

# Sai Kumar Dwivedi

Email: [saidwivedi@gmail.com](mailto:saidwivedi@gmail.com)

Webpage: [saidwivedi.in](http://saidwivedi.in) • G-Scholar: [Profile](#) • Github: [saidwivedi](#) • LinkedIn: [saidwivedi](#)

---

## Education

Max Planck Institute for Intelligent Systems

PhD in Computer Science

Supervisors: Dr. Michael J. Black, Dr. Dimitris Tzionas

Senior Collaborator: Dr. Cordelia Schmid

PhD Topic: Perceiving humans in in-the-wild scenarios

OCT 2021 - PRESENT

Tübingen, Germany

National Institute of Technology, Rourkela

Bachelors & Masters in Computer Science and Engineering

JUN 2011 - JUN 2016

Rourkela, India

---

## Experience

Meta, Reality Labs | *Research Scientist Intern*

Controllable Generation of Human-Object Interaction in 4D

JUL 2025 - DEC 2025

Zurich, Switzerland

Mercedes-Benz R&D | *Computer Vision Researcher*

2D Human Pose Estimation for Smart Car Interiors

Developed a novel real-time headpose estimation algorithm from a monocular IR camera which was deployed as Side Mirror Selection and Rear Sunblind Control in Mercedes EQS and S-Class cars.

SEP 2017 - FEB 2020

Bangalore, India

Intel Corporation | *Machine Learning Engineer*

Deep Learning Algorithms for Edge Devices

Developed Binary and Ternary CNNs similar to XNOR-Net, addressing performance issues in deeper networks by implementing innovative quantization techniques, optimizing weight representation for efficiency, particularly beneficial for edge device applications.

JAN 2016 - AUG 2017

Bangalore, India

---

## Selected Publications

- **InteractVLM: 3D Interaction Reasoning from 2D Foundational Models**  
[S.K. Dwivedi](#), D. Antić, S. Tripathi, O. Taheri, C. Schmid, M.J. Black, D. Tzionas  
CVPR 2025 — [Project Page](#)  
Introduced a novel framework that leverages large Visual-Language Models to estimate contact points on humans and objects from a single in-the-wild image, and using these contacts for joint reconstruction.
- **TokenHMR: Advancing Human Mesh Recovery with a Tokenized Pose Representation**  
[S.K. Dwivedi](#)<sup>\*</sup>, Y. Sun<sup>\*</sup>, P. Patel, Y. Feng, M.J. Black  
CVPR 2024 — [Project Page](#)  
Reformulated 3D human pose estimation as token prediction and introduced a novel Threshold-Adaptive Loss Scaling to address the paradox of declining 3D accuracy with increasing 2D precision.
- **POCO: 3D Pose and Shape Estimation using Confidence**  
[S.K. Dwivedi](#), C. Schmid, H. Yi, M.J. Black, D. Tzionas  
3DV 2024 (Oral) — [Project Page](#)  
Proposed a confidence-aware framework that extends human pose and shape regressors to estimate their own prediction uncertainty, enhancing reliability in 3D reconstructions.
- **Learning to Regress Bodies from Images using Differentiable Semantic Rendering**  
[S.K. Dwivedi](#), N. Athanasiou, M. Kocabas, M.J. Black  
ICCV 2021 — [Project Page](#)  
Introduced Differentiable Semantic Rendering to incorporate semantic clothing information into 3D human body estimation, improving accuracy in pose and shape predictions.

## Other Publications

- **SDFit: 3D Object Pose and Shape by Fitting a Morphable SDF to a Single Image**  
D. Antić, G. Paschalidis, S. Tripathi, T. Gevers, [S.K. Dwivedi](#), D. Tzionas  
ICCV 2025 — [Project Page](#)
- **PICO: Reconstructing 3D People In Contact with Objects**  
A. Cseke, S. Tripathi, [S.K. Dwivedi](#), A. Lakshminpathy, A. Chatterjee, M.J. Black, D. Tzionas  
CVPR 2025 — [Project Page](#)
- **ChatPose: Chatting about 3D Human Pose**  
Y. Feng, J. Lin, [S.K. Dwivedi](#), Y. Sun, P. Patel, M.J. Black  
CVPR 2024 — [Project Page](#)
- **Detecting Human-Object Contact in Images**  
Y. Chen, [S.K. Dwivedi](#), M.J. Black, D. Tzionas  
CVPR 2023 — [Project Page](#)
- **ProtoGAN: Towards Few Shot Learning for Action Recognition**  
[S.K. Dwivedi](#), V. Gupta, R. Mitra, S. Ahmed, A. Jain  
ICCV Workshop 2019 — [Project Page](#)
- **Out-of-Distribution Detection for Generalized Zero-Shot Action Recognition**  
D. Mandal, S. Narayan, [S.K. Dwivedi](#), V. Gupta, S. Ahmed, F. Khan, L. Shao  
CVPR 2019 — [Project Page](#)
- **Progression Modelling for Online and Early Gesture Detection**  
V. Gupta, [S.K. Dwivedi](#), S. Ahmed, A. Jain  
3DV 2019 (Oral) — [Project Page](#)

---

## Academic Internship

Max Planck Institute for Intelligent Systems   <i>Research Assistant</i> 3D Human Pose Estimation using Differentiable Semantic Rendering	JUL 2020 - AUG 2021 <i>Remote, India</i>
Tallinn University of Technology   <i>Research Exchange Student</i> Designed Neural Network based face-recognition system on FPGA	SEP 2014 - JUN 2015 <i>Tallinn, Estonia</i>
Indian Institute of Technology, Gandhinagar   <i>Research Intern</i> Developed Virtual Reality Modules for Autistic Children	APR 2014 - JUN 2014 <i>Ganadhinagar, India</i>

---

## Students Supervised

Suraj Bhor - Max Planck Institute for Intelligent Systems <i>Current Position:</i> Research Engineer, MPI-IS	MAR 2024 - MAR 2025
Sheetal Sahoo - HTW Berlin - University of Applied Sciences <i>Current Position:</i> Data Scientist, DLR Institut für Datenwissenschaften	JUN 2023 - JAN 2024

---

## Awards and Honours

IMPRS-IS Scholar - 2022	Max Planck Institute for Intelligent Systems, Germany
Research Assistant Fellowship - 2021	Max Planck Institute for Intelligent Systems, Germany
Outstanding Performer Award - 2018	Mercedes-Benz R&D, India
Employee Recognition Award - 2017	Intel Corporation, India
Student Fellowship - 2015	Erasmus Mundus European Union Program, Estonia
Student Fellowship - 2014	IIT Gandhinagar, India

---

## Academic Service

(*Outstanding Reviewer: †*)

CVPR Reviewer	2022, 2023, 2025 <sup>†</sup>
ECCV/ICCV Reviewer	2022, 2023, 2024 <sup>†</sup> , 2025
3DV Reviewer	2021, 2022, 2023