KSHAMA NITIN SHAH

Interested in developing self-supervised computer vision models that learn from multimodal sensation specifically natural language and cross modal image data.

kshama2705@gmail.com | 734-510-0744 | kshama2705.github.io/ | kshama-nitin-shah | kshama2705

OBJECTIVE

Seeking Full-time opportunities in the field of Machine Learning, Deep Learning or Computer Vision

RESEARCH AND WORK EXPERIENCE

Research Assistant

Dr. Andrew Owens' Lab, University of Michigan, Ann Arbor

May '22 – May '23 | Ann Arbor, MI

- Developed a novel self-supervised multimodal image/video registration technique for medical imaging applications by training a ResNet-50 based optical flow estimation network

Research Associate

Dr. Justin Johnsons' Lab, University of Michigan, Ann Arbor

Aug '22 – May '23 | Ann Arbor, MI

- Developed a simplified training recipe for open vocabulary instance segmentation using text supervision using a CLIP-ViT encoder and SAM Mask Decoder
- Implemented and benchmarked various CNN and Transformer-based object detection models (FCOS, Mask R-CNN, ViTDeT, DETIC, Mem ViT) for comparison
- Responsible for coding, implementation and driving the research project using distributed code, multi GPU training using Linux and SLURM.

Teaching Assistant, EECS 442/ EECS 504 - Computer Vision

University of Michigan, Ann Arbor

Fall 2022 & Winter 2023 | Ann Arbor, MI

- Responsible for teaching tutorial sessions, creating assignments and holding office hours for students

AI Algorithm Engineer

Stoneridge Inc

Aug '23 – Present | Novi, MI

- Developing algorithms for camera monitor systems specifically for perception and feature extraction using both neural networks and classical computer vision

SELECTED PROJECTS

Self Supervised Object Detection With Multimodal Image Captioning **(GitHub)**

Feb '22 – Apr '22 | University of Michigan, Ann Arbor

- Developed a novel self-supervised pipeline that uses natural language supervision as a pre-training task to localize objects in an image by generating pseudo ground truth object classes and bounding box coordinates.
- Achieved a comparable mAP of 21.57% by fine-tuning the model using only 1% of the labeled dataset, while requiring 1.5x lesser training time and compute resources compared to other state-of-the-art semi-supervised models.

Visual Question Answering using customized prompts **(GitHub)**

Aug 2022 - Dec 2022 | University of Michigan, Ann Arbor

- Developed a novel pipeline to perform zero-shot Visual Question Answering by conjoining large pre-trained models such as CLIP and T-5 transformer.
- Achieved an overall accuracy of 49.5%, which is comparable to the state-of-the-art performance in zero-shot VQA, while using 10x lesser memory and computational resources.

A Monocular Local Mapper for Urban Scenes **(GitHub)**

Aug 2021 - Dec 2021 | University of Michigan, Ann Arbor

EDUCATION

M.S. in Electrical & Computer Eng. **(Signal Processing & Machine Learning)**

University of Michigan, Ann Arbor

Aug '21 – April '23 | GPA : 3.924/4.00

B.Eng. in Electronics & Communication Engineering

Birla Institute of Technology & Science. Pilani

Aug '16 – June '20 | GPA : 9.61/10.00

COURSEWORK

Computer vision, Machine learning, Deep learning for computer vision, Natural Language Processing, Matrix Methods for signal processing, machine learning and data analysis, Probability and Random processes

SKILLS

Python | Pytorch | NumPy | TensorFlow | MATLAB | Julia | C

LEADERSHIP EXPERIENCE

Youth Entrepreneurship Program, AIESEC

Jul '17 – Aug '17

- Contributed to the United Nations' Sustainable Development Goal of Decent Work and Economic Growth by boosting sales of local SMEs and advising student entrepreneurs via AIESEC's Global Volunteer Exchange Program in Indonesia

Vice President, University Relations AIESEC in Dubai

Jan '18 – Jun '18

- Organized informational events across universities to increase awareness about AIESEC's exchange programs
- Facilitated the successful completion of several international exchange experiences for students across Dubai

Core Committee Member, IEEE, BITS Pilani, Dubai Charter

Aug 2017 – Aug 2019

- Organized and managed tech competitions and guest lectures by distinguished speakers in the university

- Developed a model that performs semantic segmentation, object detection and depth estimation simultaneously using YOLOv1 and U-Net model.
- Obtained an overall accuracy of 83% by utilizing a single model for all three tasks, reducing the number of parameters required by 2x.

Language Supervised Vision Pre-training for Fine-grained Food Classification

Mar '22 - Apr '22 | University of Michigan, Ann Arbor

- Pre-trained a downsized, memory-efficient image captioning model that used a RegNetX-800MF and a 2 layer transformer on the Food-101 dataset.
- Obtained a top 5% classification accuracy of 23.76% by doing zero-shot transfer and a classification accuracy of 20% after fine-tuning the above model on the downstream task of fine-grained food classification