

KOUTILYA PNVR

Ph.D. Graduate | Computer Vision | University of Maryland College Park

🌐 <https://koutilya-pnvr.github.io>

in [linkedin.com/in/pnvrk](https://www.linkedin.com/in/pnvrk) 🐙 github.com/koutilya-pnvr

📞 +1 202 460 6957 @ koutilya@terpmail.umd.edu

📍 4320 Rowalt Dr, Apt 202, College Park, MD-20740



I am a Ph.D. graduate from the University of Maryland, College Park, advised by **Prof. David Jacobs**. I am broadly interested in computer vision. My research topics are around : **Leveraging deep generative models such as - GANs for geometry estimation via unsupervised domain adaptation; latent diffusion models (LDMs) for text-based segmentation of real and AI generated images; LDMs for editing tasks - Virtual Try-On, Fashion clothes inpainting, Human model synthesis and Image personalization.** Some more areas of computer vision I am interested in are : Inverse Rendering; Video classification; and Inpainting. I am highly interested in collaborating with researchers/engineers from the industry to work on fascinating ideas in computer vision.

PUBLICATIONS

- Koutilya PNVR, Bharat Singh, Pallabi Ghosh, Behjat Siddiquie, David Jacobs. “**LD-ZNet : A Latent Diffusion Approach for Text-Based Image Segmentation.**” In *Proceedings of the International Conference on Computer Vision (ICCV), 2023.* - **ORAL**
- Koutilya PNVR, Hao Zhou, and David Jacobs. “**SharinGAN : Combining Synthetic and Real data for Unsupervised Geometry Estimation.**” In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020.*
- Hanson Alex*, Koutilya PNVR*, Sanjukta Krishnagopal, and Larry Davis. “**Bidirectional Convolutional LSTM for the Detection of Violence in Videos.**” *The European Conference on Computer Vision (ECCV) Workshops, 2018.*

INTERNSHIP EXPERIENCE

Dec 2022 | Amazon GO, Seattle, WA

May 2022

- Explored the utility of various text-to-image generative models for text-based segmentation of images.
- Demonstrated the use of latent diffusion models (LDMs) pretrained on large-scale internet data for text-based segmentation.
- Proposed novel ways to utilize features from the internal stages of the LDM to improve the segmentation performance by nearly **6% on real images and 20% on AI-generated images.**

Text-Based Image Segmentation Real Images AI generated Images Diffusion models Stable Diffusion

Dec 2020 | Project NEON, STAR LABS, Campbell, CA

May 2020

- Worked with team of researchers on various audio-visual and self-supervised learning techniques.
- Prototyped novel learning algorithms in large scale production system for various audio and video synthesis approaches.
- Integrated solutions in cross language technology stack consisting of Python, C++ and CUDA.

Audio-visual Computer Vision Self-supervision for Audio-Video modalities Deep Learning Pytorch

RESEARCH EXPERIENCE

Present | Ph.D. Candidate | Computer Vision, UNIVERSITY OF MARYLAND, College Park

Aug. 2017

- Text-based segmentation of real and AI generated images by leveraging text-to-image latent diffusion generative model pretrained on large-scale internet data.
- Self-training methods such as knowledge-distillation targeted for monocular depth estimation.
- Domain Adaptation between synthetic and real datasets for applications such as Monocular Depth Estimation of outdoor scenes and Face Normal Estimation.
- Violence detection in videos using a Bidirectional ConvLSTM network.
- Guided Inpainting using Generative Adversarial Networks that can enable the use of different car images as guides to edit cars in street view scenes.

Computer Vision Generative models - GANs and Diffusion models Text-Based Image Segmentation Domain Adaptation Depth and Normal Estimation Knowledge-Distillation Inverse Rendering Video classification Guided Inpainting

Aug. 2023 | Generative AI Research Scientist, Vchar.ai

Aug. 2022

- Explored the utility of generative models such as diffusion models for editing applications - Virtual Try-On, Fashion clothes inpainting, Human model synthesis, Image personalization

Generative models - GANs and Diffusion models Virtual Try-On Clothes Inpainting Model Synthesis Personalization

EDUCATION

Current	3.93	Ph.D. Computer Vision, University of Maryland, College Park
2017	3.9	M.S. Electrical and Computer Engineering, University of Maryland, College Park
2015	8.22	B.Tech. Electrical Engineering, Indian Institute of Technology, Delhi, India

SKILLS

Deep Learning frameworks Pytorch, TensorFlow, TFLearn, Keras
Other technical skills Tensorboard, OpenMPI, OpenMP, LaTeX, Linux, Visual Studio Code, Git

TECHNICAL PROJECTS

DOMAIN ADAPTATION BETWEEN SYNTHETIC AND REAL DATASETS FOR GEOMETRY ESTIMATION

MAY 2019 - DEC 2019

- Implemented a novel way of utilizing GANs to train deep networks that can minimize the domain gap between synthetic and real images for performance improvement in geometry estimation tasks.
- Demonstrated the performance gain for Monocular Depth Estimation and Face Normal Estimation tasks.
- Reduced the absolute error in testsets of KITTI by 23.77% and Make3D by 6.45% over the state-of-the-art.
- Improved the face normal prediction by 4.3% for $Acc < 20$ deg metric for the Photoface dataset over the state-of-the-art.

Domain Adaptation Monocular Depth Estimation Face Normal Estimation GANs Synthetic and Real datasets Pytorch Tensorboard

VIOLENCE DETECTION IN VIDEOS USING BIDIRECTIONAL CONVLSTM

APRIL 2018 - JUNE 2018

- Developed a novel Bidirectional Convolutional LSTM network followed by an elementwise max-pooling layer to obtain better Spatio-temporal representations for detecting violence in videos.
- Demonstrated the superiority of our method on the Hockey fights, Movies and Violent-Flows datasets over previous state-of-the-art methods.
- Signified the importance of all our modules: BiConvLSTM, elementwise maxpool, temporal encoding via ablation studies.

Violence Detection Bidirectional ConvLSTM Elementwise Max Pool Spatio-temporal encoding Pytorch

PARALLEL IMPLEMENTATION OF SMO AND MODEL SELECTION ALGORITHMS FOR SVMs

MAR 2017 - MAY 2017

- Efficiently parallelized the Sequential Minimal Optimization (SMO) and Model Selection (MS) algorithms that define an optimal SVM classifier using OpenMPI, OpenMP and Hybrid frameworks in C++ on deepthought2 cluster.
- Studied the parallel version on several cluster configurations such as nodes and threads/core and over several datasets, such as MNIST, not-MNIST, Cordna, A9a, Splice, etc.
- Model selection is performed on 900 models and the best one is chosen that gives the least validation error.
- A huge speedup of 173 is observed for the Hybrid implementation of MS-SMO on not-MNIST for 16 node configuration.

OpenMP OpenMPI Sequential Minimal Optimization (SMO) Support Vector Machines (SVM) High Performance Computing

REFERENCES

David Jacobs

PROFESSOR

DEPARTMENT OF CS AND UMIACS

University of Maryland College Park

@ djacobs@umiacs.umd.edu

☎ +1 301 405 0679

Bharat Singh

PH.D., COMPUTER VISION

SR STAFF APPLIED RESEARCH SCIENTIST

Cruise

@ bharatsingh430@gmail.com

🌐 LinkedIn Profile

Varaprasad Bandaru

ASSOCIATE RESEARCH PROFESSOR

DEPARTMENT OF GEOGRAPHICAL SCIENCES

University of Maryland College Park

@ vbandaru@umd.edu

☎ +1 301 405 3074