AVINASH KUMAR SINGH

Ph.D (Human-Robot Interaction), Post-doc (Understandable Robots)

@ avinashkumarsingh1986@gmail.com +33-758387037 Montpellier, France
www.avinashkumarsingh.co.in github.com/dravinash

EXPERIENCE

Postdoctoral Researcher

Montpellier University, France

Nov 2020 – Ongoing Montpellier, France

I am associated with the robotics lab (LIRMM) and working on the EU Project SOPHIA. As a part of the work package, I am working on human robot collaboration and my responsibilities are.

- Activity Recognition: Given the RGB-D and sensor data (kinematic and kinetic) we need to recognize the activity such as opening the door, handovering the object, grasping etc.
- Physical Interaction: Help human in physical assistance, e.g holding and passing the object in industrial environment by understanding the human gesture, facial expression etc.

Senior Solution Leader

Brane Enterprises LLP

May 2020 – Oct 2020 Hyderabad

As a lead of AI team, I was responsible for.

- To manage the team, work allocation, setting up the project deliverable, planning and execution.
- To design the architecture of deep learning and computer vision solutions in the domain object detection, scene and face recognition, image captioning etc.

Lead Consultant

Intain Technologies Limited

Jan, 2018 – May, 2020 remote

As a lead consultant my responsibilities are

- To lead the AI team, allocating and tracking work, and ensuring client deliveries.
- To create the pipeline of the machine learning solutions and to add new people into the team.

Postdoctoral Researcher

Umea University, Sweden

Feb, 2018 – Jan, 2020 Umea, Sweden

During the postdoc, I worked on humanoid robots. We train them to assist humans in the household scenarios.

- We designed a dialogue based human robot interaction system that allows humans, to talk to the robot.
- A dialogue generation system is added to translate robot's action into the natural language so that they would be understandable.

EDUCATION

Ph.D. in Human-Robot Interaction

Indian Institute of Information Technology, Allahabad, India

Nov2011 – Dec2016 CGPI: 9.0

Thesis title: On effective human-robot interaction based on recognition and association.

M.Tech. in Computer Science

KIIT University, Bhubaneswar, India

Jul2009 – Jun2011 CGPI: 8.91

Thesis title: Detection and prevention of SQL injection attacks in web applications.

M.Sc. in Information Technology

Kumaun University, Nanital, India

Jul2007 – Jun2009 Per: 76%

Thesis title: Library book management system.

B.Sc. in Mathematics

Kumaun University, Nanital, India

Jul2004 – Jun2007 Per: 65%

COURSE WORK

Soft Computing | Image Processing | Robotics & Industrial Automation | Computer Vision | Linear Algebra | Research Methodology | Numerical Methods | Pattern Recognition

SKILLS

- **Machine Learning** ●●●●○
- **Computer Vision** ●●●●○
- **Deep Leaning** ●●●○○
- **Natural Language Processing** ●●●○○

Deputy Manager

HCL Technologies

Feb, 2017 – Jan, 2018 Noida, India

I worked with HCL machine learning division. As a deputy manager my responsibilities were.

- To hire more people into the team. Work allocation, monitoring client deliveries, and identify the gaps in the current process flow to propose machine learning solutions.
- Designing the architecture of machine learning solutions.

Associate Process Manager

eClerx Services Limited

Nov, 2015 – Feb, 2017 Mumbai, India

A worked as a full stack developer and my job was to design and develop NLP, ML solutions. We used image pre-processing to improve OCR accuracy, further to integrate these solutions with RPA system.

RESEARCH PROJECTS

A Fuzzy Inference System for a Visually Grounded Robot State of Mind

Umea University, Sweden

In order for robots to interact with humans on real-world scenarios or objects, these robots need to construct a representation ('state of mind') of these scenarios that a) are grounded in the robots' perception and b) ideally should match human understanding and concepts. Using table-top settings as scenario, we propose a framework that generates a robot's ''state of mind' by extracting the objects on the table along with their properties (color, shape and texture) and spatial relations to each other. The scene as perceived by the robot is represented in a dynamic graph in which object attributes are encoded as fuzzy linguistic variables that match human spatial concepts. This project details the construction of such graph representations by combining low-level neural network based feature recognition and a high-level fuzzy inference system.

Techniques: Fuzzy Inference System, Mask-RCNN, CNN, Local Binary pattern (LBP), Multi-Layer Perceptron
Platform: Python, Choregraphe, TensorFlow.

Understandable Collaborating Robot Teams

Umea University, Sweden

We investigate understandability for teams of robots collaborating to solve a common task. We present a novel and flexible solution based on Cooperating Distributed Grammar Systems and a multi-agent algorithm for coordination of actions. The solution is implemented and evaluated on three Pepper robots collaborating to solve a task while commenting on their own and other robots' current and planned actions.

Techniques: Cooperating Distributed Grammar System, Multi-Agent System, Natural Language Generation.
Platform: Python, Choregraphe.

Criminal Identification through Human-Robot Interaction

IIIT, Allahabad, India

MOST PROUD OF

EU project H2020 SOPHIA (2020-2024)
Selected for the Post-Doc position on Deep Learning Perception for Human-Robot Collaboration at LIRMM Montpellier, France.

Postdoc funding from Kemphe Foundations, Sweden
Received 2 years of research funding to pursue my postdoctoral work in Human-Robot Interaction at Umea University, Sweden.

MHRD, India PH.D Fellowship
Selected for the Ph.D under the robotics and AI lab, IIITA

Secured 1st position in MSc
For the consecutive 2 years, I stood first in MSc at the University Level.

PROGRAMMING SKILLS

Python	●●●●●
MATLAB	●●●○○
JAVA	●●○○○
C++	●●○○○

INVITED TALKS

An Empirical Review of Calibration Techniques for the Pepper Humanoid Robot's RGB and Depth Camera

Intelligent Systems and Applications.

5th Sept 2019 London, England

Occasion: Presented conference paper in IntelliSys 2019.

Fusion of Gesture and Speech for Increased Accuracy in Human Robot Interaction

25th International Conference on Methods and Models in Automation and Robotics

24th Aug 2019 Międzyzdroje, Poland

Occasion: Presented conference paper in MMAR 2019.

Conflict Detection and Resolution in Table Top Scenarios for Human-Robot Interaction

As a supplement to mugshot detection, a new approach is proposed to capture the eyewitness's visual perception in the form of symbolic representation. It reveals physiological and facial characteristics of criminal which help in their identification. A rough set theory based technique is introduced to model those symbolic representations. This approach provides an intuitive insight to process criminal's imprecise and imperfect knowledge. We used a benchmark mug-shot dataset consisting of 300 criminals faces from the Chinese University of Hong Kong (CUHK) to study the correctness of our proposed model.

Techniques: Rough Set Theory, Nao Speech Engine.
Platform: Matlab, Choregraphe, NAOqi.

INDUSTRIAL PROJECTS

Vision for Visually Impaired

Brane Enterprises LLP

We designed a spectacle with the camera mounted in it to assist the visually impaired people. The PCB board is fabricated in house to facilitate the functionalities such as Bluetooth connection, USB charging and live feed to Android device. An app is created to offer functionality such as object detection, scene understanding, navigation, person count, etc. I was helping my team to build the deep learning solutions and their integration to the app. Together, with the team we have successfully implemented various models of object detection, image/dense captioning, currency recognition, OCR etc.

Techniques: Faster-RCNN, MobileNetV2, CNN+LSTM, YOLO
Platform: Python, TensorFlow, Keras.

Incremental Face Learning

Brane Enterprises LLP

We designed an android app to perform face recognition. We applied transfer learning where FaceNet model is used as a backbone network and then stacked a classification layer to classify the faces. Neural network suffers with the problem of catastrophic forgetting due to that you need to train the whole network (except the backbone) on the N+1 classes (here N is the number of classes, system is already trained with). In order to get rid off from this problem we applied an incremental leaning approach where we applied the distillation loss function to solve this problem.

Techniques: Transfer Leanring, FaceNet, Incremental Learning
Platform: Python, TensorFlow, Keras.

OCR in WILD

Intain Technologies

We used tesseract as a base system and then retrain it to further recognize new fonts. On other hand we used, CTPN, EAST models for text detection and then further training the CNN+LSTM model for OCR. In the another approach we trained text detection models for detecting the text and then used tesseract for OCR. In order to improve the OCR accuracy we have further used the parallel image enhancement function. Tesseract produces the word and its confidence as the output. The end word is selected based on the maximum likelihood of the word.

Computing science department, Umea University

18th June 2019 Umea, Sweden

Occasion: Poster presentation in 31st Swedish AI Society Workshop.

Deep learning and its applications

UFBI department, Umea University

15th June 2018 Umea, Sweden

Occasion: Delivered talk in Umea center for Functional Brain Imaging (UFBI) day.

Robotic Manipulator and sketch drawing

WASP-Wallenberg Foundation

11th April 2018 Stockholm, Sweden

Occasion: Presented live demo in AI4X- Collecting Ideas and Identifying Challenges for Future AI Research in Sweden.

Object Localization and Recognition

Department of Information Technology, NIT Raipur

30th Dec 2017 Raipur, India

Occasion: Delivered talk in Computational Intelligence & Cloud Computing (CICC-2017) workshop.

Artificial Intelligence and the Future of Robotics

Department of Computer Science, KIIT Bhubaneswar

11th March 2017 Bhubaneswar, India

Occasion: Delivered talk in ACM student chapter inauguration.

Sketch drawing by NAO humanoid robot

TENCON a premier international technical conference of IEEE Region 10

1st Nov 2015 Macau, China

Occasion: Presented conference paper in TENCON 2015.

Robotic Vision: Detection and Tracking

Department of Computer Science & Engineering, MNNIT Allahabad

5th July 2015 Allahabad, India

Occasion: Delivered talk in summer school on Emerging Trends in Computer (ETCS-2015).

Techniques: Tesseract, CTPN, EAST, Image Enhancement, CNN, LSTM.
Platform: Python, Keras, Opencv.

Video KYC

Intain Technologies

In the video KYC we need to process the different identity documents such as PAN, AADHAAR, Driving Licence, Voter ID and Passport and then extract the information from the given source. We used YOLO to first detect the identity document given the image and then used OCR to get the text output. Further the rule based engine is used to extract the information. On the other hand, we also trained the Mask-RCNN system to localize different fields such as in the PAN localize the name, father name, dob and pan number.

Techniques: YOLO, Mask-RCNN, NLTK.
Platform: Python, NLTK, Keras.

Invoice Processing

HCL Technologies

Given the scanned invoices coming from different vendors, we need to extract information such as invoice number, invoice date, vendor name, vendor address, total amount, etc. We used Tesseract (OCR) to first digitize the invoices and then applied machine learning to extract these information. We transformed every word into the vector space and then classified it to the respective class. Here we treated every entity as class. Since OCR involved here, we may get the OCR error in the extracted values, to correct the errors we used the domain ontology.

Techniques: Feature engineering, Multi-class classification, error correction.
Platform: Python, Scikit, Numpy, NLTK, Spacy.

SELECTED PUBLICATIONS

Journal Articles

- Singh, Avinash Kumar, Neha Baranwal, Kai-Florian Richter, et al. (2020). "Verbal explanations by collaborating robot teams". In: *Paladyn, Journal of Behavioral Robotics* 12.1, pp. 47–57.
- Singh, Avinash Kumar, Neha Baranwal, and Gora Chand Nandi (2019). "A rough set based reasoning approach for criminal identification". In: *International Journal of Machine Learning and Cybernetics* 10.3, pp. 413–431.
- Baranwal, Neha, Gora Chand Nandi, and Avinash Kumar Singh (2017). "Real-Time Gesture–Based Communication Using Possibility Theory–Based Hidden Markov Model". In: *Computational Intelligence* 33.4, pp. 843–862.
- Baranwal, Neha, Avinash Kumar Singh, and Gora Chand Nandi (2017). "Development of a Framework for Human–Robot interactions with Indian Sign Language Using Possibility Theory". In: *International Journal of Social Robotics* 9.4, pp. 563–574.
- Singh, Avinash Kumar, Neha Baranwal, and Gora Chand Nandi (2017). "Development of a self reliant humanoid robot for sketch drawing". In: *Multimedia Tools and Applications* 76.18, pp. 18847–18870.
- Singh, Avinash Kumar and Gora Chand Nandi (2017). "Visual perception-based criminal identification: a query-based approach". In: *Journal of Experimental & theoretical artificial intelligence* 29.1, pp. 175–196.

Spoofing Attacks & System Security

Arcade Business College, Patna, Bihar

31st Jan 2015 Patna, India

Occasion: Delivered talk in one day workshop on cyber security.

CERTIFICATIONS

Distributed Algorithms, Systems, and Programming

Indian Institute of Science (IISc) Bangalore, India

28th May – 8 June 2012

Occasion: Microsoft Summer School organized by Microsoft Research India.

Computing Techniques and Applications

Banaras Hindu University, Varanasi, India

1 July – 7 July 2012

Occasion: National workshop cum training program organized by Mathematics Department.

Machine Intelligence and Soft Computing

Indian Statistical Institute, Kolkata, India

21 Sept – 21 Dec 2012

Occasion: Three Months certification offered by Center for Soft Computing Research (CSCR).

Machine Learning

Indian Institute of Technology, Kanpur, India

1 July – 3 July 2013

Occasion: 1st Indian Workshop organized by computer science department.

Image Processing, Computer Vision and Pattern Recognition

National Institute of Technology Delhi, India

18th June – 22 June 2013

Occasion: Faculty Development Program organized by computer science department.

PDE Modeling and Computation

Indian Institute of Technology Madras, India

21 Oct – 25 Oct 2013

Occasion: DAAD Supported International Workshop organized by Department of Mathematics.

- Singh, Avinash Kumar and Gora Chand Nandi (2016). "NAO humanoid robot: Analysis of calibration techniques for robot sketch drawing". In: *Robotics and Autonomous Systems* 79, pp. 108–121.
- Singh, Avinash Kumar, Piyush Joshi, and Gora Chand Nandi (2014). "Face liveness detection through face structure analysis". In: *International Journal of Applied Pattern Recognition* 1.4, pp. 338–360.

Conference Proceedings

- Singh, Avinash Kumar, Neha Baranwal, and Kai-Florian Richter (2020). "A Fuzzy Inference System for a Visually Grounded Robot State of Mind". In: *Proceedings of the Europian conference on artifical intelligence, (ECAI 2020), Santiago de Compostela, Spain, 29 Aug-2 Sept 2020*. Spain: ECAI2020.
- Baranwal, Neha, Avinash Kumar Singh, and Suna Bench (2019). "Extracting primary objects and spatial relations from sentences". In: *11th International Conference on Agents and Artificial Intelligence, Prague, Czech Republic, 19-21 February 2019*.
- Singh, Avinash Kumar, Neha Baranwal, and Kai-Florian Richter (2019). "An Empirical Review of Calibration Techniques for the Pepper Humanoid Robot's RGB and Depth Camera". In: *Proceedings of SAI Intelligent Systems Conference*. Springer, pp. 1026–1038.
- Singh, Avinash Kumar, Pavan Chakraborty, and GC Nandi (2015). "Sketch drawing by NAO humanoid robot". In: *TENCON 2015-2015 IEEE Region 10 Conference*. IEEE, pp. 1–6.
- Singh, Avinash Kumar, Arun Kumar, et al. (2014). "Expression invariant fragmented face recognition". In: *2014 International Conference on Signal Propagation and Computer Technology (ICSPCT 2014)*. IEEE, pp. 184–189.
- Singh, Avinash Kumar and Gora Chand Nandi (2012). "Face recognition using facial symmetry". In: *Proceedings of the Second International Conference on Computational Science, Engineering and Information Technology*, pp. 550–554.
- Singh, Avinash Kumar and Sangita Roy (2012). "A network based vulnerability scanner for detecting sqli attacks in web applications". In: *2012 1st International Conference on Recent Advances in Information Technology (RAIT)*. IEEE, pp. 585–590.

ACADEMIC SERVICES

- Served as a chair of IEEE in IIIT- Allahabad student chapter from June 2014 to May 2015.
- Official Mentor of IEEE in IIIT-Allahabad student chapter from June 2013 to May 2014.
- Organizing committee member of "SERB sponsored 1st summer school on robotics", organized by IIIT Allahabad.
- IEEE Professional Student Member.

LANGUAGES

Hindi ●●●●●

English ●●●●●

Swedish ●○○○○

REFERENCES

Prof. G C Nandi
@ Indian Institute of Information Technology, Allahabad, India
gcnandi@iiita.ac.in

Prof. Kai-Florian Richter
@ Umeå universitet
kai-florian.richter@umu.se

Prof. Thomas Hellström
@ Umeå universitet
thomas.hellstrom@umu.se

Declaration
I hereby declare that the above mentioned particulars are true to the best of my knowledge and belief.

Date:

Place: Avinash Kumar Singh