



University of  
East London

125  
YEARS

6<sup>th</sup> IVPR

11<sup>th</sup> ICIEV

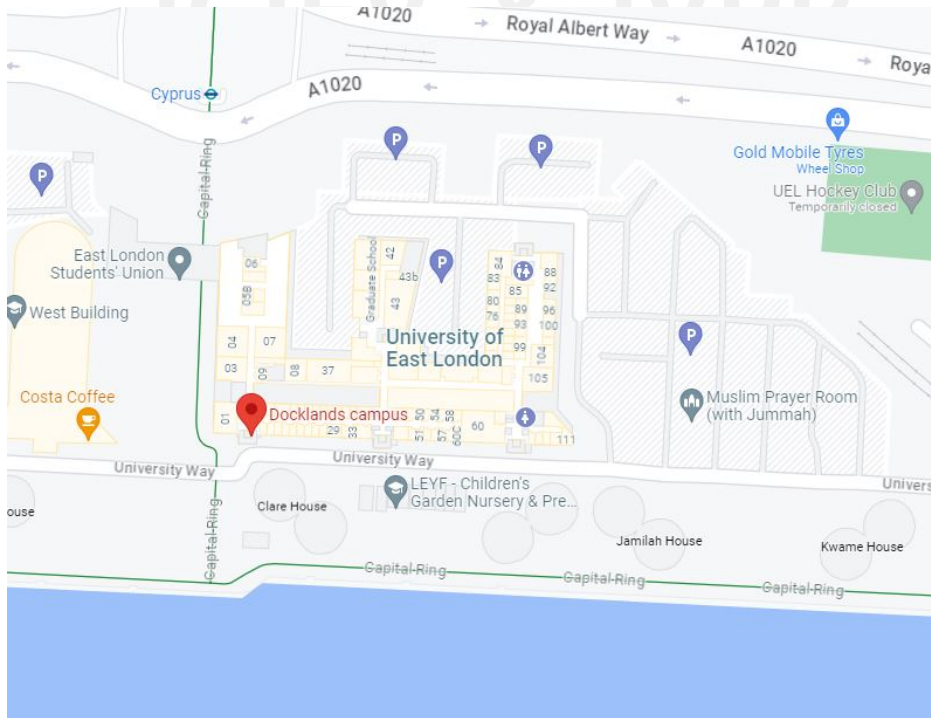
<http://cennser.org/IVPR>

<http://cennser.org/ICIEV>

6<sup>th</sup> International Conference on Imaging, Vision & Pattern Recognition (IVPR)  
11<sup>th</sup> International Conference on Informatics, Electronics & Vision (ICIEV)

## Program Book

**Venue:** Room: EB.G.08 (East Building, Ground floor, Room No. 8), Docklands, University of East London, London, E16 2RD. [(Adjacent to the Cyprus Station (DLR line))]



## Program at a glance:

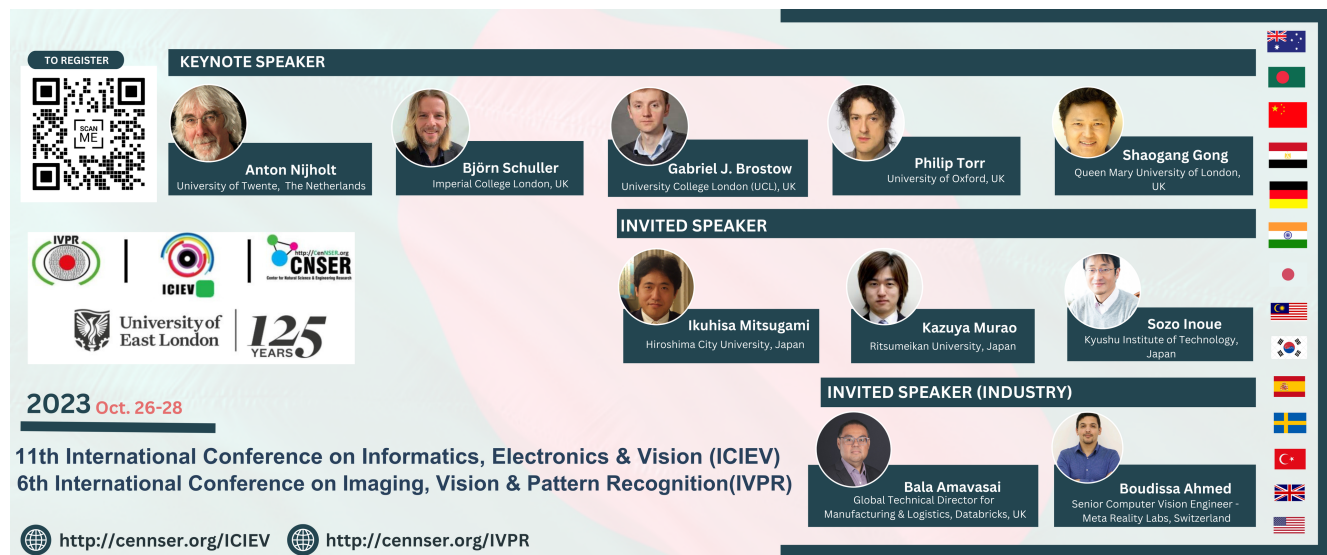
Time Zone Converter: <https://www.timeanddate.com/worldclock/converter.html>

Presentation + QA: Regular (8+7); WIP (5+5); Keynote (45); Invited (30) [in Min.]. (We wish to have a bit longer QA session for presenters - so that we can learn through more interactions)

To join ONLINE via ZOOM: <https://tinyurl.com/2md7z7cz>



Welcome to use this image as ZOOM/FB/... profile/background:



**TO REGISTER**

**KEYNOTE SPEAKER**

- Anton Nijholt**  
University of Twente, The Netherlands
- Björn Schuller**  
Imperial College London, UK
- Gabriel J. Brostow**  
University College London (UCL), UK
- Philip Torr**  
University of Oxford, UK
- Shaogang Gong**  
Queen Mary University of London, UK

**INVITED SPEAKER**

- Ikuhisa Mitsugami**  
Hiroshima City University, Japan
- Kazuya Murao**  
Ritsumeikan University, Japan
- Sozo Inoue**  
Kyushu Institute of Technology, Japan

**INVITED SPEAKER (INDUSTRY)**

- Bala Amavasai**  
Global Technical Director for Manufacturing & Logistics, Databricks, UK
- Boudissa Ahmed**  
Senior Computer Vision Engineer - Meta Reality Labs, Switzerland

**2023 Oct. 26-28**

**11th International Conference on Informatics, Electronics & Vision (ICIEV)**  
**6th International Conference on Imaging, Vision & Pattern Recognition (IVPR)**

<http://cennser.org/ICIEV> <http://cennser.org/IVPR>

## Day 1: 26 October 2023 (LONDON time)

**08.30: Registration + Breakfast**

**09.00: Welcome Session**

- **Prof. Dr. Anton Nijholt**, Program Chair, ICIEV; University of Twente, The Netherlands
- **Prof. Dr. Matthew Turk**, General Chair, ICIEV, IVPR; FACM; FIEEE; FIAPR; President, Toyota Technological Institute at Chicago (TTIC), USA
- **Dr. Aaron Kans**, Head, CDT, UEL
- **Prof. Dr. Hassan Abdalla**, Provost, UEL
- **Dr. Md Atiqur Rahman Ahad**, General Chair, ICIEV, IVPR; SMIEEE; SMOPTICA, UEL

**09.30: Keynote 1**

**Prof. Dr. Anton Nijholt**

University of Twente, The Netherlands

**Title:** Ubiquitous Augmented Reality

*Session chair: Prof. Dr. Ikuhisa Mitsugami, Hiroshima City University, Japan*

**10.15: Welcome Speech**

**Prof. Matt Bellgard**, Pro Vice-Chancellor Impact and Innovation

University of East London, UK

**10.25: Invited Talk 1**

**Prof. Dr. Sozo Inoue**

Kyushu Institute of Technology, Japan; Director, Care XDX Center, Kyutech, Japan

**Title:** Implementing Care Forecasting Services and Challenges with Generative AI

*Session chair: Prof. Dr. Ikuhisa Mitsugami, Hiroshima City University, Japan*

**10.55: Coffee Break**

**11.10: WiP Session Th-S1** : 73, 66, 90, 77, 83, 81

*Session chair: Prof. Dr. Md Sipon Miah, University Carlos III of Madrid, Spain*

**12.10: Lunch Break:** Please show your lunch coupon :)

**13.30: Regular Session Th-S2:** 10, 26, 41, 78, 101

*Session chair: Dr. Ameer Al-Nemrat, UEL, UK*

**14.45: Regular Session Th-S3 :** 31, 51, 53, 55

*Session chair: Prof. Dr. Md. Ahsan Habib, MBSTU, Bangladesh*

**15.45: Coffee Break**

**16.00: Keynote 2**

**Prof. Dr. Philip Torr FREng**

Five AI/RAEng Research Chair in Computer Vision,  
Royal Society Wolfson Research Merit Award Holder, University of Oxford, UK

**Title:** Results on theory and application of transformers

*Session chair: Prof. Dr. Anton Nijholt, University of Twente, The Netherlands*

**16.50: Regular Session Th-S4:** 91, 99, 72, 39

*Session chair: Dr. Seyed Ali Ghorashi, UEL, UK*

## Day 2: 27 October 2023

**08.30: Registration and Breakfast**

**09.00: Regular Session Fri-S5:** 54, 111, 110, 70

*Session chair: Dr. Md Roman Bhuiyan, Fraunhofer Institute for Computer Graphics Research IGD, Germany*

**10.00: Invited Talk 2**

**Prof. Dr. Ikuhisa Mitsugami**

Hiroshima City University, Japan

**Title:** Human Internal State Estimation Using CV and VR Technologies

**Invited Talk 3**

**Prof. Dr. Kazuya Murao**

Ritsumeikan University, Japan

**Title:** Wearable Sensing Security and Applications

*Session chair: Dr. Ahmed Boudissa, Meta, Switzerland*

**11.00: Coffee Break**

**11.20: WiP Session Fri-S6:** 33, 100, 50, 57, 106, 58

*Session chair: Dr. Sharif Saeed, UEL, UK*

**12.20: Lunch Break:** Please show your lunch coupon :)

**14.00: Regular Session Fri-S7:** 45, 52, 60, 61, 62

*Session chair: Dr. AbdulRazak Abba, UEL, UK*

**15.15: Keynote 3**

**Prof. Dr. Sean Gong FREng**

Turing Fellow, Alan Turing Institute; Queen Mary University of London, UK

**Title:** Multimodal Self-Supervised Learning

*Session chair: Prof. Dr. Moinul Zaber, University of Dhaka, Bangladesh; Senior Academic Fellow, United Nations University EGOV, Portugal*

**16.00: Coffee Break**

**16.15: Keynote 4**

**Prof. Dr. Gabriel J. Brostow**

University College London (UCL), UK

**Title:** Is it 2D? Is it 3D? LookOut!

*Session chair: Prof. Dr. Kazuya Murao, Ritsumeikan University, Japan*

**17.00: Regular Session Fri-S8:** 9, 24, 30, 36, 40, 29

*Session chair: Dr. Shaheen Khatoon, UEL, UK*

## Day 3: 28 October 2023

**08.30: Registration and Breakfast**

**09.00: Regular Session Sat-S9:** 44, 112, 11, 6

*Session chair: Dr. Shahera Hossain, University of Asia Pacific, Bangladesh*

**10.00: Keynote 5**

**Prof. Björn W. Schuller, Fellow, IEEE**

Imperial College London, UK; University of Augsburg, Germany; CSO, audEERING

**Title:** Computing Your Health: Turning the Vision into Reality?

*Session chair: Prof. Dr. Sozo Inoue, Kyutech, Japan*

**10.45: Coffee Break**

**11.00: WiP Session Sat-S10** : 68, 88, 109, 25, 49, 16

*Session chair: Prof. Dr. Kazuya Murao, Ritsumeikan University, Japan*

**12.00: Invited Speaker (Industry) 1**

**Dr. Bala Amavasai**

Global Technical Director for Manufacturing and Logistics, Databricks, UK

**Title:** Accelerating Industry 4.0 through Data and AI

**Invited Speaker (Industry) 2**

**Dr. Ahmed Boudissa**

Senior Computer Vision Engineer, Meta, Switzerland

**Title:** Machine Perception for the Metaverse

*Session chair: Prof. Dr. Md. Zahidul Islam, Islamic University, Kushtia, Bangladesh*

**13.00: Lunch Break:** Please show your lunch coupon :)

**14.15: Regular Session Sat-S11** : 32, 35, 56, 86

*Session chair: Dr. Mustansar Ali Ghazanfar, UEL, UK*

**15.15: Regular Session Sat-S12** : 76, 34, 4, 22

*Session chair: Dr. Sujit Biswas, UEL, UK*

**16.20: Closing and Award Ceremony**

Prof. Dr. Anton Nijholt, Program Chair, ICIEV

Prof. Dr. Björn W. Schuller, FIEEE, Program Chair, IVPR

Prof. Dr. MAS Kamal, Program Chair, ICIEV

Dr. Upal Mahbub, SMIEEE; Program Chair, IVPR

Dr. Md Atiqur Rahman Ahad, General Chair, ICIEV, IVPR

**17.00: Tutorial (Online)**

**On-Sensor Machine Learning with ST Toolchains**

**Dr. Mahesh Chowdhary**

Senior Director of MEMS software solutions, STMicroelectronics, USA

**Dr. Swapnil Sayan Saha**

Algorithm Development Engineer of MEMS software solutions, STMicroelectronics, USA

*Session chair: Dr. Upal Mahbub, SMIEEE, Qualcomm, San Diego, USA*

## Welcome Message from the Chairs

We are very pleased to welcome you to the Joint 2023 11th International Conference on Informatics, Electronics & Vision (ICIEV) and the 6th International Conference on Imaging, Vision & Pattern Recognition (IVPR) in London, UK. We are delighted to come up with a very interactive program for the participants of these conferences. The program contains 5 Keynote Talks, 5 Invited Talks, including 2 from industry experts (from Meta and Databricks), and 12 Technical Sessions. There is also a Tutorial session in our program arranged by 2 knowledgeable experts from STMicroelectronics, USA. We truly believe that this joint conference will provide the participants with a great opportunity to share their research ideas and outcomes, and expand their professional networks for future collaboration at the international level.

This joint conference has attracted a total of 107 paper submissions. After a rigorous review process, 40 papers have been included in the technical program for presentation. The review process was conducted by our expert technical committee members and a pool of highly qualified reviewers. Each paper was reviewed by at least two reviewers, and all papers were reviewed in two stages to ensure high quality (via Author Feedback in revision, then final decision). In addition to these regular papers, we have 18 Work-in-Progress (WiP) papers for interactive presentation. Authors of the presented papers are from 14 different countries around the globe namely, Australia, Bangladesh, China, Egypt, Germany, India, Japan, Malaysia, South Korea, Spain, Sweden, Türkiye, UK, and USA.

We would like to express our heartiest gratitude to all the reviewers for completing the reviews within a short time despite your busy schedules. Your reviews were very helpful in selecting good-quality papers and, thereby, maintaining the standards for the conference program. We would like to thank all speakers for your kind presence at the conferences to share your expertise with the attendees. We have no doubt that the participants will greatly benefit from your insightful talks and stimulating discussions. Finally, we would like to thank all the authors, organizing committee members, and volunteers for their great support all the way. We believe this joint conference will be another wonderful and rewarding experience in your memory.

**Program Chairs:** *Anton Nijholt, University of Twente, The Netherlands; Björn W. Schuller, FIEEE, Imperial College London, UK; MAS Kamal, SMIEEE, Gunma University, Japan; Upal Mahbub, SMIEEE, Qualcomm Technologies, Inc., USA*

**General Chairs:** *Toshio Fukuda, FIEEE; IEEE President 2020; Nagoya University, Japan; Matthew Turk, FIEEE; FIAPR; President, Toyota Technological Institute at Chicago, USA; Md Atiqur Rahman Ahad, SMIEEE, SMOPTICA; University of East London, UK*

## Day 1: 26 October 2023

### 08.30: Registration + Breakfast

### 09.00: Welcome Session

- **Prof. Dr. Anton Nijholt**, Program Chair, ICIEV; University of Twente, The Netherlands
- **Prof. Dr. Matthew Turk**, General Chair, ICIEV, IVPR; FACM; FIEEE; FIAPR; President, Toyota Technological Institute at Chicago (TTIC), USA
- **Dr. Aaron Kans**, Head, CDT, UEL
- **Prof. Dr. Hassan Abdalla**, Provost, UEL
- **Dr. Md Atiqur Rahman Ahad**, General Chair, ICIEV, IVPR; SMIEEE; SMOPTICA, UEL

### 09.30: Keynote 1

**Prof. Dr. Anton Nijholt**  
University of Twente, The Netherlands  
**Title:** Ubiquitous Augmented Reality

### 10.15: Welcome Speech

**Prof. Matt Bellgard**, Pro Vice-Chancellor Impact and Innovation  
University of East London, UK

### 10.25: Invited Talk 1

**Prof. Dr. Sozo Inoue**  
Kyushu Institute of Technology, Japan; Director, Care XDX Center, Kyutech, Japan  
**Title:** Implementing Care Forecasting Services and Challenges with Generative AI

### 10.55: Coffee Break



## 11.10: WiP Session Th-S1

### 73: A Study on Face Orientation-Controlled Smart Walker for Individuals with Parkinson's

*Lilian Picard (Kyushu Institute of Science and Technology)\*, Tomohiro Shibata (Kyushu Institute of Technology)*

This study introduces our innovative Smart Walker solution tailored for individuals with disabilities. By integrating robotic actuators powered wheels, an RGB camera, and a head position estimation algorithm using MediaPipe (Google corp), the walker seeks to enhance gait metrics, improve comfort, optimize mobility, and ensure user safety. Experimental investigations were conducted to assess the effectiveness of the four-wheel walker's control system based on face orientation intention, involving both healthy participants and a patient with Parkinson's disease.

### 66: Influence of Feedback Modality in Core Training Support System

*Keisuke Sato (Aoyama Gakuin University)\*, Guillaume Lopez (Aoyama Gakuin University)*

The COVID-19 pandemic has had a significant impact on the global economy and people's health. During these challenging times, maintaining a regular exercise routine is crucial for promoting overall well-being. However, engaging in core training exercises without proper guidance can be challenging. To address this issue, we propose a novel system called CoreMoni- , which utilizes wearable sensors and smartphones to estimate the user's pose and provide real-time feedback accurately. CoreMoni- incorporates a three-stage posture determination method and a feedback mechanism that takes into account the user's actions. By analyzing data from a single inertial motion unit, the system can provide multimodal feedback in the form of 3D skeleton images and text. This feedback helps users maintain proper form and technique during core training exercises. To evaluate the effectiveness of the CoreMoni- system, we conducted an experiment comparing it with a conventional system. Impressively, 95% of the participants preferred the new system, citing improved pose determination accuracy, enhanced user experience, and increased usability. In addition to the current capabilities, future enhancements of the CoreMoni- system include incorporating non-intrusive feedback methods, refining device placement for optimal performance, and developing long-term support content. These advancements will further enhance the system's effectiveness and user satisfaction. Overall, the proposed CoreMoni- system serves as a valuable tool for individuals seeking to maintain their health and fitness through core training, especially during periods when access to gyms and outdoor activities is limited. By providing accurate

posture evaluation and real-time feedback, the system empowers users to engage in effective and safe core exercises, promoting their overall health and well-being.

## **90: Analysis of Phishing Website Detection using Machine Learning Algorithms**

*Md Ibrahim Mamun (Northumbria University)\*, Muhammad Ali Kazmi (Northumbria University)*

The popularity of using social media such as Facebook, and Twitter creates opportunities for cybercriminals to redirect internet users to fraudulent websites. Phishing websites in this case play a vital role in misleading internet users. Internet users can easily be tricked and made to disclose their sensitive data such as date of birth, bank details, credit card details, memorable words, addresses, and more. Besides global incidents like COVID-19 create opportunities for cyber attackers to misguide web users to visit phishing websites using sensitive headlines, news, and digital content such as images and videos. Being misguided web users may share their private data with cybercriminals and that may cause them financial or strategic loss. Not only individuals but organizations also are not safe from phishing attacks. Any employee from an organization can be tricked and led to visit a phishing website. For this reason, the company may face difficulties in the competitive market. Therefore, automatic detection of phishing websites is required from the user end services. This research work aims to analyse phishing website detection using Machine Learning (ML) algorithms such as Logistic Regression, K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest Classification, Decision Tree (DT), Naïve Bayes, and Neural Network. Also, the paper analyses and compares the performance of different ML algorithms on a particular dataset which includes web attributes from both legitimate and phishing websites. In this research work, some ML algorithms produced excellent accuracy up to 97% for classifying phishing websites from legitimate websites.

## **77: Physiological Perspectives on Interpersonal Differences in Thermal Comfort**

*Kazuki K Honda (Aoyama Gakuin University)\*; Tahera Hossain (Aoyama Gakuin University); Yusuke Kawasaki (Aoyama Gakuin University); Guillaume Lopez (Aoyama Gakuin University)*

Today, the number of deaths due to heat stroke is increasing every year, with the elderly accounting for a particularly high percentage. In this research, thermal comfort is estimated using physiological indices obtained from wearable sensors. It also aims to evaluate the impact of gender and age on the accuracy of thermal comfort estimation. In the experiment, four situations were set up based on the PMV thermal comfort model: normal, slightly warm, warm, and hot. Gender

and age differences were observed during the experiment through a subjective thermal comfort questionnaire. Participants were instructed to wear the E4 wristband to collect physiological indices while performing various activities such as reading, typing, and gymnastics. These activities were performed under different thermal states, allowing for a comprehensive assessment of the participants' physiological responses in different thermal conditions. Afterward, we extracted time-domain and frequency-domain features from the data. Machine learning techniques, specifically Extra Tree (ET), Random Forest (RNF), and Gradient Boosting (GB), were applied to the datasets. For the evaluation analysis, we conducted assessments on various subsets of the data, including the overall dataset without accounting for individual differences, a dataset consisting of males only, a dataset consisting of females only, a dataset comprising individuals under 50 years old, and a dataset comprising individuals over 50 years old. By adopting this comprehensive approach, we were able to examine the performance of the machine learning models across different demographic segments. The results show that ET has the highest F1 value for all datasets. For each thermal sensation, the normal and hot conditions tended to be less misrecognized for all datasets.

### **83: A Method for Forecasting Thermal Comfort Sensations from Physiological Data**

*Manpreet kaur (University of East London)\*, Md Atiqur Rahman Ahad (University of East London)*

Heatstroke is a serious and potentially life-threatening condition caused by prolonged exposure to high temperatures. Heatstroke should be detected and prevented as early as possible in order to reduce the risk of heat-related illnesses and improve outcomes. Over the past few years, machine learning techniques have shown great promise in various healthcare applications, including disease prediction and monitoring. This research paper explores the application of machine learning algorithms in the context of heatstroke prevention by forecasting thermal comfort sensations by analyzing physiological data. According to physiological measurements, this study aims to develop a predictive model that can estimate an individual's thermal comfort level. To do this, an IEEE dataset containing physiological data such as temperature, humidity, heart rate, and subjective judgments of human thermal comfort is used. To analyze the data and uncover patterns and correlations between physiological measurements and thermal comfort perceptions, various machine learning algorithms, including regression and classification models, will be used. The suggested model will be trained on a subset of the acquired data and validated on the remainder. Performance indicators such as accuracy will be utilized to assess the model's predictive capabilities.

## 81: Dining Frames: Revealing Dining Activity using Smartphone Camera

*Huisha Yang (Aoyama Gakuin University)\*, Guillaume Lopez (Aoyama Gakuin University)*

Identifying and classifying eating-related hand movements is a challenging task that has various applications in monitoring and improving eating habits. However, existing methods for gesture recognition often require specialized sensors or devices that are not widely available or convenient to use. In this study, we propose a novel approach that uses a smartphone camera as the only input device to extract and predict eating behaviors from hand gestures. We use the Mediapipe framework, a cross-platform solution for building multimodal applied machine learning pipelines, to analyze the video frames captured by the smartphone camera. We focus on the hand and face landmarks provided by the framework, which are 3D coordinates of key points on the hand and face regions. We use these coordinates as features for four machine learning models and evaluate their performance using Leave-One-Video-Out Cross-Validation (LOVOCV). Our dataset consists of ten videos that capture different eating scenarios involving various types of food and utensils, various backgrounds, different people, and several camera distances (ranging from 0.5 to 1.5 meters). We label each video frame with one of the following five behaviors: right hand movement, taking food with right hand, left hand movement, taking food with left hand, drinking. Our results show that Random Forest is the best performing model across the dataset. We also find that camera proximity and utensil type have significant effects on model accuracy. Our study demonstrates the feasibility and potential of using smartphone camera gestures to analyze and forecast eating behaviors.

### 12.10: Lunch

### 13.30: Regular Session Th-S2

## 10: Low-profile Wideband $1 \times 2$ Array Antenna for IEEE 802.11a/h/j/n/ac/ax WLAN Applications

*Naymaa Rashid (Pabna University of Science and Technology); Tithi Rani (Rajshahi University of Engineering & Technology); Liton Chandra Paul (Pabna University of Science and Technology)\*; Sajeeb Chandra Das (Pabna University of Science and Technology); Md Ashraful haque (Daffodil International University); Sayed Shifat Ahmed (Pabna University of Science and Technology)*

In this paper, a novel low-profile wideband  $1 \times 2$  array microstrip patch antenna has been proposed

for WLAN (Wireless Local Area Network) applications. The antenna is comprised of a rectangular microstrip feed line and a directly coupled T-shaped strip, which is connected to circular patches through impedance matching technique. The material of the substrate is Rogers RT5880. The ground plane is nearly half the size of the antenna. The antenna has an overall dimension of  $40 \times 22.4 \times 0.79$  mm<sup>3</sup>. The antenna is developed step-by-step to evaluate gain and radiation pattern characteristics for optimum performance. The antenna is designed, simulated, and analyzed using CST Microwave Studio. The simulated outcomes are consistent. This antenna is suitable for IEEE 802.11a/h/j/n/ac/ax WLAN applications due to its operating frequency range of 4.72-6.32 GHz and center frequency of 5.672 GHz. Therefore, the simulated verification indicates the proposed antenna could be a suitable choice for WLAN applications.

## **26: Spectral Amplitude Coding for Secure Multichannel Communication: An Advanced investigation on AND subtraction based SAC-OCDMA**

*Fahmidul Islam (American International University-Bangladesh)\*, Mohammad Nasir Uddin (American International University-Bangladesh)*

The world is changing every day by adopting new technologies. In this developed era of science, every day new invention is coming up intending to ease the livelihoods of the human being. This vast amount of population needs a better communication system among them too. Because of this, the need for transmission bandwidth has expanded along with the growth in population. To mitigate the high demand for bandwidth installing more fiber optics may not always be a feasible option because of the cost of deployment. That is why new and advanced multiple-access solutions are introduced to overcome such situations. SAC-OCDMA, also known as spectral amplitude coding optical code division multiple access, is a recently developed multiple-access system that leverages narrowband and offers superior security for the data network. There are two users in the postulated SAC-OCDMA system design. Having a transmitted data rate of 1.6 Gbit/s, the recommended secure network architecture maintains performance up to 336 kilometers. The simulated result is gained with a minimum BER value of  $<10^{-9}$  which is the minimum acceptable range. The system is simulated using the commercially available Optisystem software.

## **41: Edge-Fog-Cloud Based Hierarchical Communication Network for Traffic Management System**

*Md. Jehad (Ahsanullah University of Science & Technology)\*, Md. Jobairul Hassan (Ahsanullah University of Science & Technology), Yousuf Hossen Hossen (Ahsanullah University of Science & Technology)*

*Technology), Chandramallika Ahmed (Ahsanullah University of Science & Technology), Shamim Akhter (Ahsanullah University of Science & Technology)*

This paper aims to investigate a real-time hierarchical communication network for traffic management systems. The suggested management program would provide effective, cost-efficient, and secure solutions to critical issues such as authentication, hierarchical processing, reduced latency, overtaking judgments, speed adjustments, and accident notifications. We describe the design stages for creating a prototype, dividing the traffic-related decision-making system into three layers (Edge, Fog and Cloud), establishing a security protocol, and comparing the performance of Edge-Cloud and Edge-Fog-Cloud-based communication networks in this article. The proposed communication network has been discovered to be more efficient and quicker, making it an excellent alternative for real-time traffic safety applications.

### **78: Gait Analysis for Cerebral Palsy Using Memory-Augmented Auto-Encoder Model**

*Ginga Sumi (Mie University)\*, Balaji Iyer (Cincinnati Children's Hospital Medical Center; University of Cincinnati), V. B. Surya Prasath (Cincinnati Children's Hospital Medical Center; University of Cincinnati; University of Cincinnati College of Medicine), Hiroharu Kawanaka (Mie University), Bruce J. Aronow (Cincinnati Children's Hospital Medical Center; University of Cincinnati College of Medicine)*

Cerebral Palsy impairs movement, which limits people's function and social participation. Treatment of patients with Cerebral Palsy requires routine and quantitative gait analysis. Unfortunately, gait analysis requires expensive motion capture systems and highly trained personnel. It leads to financial burdens for medical institutions and a considerable physical strain on patients. This study aims to establish a method for analyzing patients' gait without special equipment. This paper addressed the challenge of estimating the deviation of patients' gait from standard case, employing only an ordinary video of patients' gait and unsupervised anomaly detection methods. We used a Memory-augmented Auto-Encoder (MemAE), an extended model of Auto-Encoder (AE) based on unsupervised anomaly detection, improving the performance of AE. In the method, we train the MemAE using the data of skeleton trajectories during gait, including the two-dimensional image plane positions of joints (e.g., ankles and knees) estimated by an off-the-shelf human pose estimation algorithm from ordinary videos. The obtained correlation coefficient  $r$  between the reconstruction error and pathological gait index was -0.416 and outperformed the compared other unsupervised anomaly detection models. This result suggests that the proposed method can capture the patients' gait deviation from normality compared with the typical model.

## 101: Cervical Cancer Detection Using Multi-Branch Deep Learning Model

*Tatsuhiro Baba (The University of Aizu); Abu Saleh Musa Miah (The University of Aizu); Jungpil Shin (The University of Aizu)\*; Md. Al Mehedi Hasan (Rajshahi University of Engineering & Technology)*

Cervical cancer is a crucial global health concern for women, and the persistent infection of High-risk HPV mainly triggers this remains a global health challenge, with young women's diagnosis rates soaring from 10% to 40% over three decades. While Pap smear screening is a prevalent diagnostic method, visual image analysis can be lengthy and often leads to mistakes. Early detection of the disease can contribute significantly to improving patient outcomes. In recent decades, many researchers have employed machine learning techniques that achieved promise in cervical cancer detection processes based on medical images. In recent years, many researchers have employed various deep-learning techniques to achieve high-performance accuracy in detecting cervical cancer but are still facing various challenges. This research proposes an innovative and novel approach to automate cervical cancer image classification using multihead self-attention and convolutional neural networks (CNNs). The proposed method leverages the strengths of both multi-head self-attention mechanisms and CNNs to effectively capture both local and global features within cervical images in two streams. Multihead self-attention facilitates the model's ability to focus on relevant regions of interest, while CNN extracts hierarchical features that contribute to accurate classification. Finally, we combined the two stream features and fed them into the classification module to refine the feature and the classification. To evaluate the performance of the proposed approach, we used the SIPaKMeD dataset, which classifies cervical cells into five categories. Our model achieved a remarkable accuracy of 98.03%. This performance has high recognition accuracy of medical image classification and holds promise for its applicability in other medical image recognition tasks.

### 14.45: Regular Session Th-S3

## 31: Enhancing Glaucoma Prediction Performance in OCT Images using Machine Learning Based on HOG Features

*Mahjabin Rahman Oishe (Rajshahi University of Engineering & Technology)\*, S. M. Mahedy Hasan (Rajshahi University of Engineering & Technology), Azmain Yakin Srizon (Rajshahi Uni-*

*versity of Engineering & Technology), Anwar Hossain Efat (Rajshahi University of Engineering & Technology), Nahrin Jannat (Rajshahi University of Engineering & Technology), Mostarina Mitu (Rajshahi University of Engineering & Technology), Md. Farukuzzaman Faruk (Rajshahi University of Engineering & Technology)*

Glaucoma is a chronic eye disease characterized by progressive damage to the optic nerve, often resulting in vision loss or blindness. Manual diagnosis of glaucoma presents challenges like human error and labor-intensive processes, making technology-aided diagnosis preferable. Therefore, in this study, different feature extraction and Machine Learning (ML) approaches have been used for detecting glaucoma using Optical Coherence Tomography (OCT) images. The utilized approaches have revealed high classification performance in detecting glaucoma, as demonstrated through empirical evaluation on the ‘Acrima’ dataset. These findings will contribute to the development of more efficient techniques for glaucoma recognition and prevention, especially in settings with limited computational resources.

### **51: In-Depth Analysis of Automated Acne Disease Recognition and Classification**

*Afsana Ahsan Jeny (Bahcesehir University), Masum Shah Junayed (University of Connecticut), Md. Robel Mia (Daffodil International University), Md Baharul Islam (Bahcesehir University)\**

Facial acne is a common disease, especially among adolescents, negatively affecting both physically and psychologically. Classifying acne is vital to providing the appropriate treatment. Traditional visual inspection or expert scanning is time-consuming and difficult to differentiate acne types. This paper introduces an automated expert system for acne recognition and classification. The proposed method employs a machine learning-based technique to classify and evaluate six types of acne diseases to facilitate the diagnosis of dermatologists. The pre-processing phase includes contrast improvement, smoothing filter, and RGB to L\*a\*b color conversion to eliminate noise and improve classification accuracy. Then, a clustering-based segmentation method, k-means clustering, is applied for segmenting the disease-affected regions that pass through the feature extraction step. Characteristics of these disease-affected regions are extracted based on a combination of gray-level co-occurrence matrix (GLCM) and Statistical features. Finally, five different machine learning classifiers are employed to classify acne diseases. Experimental results show that the Random Forest (RF) achieves the highest accuracy of 98.50%, which is promising compared to the state-of-the-art methods.



### **53: Survival Analysis and Therapeutic Drug Targets Identification for Head and Neck Cancer and Chronic Lymphocytic Leukemia Cancer**

*Md. Anayt Rabbi (Jagannath University)\*, Md Manowarul Islam (Jagannath University), Md. Ashraf Uddin (Deakin University), Arnisha Akter (Jagannath University), Selina Sharmin (JNU)*

This research sought to discover relevant gene expression patterns of cancer cells having chronic lymphocytic leukemia (CLL) and head and neck cancer (HNC). We conducted a comparative genomic study to demonstrate the genetic differences between HNC and CLL. To find the differentially expressed genes in HNC and CLL related datasets obtained from the gene expression omnibus (GEO), we utilized the GEO2R program. Additionally, we used the Kyoto Encyclopedia of Genes and Genomes and wikipathway enrichment analysis to determine the biological roles of the DEGs. In order to find hub genes, we also built protein-protein networks. A protein-protein interactions network (PPI) employing differentially expressed genes revealed 64 concordant genes for both HNC and CLL samples. The CAV1, CCND1, CDKN2A, ITGA6, BCL2, ITGA3, CD8A, and PRNP hub genes have been identified in PPI networks through cytohubba plugin. The discovery of signaling pathways was extensively analyzed to offer proof of genomic responses to HNC and CLL. The hub genes were shown to be substantially expressed in cancer cells and strongly related with survival ( $p < 0.05$ ). The pathogenetic examination of signaling pathways may aid in the discovery of new drugs treatment of cancer.

### **55: Investigation of Emotional Effects on Brain Network Stimulation through EEG Signals**

*Mahfuza Akter Maria (Khulna University of Engineering & Technology), M. A. H. Akhand (Khulna University of Engineering & Technology)\*; Md Abdus Samad Kamal (Gunma University)*

Recently, research on emotion from the perspective of brain networks is becoming popular. In this study, brain functional and effective connectivity networks have been constructed from DEAP emotional EEG data to study how emotion influences patterns of this connectivity. According to the investigation results, more direct correlations are found under positive emotions than negative ones. The brain regions operate more synchronously, and there is less directed flow of information between brain regions during negative emotions. The correlation between brain regions, whether direct or inverse, is higher in the lower frequency band than in the higher frequency band. The flow of information from one brain region to another brain region increases with higher frequency and there is more synchrony between brain regions in the Gamma frequency band.

## 15.45: Coffee Break

## 16.00: Keynote 2

### **Prof. Dr. Philip Torr FREng**

Five AI/RAEng Research Chair in Computer Vision,  
Royal Society Wolfson Research Merit Award Holder,  
University of Oxford, UK

**Title:** Results on theory and application of transformers

## 16.50: Regular Session Th-S4

### **91: Cross-Lingual Transfer Learning for Arabic Signature Verification: Dataset and Baseline Evaluation**

*Tameem Bakr (Egypt-Japan University of Science and Technology)\*, Ahmed I Abdullatif (Egypt Japan University of Science and Technology), Kareem Elzaky (Egypt-Japan University of Science and Technology), Mohamed Elsayed (Egypt Japan University of Science and Technology), Rami f Zewail (Egypt-Japan University of Science & Technology)*

Writer-independent offline signature verification is a widely utilized biometric authentication method and an area of active research. However, Writer-independent offline signature verification poses unique challenges in the Arabic language context. Despite the availability of benchmarking datasets such as CEDAR, and BHSig260 for Writer-independent offline signature verification research for languages such as English, Hindi, and Bengali, there is a notable absence of the availability of Arabic datasets. In this paper, we address this gap by introducing an Arabic signature verification dataset that serves as a benchmark for cross-lingual transfer learning evaluation as a cross-validation dataset. To establish a baseline for cross-lingual transfer learning for Arabic writer-independent offline signature verification, we adopt the Siamese neural network architecture. The results indicate that the dataset is comparably performing to existing benchmarks on cross-language transfer learning scenarios setting a solid baseline for the research potential in cross-lingual transfer learning. Moreover, we explore a voting ensemble approach that further enhances the accuracy of the model, highlighting the potential for continued research and improvement in this domain.

## **99: Extraction of Clinically Validated Spatio-Spectral Propagation Markers of Schizophrenia using a Hyper-Network parameterized Explainable Model**

*Mainul Islam Labib (Khulna University), Md. Johir Raihan (Khulna University), Md Atiqur Rahman Ahad (University of East London), Abdullah-Al Nahid (Khulna University)\**

Schizophrenia, a multifaceted mental disorder marked by symptoms such as hallucinations, delusions, disorganized thinking, and altered emotional responses, presents a diagnostic challenge. A combination of clinical assessment and close examination of specific symptoms is required for the complex and frequently drawn-out process of diagnosing schizophrenia. Modern diagnostic research has been relying on non-invasive mediums such as EEG to build automatic diagnosis system but majority of them lack explainability and as a result, these improvements are not reflected back to the clinical settings. Lack of homogeneity across training and cross-validation of extracted insights also create a fake vacuum of information which do not have clinical value. Therefore, in this work, we have developed a Neural Network architecture parameterized by Hyper-Networks which uses clinically aware models to draw insights from spatial and spectral propagation data. Additionally, we have chosen a publicly available and well studied dataset with enough precursors in scientific literature such that our extracted insights can be easily cross validated. Our spatial and spectral insights, particularly related to activity in the Prefrontal and Temporal lobes during and activity, are in clear alignment with previous research and established neurobiological knowledge. In future works, we want to build this model even further to include diagnostic feature extraction schemes to make the process more reliable.

## **72: Intracranial Hemorrhage Segmentation and Application of Interpretable Transfer Learning using Grad-CAM for Classification in Computed Tomography Images**

*Tazqia Mehrub (Bangladesh University of Professionals)\*, Mosabber Uddin Ahmed (University of Dhaka)*

Intracranial Hemorrhage (ICH) being a serious condition associated with a raised risk of many other complications, might result in mortality. The absence of accurate localization or delay in subtype classification can cause incurable circumstances. Former machine learning techniques used in segmenting or classifying the CT scan images of ICH resulted in low model accuracy and low trustability due to the unavailability of proper interpretability of medical image region classification. In this study, we explore the use of transfer learning with VGG, ResNet, InceptionResNet V2, InceptionNet, and ExceptionNet models for intracranial hemorrhage CT image

classification. To improve the interpretability of our results, we use Grad-CAM to visualize the regions of the image that were most important for the classification decision. We use segmentation for intensity identification, calculating the area of the generated masks. The proposed framework for semantic image segmentation using U-Net architecture demonstrates its effectiveness in accurately identifying and labeling objects in images by predicting pixel-level segmentation masks. The performance of the model is evaluated using intersection over union (IoU) scores and the proposed model yielded a Dice coefficient of 0.91, which outperforms other supervised learning techniques irrespective of the number of labeled images present in training the model. Our results show that transfer learning with these models can achieve high accuracy for intracranial hemorrhage classification, and the segmentation can help to identify regions of interest in the image. The use of Grad-CAM provides additional insights into the decision-making process of the models, which can aid in the clinical interpretation of the results.

### **39: A biologically inspired model for perception of flicker wheel illusion**

*Keerthi S Chandran (Indian Statistical Institute)\*, Kuntal Ghosh (Indian Statistical Institute)*

Flicker Wheel illusion is a visual illusion of a wheel with spokes which appear to flicker. The perception of flicker wheel illusion is affected by physical parameters of the stimulus like number of spokes and contrast. This work generates multiple variants of flicker wheel by varying physical parameters associated with the stimulus. The stimuli are then presented on a computer screen and is labelled by a human subject into two classes on basis of whether the subject perceives flicker in the stimulus or not. DNNs are trained to classify the images the same way a human subject does. Many existing CNN Networks are shown to be able to classify the data. However, the best performance comes from the proposed bio-inspired Deep Neural Network that takes into account the microsaccades in the human eye. The proposed architecture yields accuracies close to the consistency of human subject in the psychophysics experiment.

## Day 2: 27 October 2023

**08.30: Registration and Breakfast**

**09.00: Regular Session Fri-S5**

### **54: A Low-cost IoT-based Meteorological System Using LoRaWAN and Embedded Tech**

*Norbert Dajnowski (York St John University), Aminu Bello Usman (University of Sunderland)\*, Andrew Guest (York St John University), Abdulrazaq Abba (University of East London), Saifur Rahman Sabuj (BRAC University)*

The development of meteorological stations is continuously evolving, with advanced technologies and solutions being developed to achieve the most accurate data while maintaining cost efficiency. Governments and businesses are leveraging the Internet of Things (IoT) platforms to provide hyperlocal and highly sophisticated compact meteorological stations for analysing real-time weather conditions and forecasts with unprecedented accuracy. In this study, we developed a stable, low-cost meteorological system capable of recording weather parameters that was powered with a low-cost and long-range data transfer technology to provide multiple nodes access to the internet through the LoRaWAN network server. We then developed a GUI application for meteorological systems' data visualisation to display information on the weather and stations on a mobile app to the users in the form of numbers, charts, and graphs.

### **111: Hybrid Deep Learning for Assembly Action Recognition in Smart Manufacturing**

*Abdul Matin (University of Technology Sydney)\*, Md Rafiqul Islam (University of Technology Sydney), Yeqian Zhu (University of Technology Sydney), Xianzhi Wang (University of Technology Sydney), Huan Huo (University of Technology Sydney), Guandong Xu (University of Technology Sydney)*

Deep learning algorithms have become essential in human activity recognition (HAR) for driving advancements in intelligent manufacturing. While numerous sensor systems and algorithms are developing, their real-world applicability and robustness within the manufacturing sector need validation. Artificial intelligence (AI) applications in manufacturing have gained significant traction in both academic and industrial circles. A critical aspect of intelligent manufacturing systems involves recognizing activities performed by manufacturing workers, particularly repetitive assem-

bly tasks. This recognition facilitates real-time efficiency measurement and evaluation of workers while providing augmented reality instructions to enhance their performance on the job. Despite limited research in this domain, identifying human actions in intelligent manufacturing is vital for achieving sustainable manufacturing goals aligned with Industry 4.0 principles. This paper introduces a hybrid deep-learning approach designed to monitor assembly tasks to recognize human actions within the manufacturing context. The model's performance is evaluated through simulations conducted on the HA4M dataset, comprising diverse multimodal data-capturing actions executed by various individuals constructing an Epicyclic Gear Train (EGT). Our research approach is organized into three primary domains: (i) evaluating baseline deep learning model performance on the HA4M dataset, encompassing ConvLSTM (94.61%), LSTM with VG16 (81.73%), LRCN (86.53%), 3D CNN (95.57%); (ii) crafting a hybrid deep learning model by combining ConvLSTM and 3D CNN, initially achieving 87.30% accuracy and subsequently refining it to 94.23% for enhanced effectiveness; (iii) introducing a computationally efficient deep learning architecture that leverages both 3D CNN and convLSTM to adeptly manage HA4M dataset, resulting in an impressive accuracy of 99.42%.

### **110: Convolutional Neural Network-based Secure Spectrum Sensing in CR-IoT Networks**

*Md Sipon Miah (University Carlos III of Madrid)\*, Md Mahbubur Rahman (Islamic University), Ana Garcia Armada (UC3M)*

Spectrum sensing plays a vital role in Cognitive Radio-based Internet of Things (CR-IoT) networks to better utilizing the allocated band. However, the Malicious Users (MUs) degrade the sensing performance because they report their false sensing data to the corresponding Fusion Centre (FC) through a perfect reporting channel. In this paper, we propose a Convolutional Neural Network (CNN)-based secure cooperative spectrum sensing in CR-IoT networks using multiple features. This scheme is able to accurately detect MUs from normal CR-IoT users. Moreover, we solve the probability of error minimization and the sum rate maximization problems for both primary and secondary networks using the Golden Section Search (GSS) method. Simulation results show that the proposed scheme achieves better sensing performance, lower error rate, enhanced system sum rate and security compared with other conventional schemes.

**70: Simulator-based Dataset Generator for Pothole Detection in Autonomous Driving**

*Yuga Yano (Kyushu Institute of Technology)\*, Yukiya Fukuda (Kyushu Institute of Technology), Hakaru Tamukoh (Kyushu Institute of Technology)*

Pothole detection is crucial for autonomous driving. However, the number of images and the quality of the dataset required for training a detection model are insufficient. Therefore, we propose a simulator-based RGB-D dataset generator for pothole detection. To achieve Sim2Real, a smoothing filter was applied to the depth image obtained from the simulator to make it resemble a real depth image. Additionally, obstacles other than potholes were generated in the simulator to reduce the false-positive rate. In the experiment, the proposed dataset generator generated 20,000 annotated RGB-D image pairs within 16 min. The F-score of the model trained on the proposed dataset was improved to 70%, compared to 48% for the model trained on an existing dataset.

**10.00: Invited Talk 2**

**Prof. Dr. Ikuhisa Mitsugami**

Hiroshima City University, Japan

**Title:** Human Internal State Estimation Using CV and VR Technologies

**Invited Talk 3**

**Prof. Dr. Kazuya Murao**

Ritsumeikan University, Japan

**Title:** Wearable Sensing Security and Applications

**11.00: Coffee Break**

## 11.20: WiP Session Fri-S6

### **33: Safeguarding User Privacy through Effective Information Disclosure Summary from the Software End-User License Agreements (EULAs)**

*Shamim Bin Zahid (University of Dhaka)\*, Aishwarya Ghosh Bristy (University of Dhaka), Moinul Zaber (University of Dhaka), Sarker T Ahmed Rumea (University of Dhaka)*

An End-User License Agreement (EULA), a significant precursor to modern privacy policies, are essential documents that communicate with people about how their personal information is collected, used, and protected by organizations / companies, especially through software. EULAs serve as a legal framework to govern product usage. However, the current standard of representing privacy policy documents is often composed of legal jargon, which makes them lengthy, complex, and challenging for the average user to comprehend. Not everything mentioned in a protracted EULA document is meant to be vital for a user's information safety. But proper representation of the document is crucial. There has been research on standardizing and summarizing privacy policies, A properly usable representation is yet to be standardized. This paper conducts an analysis of existing norms, policies and mechanisms to make EULAs more understandable and communicable. This serves as a vital precursor to designing intervention that can help users better understand how the applications may impact their privacy and personal security.

### **100: Personal Thermal Comfort Forecasting for Heat Stroke Prevention using Machine Learning**

*Jigar P Patel (University of East London)\*, Md Atiqur Rahman Ahad (University of East London)*

The goal of Heatstroke Prevention is to create a machine-learning model that can predict an individual's thermal comfort levels based on past experience. The training dataset comprises time-stamped observations made by 27 people over a 6-day period, with particular feature values present at specified times in each observation. For 2 days of data from 9 individuals whose thermal experiences are unknown, the goal is to develop a model that can precisely anticipate such sensations. This study examines alternative methods for choosing the dataset's most pertinent properties, as well as how various models based on activities, gender, and age might improve the precision of thermal comfort estimates. The model is trained using six days of training data, and the two days of testing data that must be used to forecast the thermal sensations make up the testing data.



## 50: IoT-based Autonomous Robot for Smart Farming

*A S M Ahsanul Sarkar Akib (Robo Tech Valley)\*, N.M.Raziul Hassan RISHAD (American International University), Md Ibrahim Shikder Mahin (Robo Tech Valley), Mahadir Islam (Robo Tech Valley), M. Firoz Mridha (AIUB), Md Shamiul Islam (Bangladesh University of Business and Technology), Asrafun Nesa Aktia (Robo Tech Valley), Shifat Hassan (Shamoloy Ideal College)*

Agriculture and smart farming get a new dimension through the Internet of Things (IoT) and artificial intelligence (AI). The adoption of a smart sensor system coupled with the Internet of Things (IoT) transformed agriculture and the AI-based crop management system. It also made it simpler to monitor agricultural fields. The three key nutrients—nitrogen (N), phosphorus (P), and potassium (K), as well as soil moisture and humidity—that are measured in agricultural fields by a robocar and sent to a cloud server are discussed in this work. We created a route path for the agricultural field using a GPS technology to develop the autonomous system. A home position is also created by the robot. Following the path, it will enter the field where its robotic arm will automatically pick up soil and store it in a chamber. We made use of the humidity, soil moisture, and NPK fertilizer sensors. After removing the soil, the device will test its settings and relay the results to an Internet of Things server. AI has suggested a suitable crop for that field based on these data. The robot will also proclaim the soil's quality, and if any parameter rises or falls, it will let the user know. The robot will return to its home position by clicking the "return home" button.

## 57: IoT-based Smart Remote Door Lock and Monitoring System using an Android Application

*Jakia Sultana (Bangladesh University of Business & Technology), Raihan Sikdar (Bangladesh University of Business & Technology), A S M Ahsanul Sarkar Akib (Robo Tech Valley)\*, Md. Anwar Hussen Wadud (Mawlana Bhashani Science and Technology University), Md. Ahsan Habib (Mawlana Bhashani Science and Technology University), Nur Nahar (ISTT)*

Today, having a completely secure home is essential. In our regular lives, we frequently forget to lock our doors, and later we suffer from confusion that whether we lock all doors perfectly or not. In this paper, we propose a cheap, cost-effective, customized Smart Door Lock System to improve our quality of life and level of security. We use an Android application and the ESP32-CAM to construct an Internet of Things (IoT) -based smart door lock system. In this project, we implement a smart door lock system, allowing the owner to use an Android application to

view the visitor's photo and lock or unlock their doors from anywhere at any time. The owner will receive a notification on his or her smartphone each time a visitor rings the doorbell at the residence. The owner can then use the Android App to view the visitor's photo and can remotely unlock the door and allow him to enter the home. They can use their smartphone to verify the door status whenever they are unsure about the Door Lock. The door lock signal in the application will display a green signal if the door is properly locked. The indicator will turn red if the door is not completely locked, at that point the owner can easily remotely lock their door from anywhere. Theft Alert is another feature of this project. When the owner leaves the house, if someone approaches the door and tries to make force entry, a theft alert message will be sent to the owner's smartphone and a loud buzzer alert will ring inside the home to alert the neighbors, who can then take appropriate action. The automatic door lock feature is also available in this system.

### **106: Efficient Overtaking of Automated Vehicles Using V2V Communication**

*Shengsheng Liu (Gunma University)\*, Srivardhini Veeraragavan (Monash University Malaysia), MAS Kamal (Gunma University), Kou Yamada (Gunma University)*

Overtaking becomes necessary when there are stopped cars, partially blocked lanes, or other obstructions on the road, which frequently causes congestion or accidents. On single-lane roads, most drivers typically find it challenging to pass stopped or slow vehicles while they encounter incoming traffic. This paper proposes an automatic overtaking coordination system utilizing vehicle-to-vehicle (V2V) communication to perform smooth and safe overtaking in various traffic scenarios. The system aims to minimize driving costs regarding traffic delays while limiting collision risks on overtaking lanes. Under the system, vehicles share the relevant information in the test environment through V2V communication, and optimal coordination directions are sent back to them to improve overtaking maneuver quality while enhancing road traffic efficiency. Simulation results show that the proposed system can significantly improve vehicle overtaking efficiency, reduce idling time, and increase the traffic speed on the lane under various constraints imposed by oncoming vehicles and obstacles while enhancing efficiency and reducing environmental pollution.

## **58: A Hybrid Deep Learning Approach for Heart Image Analysis: A Systematic Review of Recent Advances**

*Md Abu Sufian (University of Leicester)\*, Jayasree Varadarajan (Manchester Metropolitan University)*

In recent years, deep learning models have made a significant impact on the field of cardiac imaging, revolutionizing early diagnosis methods. This systematic review aims to provide a comprehensive analysis of recent advancements in hybrid deep learning techniques for cardiac image analysis, particularly focusing on the roles of Auto encoders and CNNs. The study leveraged various academic databases and publications relevant to the domain of cardiac imaging and deep learning. Peer-reviewed articles and studies published in the last five years were considered. Only those incorporating Auto encoders, CNNs, or a combination of both in the context of heart image analysis were included. For clinical studies where specific patient populations or samples are studied. Two deep learning models were developed for this review: An Autoencoder model optimized for image quality and representation and a CNN model primarily designed for image recognition. Both models were built using the TensorFlow and Keras frameworks. The Autoencoder model's performance was appraised based on accuracy metrics, while the CNN model was evaluated based on training accuracy and precision. **Synthesis Methods:** Our Autoencoder model achieved an accuracy of approximately 98.66% on test data. Concurrently, the CNN model demonstrated a training accuracy of 90% and a precision rate of 98.9%. The combined strengths of Autoencoders and CNNs present a promising direction for enhanced cardiac diagnosis. Their significant results in our study suggest immense potential for broader applications in medical technology.

### **12.20: Lunch**

### **14.00: Regular Session Fri-S7**

## **45: Evaluating Road Surface Marking Quality using Synthetic Noise-Based Reconstruction Models**

*Mehieddine Boudissa (Mie University)\*, Hiroharu Kawanaka (Mie University), Tetsushi Wakabayashi (Mie University)*

Ensuring clear and well-defined road surface markings is crucial for safe and efficient traffic management. However, the quality degradation of these markings over time poses significant challenges

to road safety. In this paper, we address this issue by proposing a new approach that leverages synthetic noise data to enhance the reconstruction of low-quality road surface markings. Our proposed solution consists of two main stages. Firstly, we employ a pre-existing segmentation model to accurately isolate road surface markings from the surrounding environment. Subsequently, we introduce synthetic noise into the segmented markings using two distinct methods: random circle drawing and a combination of erosion and Gaussian noise application. This process serves to simulate the deterioration commonly observed in real-world road markings. To accomplish the reconstruction task, we train an auto-encoder. By learning from the synthetic noise data, the trained model aims to reconstruct the low-quality road surface markings into its original state. Extensive evaluation on a synthetic validation set demonstrates promising results. However, when applied to real-world data, the reconstruction performance faces challenges, revealing a disparity between synthetic and real-world domains.

## **52: Residual Learning and Filtering Networks for End-to-End Lossless Video Compression**

*Md Baharul Islam (Bahcesehir University)\*, Afsana Ahsan Jeny (Bahcesehir University)*

Existing learning-based video compression methods still face challenges related to inaccurate motion estimates and inadequate motion compensation structures. These issues result in compression errors and a suboptimal rate-distortion trade-off. To address these challenges, this work presents an end-to-end video compression method that incorporates several key operations. Specifically, we propose an autoencoder-type network with a residual skip connection to efficiently compress motion information. Additionally, we design motion vector and residual frame filtering networks to mitigate compression errors in the video compression system. To improve the effectiveness of the motion compensation network, we utilize powerful nonlinear transforms, such as the Parametric Rectified Linear Unit (PReLU), to delve deeper into the motion compensation architecture. Furthermore, a buffer is introduced to fine-tune the previous reference frames, thereby enhancing the reconstructed frame quality. These modules are combined with a carefully designed loss function that assesses the trade-off and enhances the overall video quality of the decoded output. Experimental results showcase the competitive performance of our method on various datasets, including HEVC (sequences B, C, and D), UVG, VTL, and MCL-JCV. The proposed approach tackles the challenges of accurate motion estimation and motion compensation in video compression, and the results highlight its competitive performance compared to existing methods.

## 60: A Tri-modal Fusion Network for Object Detection Using Small Amounts of Low-Quality Data

*Yusuke Watanabe (Kyushu Institute of Technology)\*, Yuma Yoshimoto (National Institute of Technology), Hakaru Tamukoh (Kyushu Institute of Technology)*

By employing multiple sensors, a robot can improve its ability to detect objects. In this paper, we propose a tri-modal fusion network, which accepts common layout tri-modal data of RGB, depth, and thermal data as input data and is composed of three sub-networks fused into one network. We also propose a new non-random initialization method of the weights and biases of the tri-modal fusion network. Results of object detection experiments on a dataset which is composed of small amounts of low-quality data showed that the proposed tri-modal network outperformed uni-modal networks and bi-modal fusion networks. We conclude that the proposed tri-modal network can be trained on small amounts of low-quality data without diverging and achieve high performance in environments where data of multiple sensors are required for accurate object detection.

## 61: Estimate of Cat Excretion Information Using Multiple Weight Sensors

*Tetsuhiro Yamaguchi (Kyushu Institute of Technology)\*, Namiko Nakashima (Toiletta Cats Inc.), Takeshi Shigemori (Toiletta Cats Inc.), Teruki Hirahata (Toiletta Cats Inc.), Sozo Inoue (Kyushu Institute of Technology)*

The incidence of urinary system diseases in cats tends to increase with age. Chronic kidney disease, in particular, cannot be completely cured once it occurs, so it is important to detect the disease at an early stage to halt its progression. We are therefore aiming to accurately estimate the health status of cats using TOLETTA, a cat litter box that is generally available on the market, based on the cat's excretion information obtained from TOLETTA. As a first step, this paper aims to correctly extract the voiding interval from the four-weight sensor data installed in TOLETTA. The number of data is 296 urination classes. In the experiment, features were generated from the extracted voiding intervals, and the amount of urination was estimated by Random forest. As a result, compared to the existing system used in the existing TOLETTA (hereafter referred to as "the existing system"), the Random forest model this time has an RMSE score of 7.73g, while the existing system has a score of 8.46g, resulting in a 0.73g reduction in error. The R2 score was also improved by 0.09, from 0.45 for the existing system to 0.54 for the Random forest model. The challenge is that it is necessary to consider the behavior of each individual in order to correctly extract the excretion zone from the weight sensor data.

**62: Facial Expression Editing of Untrained Subject Image Using Landmarks**

*Kanade Sumino (Hiroshima City University), Shota Harada (Hiroshima City University), Ikuhisa Mitsugami (Hiroshima City University)\*, Ryusuke Sagawa (AIST)*

Many applications require methods for intuitively controlling facial expressions in images by editing the facial landmark points. However, most existing methods, such as LandmarkGAN, have limitations and can only be applied to trained subject face images. In this paper, we propose a method that can overcome the problem of those existing works. This method is based on First Order Motion Model (FOMM), which can transfer the facial expression of a person (reference video) to a single face image of another person. It uses keypoint features that describe changes in facial expressions acquired through unsupervised learning from face image datasets, but the correspondence between the values of these keypoint features and facial expressions is not clear, which makes intuitive expression editing difficult. This method estimates the change of FOMM keypoint features from that of facial landmark points acquired by OpenFace.

**15.15: Keynote 3**

**Prof. Dr. Sean Gong FREng**

Turing Fellow, Alan Turing Institute  
Queen Mary University of London, UK

**Title:** Multimodal Self-Supervised Learning

**16.00: Coffee Break****16.15: Keynote 4**

**Prof. Dr. Gabriel J. Brostow**

University College London (UCL), UK

**Title:** Is it 2D? Is it 3D? LookOut!

## 17.00: Regular Session Fri-S8

### 9: Generation of Clothing Items with Jamdani Motif Elements Using Automated Generative Adversarial Networks

*Samiur Rahman Abir (Ahsanullah University of Science and Technology)\*, Hujaiifa Islam (Ahsanullah University of Science and Technology), Md. Sakibur Rahman (Ahsanullah University of Science and Technology), Hasan Mahmud (Ahsanullah University of Science and Technology), Mohammad Shafiu Alam (Ahsanullah University of Science and Technology)*

Clothing serves as an artistic medium for humans to express their preferences, thoughts, and cultural heritage, while the application of machine learning, particularly Generative Adversarial Networks (GANs), remains largely unexplored in the realm of clothing production and design, with designers currently relying on their imaginative skills to create diverse styles. In this article, Conditional Generative Adversarial Networks (cGAN) are used to suggest an automated approach. Neural style transfer and cGAN algorithms are employed to create traditional clothing with distinctive patterns and a variety of styles. For this study, the Fashion MNIST and Jamdani Motif Dataset datasets were both employed. The conditional GAN model was used to produce several styles of apparel using the MNIST dataset. The Neural Style Transfer model is then used to combine the created picture with the Jamdani Motif pattern from the Jamdani Motif dataset. Using Otsu's image segmentation technique, the foreground, and background of the resulting picture are separated. Performance scores of this model are as follows: Inception Score is 1.3573909, Frechet inception distance is 1272.222597, Kernel Inception Distance is 636200.667, Coverage Metric is 33.79799. We polled several people on our work output, and the results are detailed in a later section.

### 24: A deep convolutional neural network based anti-aliasing technique for real-time 3D applications

*F. M. Jamius Siam (BRAC University), Zahidul Islam Prince (BRAC University), Ahmed Nafisul Bari (Personal), Jia Uddin (Woosong University)\**

In real-time 3D applications, ensuring smooth edges in different scenarios remains vital, primarily due to low-resolution output, memory, and processing-related constraints. This paper presents a deep convolutional neural network-based anti-aliasing model for solving the aliasing problem, where aliasing refers to the problem of having hard, jagged edges in the rendered scenes. A convolutional neural network is used to learn the jaggedness of an image from the given training

set and reproduce an anti-aliased output image. In the model, several convolutional neural network layers, max-pooling layers, and convolutional transpose layers are used to generate the anti-aliased images. In the experimental evaluation, we utilized a dataset that consists of demo game scenes created using the Unity game engine. The dataset consists of both raw images and super-sampled images. As a performance metric, both SSIM and PSNR scores are used to measure the proposed model's accuracy. Experimental results demonstrate that our model shows an average SSIM score of 0.9941 and a PSNR score of 42.265 for our tested dataset. The proposed model approaches, and often outperforms the other state-of-art models—DLAA, FXAA, and NFAA by exhibiting higher SSIM scores and PSNR values in specific scenarios.

### **30: The Integrity of Source Code Commenting: Benchmark Dataset and Empirical Analysis**

*Maksuda Islam (Islamic University of Technology)\*, Md Safayat Hossen (Islamic University of Technology), Ahsanul Hoque (Islamic University of Technology), Md Nazmul Haque (Islamic University of Technology), Lutfun Nahar Lota (Islamic University of Technology)*

Code comments are a vital software feature for program cognition & software maintainability. For a long time, researchers have been trying to find ways to ensure the consistency of code-comment. While doing that, two of the raised problems have been dataset scarcity and language dependency. To address both problems in this paper, we created a dataset using C# projects; there are no annotated datasets yet on C#. 9,310 code-comment pairs of different C# projects were extracted from a data pool. 4,922 code-comment pairs were annotated after removing NULL, constructor, and variable. Both method-comment and class-comment were considered in this study. We employed two evaluation metrics for the dataset, one is Krippendorff's Alpha which showed 95.67% similarity among the rating of three annotators for all the pairs & other is Bilingual Evaluation Understudy (BLEU) to validate our human-curated dataset. An ensemble machine learning model with topic modeling is also proposed, which obtained 96.2% using the performance metric AUC-ROC after fitting the model to our proposed dataset.

### **36: A Crop Recommendation System With Transformer Based Deep Learning Model**

*Nabil Sadd Sammo (Bangladesh University of Engineering and Technology)\*, Shamim Akhter (Ahsanullah University of Science and Technology)*

The development of effective and sustainable agricultural practices has become necessary in recent



years due to the rising demands on the world's food production. The incorporation of cutting-edge technology like machine learning and artificial intelligence (AI) into farming systems has drawn a lot of attention as a solution to this problem. The purpose of this study is to improve agricultural decision-making and optimize crop selection based on environmental and soil parameters. One of the most notable recent developments in deep learning and deep neural networks is the transformer model. It is mostly utilized for sophisticated applications in speech recognition, text-to-speech transformation, and natural language processing. In this study, we will develop a self-transform machine learning model that will examine the gathered soil, weather, and crop information and provide an accurate intuition regarding the best crop to grow under a given circumstance. The proposed framework has a 98.1% accuracy in training and a 97.3% accuracy in testing.

#### **40: Network Dynamics of Bangladesh Inland Waterways augmented with Complex Network Analysis of Vessels and Connectivity**

*Firoza Tabassum (BRAC University), Humaira Islam (BRAC University)\*, Matin Saad Abdullah (BRAC University), Moinul Zaber (Dnd Lab)*

The Bengal Delta is considered as world's largest delta with the three largest rivers flowing across Bangladesh. As the cheapest medium of transportation, inland waterways vehicles play the most significant role in connecting the habitats making them the most economically significant means for goods and human transportation. To understand the inland waterway structure, complex network analysis has proven to be an important tool through which we can identify port connectivity, bottlenecks, significant ports, and clusters. Several authors have conducted such analysis to ascertain various networks including transportation and port connectivity networks. However, data scientific research works on connectivity networks focusing on the developing world are scarce. In previous research, researchers analyzed on port connectivity of districts indicating the structure of the river-based connectivity looking from the administrative perspective. However, the analysis lacked more information such as vessel frequency from the ports in the districts. Vessel frequency is necessary to understand the significance of the ports. The analysis of this paper is based on a uniquely developed dataset of ports and their vessel frequency giving a more complete picture of port conditions and characteristics of inland waterways of Bangladesh. Our analysis identifies the important ports and the dependency of the network on one city- Dhaka capital of Bangladesh. These findings can be helpful in understanding and modeling how the migration of a vast number of economically challenged populations actually happens. This can also help to ascertain many other socio-economic factors that are affected by inland waterways transportation,

such as poverty, local business and trades, and transportation of goods and humans.

## **29: An Improved Framework for Classification of Skin Cancer Lesions using Transfer Learning**

*Tanjim Mahmud (Kitami Institute of Technology)\*, Koushick Barua (Rangamati Science and Technology University); Anik Barua (Rangamati Science and Technology University), Sudhakar Das (Rangamati Science and Technology University), Rishita Chakma (Chittagong University of Engineering & Technology); Nanziba Basnin (International Islamic University Chittagong), Nahed Sharmen (Chattogram Maa-O-Shishu Hospital Medical College), Mohammad Shahadat Hossain (University of Chittagong), Karl Andersson (Luleå University of Technology)*

Skin cancer is nowadays a great concerning health problem all over the world. Right now, it is becoming a common disease in the world. It will be a deadly disease if it is not identified in the early stages. Early detection of skin cancer gives the greatest chance for successful skin cancer treatment. Skin cancer most often develops on skin exposed to the sun. But this common form of cancer can also occur on areas of skin not ordinarily exposed to sunlight. The most concerning subject is that over 123,000 melanoma and 30,00,000 non-melanoma cases are found worldwide each year. As the growth rate of skin cancer patients is very high and it is necessary to identify the cancerous and non-cancerous diseases at an early stages. Consequently, we have been proposed a better model with higher accuracy with a CNN model along with few pre-trained models. Considering the inability of dermatologists who do not get an accurate diagnosis of skin cancer traditionally, the necessity to develop an automated efficient system comes to light for the diagnosis of skin cancer. This study brings an efficient automated method for skin cancer classification with better evaluation and accuracy matrices improved as compared to previous studies or expert dermatologists which are 94% for the proposed CNN model and 97% has the highest among the pre-trained models.

## Day 3: 28 October 2023

### 08.30: Registration and Breakfast

### 09.00: Regular Session Sat-S9

#### 44: Nuclei Instance Segmentation of Cryosectioned H&E Stained Histological Images using Triple U-Net Architecture

*Fardifa Fathmiul Alam (Islamic University of Technology)\*, Chowdhury Nue E Alam Siddiqi (Islamic University to Technology), Zarif Ahmed (Islamic University of Technology), Tasnim Ahmed (Islamic University to Technology), Tareque Mohmud Choudhury (Islamic University to Technology)*

Nuclei instance segmentation is a crucial step in oncological diagnosis and cancer pathology research. H&E stained images are commonly used for medical diagnosis, but pre-processing is necessary before using them for image processing tasks. Two principal pre-processing methods are formalin-fixed paraffin-embedded samples (FFPE) and frozen tissue samples (FS). While FFPE is widely used, it is time-consuming, while FS samples can be processed quickly. However, analyzing FS-derived HE stained images can be challenging due to the rapid preparation, staining, and scanning process leading to image quality degradation. This paper proposes a method that leverages the unique optical characteristics of H&E stained images. We have used a three-branch U-Net architecture, where each branch contributes to the final segmentation results. The process includes applying a Gaussian filter and watershed algorithm to separate overlapping regions and enhance accuracy. The U-Net architecture comprises an RGB branch, a Hematoxylin branch, and a Segmentation branch. The results obtained through robust experiments outperform the state-of-the-art results across various metrics. This study focuses on a novel dataset named CryoNuSeg, the first-ever FS Dataset featuring 30 FS-sectioned images of 10 human organs. The benchmark score for this dataset is AJI 52.5 and PQ 47.7. However, the proposed triple U-Net architecture achieves an AJI score of 67.41 and PQ of 50.56. The use of the three-branch U-Net model, followed by watershed post-processing, significantly surpasses the benchmark scores, showing substantial improvement in the AJI score.

## 112: Exploring Oversampling Techniques for Fraud Detection with Imbalanced Classes

*Sultan Alharbi (University of Technology Sydney)\*, Abdulrhman Alorini (University of Technology Sydney), Khaled Alahmadi (University of Technology Sydney), Hadeel Alhosaini (University of Technology Sydney), Yeqian Zhu (University of Technology Sydney), Xianzhi Wang (University of Technology Sydney)*

Credit card fraud has caused significant losses for financial institutions and individuals worldwide each year. Financial institutions must detect credit card fraud to prevent customers from being charged for products they did not order. Class imbalance has been a standing challenge for credit card transactions, as the number of fraudulent transactions is significantly lower than that of non-fraudulent transactions. In this paper, we comprehensive evaluate five oversampling techniques, namely SMOTE, Adaptive Synthetic Sampling (ADASYN), Borderline SMOTE, Random Oversampling, and SMOTE Support Vector Machine (SMOTE SVM) in combination with seven machine learning techniques (namely XGBoost, Random Forest, K-Nearest Neighbor, Naive Bayes, Support Vector Machine, LightGBM, and Convolution Neural Network). Our results show oversampling generally improves fraud detection performance and SMOTE SVM is the better oversampling method than other methods under test. Notably, it achieved an accuracy of 76.47% when used with KNN on the smaller dataset and 99.93% with CNN on the larger dataset used in our experiments.

## 11: Similarity Fusion: Camera-LiDAR Vision Similarity Fusion for 3D Object Detection

*Yaming Xu (Harbin Institute of Technology)\*; Wang Yan (Harbin Institute of Technology); Boliang Li (Harbin Institute of Technology)*

This paper proposes a novel vision similarity fusion network for fusing 2D camera images and 3D LiDAR point clouds, utilizing merged vision feature for 3D object detection. Similarity fusion network provide a non-coupled, low-complexity vision fusion framework that significantly improves the ability to connect heterogeneous vision sensor. Our fusion network fuses the proposed candidate combinations of camera and LiDAR decoders before non-maximum suppression. A vision correlation fusion law is designed using Cosine similarity, and non-coupled vision features is fused according to correlation without additional training. Finally, we designed a fusion detection head to achieve 3D detection of the object using a combination of visual features. Experiments on the challenging KITTI object detection library show that our fusion scheme significantly improves ob-

ject detection, especially for small object, compared to individual point cloud and image detection schemes. We will release our code after acceptance.

## 6: Chicken Behaviour Analysis Using Computer Vision Techniques

*Md Roman Bhuiyan (Fraunhofer Institute for Computer Graphics Research IGD)\*; Philipp Wree (Fraunhofer Institute for Computer Graphics Research IGD)*

A growing population is reliant on the essential sector of chicken breeding for subsistence. However, the chickens well-being is a major worry since bad living conditions create erratic behaviour, which impairs the flock's health and output. To monitor and improve the animals well-being, a surveillance system that records chicken behaviour and alerts farmers to potential concerns must be in place. Farmers have long depended on experts to identify and diagnose ailments in chickens. Farmers lose a lot of farmed chickens as a result of delayed diagnosis or a lack of reliable specialists. In this paper, we have solved a problem for chicken behaviour analysis from high-crowd chicken images with resolutions of  $3840 \times 2160$  using computer vision techniques. One chicken image has around 400 chickens with high resolutions images and it is really difficult to identify the chicken behaviours. Our proposed method employs a Mask-RCNN based approach the state of the art for behaviour analysis used for monitoring in chicken farms. This study aims to address a computer vision- based system for identifying and tracking chicken behaviour in poultry farms. chicken, chicken eating and chicken sitting are the three behaviours that the system classifies. To address this issue, this research aims to develop a dataset based on chicken behaviour analysis from the chicken fram. The proposed method for chicken behaviour analysis achieved an accuracy of mAP (means average precision) = 96.75% with the three classes respectively. The proposed chicken behaviour analysis result are outperformed the state-of-the-art model in most cases.

### 10.00: Keynote 5

**Prof. Dr. Björn W. Schuller, Fellow, IEEE**

Imperial College London, UK;

University of Augsburg, Germany; CSO, audEERING

**Title:** Computing Your Health: Turning the Vision into Reality?

### 10.45: Coffee Break

## 11.00: WiP Session Sat-S10

### 68: Precision fish farming to mitigate pond water quality through IoT

*Abdus Salam (Department of Fisheries, GoB); A S M Ahsanul Sarkar Akib (Robo Tech Valley)\*; Md Shamiul Islam (Bangladesh University of Business and Technology); Md. Ahsan Habib (Mawlana Bhashani Science and Technology University (MBSTU)); Nur Nahar Akter (ISTT)*

In the face of increasing environmental challenges, efficient monitoring of natural water bodies like ponds has become imperative. This work proposes a revolutionary approach by combining Internet of Things (IoT) technology with six vital sensors - Dissolved Oxygen (DO), Turbidity, Temperature, pH, Electro Conductivity (EC), and Salinity - to create a complete Pond Water Quality Monitoring System. Sensors at crucial locations collect real-time water quality data. Microcontrollers capture and send data to a cloud-based platform via defined communication protocols from these sensors. Cloud platforms store, analyze, visualize, and interact with data. Stakeholders can view the water quality dynamics of the pond on a website and Android apps. Real-time graphs, charts, and maps show temperature, pH, dissolved oxygen, turbidity, EC, and salinity changes. A comprehensive warning mechanism alerts users when any parameter violates specified levels, allowing quick actions to prevent potential hazards. The system's data analysis provides sophisticated insights. Water quality data patterns, trends, and anomalies help understand ecosystem health and pollution sources. Remote control of pumps, aerators, and dosing devices helps maintain ponds with precise, up-to-date data. IoT-based Pond Water Quality Monitoring System will change pond management. This system utilizes the Internet of Things (IoT) and sensor technologies to offer information to stakeholders, enhance environmental conservation efforts, safeguard aquatic ecosystems for future generations, and enable effective expansion of our fish farming operations by regular monitoring of pond water, yielding useful insights. The study results in the development of a pond water quality monitoring system utilizing Internet of Things (IoT) technology, which can be accessed by mobile and web applications.

### 88: Understanding the Household Electricity Bills and Perceptions in Dhaka City

*Murshed Al Amin (Data and Design Lab)\*; Moinul Zaber (University of Dhaka); Sarker T Ahmed Rumeel (University of Dhaka); Md Abu Sayed (Independent University, Bangladesh)*

Electricity consumption patterns, billing accuracy, and user perceptions play a pivotal role in understanding energy utilization behaviors and informing policy interventions. This research delves into the intricate relationship between perceived electricity consumption and actual billing data

within the context of the power distribution companies in Dhaka. Through a comprehensive analysis of data gathered from 577 customers, this study explores the convergence and disparities between users' perceptions of their electricity bill and their actual bill data. The analysis encompasses visualizations of geographic distribution, comparisons of individual bills over time, grouping consumers based on billing patterns, and the examination of alignment between perceived and mean bill ranges. The findings highlight the nuances of consumer perception accuracy and uncover potential geographic disparities. Despite certain limitations such as participant memory and data processing intricacies, this study contributes to enhancing billing systems, energy conservation, and policy formulation. By providing insights into the gap between perceived consumption habits and actual usage, this research seeks to foster more informed energy behavior and contribute to sustainable practices in electricity consumption. Future work involves deeper analysis of the comprehensive survey data to extract additional insights from a broader spectrum of users' attitudes and behaviors.

### **109: Artificial Intelligence for Seismic Structural Damage Assessment: The Future Prospects**

*Kaveesh G Abey Suriya (University of East London)\*; Kaveesh G Abey Suriya (University of East London); Mihaela Ciupala (University of East London); Seyed Ali Ghorashi (University of East London); Alper Illki (Istanbul Technical University)*

”Evaluating the structural integrity of reinforced concrete (RC) structures following an earthquake is a time-consuming but critical process that enables stakeholders to make important decisions about occupancy safety and rectifications. Engineers conduct these assessments by thoroughly examining structural damages based on seismic design codes and guidelines, primarily through visual inspections, and occasionally by structural analysis based on measurement data. Artificial intelligence (AI) techniques, including machine learning (ML) and deep learning (DL) techniques, have been used recently by researchers to detect damage in RC buildings. These data-driven approaches have successfully identified cracks, spalling, and crushing in structural components by automatically extracting damage-sensitive features from images. This paper examines recent advancements and research gaps in utilising AI in seismic structural damage assessment to offer a framework for identifying damage states swiftly and accurately. Convolutional neural networks (CNN) and transfer learning (TL) have shown unparalleled accuracy and capability of classifying seismic damage in reinforced concrete components of a building structure. Transfer learning retrains already trained deep networks to classify structural damage by fine-tuning with a limited

relevant dataset, as such in the case of civil engineering applications, enhancing generalisation and preventing overfitting. Therein, this paper aims to explore the gaps in existing research body to establish solid grounds to develop a robust deep learning-based model incorporating transfer learning techniques using seismic damage images to identify local element failure modes and classify the damage state of the structure per the criteria delineated in seismic design codes and guidelines.

## **25: An Information Technology-enabled Query-Based Approach to Mitigating the Shortcomings of Widely Used Learning Methods**

*Khandoker Ashik Uz Zaman (Independent University, Bangladesh)\*; Md Abu Sayed (Independent University, Bangladesh); Yusuf Mahbubul Islam (Independent University, Bangladesh); Moinul Zaber (University of Dhaka)*

This technology-driven study begins with a concise review of popular existing learning methods, identifying the gaps that hinder their effectiveness. To address these gaps, this paper proposes a hybrid method named “Query-Based Access to Neurons (QuBAN)” that leverages Information Technology to establish a connection between the learner’s natural curiosity and the knowledge acquisition process using questions. As a key component of the QuBAN method, this paper introduces a progressive-web application concept aimed at enhancing learner’s ability to ask well-developed questions and promote deeper learning experiences. The proposed application features a questioning platform that will employ Bloom’s Taxonomy to rank questions and assign scores based on the level of inquiry students engage in on their classroom topic. The foundation of QuBAN lies in recognizing the paramount role of questions in influencing synaptic plasticity – the brain’s mechanism for learning and retaining memory. The prevailing challenge of passive information transmission between educators and students in traditional learning methods has led to limited classroom interactivity and over-reliance on rote memorization. While non-traditional learning methods like Inquiry-Based Learning (IBL) and Socratic Learning Method (SLM) have shown promise in promoting active learning, they still struggle to sustain high levels of student engagement and motivation throughout the entire session. In contrast, the proposed QuBAN method presents a promising hybrid learning approach that effectively harnesses the power of questions and information technology. Ultimately, this study explores the potential of the QuBAN approach through the use of technology. The proposed application should transform education and empower learners to become active participants in their knowledge-acquisition process through the art of questioning.



#### **49: Fusion of statistical models at denoiser to restore local features from high density SAP noise**

*Golam Moktader Daiyan (Mawlana Bhashani Science and Technology University); Md. Ahsan Habib (Mawlana Bhashani Science and Technology University)\**

The presence of salt and pepper noise in images affects the image quality. Images that are contaminated by the noise require an appropriate filter. Several filtering approaches have been improved, yet local feature and edge restoration remain a challenge. We aim to improve the efficiency of the filtering approach at high-density of noise. A statistical estimated filter performs a noise elimination through ranking features in the following steps: detecting exact noise, counting median without noise at lower noise density, mean at higher noise density, and the approximation of discontinuity of the processing pixel. Local image data are only considered while performing filter operations on the distorted images. The maximum 255 or the minimum 0 pixel value for black and white image is not considered SAPN. The filter utilizes the simple mathematical logic at  $3 \times 3$  mask and deploys complex logic with increased noise density. The Visual experiment shows that the filter can recover most image information from higher noise density. Performance analysis matrices such as the peak signal to noise ratio (PSNR), image enhancement factors(IEF), and the structural similarity index measure (SSIM) show that the proposed filter effectively removes the high- density of noise from the images without any significant loss of image details.

#### **16: Machine Learning Algorithms for Passive OS Fingerprinting: A Study with Real-Time Data and Ensemble Techniques**

*Fahmina Nur Salma (Bowie State University)\*; Ayen Kuol (Bowie State University); Timothy Oladunni (Morgan State University); Hoda El-Sayed (Bowie State University)*

Detecting operating system (OS) class, such as Windows, Ubuntu, or Linux, is crucial for identifying threats and vulnerabilities in a system, contributing to network security. Traditional OS fingerprinting, which involved manual identification of network packet transmissions, has evolved into an automated approach using Machine Learning (ML) techniques to generate and detect OS signatures. In this study, we compare the prediction accuracy, CPU execution time, CPU utilization, and memory usage of various ML algorithms. Real-time data from the CICIDS2017 dataset is utilized, and different ML models are implemented to determine the most efficient approach. Additionally, OS system resources are monitored during the testing phase of each experiment. The research provided insights into the trade-offs between time, CPU utilization, and memory

usage for algorithm selection in OS fingerprinting.

### 12.00: Invited Speaker (Industry) 1

**Dr. Bala Amavasai**

Global Technical Director for Manufacturing and Logistics, Databricks, UK

**Title:** Accelerating Industry 4.0 through Data and AI

### Invited Speaker (Industry) 2

**Dr. Ahmed Boudissa**

Senior Computer Vision Engineer, Meta, Switzerland

**Title:** Machine Perception for the Metaverse

### 13.00: Lunch

### 14.15: Regular Session Sat-S11

#### 32: An Experimental Study on Speech Emotion Recognition for Bangla Language

*Moinul Zaber (University of Dhaka); Sarker T Ahmed Rumeel (University of Dhaka)\*; Md. Mehedi Hasan (University of Dhaka)*

Embedding speech recognition has increased the usability of artificial intelligence (AI) applications (e.g., smart home, chatbots, etc.). Recent research works have focused on increasing usability further enabling personalized applications. Speech emotion recognition (SER) is one of the building blocks of speech-enabled personalized systems. However, recognition of emotion from speech is complex as the intensity of the emotion, also known as arousal, can vary from person to person and change within a very short time. Data is another major challenge, especially for low-resource languages such as Bangla. To our knowledge, only two datasets are available open-sourced for Bangla speech emotion recognition. Considering the gaps, we have decided to focus on this problem. Our goal is to study the existing system, experiment, and evaluate the usability of available resources. In this work, we are focusing on applying deep learning techniques to available data resources in the Bangla language and evaluating their performance to find suitable techniques for Bangla SER. We have utilized two publicly available datasets, SUBESCO and BanglaSER.

We have used a set of 11 different features as the audio representation. Convolutional neural networks and artificial neural networks have been used as deep learning techniques for this study. We have performed three types of experimental study on three different setups of datasets. We have achieved more than 70% accuracy for all setups without cross-validation. After cross-validation, we have achieved 68.08% average accuracy on setup 3 (only BanglaSER). After regularization with cross-validation, we have achieved 62.62% average accuracy on setup 3. We have evaluated the performance of SUBESCO and BanglaSER for other setups and approaches. This work will summarize the current state of Bangla SER which will help research communities to improve the existing system accordingly.

### **35: Evaluating Performance of Machine and Deep Learning Techniques in Analyzing Multi-class Bangla Public Sentiment Using word2vec and doc2vec Embeddings**

*Md Tazimul Hoque (United International University); Syed Tangim Pasha (Daffodil International University); Rubaiya Khanam (American International University-Bangladesh); Ashraful Islam (Independent University, Bangladesh)\*; Md Zahangir Alam (Independent University, Bangladesh); Mohammad Nurul Huda (United International University)*

The advent of representing textual documents as vectors or numerical data has greatly influenced the domain of natural language processing (NLP). Embedding methods, e.g., word2vec and doc2vec are pivotal in this transformation. While word2vec generates vectors for individual words, doc2vec escalates this process by producing vectors for entire sentences or documents. This comprehensive representation allows for comparisons of large volumes of words or sentences concurrently, thus optimizing computational resources and bandwidth. This study aims to train doc2vec and word2vec models on a corpus comprising 10,500 Bangla Facebook posts, with the objective of classifying sentiments as positive, negative, or neutral, in accordance with different public reactions (e.g., love, wow, haha, etc.) to these posts. The sentiment classification leverages nine different ML and DL techniques, including Logistic Regression (LR), Support Vector Machine (SVM), Decision Tree (DT), K-nearest Neighbor (KNN), Linear Discriminant Analysis (LDA), Gaussian Naive Bayes (GNB), Sequential Model (SM), Long Short-term Memory (LSTM), and Bidirectional Long Short-term Memory (BiLSTM). The performance of these techniques is evaluated through two 10-Fold cross-validation procedures, where document vectors are directly sourced from the doc2vec model, and TF-IDF averaged document vectors are derived from the word2vec model. The BiLSTM classifier is ranked as the top performer by achieving a word2vec accuracy of 77.34% and a doc2vec accuracy of 75.56%, closely followed by the LSTM model with

a word2vec accuracy of 76.71% and a doc2vec accuracy of 74.97%. However, the DT classifier performs the worst with the lowest accuracy scores of 57.9% and 51.02% for word2vec and doc2vec, respectively.

## **56: Empowering Bengali Language in Drone Control with Artificial Neural Networks**

*Sajjad Hossain Talukder (Premier University); Tanjim Mahmud (Kitami Institute of Technology)\*; Nanziba Basnin (International Islamic University Chittagong); Rishita Chakma (CUET); Koushick Barua (Rangamati Science and Technology University); Mohammad Shahadat Hossain (University of Chittagong); Karl Andersson (Luleå University of Technology)*

The demand for intuitive and efficient methods of controlling drones is increasing with the advancement of drone technology. Traditionally, a drone is controlled by a remote controller, which requires a trained operator to pilot the drone. This paper presents a novel approach that utilizes a deep learning model on a Raspberry Pi to enable the control of drones, allowing users to pilot drones quickly and easily since they would not require any knowledge about remote controller devices. It's all about picking up the remote, giving any random Bengali voice commands, and effortlessly piloting drones. In the proposed study, a model based on artificial neural networks is outlined for controlling a drone system by using random Bengali voice commands in the real environment. A remote, which is basically a Raspberry Pi, employs the trained model that was trained on a compact dataset, encompassing 15 distinct commands, leveraging the Mel-frequency Cepstral Co-efficients (MFCCs) feature of audio data. “Take Off/ ”, “Turn Right/ ”, “Go Up/ ”, “Move Forward/ ”, “Stop/ ” are some examples of commands that we have harnessed to train the model. We have achieved an impressive accuracy of 73% through extensive experimentation, highlighting the robustness of our model in classifying commands.

## **86: Arabic Music Classification and Generation using Deep Learning**

*Mohamed I Elshaarawy (Egypt-Japan University of Science and Technology); Ashrakat S Elfawal (Egypt-Japan University of Science and Technology)\*; Mariam A Sheta (Egypt-Japan University of Science and Technology); Assem Mohamed Bakr (Egypt-Japan University of Science and Technology); Abdelrahman Said Ahmed (Egypt-Japan University of Science and Technology); Omar B Sayed El Ahl (Egypt-Japan University of Science and Technology); Walid Gomaa (Egypt-Japan university of science & technology)*

This paper proposes a machine learning approach for classifying classical and new Egyptian music by composer and generating new similar music. The proposed system utilizes a convolutional neural network (CNN) for classification and a CNN autoencoder for generation. The dataset used in this project consists of new and classical Egyptian music pieces composed by different composers. To classify the music by composer, each sample is normalized and transformed into a mel spectrogram. The CNN model is trained on the dataset using the mel spectrograms as input features and the composer labels as output classes. The model achieves 81.4% accuracy in classifying the music by composer, demonstrating the effectiveness of the proposed approach. To generate new music similar to the original pieces, a CNN autoencoder is trained on a similar dataset. The model is trained to encode the mel spectrograms of the original pieces into a lower-dimensional latent space and then decode them back into the original mel spectrogram. The generated music is produced by sampling from the latent space and decoding the samples back into mel spectrograms, which are then transformed into audio. In conclusion, the proposed system provides a promising approach to classifying and generating classical Egyptian music, which can be applied in various musical applications, such as music recommendation systems, music production, and music education.

### 15.15: Regular Session Sat-S12

#### **76: BlockCampus: A Blockchain-Based DApp for Enhancing Student Engagement and Reward Mechanisms in an Academic Community for E-JUST University**

*Mariam A. M. Mahmoud (Egypt-Japan University of Science and Technology)\*; Walid Gomaa (Egypt-Japan University of Science & Technology); Youssef Elharty (Egypt-Japan University Of Science & Technology); Ahmed Fathi Heshmat (EJUST); Ahmed Hagag Abdallah (Egypt-Japan University of Science and Technology); Ahmed H Rashed (Egypt-Japan University of Science and Technology); Omar Kilej (Egypt-Japan University for Science and Technology)*

In today's digital age, online communities have become an integral part of our lives, fostering collaboration, knowledge sharing, and community engagement. Higher education institutions, in particular, can greatly benefit from dedicated platforms that facilitate academic discussions and provide incentives for active participation. This research paper presents a comprehensive study and implementation of a decentralized application (DApp) leveraging the blockchain technology to address these needs specifically for E-JUST (Egypt-Japan University of Science and Technology) students and academic staff.

### **34: Exploration of Hyperledger Besu in Designing Private Blockchain-based Financial Distribution Systems**

*Md. Raisul Hasan Shahrukh (University of Liberal Arts Bangladesh)\*; Md. Tabassinur Rahman (University of Liberal Arts Bangladesh); Nafees Mansoor (ULAB)*

Blockchain, a decentralized technology that provides unrivaled security, transparency, and process validation, is redefining the operational landscape across numerous industries. This article focuses on the development of an innovative consortium blockchain-based financial distribution application. This paper illuminates the transformative role of blockchain technology in a variety of sectors by drawing on a plethora of academic literature and current industry practices. It demonstrates the diverse applications of blockchain, ranging from remittances, lending, and investments in finance to data administration in healthcare and supply chain tracking. The paper reveals the design and potential of a consortium blockchain-based application for financial distribution. Utilizing the capabilities of Hyperledger Besu, the application is tailored to improve security, scalability, and interoperability, thereby contributing to a more integrated financial ecosystem. The investigation sheds light on the combination of consortium blockchain's controlled access and Besu's comprehensive functionality, proposing a secure, transparent, and efficient financial transactions environment. The investigation serves as a resource for academics, industry professionals, and policymakers alike, highlighting the vast potential of blockchain technology, enabled by platforms such as Hyperledger Besu, in accelerating the evolution of traditional systems toward a more decentralized, secure, and efficient future.

### **4: An Ensemble Learning Classifier to Predict Net Electricity Generation from Nuclear Power Plants**

*Mushfiqur Rashid Khan (Bangladesh University of Engineering and Technology)\*; Faiyaz Fahim (Bangladesh Army University of Engineering and Technology); Nahid Hasan (BUET); Md. Parveq Plaban (Bangladesh University of Engineering and Technology)*

A nuclear power plant (NPP) is an important source that can generate low-emission electricity. It is important to make appropriate predictions about electricity generation to operate an NPP very effectively, which is an essential statistic for assessing the efficiency of an NPP and its performance in terms of its actual impact on the generation of energy. The net electricity generation from NPPs can change on a monthly basis. In this research, a time-series dataset containing the net electricity generation in Thousand Megawatt Hours (TMWHs) from January 2001 to March 2023 is utilized

for the prediction of net electricity generation from NPPs in the United States. The dataset, containing 267 data points, ranges from 54547 TMWH to 74649 TMWH. Data preprocessing is performed, and some ensemble learning classifiers such as Random Forest, Voting Ensemble, Bagging, and AdaBoost are applied for the experiment. The Random Forest classifier achieves the highest performance based on its 99.82% accuracy, 0.0018 root mean square error (RMSE), 0.0633 mean absolute percentage error (MAPE), and a minimum training time of 42.4 seconds. This research can aid in preventing system loss and ensuring electricity generation.

## **22: Towards Sustainable AI: Investigating Energy Efficiency and Emissions in Deep Learning Optimizers**

*Md Asif Mahmud Tusher Siddique (Leeds Beckett University); MD. Sakibul Islam (Islamic University of Technology)\*; Ah-Lian Kor (Leeds Beckett University); Rashedul Kabir (University of Asia Pacific); Nusrath Jahan Happy (University of Asia Pacific)*

Optimizers play a pivotal role in the efficacy of deep neural network models, directing the rate of training and ultimate prediction performance. Due to the computational costs and energy consumption, these models face sustainability challenges. Despite the growing literature on the comparative accuracy of various deep learning optimizers, the lack of empirical research and theoretical guidance regarding their environmental footprint and energy consumption often results in their oversight during the model selection process. So it is crucial to identify the optimal optimizer that maximizes accuracy while also prioritizing sustainability. This research, therefore, conducts rigorous comparative and statistical scrutiny of the energy consumption and associated environmental impacts of nine contemporary deep learning optimizers, aiming to discern the most ecologically friendly alternative. The experimental results indicate that Stochastic Gradient Descent (SGD) markedly surpasses other optimizers in terms of energy and carbon efficiency. Additionally, commonly used optimizers such as Adam, Adagrad, and AdaMax emerge as viable choices in the construction of deep learning architectures. The research work will provide invaluable insights for future endeavors in the conceptualization of energy-efficient deep learning implementations for sustainable AI development and reducing the ecological impact of advanced machine learning models.

**16.20: Closing and Award Ceremony**

**17.00: Tutorial (Online)**

**On-Sensor Machine Learning with ST Toolchains**

**Dr. Mahesh Chowdhary**

Senior Director of MEMS software solutions, STMicroelectronics based in Santa Clara, CA, USA

**Dr. Swapnil Sayan Saha**

Algorithm Development Engineer of MEMS software solutions, STMicroelectronics based in Santa Clara, CA, USA



October 28, 5.00 pm BST (9.00 am PDT)

**Tutorial: On-Sensor Machine Learning with ST Toolchains**

**Mahesh Chowdhary  
Swapnil Sayan Saha**

Register: Scan QR code or visit  
<https://forms.office.com/e/ykwnGHmAs>  
Tutorial will be online (virtual)



**2023 11<sup>th</sup> International  
Conference on Informatics,  
Electronics and Vision**

**2023 6th International  
Conference on Imaging,  
Vision & Pattern Recognition**

October 26 – 29, London, United Kingdom  
<http://cennser.org/ICIEV>  
<http://cennser.org/IVPR>



## Call for Paper:

6<sup>th</sup> International Conference on Activity and Behavior Computing (ABC)

May 29-31, 2024, Japan (Hybrid)



<https://autocare.ai/abc2024>

**Contact:** Send an empty email to be updated: [abc2024@sozolab.jp](mailto:abc2024@sozolab.jp)

ICIEV & IVPR

### General Chair:

Sozo Inoue, Kyushu Institute of Technology / AUTOCARE LLC, Japan  
Md Atiqur Rahman Ahad, University of East London, UK

### Advisory Board:

Anind Dey, University of Washington, USA  
Björn W. Schuller, Imperial College London, UK and University of Augsburg, Germany  
Jeffrey Cohn, University of Pittsburgh, USA  
Kristof Van Laerhoven, University of Siegen, Germany  
Michael Beigl, Karlsruhe Institute of Technology, Germany  
Paul Lukowicz, RPTU / DFKI, Germany  
Thomas Plötz, Georgia Institute of Technology, USA

## Call for Paper:

12<sup>th</sup> International Conference on Informatics, Electronics & Vision (ICIEV)

22-25 Aug., 2024

Kitakyushu, Japan

<http://cennser.org/ICIEV>

ICIEV & IVPR

## Call for Paper:

7<sup>th</sup> International Conference on Imaging, Vision & Pattern Recognition (IVPR)

22-25 Aug., 2024

Kitakyushu, Japan

<http://cennser.org/IVPR>

*Together - we can always do better.  
See you all again at future events.*

## **Journal: Call for Paper:**

International Journal of Computer Vision & Signal Processing  
(ISSN: 2186-1390)

- Open Access but **NO publication fee** - it is **FREE to publish/access**, hence, submit your good papers.
- Details the development of various aspects of vision, machine learning, image & signal processing researches.
- Presents key engineering and industrial developments on various computer vision & signal processing challenges.
- Offers concrete peer-reviewed publications of good works.
- Welcomes original works, comprehensive survey, and review works, which cover new insights based on in-depth analysis of the state-of-the-art.
- Allows publication and download without any charge.

Share & download the FREE published papers.

Visit <http://cennser.org/IJCVSP>

**Thanks a lot indeed!**

**Together - we can always do better.**

**Let's support & encourage each other for a better future!**