Vineet Gandhi

Associate Professor, Center for Visual Information Technology (CVIT)
Kohli Center for Intelligent Systems (KCIS)
International Institute of Information Technology, Hyderabad
(+91) 9052244425, vgandhi@iiit.ac.in, https://faculty.iiit.ac.in/ vgandhi/

RESEARCH INTERESTS

I am interested in applied machine learning for computer vision and multimedia applications. In recent times, the primary threads of my research are around creative AI, multimodal Machine Learning, and improving the generalization abilities of the deep learning models.

EDUCATION

Institut national de recherche en sciences et technologies du numérique (INRIA), Grenoble, Rhône-Alpes, France   2011 - 2014
PhD, Applied Mathematics and Computer Science
Advisor: Remi Ronfard
Committee: James L. Crowley, Patrick Perez, Frederic Jurie, Alexander Sorkine-Hornung, Michael Gleicher

European Master CIMET   2009 - 2011
University Jean Monnet (France), University of Granada (Spain), Norwegian University of Science and Technology (Gjøvik, Norway), INRIA (France)
M.Sc., Optique, Vision and Computer Science (GPA 3.6/4, Rank 1/22)

Indian Institute of Information Technology, Design and Manufacturing, Jabalpur, India   2005 - 2009
B.Tech., Electronics and Communication Engineering (CGPA 8.2/10)

EXPERIENCE

IIIT Hyderabad, India   April 2015 - present
Assistant Professor

Animaker Inc., California   January 2020 - present
Chief Scientific Advisor. Envisioning and leading their machine learning efforts. Animaker is one of largest DIY animation company with over 15 million users. I am helping the company transform from a pure DIY company to a ML focused company with user in the loop.

TCS Research, India   July 2022 - July 2023
Adjunct Faculty

INRIA, Grenoble France   Jan 2011 - July 2011
Research intern. Worked at the Perception team, under the supervision of Radu Patrice Horaud and Jan Cech.

Xerox Research Center Europe, Grenoble France   June 2010 - Aug 2010
Research intern. Worked on large scale image classification, under the supervision of Florent Perronnin and Mathieu Chuat.

Emantras Media Lab, Gurgaon   Apr 2008 - July 2008
Worked at the animation and game design group under the supervision of Aakash

Kumar (now head of design at makemytrip)

HONORS and AWARDS	Best student paper award at ICMI 2022
	Outstanding reviewer award: ICML 2022
	Qualcomm Innovation Fellowship Winner 2020 and Super Winner 2021.
	Early Career Research Award, Science and Engineering Board (SERB), India.
	Best paper award at Digital Heritage conference 2015, Spain
	CIBLE scholarship by Region Rhone Alpes, 2011-2014
	Awarded Erasmus Mundus M.Sc. scholarship (42000 Euros) 2009-2011.

PATENTS

[1] Vineet Gandhi, Moneish Kumar, Remi Ronfard, Michael L Gleicher. **System and method for automatically generating split screen for a video of a dynamic scene**. *US Patent 10,084,970*

[2] Vineet Gandhi, S R Rajendran. **Methods and systems of automatically generating video content from scripts/text**. *US Patent 11,158,349*

[3] Vineet Gandhi, Remi Ronfard, Michael L Gleicher. **Device for Generating Cinematographic rushes by video analysis**. *French and European Patent FR3028701B1, Worldwide Application*

[4] Vineet Gandhi, S R Rajendran. **Methods and systems of automatic one click virtual button with ai assist for diy animation**. *US Patent 11481238*

[5] Vineet Gandhi, Swetanjal Murali Dutta. **System and method for optical tracking of one or more players in a sports event**. *Indian Patent App. 202041040579*

PUBLICATIONS Google Scholar: http://scholar.google.fr/citations?user=PVlBz8oAAAAJhl

[1] Darshana Saravanan, Naresh Manwani, Vineet Gandhi **Pseudo-labelling meets Label Smoothing for Noisy Partial Label Learning**, *Workshop on Fine-Grained Visual Categorization, CVPR 2025*

[2] Darshana Saravanan, Makarand Tapaswi, Vineet Gandhi **Investigating Mechanisms for In-Context Vision Language Binding**, *Mechanistic Interpretability for Vision, CVPR workshop 2025*

[3] Aishwarya Agarwal, Srikrishna Karanam and Vineet Gandhi **TIDE: Training Locally Interpretable Domain Generalization Models Enables Test-time Correction**, *CVPR 2025*

[4] Darshana Saravanan, Varun Gupta, Darshan Singh, Zeeshan Khan, Vineet Gandhi, and Makarand Tapaswi **VELOCITI: Benchmarking Video-Language Compositional Reasoning with Strict Entailment**, *CVPR 2025*

[5] Rohit Girmaji, Bhav Beri, Ramanthan Subramanian and Vineet Gandhi. **EditIQ: Automated Cinematic Editing of Static Wide-Angle Videos via Dialogue Interpretation and Saliency Cues**, *IUI 2025*

[6] Kawshik Manikantan, Makarand Tapaswi, Vineet Gandhi, Shubham Toshniwal. **IdentifyMe: A Challenging Long-Context Mention Resolution Benchmark**, *NAACL 2025 (short)*

[7] Neil Shah, Ayan Kashyap, Shirish Karande and Vineet Gandhi. **MRI2Speech: Speech Synthesis from Articulatory Movements Recorded by Real-time MRI**, *ICASSP 2025*

[8] Ayan Kashyap, Neil Shah and Vineet Gandhi. **Prompt-to-Correct: Automated Test-Time Pronunciation Correction with Voice Prompts**, *ICASSP 2025*

[9] Neil Shah, Shirish Karande and Vineet Gandhi. **Advancing NAM-to-Speech Conversion with Novel Methods and the MultiNAM Dataset**, *ICASSP 2025*

[10] Rohit Girmaji, Siddharth Jain, Bhav Beri, Sarthak Bansal and Vineet Gandhi. **Minimalistic Video Saliency Prediction via Efficient Decoder Spatio Temporal Action Cues**, *ICASSP 2025*

[11] Kawshik Manikantan, Shubham Toshniwal, Makarand Tapaswi, Vineet Gandhi. **Major Entity Identification: A Generalizable Alternative to Coreference Resolution**, *EMNLP 2024*

[12] Neil Shah, Neha Sahipjohn, Vishal Tambrahalli, Ramanathan Subramanian, Vineet Gandhi. **StethoSpeech: Speech Generation Through a Clinical Stethoscope Attached to the Skin**, *Ubicomp* 2024

[13] Kanishk Jain, Shyamgopal Karthik and Vineet Gandhi. **Test-Time Amendment with a Coarse Classifier for Fine-Grained Classification**, *NeurIPS* 2023

[14] Saiteja Kosgi, Neil Kumar Shah, Vishal Tambrahalli, Neha Sherin, Vineet Gandhi. **ParrotTTS: Text-to-Speech synthesis by exploiting self-supervised representations**, *EACL (findings)* 2023

[15] Neha Sahipjohn, Neil Shah, Vishal Tambrahalli and Vineet Gandhi. **RobustL2S: Speaker-Specific Lip-to-Speech Synthesis exploiting Self-Supervised Representations**, *APSIPA* 2023

[16] Ritu Srivastava, Saiteja Kosgi, Sarath Sivaprasad, Neha Sahipjohn and Vineet Gandhi. **Adversarial Robustness of Mel Based Speaker Recognition Systems**, *APSIPA* 2023

[17] Laksh Nanwani, Anmol Agarwal, Kanishk Jain, Raghav Prabhakar, Aaron Monis, Aditya Mathur, Krishna Murthy, Abdul Hafez, Vineet Gandhi, K Madhava Krishna. **Instance-Level Semantic Maps for Vision Language Navigation**. *ROMAN*, 2023

[18] Kanishk Jain, Varun Chhangani, Amogh Tiwari, K. Madhava Krishna and Vineet Gandhi. **Ground then Navigate: Language-guided Navigation in Dynamic Scenes**. *ICRA*, 2023

[19] Sarath Sivaprasad, Akshay Goindani, Mario Fritz, Vineet Gandhi. **Classwise Domain Generalization: A Novel Framework for Evaluating OOD Performance**. *Neurips Workshop on Distribution Shifts*, 2022

[20] Saransh Dave, Ritam Basu, Vineet Gandhi. **Cross-Domain Class-Contrastive Learning: Finding Lower Dimensional Representations for Improved Domain Generalization**. *ICVGIP*, 2022

[21] Shreyank Jyoti, Ritvik Agrawal, Rohit Girmaji, Sarath Sivaprasad, Vineet Gandhi. **Salient Face Prediction without Bells and Whistles**. *DICTA*, 2022

[22] Ritvik Agrawal, Shreyank Jyoti, Rohit Girmaji, Sarath Sivaprasad, Vineet Gandhi. **Does Audio help in deep Audio-Visual Saliency prediction models?**. *ICMI*, 2022 (Best Student Paper Award)

[23] Saiteja Kosgi, Sarath Sivaprasad, Niranajan Padnekar, Anil Nelakanti, Vineet Gandhi. **Empathic Machines: Using Intermediate Features as Levers to Emulate Emotions in Text-To-Speech Systems**. *NAACL*, 2022

[24] Kanishk Jain, Vineet Gandhi. **Comprehensive Multi-Modal Interactions for Referring Image Segmentation**. *Findings of ACL*, 2022

[25] Pratikkumar Bulani, Jayachandran S, Sarath Sivaprasad, Vineet Gandhi. **Framework to Computationally Analyze Kathakali Videos**. *Workshop on Intelligent Cinematography and Editing (WICED)*, 2022

[26] Sarath Sivaprasad, Akshay Goindani, Vaibhav Garg, Vineet Gandhi. **Reappraising Domain Generalization in Neural Networks**, *arXiv preprint* 2021

[27] Jeet Vora, Swetanjal Dutta, Shyamgopal Karthik, Vineet Gandhi. **Bringing Generalization to Deep Multi-view Detection**, *arXiv preprint* 2021

[28] Nivedita Rufus, Kanishk Jain, Unni Krishnan R Nair, Vineet Gandhi, K Madhava Krishna. **Grounding Linguistic Commands to Navigable Regions**. *IROS* 2021

[29] Samyak Jain, Pradeep Yarlagadda, Shreyank Jyoti, Shyamgopal Karthik, Ramanathan Subramanian, Vineet Gandhi. **ViNet: Pushing the limits of Visual Modality for Audio-Visual Saliency Prediction**. *IROS* 2021

[30] Sarath Sivaprasad, Saiteja Kosgi, Vineet Gandhi. **Emotional Prosody Control for Speech Generation**. *INTERSPEECH* 2021

[31] Sarath Sivaprasad, Ankur Singh, Naresh Manwani, Vineet Gandhi. **The Curious Case of Convex Neural Networks**. *ECML-PKDD* 2021

[32] Shyamgopal Karthik, Ameya Prabhu, Puneet Dokania, Vineet Gandhi. **No Cost Likelihood Manipulation at Test Time for Making Better Mistakes in Deep Networks**. *International Conference on Learning Representations (ICLR)* 2021

[33] Nivedita Rufus, Unni Krishnan Nair, K Madhava Krishna, Vineet Gandhi. **Cosine meets softmax: A tough-to-beat baseline for visual grounding**. *ECCV Workshops*, 2020

[34] Shyamgopal Karthik, Ameya Prabhu, Vineet Gandhi. **Simple Unsupervised Multi-Object Tracking**. *arXiv preprint* 2020

[35] Aasheesh Singh, Aditya Kamireddypalli, Vineet Gandhi, K Madhava Krishna. **LiDAR guided Small obstacle Segmentation**. *IROS* 2020

[36] Navyasri Reddy, Samyak Jain, Pradeep Yarlagadda, Vineet Gandhi. **Tidying Deep Saliency Prediction Architectures**. *IROS* 2020

[37] K L Bhanu Moorthy, Moneish Kumar, Ramanathan Subramanian, Vineet Gandhi. **GAZED – Gaze-guided Cinematic Editing of Wide-Angle Monocular Video Recordings** *ACM CHI (Conference on Human Factors in Computing Systems)* 2020

[38] Shyamgopal Karthik, Abhinav Moudgil, Vineet Gandhi. **Exploring 3 R's of Long-term Tracking: Re-detection, Recovery and Reliability**. *WACV* 2020

[39] Murtuza Bohra, Vineet Gandhi. **ColorArt: Suggesting Colorizations For Graphic Arts Using Optimal Color-Graph Matching**. *Graphics Interface* 2020

[40] Murtuza Bohra, Sajal Maheshwari, Vineet Gandhi. **TextureToMTF: predicting spatial frequency response in the wild**. *Signal, Image and Video Processing* 2020

[41] Sudheer Achary, Syed Ashar Javed, Nikita Shravan, K L Bhanu Moorthy, Vineet Gandhi, Anoop Namboodiri. **CineFilter: Unsupervised Filtering for Real Time Autonomous Camera Systems**. *Workshop on Intelligent Cinematography and Editing (WICED)* 2019

[42] N. N. Sriram, Maniar Tirth, Kalyanasundaram Jayaganesh, Gandhi Vineet, Bhowmick Brojeshwar, Krishna Madhava K. **Talk to the Vehicle: Language Conditioned Autonomous Navigation of Self Driving Cars**. *IROS* 2019

[43] Javed Syed Ashar, Saxena Shreyas, Gandhi Vineet. **Learning Unsupervised Visual Grounding Through Semantic Self-Supervision**. *IJCAI* 2019

[44] Gupta Aryaman, Thakkar Kalpit, Gandhi Vineet and Narayanan P J. **Nose, Eyes and Ears: Head Pose Estimation By Locating Facial Keypoints**. *ICASSP* 2019

[45] Moudgil Abhinav and Gandhi Vineet. **Long-Term Visual Object Tracking Benchmark**. *ACCV* 2018

[46] Gupta Krishnam, Javed Syed Ashar, Gandhi Vineet and Krishna Madhava K. **MergeNet: A Deep Net Architecture for Small Obstacle Discovery**. *ICRA* 2018

[47] Kumar Kranthi, Kumar Moneish, Gandhi Vineet, Subramanian Ramanathan. **Watch to Edit: Video Retargeting using Gaze**. *EUROGRAPHICS* 2018

[48] Rai Pranjal Kumar, Maheshwari Sajal, and Gandhi Vineet. **Document Quality Estimation using Spatial Frequency Response**. *ICASSP* 2018

[49] Shah Vatsal, and Gandhi Vineet. **An Iterative approach for Shadow Removal in Document Images**. *ICASSP* 2018

[50] Sharma Rahul Anand, Bhat Bharath, Gandhi Vineet, Jawahar CV. **Automated Top View Registration of Broadcast Football Videos**. *WACV* 2018

[51] Rahul Anand Sharma, Vineet Gandhi, Visesh Chari, C V Jawahar. **Automatic analysis of broadcast football videos using contextual priors**. *Signal, Image and Video Processing*, 2017

[52] Rai Pranjal Kumar, Maheshwari Sajal, Mehta Ishit, Sakurikar Parikshit and Gandhi Vineet. **Beyond OCRs for Document Blur Estimation**. *ICDAR* 2017

[53] Sheethal Reddy, Vineet Gandhi, K Madhava Krishna. **3D Region Proposals For Selective Object Search.**. *VISIGRAPP*, 2017

[54] Kumar Moneish, Gandhi Vineet, Ronfard Remi and Gleicher Michael. **Zooming On All Actors: Automatic Focus+Context Split Screen Video Generation**. *EUROGRAPHICS* 2017

[55] K Gupta, S Upadhyay, V Gandhi, KM Krishna. **Small obstacle detection using stereo vision for autonomous ground vehicle**. *Proceedings of the Advances in Robotics*, 2017

[56] Maheshwari Sajal, Rai Pranjal, Sharma Gopal and Gandhi Vineet. **Document Blur Detection using Edge Profile Mining**. *ICVGIP* 2016

[57] Gandhi Vineet and Ronfard Remi. **A Computational Framework for Vertical Video Editing**. *Eurographics workshop on intelligent cinematography and editing (WICED)* 2016

[58] Rémi Ronfard, Benoit Encelle, Nicolas Sauret, P-A Champin, Thomas Steiner, Vineet Gandhi, Cyrille Migniot, Florent Thiery. **Capturing and indexing rehearsals: the design and usage of a digital archive of performing arts**. *Digital Heritage* 2016 (Best Paper Award)

[59] Gandhi Vineet, Ronfard Remi, Gleicher Michael. **Multi-Clip Video Editing from a Single Viewpoint**. *European Conference on Visual Media Production (CVMP)* 2014

[60] Gandhi Vineet, Ronfard Remi. **Detecting and naming actors in movies using generative appearance models**. *Computer Vision and Pattern Recognition (CVPR)* 2013

[61] Ronfard Remi, Gandhi Vineet and Boiron Laurent. **The Prose Storyboard Language: A Tool for Annotating and Directing Movies**. *Workshop on Intelligent Cinematography and Editing (WICED)* 2013

[62] Gandhi Vineet, Cech Jan and Horaud Radu. **High-Resolution Depth Maps Based on TOF-Stereo Fusion**. *ICRA* 2012

PROFESSIONAL SERVICES

PC Member/Conference Reviewing

ICML	2022,2023
ICLR	2021-2024
NeurIPS	2020-2025
SIGGRAPH, SIGGRAPH ASIA	2017,2019,2025
CVPR, ECCV, ICCV	2017-2024
ACM Multimedia	2018-2022
ACM Multimedia (area chair)	2024-2025
AAAI	2018-2022
IJCAI	2019,2020
WACV	2018-2024
BMVC, ACCV	2018-2023
ICRA, IROS	2019-2024

Journal Reviewing

IEEE Transactions on Pattern Analysis and Machine Intelligence
Transactions on Graphics
Computer Graphics Forum
IEEE Transactions on Image Processing
IEEE Transactions on Multimedia

Others

Since 2014, I have played an active role in the Workshop on Intelligent Cinematography and Editing (WICED) organization. The workshop was initially held with AAAI, and for the last few years, it has happened with Eurographics. I was the Program Chair of WICED 2020, jointly with Tsai-Yen-Li.

TEACHING

Statistical Methods in Artificial Intelligence	Monsoon 2017, Spring 2020-2025
Digital Image Processing	Monsoon 2015,2016
Digital Signal Analysis and Applications	Spring 2016-2018
Computer Programming	Monsoon 2018,2019
Advanced Data Structure and Algorithms for Problem Solving	Monsoon 2020-2023
Human Values (part of collective faculty group)	2015-2020
PG Certification in AI and Machine Learning	Since 2018
(a program by IIIT outreach for Industry Professionals)	

Languages   English(fluent), Hindi(native), French(functional), Spanish(beginner)