulsion Complex (IPRC) in Mahendragiri, Tamil Nadu, we witness the transformative impact of advanced technologies in aerospace engineering. The theme of this conference is pivotal in shaping the future of scientific research and application.

I wish engaging discussions, insightful presentations, and meaningful collaborations during NCVPRIPG-2024. Let us harness this opportunity to exchange ideas, foster partnerships, and explore new avenues for research and development.

Wishing everyone a productive and inspiring conference experience.

(J. ASIR PACKIARAJ)

भारत सरकार
अंतरिक्ष विभाग
इसरो जड़त्वीय प्रणाली यूनिट
वट्टियूरकाव, तिरुवनंतपुरम-695 013
दूरभाष : +91(471)2569333/2569666
ई-मेल : iisu_director_office@vssc.gov.in

सत्यमेव जयते

Government of India
Department of Space
ISRO Inertial Systems Unit
Vattiyoorkavu, Thiruvananthapuram-695 013
Phone : +91(471)2569333/2569666
e-mail : iisu_director_office@vssc.gov.in

पद्मकुमार ई एस/Padmakumar ES
विशिष्ट वैज्ञानिक/Distinguished Scientist
निदेशक/Director

MESSAGE

It is heartening to note that the 9th National Conference on Computer Vision, Pattern Recognition, Image Processing, and Graphics (NCVPRIPG) is being organised by IIST Trivandrum in collaboration with the Indian Unit for Pattern Recognition and Artificial Intelligence (IUPRAI) during 18th to 20th July 2024. The topics cover some of the most important and booming areas of research interest in current times.

The capability to make machines sense and understand the world around them visually, just the way humans do, is one of the most challenging aspects of current human endeavour. In the last two decades spurred by the availability of very high computing capabilities across the platforms, this capability has witnessed significant progress. Self-driving cars, various kinds of Autonomous Robots, Medical diagnosis, Manufacturing, and Remote sensing applications are some of the highlights we all will immediately relate to. But we know the potential goes far beyond these few examples. I am sure that the extensive progress in Computer vision, duly supported by Artificial Intelligence and Machine Learning, will truly transform, in the near future, the world that we live in. Hence it is really imperative to debate, formulate, and pursue the most potent and reliable research ideas on computer vision algorithms.

I am sure that this National conference will generate a significant impetus in this direction, especially for the space sector, providing a forum where engineers, researchers, and academicians can network and exchange and also set certain goals to be achieved immediately.

I wish the conference a great success and convey my best wishes to all delegates for fruitful interactions and deliberations.

Thiruvananthapuram
17-July-2024

(पद्मकुमार ई एस /
Padmakumar ES)

Prof. P.J.Narayanan
DIRECTOR

International Institute of Information Technology, Hyderabad
A Research University

15 July 2024

To

Prof. Deepak Mishra,
Professor,
Indian Institute of Space Science and Technology,
Trivandrum.

I welcome you all to the National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG-2024), hosted by the Department of Avionics, IIST, as its General Co-Chair. This is being conducted in collaboration with the Indian Unit of Pattern Recognition and Artificial Intelligence (IUPRAI). I am sure all of you will have wonderful moments in this conference.

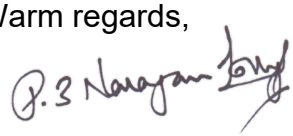
I am delighted to witness the convergence of leading researchers, practitioners, and scholars in the fields of computer vision, pattern recognition, image processing, and graphics. NCVPRIPG and ICVGIP are major forums for Indian researchers, students, and others to meet and exchange ideas in these ever-important areas. This conference serves as a pivotal platform for sharing cutting-edge research, fostering collaboration, and exploring new avenues of innovation.

I extend my gratitude to all participants for their valuable contributions towards advancing scientific knowledge and technological advancements. Your dedication and expertise are crucial in addressing the complex challenges and opportunities in our rapidly evolving world.

I encourage you to engage actively in the diverse sessions, including plenary talks, invited research sessions, student research symposium, and poster presentations. Let us seize this opportunity to exchange ideas, forge new partnerships, and collectively envision the future of our disciplines.

I wish you all a fruitful and inspiring conference experience. May NCVPRIPG-2024 be a catalyst for transformative insights and impactful discoveries.

Warm regards,



(P. J. Narayanan)
General Chair

Prof. C. R. Rao Road, Gachibowli, Hyderabad - 500 032, Telangana State, India
Phone: +91-40-6653 1144, Fax: +91-40-6653 1413 e-mail: director@iiit.ac.in URL: http://www.iiit.ac.in

सत्यमेव जयते

भारतीय अंतरिक्ष विज्ञान एवं प्रौद्योगिकी संस्थान
(वि.अ.आयोग अधिनियम 1956 की धारा-3 के अधीन मानित विश्वविद्यालय घोषित)
भारत सरकार, अंतरिक्ष विभाग, वलियमला पोस्ट, तिरुवनंतपुरम 695 547 भारत

INDIAN INSTITUTE OF SPACE SCIENCE AND TECHNOLOGY
(A Deemed to be University u/s 3 of the UGC Act, 1956)
Government of India, Department of Space
Valiamala P. O, Thiruvananthapuram 695 547 India

www.iist.ac.in

MESSAGE

This year's conference is filled with opportunities for learning, collaboration, and innovation. Our robust technical program includes three challenges, five workshops, four tutorials, numerous invited talks, a student research symposium, and industry panel discussions. These events offer ample occasions for individual and group discussions on recent advancements in Computer Vision, Image Processing, Graphics, and AI & Machine Learning applications.

We receive overwhelming participation both from students, academicians and industry delegates. As researchers immersed in the fields of computer vision, robotics, and augmented reality, we eagerly anticipate the discussions and innovations that will arise from this conference. These fields are at the cutting edge of technological progress, and we are confident that our collective expertise will drive significant contributions and breakthroughs.

Thank you for being part of NCVPRIPG-2024. We look forward to a successful and inspiring conference.

Warm Regards,

(Prof. Kuruvilla Joseph)
Local Organizing Committee Chair

Welcome Message from the Vice Chancellor of Digital University Kerala

Dear Participants and Esteemed Guests,

It is my great honor and pleasure to welcome you to the ninth edition of the National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG-2024). This prestigious event, organized by the Indian Unit of Pattern Recognition and Artificial Intelligence (IUPRAI) and hosted by the Department of Avionics, IIST, serves as a vital platform for the exchange of innovative ideas and the exploration of cutting-edge research in these dynamic fields.

As the Vice Chancellor of the Kerala University of Digital Sciences, Innovation and Technology (Digital University Kerala), I am particularly proud to witness the convergence of such brilliant minds dedicated to pushing the boundaries of computer vision, pattern recognition, image processing, and graphics. These disciplines are not only transforming technology but also driving forward a multitude of applications that impact our daily lives and global progress.

The themes of this year's conference—ranging from future cameras and diffusion model generation to machine intelligence—are of paramount importance in today's digital age. The insights and advancements shared here will undoubtedly contribute to the broader goals of innovation and sustainable development.

I extend my heartfelt appreciation to all the distinguished speakers, researchers, and participants for their invaluable contributions. Your dedication and hard work are the backbone of this conference's success and the driving force behind the advancements in our fields.

I am confident that NCVPRIPG-2024 will be an inspiring and enriching experience for all attendees. May this conference spark new ideas, foster meaningful collaborations, and pave the way for groundbreaking discoveries.

Welcome to NCVPRIPG-2024, and I wish you all a productive and memorable conference.

Warm regards,

Saji Gopinath
Vice Chancellor
Kerala University of Digital Sciences, Innovation and Technology (Digital University Kerala)

सत्यमेव जयते

भारतीय अंतरिक्ष विज्ञान एवं प्रौद्योगिकी संस्थान
(वि.अ.आयोग अधिनियम 1956 की धारा-3 के अधीन मानित विश्वविद्यालय घोषित)
भारत सरकार, अंतरिक्ष विभाग, वलियमला पोस्ट, तिरुवनंतपुरम 695 547 भारत

INDIAN INSTITUTE OF SPACE SCIENCE AND TECHNOLOGY
(A Deemed to be University u/s 3 of the UGC Act, 1956)
Government of India, Department of Space
Valiamala P. O, Thiruvananthapuram 695 547 India

www.iist.ac.in

FOREWORD

The Department of Avionics, IIST is hosting the ninth edition of the National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG-2024), organized annually by the Indian Unit of Pattern Recognition and Artificial Intelligence (IUPRAI). The NCVPRIPG series began with the "National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)-2008". After the success of the initial conference, subsequent editions have been hosted by various esteemed institutes across the country. The current conference, NCVPRIPG - 2024, focuses on the theme 'Computer Vision, Pattern Recognition, Image Processing and Graphics'. The conference aims to present current research in these areas, highlighting the latest advances and strategies for the design and development of novel algorithms and applications. Three plenary talks, covering topics *viz.*, future cameras, controlling diffusion model generation to user constraints, and machine intelligence, will be delivered by eminent experts discussing recent trends in these areas. Additionally, there will be invited research talk sessions and a student research symposium. A few student delegates will be given the opportunity for oral presentations as well.

I sincerely thank the Director, IIST, for the support rendered towards hosting this conference. I thank all the members of IUPRAI for their encouragement in organizing the conference. I take the opportunity to thank Registrar, IIST and all my colleagues from IIST for embracing the concept of the conference and providing wholehearted support and participation in organizing the event. I am confident that NCVPRIPG-2024 will be a memorable and rewarding experience for all the delegates.

Coordinator, NCVPRIPG-2024

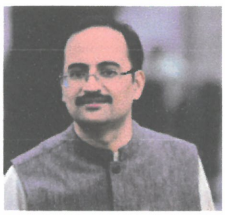
विभागाध्यक्ष / Head of the Department
एविओनिकी विभाग / Department of Avionics
भारतीय अंतरिक्ष विज्ञान एवं प्रौद्योगिकी संस्थान
Indian Institute of Space Science and Technology
अंतरिक्ष विभाग, भारत सरकार
Dept. of Space, Govt. of India
तिरुवनंतपुरम / Thiruvananthapuram - 695547

सत्यमेव जयते

भारतीय अंतरिक्ष विज्ञान एवं प्रौद्योगिकी संस्थान
(वि.अ.आयोग अधिनियम 1956 की धारा-3 के अधीन मानित विश्वविद्यालय घोषित)
भारत सरकार, अंतरिक्ष विभाग, वलियमला पोस्ट, तिरुवनंतपुरम 695 547 भारत

INDIAN INSTITUTE OF SPACE SCIENCE AND TECHNOLOGY
(A Deemed to be University u/s 3 of the UGC Act, 1956)
Government of India, Department of Space
Valiamala P. O, Thiruvananthapuram 695 547 India

www.iist.ac.in



Dear Colleagues, Participants, and Esteemed Guests,

It is our pleasure to welcome you all to the NCVPRIPG-2024 Conference. We are excited to have you join us for what promises to be an enlightening and enriching experience.

This year, our conference is packed with various events designed to foster learning, collaboration, and innovation. We have a highly enriching technical program that includes three challenges, five workshops, four tutorials, multiple invited talks, a student research symposium, and industry panel discussions. There are multiple occasions for one-on-one and group deliberations on recent topics in Computer Vision, Image Processing, Graphics, and AI & Machine Learning applications.

With around 250 participants and 70 speakers, this conference is an excellent opportunity to expand your knowledge, exchange ideas, and network with peers from various parts of the country. As researchers deeply engaged in the fields of computer vision, robotics, and augmented reality, we are particularly excited about the discussions and innovations that will emerge from this conference. These areas are at the forefront of technological advancement, and we are confident that our collective expertise will lead to significant contributions and breakthroughs.

We are grateful to all the participants from various institutions, including IIST, VSSC, LPSC, IPRC, IISU, DUK, IUPRAI, and IISER-TVM, who partnered together to organize and participate in this event jointly. We also express our gratitude to all the sponsors for their support.

Thank you for being a part of NCVPRIPG-2024. We look forward to a successful and inspiring conference.

Warm Regards,

Deepak Mishra (IIST Trivandrum) & Avinash Shrama (IIT Jodhpur)
Program Chairs,
The 9th National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 2024

About NCVPRIPG 2024

Welcome Message from General Chairs and Program Chairs

We are delighted to extend a warm welcome to all participants of the Ninth National Conference on Computer Vision, Pattern Recognition, Image Processing, and Graphics (NCVPRIPG 2024)! The conference will take place from 18 to 20 July 2024, amidst the picturesque surroundings of the IIST campus in Thiruvananthapuram.

NCVPRIPG 2024, organized by the Indian Institute of Space Science and Technology Thiruvananthapuram in collaboration with the Indian Unit for Pattern Recognition and Artificial Intelligence (IUPRAI), promises to be a remarkable event. This conference serves as a platform to bring together researchers and practitioners from diverse domains, including computer vision, graphics, image processing, and pattern recognition. By facilitating community-wide discussions and the exchange of innovative ideas, NCVPRIPG aims to foster continued research and development in this rapidly evolving field.

Throughout the years, NCVPRIPG has grown into a vibrant national conference, attracting enthusiastic participation from numerous students and researchers. This event not only provides an opportunity to present cutting-edge research but also encourages fruitful collaborations and networking among like-minded professionals. We are confident that NCVPRIPG 2024 will be an enriching experience for all attendees. Engage in insightful discussions, attend thought-provoking sessions, and explore the latest advancements in computer vision, pattern recognition, image processing, and graphics and do not forget to explore beautiful city of Thiruvananthapuram and its famous tourist destinations such as Ponmudi Hills, Koyikkal Palace, Kovalam Beach to name a few.

On behalf of the organizing committee, we express our sincere appreciation to all the participants, speakers, sponsors, and volunteers who have contributed to making this conference possible. Your presence and active involvement will undoubtedly contribute to the success of NCVPRIPG 2024.

Once again, welcome to the Ninth National Conference on Computer Vision, Pattern Recognition, Image Processing, and Graphics. May your time here be filled with enlightenment, fruitful discussions, and memorable experiences.

General Chairs: Dr S. Unnikrishnan Nair (Director, IIST/ VSSC), Dr. P. J. Narayanan (Director, IIIT Hyderabad)
Program Chairs: Dr. Deepak Mishra (IIST Thiruvananthapuram), Dr. Avinash Sharma (IIT Jodhpur)

Sponsors and Partners

Premier Sponsors

Diamond Sponsors

Sponsors and Partners

Silver Sponsors

NUGENIX

vehant TECHNOLOGIES

Bronze Sponsors

ASP ALLIED SCIENTIFIC PRODUCTS

IEEE GRSS Geoscience and Remote Sensing Society

Conference Partner

IUPRAI

Sponsors and Partners

Host

Organizing Partners

Sponsors and Partners

Financial Partner

FEDERAL BANK
YOUR PERFECT BANKING PARTNER

Committee

Steering Committee Patrons

Dr. S. Somanath, Secretary DoS, Chairman ISRO
Dr. B. N. Suresh, Chancellor IIST

National Organizing Committee

S. Unnikrishnan Nair, Director, IIST, VSSC
V. Narayanan, Director, LPSC
E. S. Padmakumar, Director, IISU
J. Asir Packiaraj, Director, IPRC
Jarugu Narasimha Moorthy, Director, IISER Thiruvananthapuram
Saji Gopinath, Vice Chancellor, Digital University Kerala

General Chairs

S. Unnikrishnan Nair, Director, IIST, VSSC
P.J. Narayanan, Director, IIIT Hyderabad

Program Chairs

Deepak Mishra, IIST
Avinash Sharma, IIT Jodhpur

Technical Program Chair

Manoj B.S., IIST

Publication Chairs

Arvind Selwal, Central University Jammu
Shanmuganathan Raman, IIT Gandhi Nagar
Rama Krishna Sai Gorthi, IIT Tirupati

Tutorial Chairs

Sinnu Susan Thomas, Digital University Kerala
Vaanathi Sundaresan, IISc Bangalore

Committee

Workshop Chairs

Puneet Gupta, IIT Indore
Konda Reddy Mopuri, IIT Hyderabad

Challenge Chairs

Minesh Mathew, Wadhawani AI
Ravi Kiran Sarvadevabhatla, IIIT Hyderabad

Corporate and Start-up Relation Chairs

Anoop M. Namboodiri, IIIT Hyderabad
Kanishka Tyagi, UHV Technologies, Inc.

Finance Chair

Palash Kumar Basu, IIST

AI-CV in Space Science and Technology Chair

Jishy Samuel, VSSC

Student Research Symposium Chairs

Pratik Mazumder, IIT Jodhpur
Puneet Goyal IIT Ropar & NIMS University
Rajendra Nagar, IIT Jodhpur

Publicity & Website Chairs

Prerana Mukherjee, JNU
Anand Mukhopadhyay, Mathworks
Mukesh Mann, IIIT Sonepat

Local Organizing Chairs

Kuruvilla Joseph, IIST
A. Chandrasekar, IIST
Raju K. George, IIST
Narayanamurthy C. S., IIST

Committee

AI-CV in Space Science and Technology Committee

Chairperson	Jishy Samuel, CGSE, VSSC
Convener	Renjith S, MSA, VSSC
Members	Durai Raj R, IISU
	Deepak Mishra, IIST
	Vidya L, SR, VSSC
	Umasankari S S, SR, VSSC
	Viji H, MSA, VSSC
	Priya P, AVN, VSSC

Institute Steering Committee

Chairperson	Dr. Kuruvilla Joseph
Co-Chairperson	Dr. N. Selvaganesan
Convener	Dr. Deepak Mishra
Members	Dr. Raju K. George
	Dr. A. Chandrasekar
	Dr. C. S. Narayanamurthy
	Shri. Harish C. S.

Cultural Committee

Chairperson	Dr. Seena V
Co-Chairperson	Dr. Anand Narayanan
Convener	Aswathy P
Members	Dr. Raju K. George
	Dr. Harsh Simha
	Minha Mubarak
	Shilpa A S

Committee

Finance/Purchase Committee

Chairperson	Dr. Palash Babu
Co-Chairperson	Dr. Anoop C S
Convener	Vijay Joshi
Members	Dr. Basudev Majumdar
	Reny Thomas
	Subhash Chandran M B
	Nibin Raj
	Nusrat Praween
	Thomaskutty Mathew

Printing & Publishing/Publicity/Outreach Committee

Chairperson	Dr. Basudeb Ghosh
Co-Chairperson	Dr. Vineeth B S
Convener	Krishnendu Lalaji
Members	Dr. Sourav Bhaumick
	Abdunasar
	Anjana K
	Meera k

Technical Program Committee

Chairperson	Dr. B S Manoj
Co-Chairperson	Dr. Priyadarshan
Convener	Jayakumar T V
Members	Dr. Vani Devi
	Dr. Sam K Zacharia
	Olivia Zacharia
	Dinesh Nayak

Program & Invitation Committee

Chairperson	Dr. S Moosath
Co-Chairperson	Dr. Chris Prema
Convener	Abdul Rahim V C
Members	Nirmal Ninan
	Komal Gaikwad
	Preethi Yadav
	Abhilash

Committee

Food/Transport/Accomodation Committee

Chairperson	Dr. Sooraj R
Co-Chairperson	Dr. J Mary Gladis
Convener	Krishnanunni R A
Members	Vinod Kaimal
	Bindhya K R
	Pradeep Kumar K R

Web & Registration/Venue Management Committee

Chairperson	Dr. Sheeba Rani
Co-Chairperson	Dr. Ramiya A M
Convener	Mahima
Members	Dr. Rajesh Joseph Abraham
	Mohan Sukumar
	Sreeja
	Ananthalakshmi
	Ramzan Basheer
	Suganya M S
	Nirmala Murali

Sponsorship/Funding/Exhibition Committee

Chairperson	Dr. Chimoy Saha
Co-Chairperson	Dr. Immanuel Raja
Convener	Gopika R
Members	Dr. Sudharsan Kaarthik
	Dr. Anindhya Dasgupta
	Dr. Lakshminarayanan
	Shiv Charan Bhoi

Committee

Student Volunteers

Adithya Dileep
Ahna S. Fathima
Amala K
Amba Agnes R P
Arghya Roy
Ashwin Bhasi
C Subbareddy
Deeksha Bhule
Dwaipayan Dutta
Jaydev Kumar Jha
Jeffin George
Jibanjyoti Moharana
Jyothikrishna K V
M D Sohail Ansari
Naufal N
Neethu Sajeev
Nidhikumari
Prabhash Kumar Abhishek
Rohan Swain
Roshan Jabeen K
Salihath P
Sreekavya M K
Srishty Shreya
Sruthi N T
Suganthan C
Vishnu S Varma
Gnaneshwaran A S
K Sahaj Rao
Thejas T
Palaniraj I

Schedule

DAY 1: July 18, 2024 (Thursday)

Time	Event		
07:30 - 08:30	Breakfast		
08:00 - 09:00	Registration (Venue: SAC, Recreation Room)		
09:30 - 10:45	Inauguration (Venue: MPH, SAC)		
10:45 - 11:00	Tea Break		
11:00 - 11:50	Plenary talk I: Algorithms for Single Cell Data Analysis *Prof. Sanghamitra Bandyopadhyay, ISI Kolkata* (Venue: MPH, SAC)		
12:00 - 13:30	Challenge: Automatic Assessment of True/-False Questions in Handwritten Answer Sheet (Venue: MPH, SAC)	Challenge: Smart City Surveillance: Unveiling Indian Person Attributes in Real Time (Venue: Avionics Swayam Prabha Hall)	Panel Discussion: AI4Health @ NCVPRIPG (Venue: Aerospace Conference Hall)
13:30 - 14:30	Lunch		
14:40 - 17:30 (15:30 - Tea Break)	Tutorial: Transforming Human Perception: A Tutorial from Transformer to LLM Architectures (Venue: MPH, SAC)		Tutorial: Deep Learning Models for Land Cover Mapping from High Resolution Satellite Images (Venue: Council Hall)
17:30 - 18:45	Student Research Symposium I (Venue: MPH, SAC)		Student Research Symposium II (Venue: Council Hall)
19:00 - 19:50	Cultural Events (Venue: MPH, SAC)		
19:50 - 20:50	Gala Dinner		

DAY 2: July 19, 2024 (Friday)

Time	Event	
07:30 - 08:30	Breakfast	
09:15 - 10:00	Plenary talk II: Future Cameras *Prof. Shree K Nayar, Columbia University* (Venue: MPH, SAC)	
10:10 - 13:10 (11:25 - Tea Break)	Workshop: High-throughput Vision based Phenotyping (HTVP'24) (Venue: MPH, SAC)	Workshop: AI/ML-CV in Space Science and Technology (Venue: Council Hall)
13:15 - 14:15	Lunch Break	
14:15 - 17:15 (15:30 - Tea Break)	Tutorial: Use of LLMs and multimodal LLMs for information extraction from visually rich PDFs and document images (Venue: MPH, SAC)	Tutorial: Conditional Generative Models (Venue: Council Hall)
17:15 - 18:45	Invited Research Talks Session I (Venue: MPH, SAC)	Challenge: Detection of Pond Fish Challenge-2024 (DePondFi'24) (Venue: Council Hall)
19:00 - 20:00	Dinner	

DAY 3: July 20, 2024 (Saturday)

Time	Event		
07:30 - 08:30	Breakfast		
09:10 - 10:05	Plenary talk III: Controlling diffusion model generation to user constraints *Dr. Balaji Vasan Srinivasan, Adobe* (Venue: MPH, SAC)		
10:05 - 11:25	Workshop: Synthetic Data for Advancements in Contactless Hand Biometrics and Gesture Analysis (Syn-CHBGR 2024) (Venue: MPH, SAC)		
11:25 - 11:30	Tea Break		
11:30 - 13:00	Workshop: Synthetic Data for Advancements in Contactless Hand Biometrics and Gesture Analysis (Syn-CHBGR 2024) (Venue: MPH, SAC)	Industry Panel Discussion: The Rise of Language Models: Opportunities and Challenges for the Job Market in India (Venue: Council Hall)	Workshop: Unveiling Human Experience: Exploring Macro and Micro Emotion Recognition (Venue: Avionics Swayam Prabha Hall)
13:00 - 14:00	Lunch		
14:00 - 16:00	Invited Research Talks Session II (Venue: MPH, SAC)	Workshop: AI for 6G Networks (Venue: Council Hall)	Workshop: Unveiling Human Experience: Exploring Macro and Micro Emotion Recognition (Venue: Avionics Swayam Prabha Hall)
16:00 - 16:35	Valedictory Function (Venue: MPH, SAC)		
16:35 - 16:45	High Tea		

Plenary Talks

Plenary Talk I

Algorithms for Single Cell Data Analysis

Speaker: Dr. Sanghamitra Bandyopadhyay

Director and Professor, ISI Kolkata

Abstract: With advances in technology, it has become possible to extract valuable biological information at the granularity of single cells. Single cell data for up to a million observations and with thousands of measurements per cell are now available. Single cell data sets are inherently noisy and have significantly high missing values, thus challenging algorithm designers and data analysts. Nevertheless, proper analysis of this data has the potential to uncover hitherto unknown information on cell types, cellular signatures and multiple cellular processes. In this talk, we will first provide brief introductions to basic molecular biology and single cell RNA-seq data. This will be followed by a discussion on some recent algorithms for single cell data analysis developed in our group, with particular emphasis on clustering single cell RNA-seq data. Finally, some results and their biological insights will be described for a few real-life data sets.

Biography: Sanghamitra Bandyopadhyay joined the Machine Intelligence Unit of the Indian Statistical Institute as a faculty member in 1999, after completing her PhD from the same Institute. She was the Director of the Institute from August 2015 to July 2020, and is currently on her second tenure as Director from September 2020 onwards. Sanghamitra has worked in various Universities and Institutes world-wide including in USA, Australia, Germany, France, Italy, China, Slovenia and Mexico, and delivered invited lectures in many more countries. She has received several awards and fellowships including the Bhatnagar Prize, Infosys award, TWAS Prize, DBT National Women Bioscientist Award (Young), INAE Silver Jubilee Prize, Young scientist/engineer medals of INSA, INAE and Science Congress, JC Bose Fellowship, Swarnajayanti Fellowship and Humboldt Fellowship. She is a Senior Associate of ICTP and Fellow of INSA, INAE, NASI and IEEE. She is currently a member of the Science, Technology and Innovation Advisory Council of the Prime Minister of India (PM-STIAC).

The current research interests of Prof. Sanghamitra Bandyopadhyay include artificial intelligence, machine learning, soft and evolutionary computation, pattern recognition, data mining and various other applications. She has authored/co-authored more than 300 research article in international journals, conferences and book chapters, and published six authored and edited books from publishers like Springer, World Scientific and Wiley. She has also edited journals special issues in the area of soft computing, data mining, and bioinformatics.

Plenary Talk II

Future Cameras

Speaker: Shree K Nayar

Professor, Columbia University

Abstract: Computational imaging uses new optics to capture a coded image, and an appropriate algorithm to decode the captured image. This approach has enabled mobile devices to produce images that are rich, immersive and interactive. In this talk, we will show examples of computational cameras that are transforming the way visual information is captured, communicated and used by both humans and machines.

Biography: Shree K. Nayar is the T. C. Chang Professor of Computer Science at Columbia University. He heads the Columbia Vision Laboratory (CAVE), which develops computational imaging and computer vision systems. His research is focused on three areas - the creation of novel cameras that provide new forms of visual information, the design of physics-based models for vision and graphics, and the development of algorithms for understanding scenes from images. His work is motivated by applications in the fields of imaging, computer vision, robotics, virtual reality, augmented reality, visual communication, computer graphics, and human-computer interfaces.

Nayar received his PhD degree in Electrical and Computer Engineering from the Robotics Institute at Carnegie Mellon University. For his research and teaching, he has received several honors including the David Marr Prize (1990 and 1995), the David and Lucile Packard Fellowship (1992), the National Young Investigator Award (1993), the NTT Distinguished Scientific Achievement Award (1994), the Keck Foundation Award for Excellence in Teaching (1995), the Columbia Great Teacher Award (2006), the Carnegie Mellon Alumni Achievement Award (2009) , Sony Appreciation Honor (2014), the Columbia Engineering Distinguished Faculty Teaching Award (2015), the IEEE PAMI Distinguished Researcher Award (2019), and the Funai Achievement Award (2021).

For his contributions to computer vision and computational imaging, he was elected to the National Academy of Engineering in 2008, the American Academy of Arts and Sciences in 2011, and the National Academy of Inventors in 2014.

Plenary Talk III

Controlling diffusion model generation to user constraints

Speaker: Dr. Balaji Vasan Srinivasan

Principal Scientist, Adobe

Abstract: Recent advancements in text-guided image synthesis have revolutionized creative workflows, opening up a lot of possibilities in the realms of image generation. However, in many scenarios, the user would like to control specific aspects of the generation. In this talk, I will cover a few practical scenarios that gives raise to such constraints/controls and present some of our explorations that tailor diffusion models to provide such control and flexibility. I will begin with EMILIE, that allows iterative image generation where a user can retain parts of a generated image while customizing other parts to their needs. To allow for exemplar based image control, I will introduce MATTE, a textual inversion based framework and TINTIN, a T2I adaptor based framework that provides user to control generation on aspects like color, style from an example image(s). To address the computational costs associated with such custom generation, I will finally introduce our recent exploration on zero-shot customization of the generation by leveraging the latents.

Biography: Balaji Vasan Srinivasan is a principal scientist in Adobe Research focused on next generation technologies for enabling several content creation workflows. His current areas of research include multimodal content understanding, autogeneration of single-page graphic design, automatic translations of content across different modalities and informing content autogeneration with user-behavior data. In the past, he has worked on problems in text mining, social data analytics, high performance computing, scalable machine learning and speaker recognition. He has published several papers published in top NLP, Vision, HCI and AI conferences; and has several issued patents in related areas.

Balaji completed his Ph.D. in computer science at the University of Maryland in September 2011. His thesis was on Scalable Learning Methods for Speaker Recognition and Geostatistics. He completed his M.S. in electrical engineering from University of Maryland in 2008 and B.E. in electrical engineering from Anna University (India) in 2006. His work experience includes research internships at National Institutes of Health, Bethesda, MD (May – Aug 2007) and Xerox Research Center, Webster, NY (May – Aug 2011).

Invited Talks

Invited Talks

Ms. Siddhi Brahmbhatt, Wadhwani AI

Title: AI-Powered Pest Detection in Agriculture

Dr. Nikita Pinto, MathWorks

Title: Practical Approaches for Researchers: Overcoming challenges in AI and Computer Vision

Dr. Mohit Sharma, IISc

Title: Topological Structures and Operators for Bivariate Data Visualization

Mr. Nitish Kumar, ISRO

Title: AI Agents: How Intelligent can AI Agents Get?

Dr. Nandini J, IIIT-K

Title: Kathakali Mudra Recognition: Leveraging Pose Estimation and Pattern Recognition Techniques

Mr. Pavan Reddy, TCS Research

Title: Inverse Problems and Medical Imaging

Industry Panel Discussion

NCVPRIPG 2024
18th - 20th July
Indian Institute of Space Science and Technology-Thiruvananthapuram
Jointly Organized by:

Industry Panel Discussion

The Rise of Language Models: Opportunities and Challenges for the Job Market in India

Dipinder Sekhon

Founder, Kites.ai

Sharada Acharya

Machine Learning Engineer, YouTube,Google

Raghavan Muthuregunathan

Senior Engineering Manager, Linkedin

Pranjul Yadav

Machine Learning Lead, Google

Animesh Nayan

Senior Staff Software Engineer, Google

Tutorials

Tutorials

Tutorial 1

Transforming Human Perception: A Tutorial from Transformer to LLM Architectures

This tutorial aims to provide participants with a comprehensive understanding of transformer models and their evolution into Large Language Models (LLMs), highlighting their profound impact on NLP and other AI domains. Starting with an overview of the foundational principles of transformers as outlined in "Attention is All You Need" and "BERT," the tutorial will delve into their architectural design, operational mechanisms, and their diverse applications in fields such as computer vision, medical image analysis, natural language understanding, and beyond.

Furthermore, the tutorial will explore recent advancements and extensions of transformer models, including variants like GPT (Generative Pre-Trained Transformer). Participants will gain insights into how these models have revolutionized tasks such as text generation, language translation, question answering, and summarization. Additionally, the tutorial will discuss challenges and future directions in transformer research, such as improving model interpretability, handling multimodal data, and scaling up to even larger datasets and models. A brief plan is given below:

- Introduction to Transformers
- Transformer Architectural Insights and Training Techniques
- Delving into ViT and ViViT Models
- Application of Transformers for LLMs

Speakers/Organizers:

Sumantra Dutta Roy (IIT Delhi) , Aditya Nigam (IIT Mandi) , Arnav Bhavsar (IIT Mandi) , Gaurav Jaswal (TIH IIT Mandi)

Tutorial 2

Use of LLMs and multimodal LLMs for information extraction from visually rich PDFs and document images

Many documents' businesses and individuals deal with daily are document images (digital images of physical documents). On the other hand, born-digital PDFs (also called native PDFs) are no better than document images in terms of machine-readability. These documents contain multimodal content, and understanding natural language alone is not sufficient to comprehend them. Research in multimodal LLMs explores how text, visual, and layout information

Tutorials

can be combined to understand visually rich documents holistically. In this tutorial, we will cover significant works that explore the problem of information extraction from document images and demonstrate popular approaches using both text-only and multimodal LLMs. Below is a brief plan:

- LLMs and multimodal LLMs for document images
- Retrieval Augmented Generation
- Transform your visually rich PDFs or document images into a RAG-ready format
- Transformed documents with RAG + text-only LLMs for information extraction
- Multimodal LLMs (with text, layout, and vision capabilities) for information extraction

Speakers/Organizers:

Mihir Goyal (Wadhwani AI) , Minesh Mathew (Wadhwani AI)

Tutorial 3
Conditional Generative Models

Generative models generate new samples of a specific type of data. These models have had an irreversible impact on the industry and, most significantly, on daily life. To better understand this domain we present the tutorial on Conditional Generative models. The generative models need to be controllable, to that end they are "conditional" generative models. We provide sufficient examples to cover the essential techniques of the present and potentially in the future as well. We look through Autoencoders (AE), Variational Autoencoders (VAE), Generative Adversarial Networks (GAN), and Diffusion Models used in image generation tasks. We then focus on a relatively newer generative technique inspired from reinforcement learning: Generative flow networks(GFlowNets) used in molecular generation. Below is a brief plan:

- Generative Models
- Autoencoders/ Variational Encoders
- Generative Adversarial Networks
- Diffusion Models
- Reinforcement Learning
- GFlowNets
- Applications of GFlownets

Speakers/Organizers:

Sinnu Susan Thomas (Digital University Kerala) , Vineeth B.S. (IIST Thiruvananthapuram)

Tutorials

Tutorial 4

Deep Learning Models for Land Cover Mapping from High Resolution Satellite Images

Land cover mapping using remote-sensing imagery has attracted significant attention in recent years. Classification of land use and cover is an advantage of remote sensing technology, which provides all information about the land surface. Over the past decade, numerous studies have investigated land cover classification using a broad array of sensors, resolution, feature selection, classifiers, and other techniques of interest. Pixel-based and image-based classification techniques are used for land cover classification from remote sensing images. Accurate and real-time land use/land cover (LULC) maps are essential for dynamic monitoring, planning, and management of the Earth. With the advent of cloud computing platforms and machine learning classifiers, new opportunities arise for more accurate and large-scale LULC mapping from high-resolution remote sensing images. Deep learning-based segmentation of high-resolution satellite images provides valuable information for various geospatial applications, specifically for land use/land cover (LULC) mapping. The segmentation task becomes more challenging with the increasing number and complexity of LULC classes. This tutorial session presents a detailed introduction and implementation of deep learning algorithms for land use/land cover (LULC) mapping from high-resolution remote sensing images. Below is a brief plan:

- Introduction and Applications of Land Cover Mapping
- Deep Learning Models for Land Cover Mapping from High Resolution Satellite Images
- Understanding 2D U-Net and Attention U-Net CNN Models for land use/land cover (LULC) mapping
- Introduction to Deep Learning with TensorFlow
- Implementation of 2D U-Net and Attention U-Net CNN Models for land use/land cover (LULC) mapping using Python
- Challenges in land use/land cover (LULC) mapping from High Resolution Satellite Images
- Advanced deep learning algorithms for land use/land cover (LULC) mapping

Speakers/Organizers:

Shyam Lal (NITK Surathkal)

Workshops

Workshops

Workshop 1

Workshop on High-throughput Vision based Phenotyping (HTVP'24)

The application of methodologies to measure specific organism's (e.g. plant, insect etc.) traits (morphology, growth etc.) related to its structure and function is termed as phenotyping. With the emergence of low-cost and high-resolution multi-modal cameras, acquisition of 2D and 3D data permits high-throughput micro and macro analysis. This is a rapidly growing field at the interface of biology and computer vision (CV) termed-High-throughput Vision based Phenotyping. Agriculture Automation & Information systems nowadays extensively use high speed image sensors along with multi-scale and hyperspectral imaging methods to effectively capture plant traits, facilitating comprehensive monitoring and analysis. Vision-based plant phenotyping encounters difficulties in handling large and intricate data, coping with unpredictable environmental changes, and combining different types of information. Overcoming these challenges requires advanced methods to make sense of the data and help improve crops and farming practices effectively.

In contrast to common datasets like ImageNet, MSCOCO, PASCAL VOC datasets, organisms present unique challenges due to their dynamic nature and inherent variability. These challenges include tracking deformable objects like microbes in microscopy, multi-label segmentation of self-similar objects such as leaves in plants and fish, and 3D reconstruction in complex environments with overlapping surfaces, such as plant structures. Natural conditions, such as those found in agricultural fields, greenhouses, forests, and marine ecosystems, further complicate image analysis. This workshop will spotlight the challenges of applying computer vision to plant/animal phenotyping and agricultural research, aiming to showcase advanced methodologies, identify critical unresolved issues, and engage computer scientists interested in this domain. As effective plant phenotyping is crucial for sustainability, fostering community involvement and welcoming computer vision experts into this field is of paramount importance.

Topics of Interest :

Poster Showcase Topics (not limited to):

- Farmland pattern classification, detection, and segmentation from agricultural/phenotyping imagery.
- Resources and dataset benchmarks for agricultural imagery based pattern analysis.
- Data fusion of multi/hyper-spectral image data and multi-modal data sources
- Self, semi, and weakly supervised methods for agricultural/phenotyping imagery
- Transfer learning and domain adaptation
- Generative AI for plant/animal/microorganisms phenotyping
- 3D modeling and segmentation, UAV based field phenotyping
- Efficient data sampling methods and learning with limited training data or in presence of noisy, sparse, and imbalanced annotations.

- Computer vision applications which promote the study or adoption of sustainable agriculture

Organizers:
Brejesh Lall (IIT Delhi)
Prerana Mukherjee (JNU)
Vinay Kaushik (IIIT Sonipat)
Manoj Sharma (Bennett University)
Swati Bhugra (IIT Delhi)

Workshops

Workshop 2

Workshop on Synthetic Data for Advancements in Contactless Hand Biometrics and Gesture Analysis (Syn-CHBGR 2024)

This workshop covers recent developments in contactless hand biometrics and gesture recognition, spanning various practical applications. Recent advancements in generative models within the realms of computer vision and artificial intelligence have revolutionized the way researchers approach data-driven tasks. The advent of sophisticated generative models, such as GANs (Generative Adversarial Networks), VAEs (Variational Autoencoders), or more recently, diffusion models, has empowered practitioners to create synthetic data that closely mirrors real-world scenarios. These models enable the generation of high-fidelity images and sequences, laying the foundation for groundbreaking applications in hand biometrics and gesture analysis. The significance of these generative models lies in their ability to produce synthetic data that is remarkably realistic, thereby mitigating challenges associated with data scarcity and privacy concerns. As a result, the utilization of synthetic data has become increasingly prevalent in biometric domains, offering a versatile and ethical alternative for training and testing the latest deep learning algorithms.

The second edition of the workshop aims to delve into the diverse applications of synthetic data in the realm of hand biometric traits and gesture analysis. Participants will explore how synthetic datasets have been instrumental in training physiological and behavioral biometric recognition systems (fingerprint, finger knuckle image, palmprint, gait, gestures, finger nail image, dorsal palm, finger vein, palm vein), enhancing feature detection and refining recognition algorithms. The workshop will showcase exemplary use cases where the integration of synthetic data has not only overcome data limitations but has also fostered the development of more robust and accurate models. To promote the utilization of synthetic data in biometrics, the second edition of CHBGR is centering its focus on Synthetic Data, aiming to support its adoption and advancement within the field. Topics related to the theme of this workshop continue to draw widespread attention from local, national and international researchers.

Topics of Interest :

- Synthetic Data for hand biometrics analysis: hand geometry, finger knuckle, finger nail, fingerprint, finger photo, palm print, palm vein, finger vein, hand expression, signature, keystroke dynamics.
- Synthetic data for Multi-biometrics
- Synthetic data for Behavioural Biometrics
- Synthetic Data for Medical biometrics: EEG, ECG, PPG

- Synthetic Data for the Mitigation of Demographic Biases in biometrics
- Cross domain biometric recognition: fingerprint Vs finger photo
- EMG/ mm-wave Radar hand gesture synthesis and recognition
- Hand gesture recognition: cross sensor, cross modality
- Demographic effects on physical and behavioral biometric traits
- Person identification and re-identification
- Presentation attacks and detection in biometrics verification and identification
- Morphing attacks and detection
- Applications: Forensics, HCI, Healthcare, Image generation
- Databases, benchmarking and new protocols, robustness analysis

Organizers:
Aditya Nigam (IIT Mandi)
Arnav Bhavsar (IIT Mandi)
Gaurav Jaswal (IIT Mandi)

Workshops

Workshop 3

Workshop on Unveiling Human Experience: Exploring Macro and Micro Emotion Recognition

This workshop focuses on the recognition and analysis of macro and micro expressions, providing valuable insights into human behaviour and cognition. Macro expressions are overt and visible manifestations of emotions, conveyed through facial expressions, body language, and vocal cues, and are easily recognizable. Micro expressions, however, are fleeting, involuntary facial movements that reveal concealed emotions. The workshop will discuss the challenges in collecting and labelling datasets for these expressions and explore advanced analytical techniques and algorithm development for accurate detection. It will delve into technological advancements in computer vision, machine learning, and affective computing, emphasizing the integration of multimodal learning using multisensory data to enhance emotion recognition systems. Additionally, the workshop will explore the neural correlates of emotions, using methodologies like functional Magnetic Resonance Imaging (fMRI) to understand the brain regions activated during emotional experiences, thereby providing deeper insights into the cognitive processes behind these expressions. The practical applications of these advancements in fields such as human-computer interaction, virtual reality, gaming, and healthcare will also be examined.

Topics of Interest :

- Human Emotion Recognition: Understanding the fundamentals of human emotion recognition and its significance.
- Macro and Micro Emotions: Exploring the differences between macro and micro emotions and their respective characteristics.
- Challenges in Dataset Creation: Identifying the difficulties in creating and labeling datasets for emotion recognition, particularly for micro expressions.
- Neural Correlates of Emotions: Examining the neural mechanisms underlying emotional experiences and how they are studied using techniques like fMRI.
- Algorithm Design and Development: Discussing the challenges and advancements in designing and developing robust algorithms for emotion recognition.
- Multimodal Data Integration: Exploring the integration of multisensory data in emotion recognition systems and the challenges associated with this approach.

Organizers:
Santosh Kumar Vipparthi (IIT Ropar)
Priyanka Jain (C-DAC Delhi)
Rajkumari Reddy (NIMHANS, Bangalore)

Workshops

Workshop 4
DST SERB Workshop on AI for 6G Networks

Artificial Intelligence (AI) will greatly improve 6G networks by making them more efficient, reliable, and tailored to individual needs. 6G, the next step after 5G, will support extremely fast data speeds, very low delays, and connections for billions of devices. AI will help manage these networks automatically, predicting needs and fixing issues quickly. It will use machine learning to smartly allocate resources, direct signals, and adjust network sections, ensuring smooth connections for everything from self-driving cars to augmented reality. AI will also boost network security, protecting against new cyber threats and keeping data safe. In short, AI in 6G networks will transform digital interactions, providing better connectivity and enabling new, exciting technologies. AI in 6G networks promises to revolutionize the way we connect and interact with technology, offering faster, more reliable, and secure communication tailored to individual needs.

Topics of Interest :

- Resource allocation problems in modern wireless networks
- Introduction to AI and its Applications
- Integrated Sensing and Communication Systems
- Evolution of Localization Techniques from Sensor Networks to 6G
- Semantic Communication in random access: Frame less ALOHA and age of information
- Semantic Communication for modern networks: Age of information for multihop networks
- On the Road to 6G: Trends, Technologies and Challenges
- Handover Management in 6G Integrated Satellite Networks

Organizers & Speakers:
B. S. Manoj (IIST Thiruvanathapuram)
Sumitra S. (IIST Thiruvanathapuram)
Vineeth B. S. (IIST Thiruvanathapuram)
Deepak Mishra (IIST Thiruvanathapuram)
Vani Devi M. (IIST Thiruvanathapuram)
Bodhibrata Mukhopadhyay (IIT Roorkee)
Minha Mubarak (IIST Thiruvanathapuram)
Nibin Raj (IIST Thiruvanathapuram)
Sruthi N.T. (IIST Thiruvanathapuram)
Neethu Sajeev (IIST Thiruvanathapuram)

Workshops

Workshop 5
Workshop on AI/ML-CV in Space Science and Technology

Talk 1: Computer Vision in planetary landing missions with emphasis on Chandrayaan Missions

Landing on any celestial body with stringent engineering and mission constraints poses an inordinate challenge on any landing mission. Hazard Detection and Avoidance (HDA) becomes the crucial technology to be employed in landers for exploring scientifically interesting and landing near hazardous regions. Recently landing missions to Mars and Moon fully relied on computer vision techniques for safe and pin-point landing on surface. Chandrayaan-3 Vikram Lander landed on the lunar surface with the help of a computer vision-based HDA approach employed onboard for safe landing near the south pole. In this talk, we will discuss, computer vision techniques adopted in landing missions and with a focus on Chandrayaan Missions.

Talk 2: Satellite Remote Sensing Data Analysis using Artificial Intelligence & Machine Learning

AI and ML models have great success in many fields related to obtaining large amounts of image data to aid in pattern recognition and create algorithms through computer systems. AI can improve in the analysis of large areas of interest, to classify objects, detect and monitor land use, data fusion, cloud removal, and spectral analysis of environmental changes from satellite or aerial imagery. AI can aid in data collection, processing, and understanding using neural networks and deep learning through Computer Vision models to allow data users to better understand and handle data more efficiently in a timely manner, at multiple spatial resolutions. However, despite the numerous opportunities presented by AI and ML in Earth observation, there are also challenges that need to be addressed like issues related to data quality, scalability, interpretability of AI models, and ethical considerations surrounding the use of AI in decision-making processes. This talk holistically presents the existing satellites in orbit, future satellite missions, current technology, AI&ML use cases on various geospatial applications, challenges, and opportunities and overall big picture about Earth AI.

Project Presentation 1: Vision & DNN based pose estimation techniques for robotics applications and RLVs

The talk will explore the advanced vision systems developed for the ISRO Inertial Systems Unit (IISU) and their crucial roles in various robotic applications. These systems are utilized in humanoid robots, flying robots, robotic manipulators, and horizontal and vertical landing missions, enabling autonomous navigation, precise manipulation, and safe landings. The challenges addressed by these systems, including varying lighting conditions, surface texture recognition, and real-time obstacle avoidance, will be discussed in detail. Innovative solutions and technologies, such as image processing, machine learning, and sensor fusion, will be presented to demonstrate how these challenges were overcome. The transformative impact of vision-based

robotics in the space exploration and aerospace sectors will be emphasized, providing a comprehensive understanding of their current state and future prospects.

Project Presentation 2: Quality Inspection of Crimp Joints using CV & AI

Wire crimp visual inspection for quality is a very critical operation in the realization of aerospace systems. Being a process done in very large numbers on daily basis in aerospace manufacturing, the detection of the defects in wire crimp by manual visual inspection can be a daunting task. This talk proposes the use of a computer vision based solution for the quality inspection of each wire crimp joint.

Project Presentation 3: Automated Sign Check verification using Computer Vision

As part of subsystem level and vehicle level testing of a launch vehicle, each actuator is commanded with positive and negative command and the corresponding movements of actuators are verified against the expected direction of movement. Currently this verification is done manually in the presence of a group of people who witness the test. This method needs a substantial amount of man hours and is prone to errors. This talk proposes the use of a computer vision based solution for automated movement verification of actuators and nozzles in hardware in loop (HIL) simulation of a launch vehicle. A raspberry pi based computer with a camera captures the images. Object tracking algorithms are used to identify and find the position of an AruCo marker placed on the object of interest. The velocity of the movement of the marker is calculated and the information is displayed on the video for further processing or broadcast. This setup can also be extended to track the exact motion of the body by calibrating the camera as well.

Speakers/Organizers:
Suresh K. (SAC, ISRO)
Rashmit Singh Sukhmani (SatSure Analytics)
P. Manjusree (NRSC, ISRO)
Jyothish M. (IISU, ISRO)
Sujo Joseph K. (VSSC, ISRO)
Mijaz Mukundan (VSSC, ISRO)

Challenges

Challenges

Challenge 1
Automatic Evaluation of Handwritten Answer sheets

Given a set of handwritten answer sheets in a predetermined layout, the task is to automatically evaluate them and return the total marks scored by the student based on the correct answers. The task of Auto Evaluation presents a unique challenge. In order to achieve high performance, however, the underlying models need to contend with inconsistencies in spelling, incomplete words, variations in student penmanship (cursive vs print, letter size) etc. Furthermore, differentiating intended answers from markings (scribbles, underlines) adds another layer of complexity.

Organizers:
Anand Mishra (IIT Jodhpur)
Gyan Prabhat (IIT Jodhpur)
Abhirama Penamakuri (IIT Jodhpur)
Devesh Sharma (IIT Jodhpur)
Uday Agarwal(IIT Jodhpur)
Shreyas Vaidya (IIT Jodhpur)
Arvind Kumar Sharma (IIT Jodhpur)

Challenge 2
Detection of Pond Fish 2024 Challenge

DePondFi'24 aims to detect fish key points from underwater images captured in real-time pond environment. Fish key point detection supports the intelligent aquaculture system in fish identification and biomass estimation. In today's maritime era, aquaculture stands as a cornerstone of our nation's economic vitality. Fish biomass estimation is a significant challenge faced by the aquaculture industry. Hence, aquafarms are highly in need of smart intelligent systems to estimate the fish mass automatically. Intelligent system requires fish key point detection in underwater images as the first step for Fish biomass estimation. Fish key point detection refers to the identification and localization of specific anatomical landmarks or key points on fish bodies. These key points could include features such as the eye, dorsal fin, anal fin, and tail. However, this task poses significant challenges due to the complex underwater environment, including varying lighting conditions, water turbidity, fish occlusions, and diverse fish species. The objective of this challenge is to develop robust computer vision algorithms capable of accurately detecting fish key points in underwater images captured in wild environments. Key points provide valuable information about fish biomass, behavior, health, and growth, aiding in disease detection, feeding optimization, and overall aquaculture management. Improving the efficiency and efficacy of aquaculture monitoring directly supports the objective of SDG 14. SDG14: Improving the efficiency and efficacy of aquaculture monitoring directly aligns with Sustainable Development Goal 14(SDG 14), which focuses on conserving and sustainably using the oceans, seas and marine resources.

Challenges

Organizers:
Sasithradevi A (VIT Chennai)
P Prakash (Anna University)
Sabarinathan S (Couger Inc.)
S Md Mansorr Roomi (TCE, Madurai)
R Suganya (VIT Chennai)
Vijayalakshmi M (VIT Chennai)

Challenge 3
PestVision 2024

Pests pose a significant threat to agriculture, endangering crop yields and global food security. Detecting pests early is crucial for farmers, yet the scarcity of comprehensive pest datasets present a formidable challenge in training effective pest detection models.In this competition, participants are invited to tackle this challenge by leveraging innovative approaches such as large-scale synthetic data pre-training and low-data ML techniques. The goal is to develop robust pest detection solutions capable of empowering farmers and safeguarding their crops effectively.

Significance:

- **Addressing a Critical Need:** By developing accurate pest detection models, participants contribute to mitigating crop losses caused by pest infestations, thereby enhancing food security.
- **Encouraging Innovation:** This challenge encourages participants to think creatively and experiment with novel approaches to synthetic data generation and model training, fostering innovation in agricultural technology.
- **Real-world Impact:** The solutions developed in this competition have the potential to make a tangible impact on agriculture, benefiting farmers and communities worldwide.

Organizers:
Siddhi Brahmbhatt (Wadhwani AI)
Chandan Agrawal (Wadhwani AI)
Ashish Papanai (Wadhwani AI)
Soma Dhavala (Wadhwani AI)
JP Tripathi (Wadhwani AI))
Mohammad Salman (Wadhwani AI)
Sonali Ghike (Wadhwani AI
Jatin Agrawal (Wadhwani AI)

Challenges

Challenge 4

Smart City Surveillance: Unveiling Indian Person Attributes in Real Time

Given a set of handwritten answer sheets in a predetermined layout, the task is to automatically evaluate them and return the total marks scored by the student based on the correct answers. The task of Auto Evaluation presents a unique challenge. In order to achieve high performance, however, the underlying models need to contend with inconsistencies in spelling, incomplete words, variations in student penmanship (cursive vs print, letter size) etc. Furthermore, differentiating intended answers from markings (scribbles, underlines) adds another layer of complexity.

Organizers:
Shikha Gupta (Vehant)
Renu M. Rameshan (Vehant)
Shivam Nigam (Vehant)
Abhay Kumar (Vehant)
Swathy Pandey (Vehant)

Oral Presentations

Oral Presentations

Characterization of Down Syndrome in Fetal Ultrasound Using Deep Learning Models

Janavi T. S., Keerthana V., Ayesha Jumana, Dr. G. Kavitha
College of Engineering Guindy Anna University
janavi.ts06@gmail.com, keeruv2002@gmail.com, ayeshajumana02@gmail.com, kavithag_mit@annauniv.edu

Down syndrome is a chromosomal disorder that affects approximately 1 in 800 live births worldwide (WHO 2023 reports), resulting in intellectual disability and distinct physical features. Early detection is imperative for informed decision-making and management. By offering an improved non-invasive risk analysis method, this study presents a safer alternative to invasive procedures like amniocentesis, Chorionic villus sampling, and Percutaneous umbilical blood sampling which pose a 3% risk of miscarriage. This study proposes the utilization of deep learning models, specifically Vision Transformer, for the characterization of Down syndrome. The dataset comprises fetal ultrasound images [1500 patient data]. The process begins with preprocessing the images to remove noise. Biomarkers indicative of Down syndrome, including Nuchal Translucency, Nasal tip, Nasal bone, Nasal skin, Intracranial translucency, Cisterna Magna, Mouth palate, Midbrain, and Thalami are segmented using Variational level set methods and are validated with over 90% accuracy. Extraction of geometric and texture features from these regions are performed. The Vision Transformer, known for its ability to learn global image features and reduced sensitivity to data augmentation, is trained on these features. The model classifies Down syndrome, successfully when trained with all nine biomarkers. This model enhances prenatal screening and management for affected individuals and their families.

NCVPRIPG 2024
18th - 20th July
Indian Institute of Space Science and Technology-Thiruvananthapuram
Jointly Organized by:

Oral Presentations

Marathi Words Classification for Brain-Computer Interface Using Machine Learning

Shubham Tayade, Manoj B.S, S. Sumitra
Indian Institute of Space Science and Technology, Thiruvananthapuram
innovativeshubham369@gmail.com, bsmanoj@gmail.com, sumitra@iist.ac.in

This research paper explores the implementation of Machine Learning (ML) techniques within Brain-Computer Interface (BCI) systems to recognize Marathi words, aiming to alleviate communication hurdles for Marathi-speaking individuals with neurodegenerative conditions. Using an 8-channel EEG device, EEG signals were collected and underwent preprocessing, including filtering with Butterworth-low-pass and Moving Average filters. Time-domain features, such as Hjorth parameters and AMC Feature Set, were then extracted from each channel. Classification tasks were conducted using Support Vector Machine (SVM) and Naive Bayes algorithms, with performance metrics such as average training accuracy and testing accuracy analyzed comprehensively. Results revealed promising accuracy rates: Naive Bayes achieved 95.11% and 95% accuracy for Marathi vocal and sub-vocal datasets in two-word classification tasks, respectively, while English vocal data attained 90% accuracy.

For three-word classification tasks, Naive Bayes utilizing the AMC feature set achieved testing accuracy of 70% and 63.33% for Marathi and English datasets, respectively. SVM (linear) with Hjorth parameters achieved a testing accuracy of 73% for English vocal data. This study lays the groundwork for the development of a Marathi language BCI system, offering potential benefits in enhancing communication between caregivers and Marathi-speaking individuals with neurodegenerative disorders.

Oral Presentations

Weed Detection and Localization in Agricultural Environment Using YOLO Algorithm

Marala Bhavvyya sree, Deepa S. N., Athkuri Sai Saketha Chandra
National Institute of Technology Calicut, Kozhikode, Kerala
bhavvyya.marala@gmail.com, bhavvyya_m230556ee@nitc.ac.in,
deepapsg@gmail.com, sndeepa@nitc.ac.in, saisakethachandra@nitc.ac.in

A weed is an undesirable plant which grows along with the essential plants in a field which results in reduced crop yields and some weeds may transfer plant diseases to crops. It is necessary to identify the weeds and localizing the weeds species within crop fields and remove the weeds without harming the crop. This involves the usage of technologies, which are deep learning neural networks, machine learning techniques for creating a weed detection model using YOLO. The YOLO (you only look once) is a single-shot object detection algorithm. YOLO variant is proposed for weed detection is a reliable methodology of agricultural image analysis. Then we gathered and prepared a dataset with different types of weeds along with plants. Then used the YOLO model, which is designed to detect weeds, and improved its performance by transforming the dataset to make the model more versatile for identification and localizing the weeds. The performance of the algorithm in identifying the weeds and localizing the weeds within crop fields was evaluated using many datasets of crops with weeds species. Employing the proposed YOLO variant, the algorithm achieved a good accuracy, mean average precision in weed identification in the real time agricultural fields or environment, with minimum false positives. Classification accuracy rate of 93.7% is achieved and is better than other deep learning models from literatures like AlexNet, ConvNet, ResNet 50, Centernet, R-CNN. As YOLO is a single-shot object detection algorithm it exhibits excellent processing speeds, making it suitable for practical field applications. The results show the potential of deep learning techniques for real-time weed detection in agricultural fields. The weed management practices can be benefited from the help of machine vision and deep learning techniques and contribute towards sustainable development.

Oral Presentations

Digital Photogrammetry and Computer Vision Based Technique for Advance Studies in Structural Geology

Aastha Mishra, Yengkhom Kesorjit Singh
School of Earth, Ocean and Climate Sciences, IIT Bhubaneshwar, Khordha 752050
22gg05007@iitbbs.ac.in, yksingh@iitbbs.ac.in

With the advancements in science and technology, remote sensing has emerged as a crucial tool in geosciences, expedite scientific investigating and conclusions. It has greatly facilitated in reconnaissance studies and mapping large inaccessible areas in terms of time and cost efficiency. Photogrammetry, a technique for constructing measurable three-dimensional models using multiple photographs captured by RGB commercial camera. Modern photogrammetric software incorporated Structure from Motion (SFM) algorithms, similar to conventional photogrammetry but enhanced through computer vision algorithms. It identifies significant points in individual photos, to derive the photographic parameters, and match common recognizable points in multiple photographs to determine the spatial coordinates. Following the image matching and sparse reconstruction, Multi-view Stereo Reconstruction (MVS) is employed to produce a dense point cloud by increasing the number of points. The present study is done using digital photogrammetry-based 3-D reconstruction of outcrops for extracting structural elements from multiphase deformation, rock exposures in the Ambaji basin, Gujarat. The Ambaji basin (AB) is a Neoproterozoic basin in the South Delhi Terrane of Proterozoic Aravalli Delhi Mobile Belt (ADMB). Throughout the evolution of ADMB, the basin had undergone multiple stages of deformation, metamorphism and magmatism, resulting in a very complex deformational geometry. Establishing the deformational and metamorphic history of Ambaji basin thus proves challenging and tedious. This work aimed to facilitate the identification of different generations of deformational structures to establish chronological order of geological events on macro scale, thereby correlating it with regional structures. Utilizing digital photogrammetry technique, 3-D reconstructions of the outcrop, covering both macro (six locations) and a regional scale structure is done. This approach reduces the time to survey outcrops and offers varied field of view for observation. On macro scale, photographs were acquired using DSLR camera and UAV drone, while on regional scale photographs are obtained using hillshade map processed from ALOS PALSAR DEM data. The 3-D reconstructed models of the outcrops were generated using SFM software packages, AgiSoft Metashape. Based on detailed fieldwork, petrography and structural attributes extracted from Virtual Reality for Geosciences (VRGS) software it can be concluded that macro scale outcrop depicts that AB has undergone three generations of deformation. The first generation is represented by rootless, interfolial, tight isoclinal F1 folding, refolded along with S0——S1 plane by a second generation of folding (F2). F1 an F2 trending NE-SW and coaxial. Superposing of F1 and F2 produced Type - 2 Hook-shape interference pattern. The F2 folds are upright, open to tight and plunging folds. An open F3 folding outcropping along with F1 and F2 has been newly identified on macro scale outcrop. F3 is trending NW-SE. Superposing of F1 on F3 produced Type-2 Mushroom-shaped interference pattern.

NCVPRIPG 2024
18th - 20th July
Indian Institute of Space Science and Technology-Thiruvananthapuram
Jointly Organized by:

Oral Presentations

Enhanced Panoramic Dental Radiography Using Specialized Kernels and Advanced YOLO Models for Improved Tissue Visualization and Pathology Detection

Vaishali Vijayarajan, S.Mohamed Mansoor Roomi
Thiagarajar College of Engineering
vaishalivijayarajan@gmail.com, smmroomi@tce.edu

Digital dental radiography is pivotal in the diagnosis and treatment planning in dentistry, providing essential and objective information that is not visible during clinical examinations. Panoramic radiography is widely used in dental practices as it offers a comprehensive view of the entire mouth in a single image, revealing anatomical structures in the teeth, jaws, temporomandibular joints, sinuses, bone structures, and developmental anomalies such as cysts and tumors. However, a significant limitation of panoramic radiography is its inability to accurately depict the internal structures of soft and hard tissue regions. Enhancing the visualization of both soft and hard tissues in panoramic radiographs is crucial for precise segmentation of regions such as the mandible, maxilla, alveolar bone, and the detection of pathological abnormalities. To address this limitation, we propose the use of two distinct kernels (filters): one specifically designed to sharpen the details of the soft tissue regions, and another to enhance the hard tissue regions, both operating in the spatial domain. Through comprehensive quantitative and qualitative analyses, our approach has demonstrated superior results in enhancing the details of both hard and soft tissues, thereby supporting its utility for precise segmentation of the inferior alveolar canal (IAC) and effective detection of cysts and tumors. We also introduce an attention-based YOLO model capable of segmenting the IAC, along with a vision transformer-based YOLO model for the concurrent detection of cysts and tumors in a single panoramic radiograph. This integrated approach promises to improve the accuracy and efficiency of diagnostic procedures in dental radiography.

Oral Presentations

3D-CNN Based Milk Quality Inspection Using Hyperspectral Imaging Method

P.Padmasri, B.Sathya Bama
Thiagarajar College of Engineering
padmasrip@student.tce.edu, sbece@tce.edu

Due to rapid population growth and increased demand for milk, economically driven motives have led to the widespread adulteration of milk with various edible, chemical additives, and preservatives. The impact of adulteration on health highlights the need for effective technologies to combat this issue promptly. This research work investigates the viability of harnessing hyperspectral imaging technologies, coupled with deep learning 3D models, to enhance the identification and classification of common adulterants such as water and starch, presence in milk. Initially, the milk samples are prepared at different concentrations levels with starch, considered as low 0.5g/50 ml, mid 0.75g/50 ml and 1g/50 ml and with water, the levels are 25%, 50% and 100% amount of water to 50ml milk. The hyperspectral image (HSI) dataset is generated using a Resonon Pika-L hyperspectral imager with 300 spectral channels across a range of 400 nm to 1000 nm. However, the problem of spectral feature redundancy is solved by using principal component analysis.

In recent years, 3D Convolution Neural Network models have gained more attention due to its ability to automatically learn nonlinear features in HSI dataset. This research experiments and demonstrates the 3D-Resnet-SE, 3D-EfficientNet, and 3D-VGGNet models to classify the milk adulterants based on the spectral-spatial information of HSI dataset and the accuracies are 99%, 91%, and 89% respectively. Finally, the 3D-Resnet-SE architecture provides better results in classifying various milk adulterants compared to other state-of-the-art deep learning models.

Oral Presentations

Transfer Learning-based Classification of Soybean Kernel Damages Using Computer Vision

Prerna Saurabh, Husain Kanchwala
School of Interdisciplinary Research (SIRe) - IIT Delhi, Centre for Automotive Research and Tribology (CART) - IIT, Delhi

The agricultural industry faces significant challenges due to insufficient crop management techniques, including suboptimal productivity. Detecting these damages accurately is crucial for maintaining farm productivity and sustainability. The proposed solution will involve integrating transfer learning techniques using pre-trained VGG-16 to enhance adaptability. It combines Artificial Intelligence (AI) with advanced computer vision algorithms to improve early leaf disease detection in plants. This approach supports better decision-making, potentially increasing agricultural yield and efficiency through precision farming. The integration of edge computing into precision agriculture allows for collecting and analyzing real-time data, providing immediate insights into crop health. Since these damages display distinct and vibrant color characteristics, extracting pertinent features through color space analysis can be effectively achieved using the LAB color transformation.

Oral Presentations

Enhancing Marine Conservation: Advanced Underwater Waste Detection Using Deep Learning

Akshat Singh, Krishna Rochani, Sasithradevi A., Vijayalakshmi M., P. Prakash
School of Computer Science and Engineering - VIT, Chennai, Centre for Advanced Data Science - VIT, Chennai, School of Electronics Engineering - VIT, Chennai, Department of Electronics Engineering - MIT, Chennai
akshat05022003@gmail.com, rochanikrishna.pradeep2021@vitstudent.ac.in, sasithradevi.a@vit.ac.in,Vijayalakshmi.m2021@vitstudent.ac.in, prakashp_mit@annauniv.edu

Escalating pollution from plastics and other non-biodegradable materials in marine ecosystems necessitates advanced automated monitoring systems. Underwater imaging faces significant challenges, including poor visibility, variable lighting, and color distortion due to light absorption and scattering. This paper focuses on detecting waste in pond environments using computer vision, leveraging sophisticated monitoring technologies to address these complexities. A vision-based approach is employed to differentiate between biodegradable and non-biodegradable waste.

A novel dataset was developed to enhance waste detection across diverse underwater environments. The dataset consists of 1181 images with implementing augmentation. An extensive evaluation was conducted using various deep learning-based object detection models. The study specifically focused on different configurations of the YOLO series, which were fine-tuned through extensive hyperparameter adjustments. Our findings indicate that YOLOv9 achieved a Mean Average Precision (mAP)@50 of 90.3%, with precision and recall rates of 89.8% and 83.9%, respectively. Future research will aim to refine these models by integrating multi-sensor data and implementing algorithmic improvements. These enhancements seek to boost detection accuracy under varying conditions, reduce false positives, and thus enhance the robustness of underwater trash detection. Such advancements are expected to substantially contribute to marine conservation efforts, providing more reliable tools for monitoring and mitigating the impacts of underwater waste.

Oral Presentations

Development of an Efficient Plant Species Identification System for South Indian Flora Using Attention-Based Vision Transformers

Dr. P.Uma Maheswari[1], V.Karthiga[1], S. Deepthi Sri[1], K. Rama Lakshmi[1]
Dr.S.Mohamed Mansoor Roomi[2]
Velammal College of Engineering and Technology[1], Thiagarajar College of Engineering[2]
umamahes.p@gmail.com, vkarthigavadivel@gmail.com, deepthisrikavitha@gmail.com, mithrakannan@gmail.com, smmroomi@tce.edu

The complexity and diversity of plants, with millions of species yet to be fully discovered and studied, present significant challenges for botanists. Effective categorization and study of plant species are essential for monitoring and disseminating botanical knowledge. However, the extensive diversity of plant species and varying geographical conditions have resulted in a lack of comprehensive databases for indoor and outdoor plant species. To address this gap, we have compiled a robust database of images for indoor and outdoor plant varieties native to Tamil Nadu, South India, focusing on the regions of Madurai and Tirunelveli. To build this database, we employed two high-quality imaging devices: the Vivo 21e, with a 64MP primary sensor, 8MP ultrawide lens, and 2MP macro lens, and the iPhone 13, featuring a dual 12MP camera system with advanced image stabilization. The datasets were gathered from various latitudes and longitudes, across different times and occasions, ensuring a diverse range of plant specimens. We sourced plants from several nurseries, including Velan Nursery in Tirunelveli, which provided species such as Thuja, Swiss Cheese Plant, and Dahlia pinnata, as well as nurseries in Madurai, such as Nikhilesh Nursery, which supplied species like Scindapsus, Parodia magnifica, and Florist Kalanchoe. Our dataset comprises 13,506 images encompassing 32 different plant varieties, including Radermachera sinica, Araucaria, Peregrina, Chlorophytum comosum, Jatropha, and Trailing Lantana. Motivated by the need for a reliable and efficient plant species identification system, we developed a method utilizing an attention-based vision transformer to classify the compiled plant varieties.

This method was rigorously evaluated using standard metrics and state-of-the-art approaches, achieving a high classification accuracy of 90.2%. The findings of this research significantly advance the development of plant species identification systems, with important implications for agriculture, ecology, and environmental monitoring. The proposed attention-based vision transformer systems offer a practical and efficient solution, aiding landscapers, plant enthusiasts, and gardeners in making informed decisions. This ensures the selection of appropriate plants for specific environments, promoting healthier plants and more sustainable gardening practices.

Oral Presentations

RAFGIC: Retrieval Augmented Fine Grained Image Classification

Varun Sampath Kumar, Anrutha J K
PSG College Of Technology
varunkavin5@gmail.com

This research explores the potential of Retrieval-Augmented Fine-Grained Image Classification (RA-FGIC) for overcoming limitations in traditional FGIC methods. Fine-grained image classification (FGIC) presents a significant challenge due to subtle visual variations within categories. This work proposes a novel RA-FGIC approach that leverages a CLIP encoder for semantic image representation and a Convolutional Neural Network (CNN) for feature extraction. During classification, the model retrieves similar images from a CLIP-encoded vector database for a given image. To capture the interaction between the target image and retrieved images, we propose a cross-product operation between the retrieved image embeddings and the last layer output of the pretrained CNN of target image. This enriched feature representation is then fed to a classification layer for fine-grained category prediction. We evaluate our approach on benchmark dataset such as Stanford Cars and investigate the effectiveness of the cross-product strategy compared to alternative feature combination methods. We evaluate our approach on these datasets and achieve significant improvements in classification accuracy, surpassing a pretrained EffecientNet CNN model by 11%. Our work demonstrates the effectiveness of RA-FGIC with CLIP and cross-product feature interaction for fine-grained image classification tasks.

Oral Presentations

AI/ML for Onboard Remote Sensing

Sandip Paul, Deepak K Mishra, Jimit G., Anirban Paul
Indian Institute of Space Science and Technology, Thiruvananthapuram
sandippaul7204@gmail.com

Remote sensing applications uses multiple observation instruments on satellites to manage and monitor natural resources, infrastructure and hazards. Data rate of Electro-optical payload increases with increase in the spatial resolution, input dynamic range (bits per pixel), wider swath and number of spectral bands. Demand for sub-meter resolution imagery along with shift towards multi-spectral and Hyper-spectral data lead to exponential increase in data-rate. However, downlink bandwidth has not increased proportional to the data-rate generated and is increasingly becoming a bottleneck for the system. This requires novel approaches to efficiently utilize available downlink bandwidth of satellite. AI/ML based on-board intelligence can discard potentially irrelevant and redundant data, transmit useful data and facilitate efficient utilization of downlink bandwidth.

Recent advancements in semiconductor technology introduced high performance space grade FPGAs, which now enable deployment of compute intensive AI/ML algorithms on-board. On-board AI/ML algorithm deployment calls for high throughput real time AI/ML inference hardware in a SWaP (size, weight and power) constrained satellite environment. Additionally, real time on-board processing mandate use of RAW satellite images which inherently have radiometric and geometric distortions. These impose unique challenges for onboard deployment of AI/ML algorithm in satellite payload. This paper attempts to elaborate the use of AI/ML based on-board intelligence for efficient utilization of downlink bandwidth, challenges which are unique to satellite for deployment of AI/ML inference using raw high-speed data, hardware realization options, and implementation aspects in space-grade hardware.

Poster Presentations

NCVPRIPG 2024
18th - 20th July
Indian Institute of Space Science and Technology-Thiruvananthapuram
Jointly Organized by:

Poster Presentations

Computer Vision Enhanced Multiphysics Simulations via GAN Integration

Jaisimha Manipatruni
Centre for Nanotechnology, Department of Electrical Engineering, Andhra University
jaisimha34@gmail.com

The integration of Generative Adversarial Networks (GANs) with Multiphysics simulations presents a transformative approach to modelling complex physical phenomena. Traditional Multiphysics simulations encompass fluid mechanics, electromagnetics, thermodynamics, structural mechanics, and related fields, known for their computational expense and time consumption. Physics Informs GANs offer a promising solution by generating high-accuracy data points and simulations efficiently. This work focuses on training GANs and integrating them into Multiphysics simulation tools to produce high-fidelity simulations and reduce computational costs. The training begins by collecting and preprocessing a comprehensive dataset comprising simulation results and experimental data. The GANs are trained on this dataset with to learn underlying physics and generate high-fidelity simulations of complex physical phenomena at lower computational cost.

NCVPRIPG 2024
18th - 20th July
Indian Institute of Space Science and Technology-Thiruvananthapuram
Jointly Organized by:

Poster Presentations

Detection of Driver Behavior in Public Transportation Using Adaptive Spatial Attention in YOLO Model

Srilakshmi Ravali M., Dr. S. Mohamed Mansoor Roomi
Thiagarajar College of Engineering
msrilakshmi@stanley.edu.in, smmroomi@tce.edu

Public transportation is integral to the daily lives of individuals commuting to school, workplaces, hospitals, and other destinations due to its subsidized rates, eco-friendly nature, and accessibility. Given its importance, enhancing the safety of public transportation is crucial, particularly through the reduction of road accidents. Developing information systems to detect driver behavior and distraction is essential in achieving this goal. The objective of this research is to estimate driver behavior by analyzing various aspects such as aggressiveness, lane deviation, vehicle stopping behavior, and driver status while operating a bus, using benchmark datasets like UAH-DriveSet, SPMD, and the Driving Dataset. To effectively predict driving behavior, we propose a novel YOLO model that leverages an adaptive spatial attention mechanism. This approach aims to improve the accuracy and efficiency of driving behavior prediction by dynamically adjusting the model's focus based on the most relevant spatial features and patterns identified during the analysis phase. By integrating spatial attention mechanisms into the deep learning framework, the proposed model can intelligently capture and interpret subtle nuances in driver behavior that traditional methods may miss. This innovative approach enhances the model's capability to detect and analyze complex driving behaviours, ultimately contributing to safer public transportation systems.

NCVPRIPG 2024
18th - 20th July
Indian Institute of Space Science and Technology-Thiruvananthapuram
Jointly Organized by:

Poster Presentations

Object Detection in real-time environment for visually impaired using YOLO model

Polepaka Pranay Raj, Deepa S. N., Athukuri Sai Saketha Chandra
National Institute of Technology, Calicut
pranay_m230815ee@nitc.ac.in, sndeepa@nitc.ac.in, saisakethachandra@nitc.ac.in

People who are visually impaired or blind often face many challenges navigating indoor and outdoor environments on their own. These challenges include lack of spatial awareness, obstacles in their path, and limited access to visual information. So object detection is a crucial technology in various applications including assistive devices for visually impaired. Although some traditional mobility aids are present but they are not very helpful for navigating in outdoor environment. By developing the high speed and highly accurate object detection model especially for visually impaired will help them to navigate in all situations. Employing the ability of Deep learning for developing object detection model using YOLO (You Only Look Once) which is a real-time object detection system, it provides real-time guidance for visually impaired individuals. And its performance is improved on diverse datasets consisting of various types of obstacles including static and dynamic objects, the YOLO model demonstrates exceptional performance in real world scenarios. The developed object detection model using YOLO produces promising results, achieving better accuracy and speed suitable for real-world applications, making it suitable for applications which requires instant feedback. Performance metrics are obtained such as mean average precision, precision with a value of and recall, they are utilized to evaluate the model's effectiveness in detecting objects. The value of mean average precision settled at 85.4% which states that this method is accurate when compared to single shot detection, R-CNN.

NCVPRIPG 2024
18th - 20th July
Indian Institute of Space Science and Technology-Thiruvananthapuram
Jointly Organized by:

Poster Presentations

Unveiling Facial Expressions in Indian Classical Dance with Attention-Driven CNN

Sajitha V. Raj, Dr Sunil T. T.
College of Engineering Attingal
sajithagodwin@gmail.com, vu2swx@gmail.com

Dance and music have significantly shaped human civilizations, particularly in India, where they offer a rich panorama of expression, storytelling, and devotion. The Sangeet Natak Akademi recognizes eight traditional styles as classical dances, each involving complex mudras (hand gestures), facial expressions, and body movements. Capturing and analyzing these elements accurately using computer vision techniques is challenging. This paper aims to enhance facial expression recognition in Indian classical dance using computer vision techniques. We propose integrating the Convolutional Block Attention Module (CBAM) into a Convolutional Neural Network (CNN) to improve accuracy. CBAM enhances the model's focus on relevant features by inferring attention maps along channel and spatial dimensions. Our model, trained on a curated dataset of facial expressions from Indian classical dance performances, achieved a reasonable improvement in accuracy.

This study demonstrates the effectiveness of incorporating attention mechanisms like CBAM in CNNs for facial expression recognition in Indian classical dance, with potential applications in digital archiving, automated annotation, and educational tools. Future work will optimize the architecture and expand the dataset to include more diverse expressions from various dance forms.

NCVPRIPG 2024
18th - 20th July
Indian Institute of Space Science and Technology-Thiruvananthapuram
Jointly Organized by:

Poster Presentations

Advanced Basketball Tactics Analysis Using YOLO

K. Purnanandh, S. Meghanath Reddy, Sasithradevi A., Vijayalakshmi M., Keerthika N.
VIT, Chennai
purnanandh.k2022@vitstudent.ac.in, meghanathreddy.s2022@vitstudent.ac.in, Sasithradevi.a@vit.ac.in, Vijayalakshmi.m2021@vitstudent.ac.in, keerthika.n2021a@vitstudent.ac.in

This project uses deep learning to revolutionize how we analyse basketball game strategies. Our goal is to develop a smart system that can identify and track players and the ball in videos, helping coaches, players, and analysts understand team tactics better. We used the YOLO (You Only Look Once) v8 object detection framework for this purpose, thanks to its high accuracy and speed. To create our dataset, we took video footage from basketball games and extracted one frame per second. YOLO v8 was chosen because it handles the fast-paced and complex nature of basketball well. Even with challenges like players being blocked from view and different lighting conditions, our model performed impressively, achieving a precision score of 0.87 and a mean Average Precision (mAP) score of 0.640 at an IoU threshold of 0.5. Our system analyses movement patterns and trajectories to interpret team formations. The results show significant improvements in understanding basketball tactics, highlighting how AI can transform sports analytics. This research not only improves basketball analysis but also has potential applications in other team sports. Looking ahead, we plan to refine our model for even more detailed tactical analysis and expand its use to other sports. This project demonstrates the power of combining deep learning with real-time analytics to enhance strategic planning and on-court performance. By offering a detailed look at player movements and interactions, our system helps teams improve their game strategies and performance.

NCVPRIPG 2024
18th - 20th July
Indian Institute of Space Science and Technology-Thiruvananthapuram
Jointly Organized by:

Poster Presentations

Decentralized Defense for Adversarial Robustness in Federated Learning

Cynara Justine[1], Sathyanarayanan Manamohan[2], Linu Shine[1], Jiji C.V.[2]
College of Engineering Trivandrum[1], Shiv Nadar University, Chennai[2]
cynara@cet.ac.in, sathya@hpe.com, linu.shine@cet.ac.in, jijicv@gmail.com

Adversarial poisoning attacks are growing threats to federated deep learning systems, as models can be poisoned from any of the compromised participants. Existing solutions attempt to detect and defend these attacks by assuming that there is an honest central parameter aggregator on which the detection and defense techniques can be applied.

This assumption can give rise to a single point of failure (SPOF) and an easy target for attackers to compromise the whole federated learning system. They also lack a robust mechanism to isolate the malicious node so as to protect the rest of the federated system. We propose SL-Defender, a novel solution that overcomes these limitations by using a decentralized method to detect and defend against attacks based on Swarm Learning. Our solution uses the selected layers from each of the federated learning participants and projects their parameters into a lower dimensional subspace and performs clustering analysis in this subspace. The proposed methodology adds the ability to act on identified malicious participants and contain their impact on the federated learning process or remove them if needed. The proposed SL-Defender thus eliminates the need for honest central aggregator and distributes computation to peers and dynamically monitors them and defends them from the federating poisoned parameters without interrupting the learning process. With the support of experimental results on MNIST Dataset, we prove that SL-Defender will offer a unique position to improve the reliability and trustworthiness of Swarm Learning and Federated Machine learning in general.

NCVPRIPG 2024
18th - 20th July
Indian Institute of Space Science and Technology-Thiruvananthapuram
Jointly Organized by:

Poster Presentations

LLM Counselling Chatbot Using TinyLlama model and Ragging techniques

Rohan Majo Mathew, Annmaria Mathew Shibu, Ryan Binny Mathews, Alphin Tom, Jacob Thomas
St Joseph's College of Engineering and Technology, Palai
rohanmajomathew@gmail.com

A chatbot for counseling specializes in applying artificial intelligence (AI) and machine learning (ML) to the provision of counseling services in order to improve and expedite mental health assistance. With the increasing prevalence of mental health issues globally, there is a growing need for accessible and efficient counselling solutions. The proposed AI-enhanced counselling platform aims to bridge this gap by integrating cutting edge technologies to provide personalized and timely support to individuals in need. The project entails creating an intuitive online or mobile application that makes use of natural language processing (NLP) algorithms to have meaningful conversations with users and LLM techniques to deliver responses that are more comprehensive and comprehensible. The AI counsellor will be designed to understand and respond to users' emotions, concerns, and queries,offering empathetic and non-judgmental responses with Rag techniques. The platform will also incorporate machine learning algorithms to continuously adapt and improve its counselling capabilities based on user interactions. This mini project aims to showcase the feasibility and effectiveness of integrating AI into counseling services, offering a scalable solution that complements traditional mental health support systems Index Terms—Natural Language Processing, Ragging Techniques, Artificial Intelligence, Large language Model(LLM), TinyLlama Model.