Soumya Shamarao Jahagirdar

Website: soumyasj.github.io

Contact No.: +91-9591163792
Email ID: soumyasj22@gmail.com
LinkedIn: Soumya Jahagirdar

RESEARCH INTERESTS
Computer Vision, Deep Learning, Multi-modal learning: Vision + Language, Natural Language Processing.

EDUCATION
2021 - Present	International Institute of Information Technology, Hyderabad, (IIIT-H).
	M.S. by Research, Computer Science and Engineering, CGPA: 8.4/10 (3rd sem)
2017 - 2021	KLE Technological University, Hubli
	BE, Computer Science and Engineering, CGPA: 9.01/10
2015 - 2017	Mahesh PU College, Belagavi.
	Percentage: 91%

EXPERIENCE
- **Computer Vision Center, CVC, UAB, Barcelona, Spain. (September 2022 - November 2022)**
  Completed an internship at CVC under the guidance of Prof. Dimosthenis Karatzas and Prof. Ernest Valveny. Worked on Multimodal learning, specifically video question answering on the videos containing text.
- **Computer Vision and Graphics Laboratory, Hubballi, India. (December 2019 - July 2021)**
  Was an active member of CVG Research Lab working towards solving problems in Computer Vision and 3D processing.
- **Think & Ink Education and Research Foundation, Bangalore, India. (February 2020 - July 2020)**
  Worked as Research Assistant under the guidance of Prof. B. A. Patil, Founder and CEO of Think & Ink Education and Research Foundation, on magnitude spectrum and descriptive statistics-based soil classification. Also developed GUI for the same application for the farmers to use.
- **Indian Institute of Delhi (IIT-D), Delhi, India. (June 2019 - July 2019)**
  Worked with Dr. Prem Kumar Kalra and his Ph.D. students on the project "Activity recognition for Suturing by Neuro Surgeon" which was a collaborative project with AIIMS, Delhi.

PROJECT DETAILS
- **Text-based Video Retrieval with Multilingual Capabilities (January 2023 - Present)**
  Name of the Institution: IIIT-Hyderabad
  Summary: Set up a server and curated a large-scale dataset of videos from domains such as educational videos and news videos. Built a simple yet effective TF-IDF based search engine using OCR tokens present in the videos. It also incorporates queries in Multiple Indian languages. The dataset collection also contains videos in Indian languages such as Telugu, Kannada, Malayalam, Hindi, etc.
- **Video Question Answering on videos containing text (August 2021 - Present)**
  Name of the Institution: IIIT-Hyderabad
  Summary: Presently working on multimodal learning: video question answeirng on the videos containing text. As textual content plays an important role in scene understanding, we put forward for the first time a new task of text-based video question answering, where given a video and a question, the deep learning models have to read the text present in the videos to answer the question.

- **Text Based Visual Question Generation (September 2020 - July 2021)**
  Name of the Institution: KLE Technological University
  Summary: Worked on a novel task of Visual Question Generation leveraging textual content in the image (TextVQG). This task finds its application in self-supervised based data generation for Visual Question Answering (VQA) and better self-learning systems in general.
- **Depth Densification using RGB and Sparse Sensing (February 2020 - February 2021)**
  Name of the Institution: Samsung Research India Bangalore (SRIB) + KLE Technological University
  Summary: Proposed novel DRS (Dense-Residual-Skip) Autoencoder for the task of depth densification from RGB image and sparse sensing. Introduced novel Gradient-Aware-Mean-Square-Error (GAMSE) Loss function for better depth predictions. Also developed android ARCore application to obtain Sparse depth data.
- **Large Scale Densification of Sparse Point Cloud towards 3D reconstruction (August 2019-December 2019)**
  Name of the Institution: KLE Technological University
  Summary: Proposed enhanced version of Patch-Based-Multi-View-Stereo (PMVS) for large scale depth densification of Sparse point cloud.

PUBLICATIONS
- **Soumya Jahagirdar**, Minesh Mathew, Dimosthenis Karatzas, C.V. Jawahar, **"Watching the News: Towards VideoQA Models that can Read",** WACV 2023.
- Shamanthak Hegde, **Soumya Jahagirdar**, Shankar Gangisetty**, "Making the V in Text-VQA Matter",** ODRUM-W, CVPR, 2023.
- Charani Alampalle, Shamanthak Hegde, **Soumya Jahagirdar**, Shankar Gangisetty**, "Weakly Supervised Visual Question Answer Generation",** ODRUM-W, CVPR, 2023.
- **Soumya Jahagirdar**, Shankar Gangisetty, Anand Mishra, **"Look, Read and Ask: Learning to Ask Questions by Reading Text in Images",** ICDAR (Oral) 2021.
- Girish Hedge, Tushar Pharale, **Soumya Jahagirdar,** Vaishakh Nargund, Ramesh Tabib, Uma Mudenagudi, Basavaraja Vandrotti, Ankit Dhiman, **"DeepDNet: Deep Dense Network for Depth Completion Task"**, WiCV-W, CVPR 2021.

PATENTS
- US PATENT: Uma Mudenagudi, Girish Hegde, Ramesh Tabib, **Soumya Jahagirdar**, Tushar Pharale, Basavaraj Vandrotti, Ankit Dhiman, Vaishakh Nargund, "**Method and Device of Depth Densification using RGB Image and Sparse Depth".** Patent number: WO2022103171A1

TRAININGS AND CERTIFICATIONS
- Participated in 5th Summer School of AI, IIIT-Hyderabad. (2021)
- Undergone a short course on 3D Computer Vision in IIIT Hyderabad (2020).
- CVG Winter Workshop on Image Processing, Machine Learning, Deep Learning and Neural Networks (2019).
- IIT Techfest BOMBAY-Workshop on Deep Learning (2018).
- Completed certified course on Deep Learning with Computer Vision by National Research University Higher School of Economics offered through Coursera.
- Completed certified course on Improving deep neural networks: Hyperparameter tuning, regularization and optimization by deeplearning.ai offered through Coursera.

TECHINCAL SKILLS
- Deep learning: Pytorch, Tensorflow, NTLK and MMF.
- Programming/Scripting: Python, C++, C, Java, R, Linux, HTML, CSS

ACHIEVEMENTS/CO-CIRRICULAR ACTIVITES

- Organised and conducted tutorial sessions in 6th Summer School of AI, IIIT-H. (2022)
- Volunteer of ICFHR-2022 Conference.
- Finalist of PRISMGlanz, SRIB, Bangalore, India. (2020)
- Brand Ambassador of Smart Knower. (2020)
- Secured 3rd position in Buildathon, product building event at KLETech. (2019)
- Event organizer: TECHNOPHIILIA (Coding Event Organizing committee-2019).
- Conducted winter workshop on Machine Learning and Computer Vision. (2019)
- Department Co-curricular committee member (2018).
- Regional Level Essay Writer. (2016)
- Presented Research Study on Poultry Farming in National Children Science Congress (NCSC) KV IISc, Bangalore. (2015)
- Regional level Debater (KV Air-Force Station Yelahanka, Bangalore- 2015).
- Participated in Regional German Camp by Goethe Institute in Bangalore. (2015)

PERSONAL DETAILS

Date of Birth : February 22, 2000.
Address   : 217-Parijat main block, IIIT-Hyderabad, Gachibowli, Hyderabad.
Languages Known : English, Hindi, Kannada, Sanskrit and Spanish.
Hobbies   : Singing, Reading, Sketching, Writing and Cooking.
Place   : Hyderbad, Telangana, India.