# Vicente Ordóñez-Román (publishes as Vicente Ordonez)

## Contact Information

| Detailed Information                                      |  |
|-----------------------------------------------------------|---|
| Duncan Hall 2080                                          | vicenteor@rice.edu |
| 6100 Main St                                              | https://www.cs.rice.edu/~vo9 |
| Houston, TX 77005-1827                                     | https://vislang.ai |

## Research Interests

My research lies at the intersection of computer vision, natural language processing, and machine learning. I am interested in developing machine learning models that can understand the real world through multiple modalities and can learn naturally from human guidance. I am generally interested in building efficient visual recognition models that can perform high-level perceptual tasks and doing so in a way that is fair, transparent, and interpretable.

## Current Positions

- **Associate Professor (with tenure)** 2021 - Present
  - Rice University, Houston, Texas.
  - Department of Computer Science, George R. Brown School of Engineering.
  - Leading the Vision, Language and Learning Lab – [https://www.vislang.ai](https://www.vislang.ai).
- **Amazon Visiting Academic** 2024 - Present
  - Amazon AGI Foundations
  - Working in developing the next generation of intelligent multimodal models.

## Honors and Awards

- Google Award for Inclusion Research, 2022.
- U.S. National Science Foundation CAREER Award, 2021
- Research Highlight of the Communications of the ACM, December 2020 Issue.
- Facebook Research Award, 2020.
- Google Faculty Research Award, 2017.
- IBM Faculty Award, 2017.
- Best Long Paper Award, 2017.
  - Intl. Conf. on Empirical Methods in Natural Language Processing (EMNLP), 2017.
- Research Highlight of the Communications of the ACM, March 2016 Issue.
- Allen Institute for Artificial Intelligence Hackathon 2015 – *Peer Favorite Award*.
- Best Paper Award – IEEE Marr Prize 2013.
  - International Conference on Computer Vision (ICCV), 2013.
- Yahoo! Key Scientific Challenges Award, 2012.
- Renaissance Technologies Fellowship, 2009 - 2011.
- Philanthropic Society Medal, Guayaquil, Ecuador, 2007.

## Education

| University/Institution                                      |  Years |
|-----------------------------------------------------------|--------|
| The University of North Carolina at Chapel Hill (UNC)     | 2013 - 2015 |
| Doctor of Philosophy in Computer Science                  |        |
| Thesis: Language and Perceptual Categorization in Computational Visual Recognition Advisor: Tamara L. Berg |        |
| Stony Brook University, The State University of New York (SUNY) | 2009 - 2013 |
| Master of Science in Computer Science                     |        |
| Escuela Superior Politécnica del Litoral (ESPOL), Ecuador. | 2003 - 2008 |
| Computer Engineering Degree (GPA: 9.22/10.0)              |        |
**Academic Experience**

University of Virginia (UVA), Charlottesville, Virginia. 2016 - 2021
Assistant Professor, Department of Computer Science
School of Engineering and Applied Science.
Led for five years the Vision, Language and Learning Lab (vislang.ai)

The University of North Carolina at Chapel Hill 2013 - 2015
Research Assistant, Department of Computer Science

Stony Brook University (SUNY), Stony Brook, New York. 2010 - 2013
Research Assistant, Vision and Digital Media Lab

Center for Information Technologies (ESPOL), Guayaquil, Ecuador. 2006 - 2009
Research Assistant, Technology Enhanced Learning Group

**Other Experience**

Amazon Alexa AI, San Jose, California (remote) July 2021 - December 2023
Amazon Visiting Academic
Contributing to the Create with Alexa skill in the Amazon Alexa AI division.
https://www.aboutamazon.com/news/devices/what-is-create-with-alexa

Yarn Labs Inc, Cambridge, Massachusetts September 2019 - March 2020
Independent Contractor
Research on the state of the field and regulation of facial recognition technologies, with funding from the MacArthur Foundation.

Adobe Research, College Park, Maryland. Summer 2019
Visiting Professor, Document Intelligence Lab (DIL)
Pursuing projects at the intersection of Computer Vision and NLP for documents.

Allen Institute for Artificial Intelligence (AI2), Seattle, Washington. 2015 - 2016
Visiting Research Fellow, Computer Vision Group (now PRIOR)
Worked at the intersection of Vision and Language in the Computer Vision group.

Microsoft Research, Cambridge, Massachusetts. Summer 2014
Research Intern, Computer Vision Group
Large scale data-driven scene parsing using deep learning features.
Mentors: Ce Liu and Michael Rubinstein.

eBay Research Labs, San Jose, California. Summer 2013
Research Intern, Computer Vision Group
Analysis of furniture images as part of the Computer Vision group.
Mentors: Robinson Piramuthu and Vignesh Jagadeesh.

Google, Mountain View, California. Summer 2011
Software Engineering Intern, Android Multimedia Content Analysis Group
Automatic organization of personal image collections using visual features.
Mentors: Rodrigo Carceroni and Wei Hua.

Google, Mountain View, California. Spring 2008, Summer 2008
Software Engineering Intern, Google Earth
Automated quality analysis of satellite images for Google Earth and Google Maps.
Mentor: Rodrigo Carceroni.
Facial Recognition Technologies in the Wild: A Call for a Federal Office.
Erik Learned-Miller, Vicente Ordóñez, Jamie Morgernstern, Joy Buolamwini.
May 2020. https://www.ajlunited.org/federal-office-call

Facial Recognition Technologies: A Primer.
Joy Buolamwini, Vicente Ordóñez, Jamie Morgernstern, Erik Learned-Miller.
May 2020. https://www.ajlunited.org/federal-office-call

PropTest: Automatic Property Testing for Improved Visual Programming. Jaywon Koo, Ziyan Yang, Paola Cascante-Bonilla, Baishakhi Ray, Vicente Ordonez. arXiv:2403.16921. March 2024. https://arxiv.org/abs/2403.16921

Learning from Models and Data for Visual Grounding. Ruozhen He, Paola Cascante-Bonilla, Ziyan Yang, Alexander C. Berg, Vicente Ordonez. arXiv:2403.13804. March 2024. https://arxiv.org/abs/2403.13804

Grounding Language Models for Visual Entity Recognition. Zilin Xiao, Ming Gong, Paola Cascante-Bonilla, Xingyao Zhang, Jie Wu, Vicente Ordonez. arXiv:2402.18695. February 2024. https://arxiv.org/abs/2402.18695

Characterizing Video Question Answering with Sparsified Inputs. Shiyuan Huang, Robin-son Piramuthu, Vicente Ordonez, Shih-Fu Chang, Gunnar A. Sigurdsson. arXiv:2311.16311. November 2023. https://arxiv.org/abs/2311.16311

ViC-MAE: Self-Supervised Representation Learning from Images and Video with Contrastive Masked Autoencoders. Jefferson Hernandez, Ruben Villegas, Vicente Ordonez. arXiv:2303.12001. March 2023. https://arxiv.org/abs/2303.12001

On the Transferability of Visual Features in Generalized Zero-Shot Learning. Paola Cascante-Bonilla, Leonid Karlinsky, James Seale Smith, Yanjun Qi, Vicente Ordonez arXiv:2211.12494. November 2022. https://arxiv.org/abs/2211.12494

Chair Segments: A Compact Benchmark for the Study of Object Segmentation. Leticia Pinto-Alva, Ian K. Torres, Rosangel Garcia, Ziyan Yang, Vicente Ordonez arxiv:2012.01250. December 2020. https://arxiv.org/abs/2012.01250

Moviescope: Large-scale Analysis of Movies using Multiple Modalities. Paola Cascante-Bonilla, Kalpathy Sitaraman, Mengjia Luo, Vicente Ordonez arXiv:1908.03180. August 2019. https://arxiv.org/abs/1908.03180

Improved Visual Grounding through Self-Consistent Explanations. Ruozhen He, Paola Cascante-Bonilla, Ziyan Yang, Alexander C. Berg, Vicente Ordonez. Conf. on Computer Vision and Pattern Recognition. CVPR 2024. June 2024. Seattle, WA.

ElasticDiffusion: Training-free Arbitrary Size Image Generation. Moayed Haji-Ali, Guha Balakrishnan, Vicente Ordonez. Conf. on Computer Vision and Pattern Recognition. CVPR 2024. June 2024. Seattle, WA.

SCoRD: Subject-Conditional Relation Detection with Text-Augmented Data. Ziyang Xiang, Kushal Kafle, Zhe Lin, Scott Cohen, Zhihong Ding, Vicente Ordonez. Winter Conference on Applications of Computer Vision. WACV 2024. January 2024. Waikoloa, HI
Variation of Gender Biases in Visual Recognition Models Before and After Finetuning.
Jaspreet Ranjit, Tianlu Wang, Baishakhi Ray, Vicente Ordonez.
Workshop on Algorithmic Fairness through the Lens of Time at NeurIPS 2023.
December 2023. New Orleans, LA.

Going Beyond Nouns With Vision & Language Models Using Synthetic Data..
Paola Cascante-Bonilla, Khaled Shehada, James Seale Smith, Sivan Doveh, Donghyun Kim,
Rameswar Panda, Gül Varol, Aude Oliva, Vicente Ordonez, Rogerio Feris, Leonid Karlinsky.
International Conference on Computer Vision. ICCV 2023.
October 2023. Paris, France.

Improving Visual Grounding by Encouraging Consistent Gradient-based Explanations.
Ziyan Yang, Kushal Kafle, Franck Dernoncourt, Vicente Ordonez
Conf. on Computer Vision and Pattern Recognition. CVPR 2023.
June 2023. Vancouver, Canada.

Estimating and Maximizing Mutual Information for Knowledge Distillation.
Aman Shrivastava, Yanjun Qi, Vicente Ordonez
Workshop on Fair, Data Efficient and Trusted Computer Vision at CVPR 2023.
June 2023. Vancouver, Canada.

CLIP-Lite: Information Efficient Visual Representation Learning from Textual Annotations.
Aman Shrivastava, Ramprasaath R. Selvaraju, Nikhil Naik, Vicente Ordonez
Artificial Intelligence and Statistics Conference. AISTATS 2023. April 2023.

SimVQA: Exploring Simulated Environments for Visual Question Answering.
Paola Cascante-Bonilla, Hui Wu, Letao Wang, Rogerio Feris, Vicente Ordonez
Conf. on Computer Vision and Pattern Recognition. CVPR 2022. July 2022.

Towards Understanding Gender-Seniority Compound Bias in Natural Language Generation.
Samhita Honnavalli, Aesha Parekh, Lily Ou, Sophie Groenwold, Sharon Levy, Vicente Ordonez,
William Yang Wang.
Language Resources and Evaluation Conference. LREC 2022. June 2022.

Backpropagation-Based Decoding for Multimodal Machine Translation.
Ziyan Yang, Leticia Pinto-Alva, Franck Dernoncourt, Vicente Ordonez.
Frontiers in Artificial Intelligence. January 2022. (Journal Paper)

Evolving Image Compositions for Feature Representation Learning.
Paola Cascante-Bonilla, Arshdeep Sekhon, Yanjun Qi, Vicente Ordonez
British Machine Vision Conference. BMVC 2021. November 2021.

Visual News: Benchmark and Challenges in Entity-aware Image Captioning.
Fuxiao Liu, Yinghan Wang, Tianlu Wang, Vicente Ordonez.
Conf. on Empirical Methods in Natural Language Processing. EMNLP 2021. Punta Cana,
Dominican Republic. (Oral Presentation)

Instance-level Image Retrieval using Reranking Transformers.
Fuwen Tan, Jiangbo Yuan, Vicente Ordonez
International Conference on Computer Vision. ICCV 2021.
MEDIRL: Predicting the Visual Attention of Drivers via Maximum Entropy Deep Inverse Reinforcement Learning. Sonia Baee, Erfan Pakdamanian, Inki Kim, Lu Feng, Vicente Ordonez, Laura Barnes. International Conference on Computer Vision. ICCV 2021.

General Multi-label Image Classification with Transformers.
Jack Lanchantin, Tianlu Wang, Vicente Ordonez, Yanjun Qi. Conf. on Computer Vision and Pattern Recognition. CVPR 2021.

Black-box Explanation of Object Detectors via Saliency Maps.
Vitali Petsiuk, Rajiv Jain, Varun Manjunatha, Vlad I. Morariu, Ashutosh Mehra, Vicente Ordonez, Kate Saenko. Conf. on Computer Vision and Pattern Recognition. CVPR 2021. (Oral Presentation)

Curriculum Labeling: Revisiting Pseudo-Labeling for Semi-Supervised Learning.
Paola Cascante-Bonilla, Fuwen Tan, Yanjun Qi, Vicente Ordonez. The Thirty-Fifth AAAI Conference on Artificial Intelligence. AAAI 2021.

Enabling AI at the edge with XNOR Networks.
Mohammad Rastegari, Vicente Ordonez, Joseph Redmon, Ali Farhadi. Communications of the ACM. December 2020 Vol 63, No. 12. CACM 2020. (Research Highlight, Invited Paper)

Using Visual Feature Space as a Pivot Across Languages.
Ziyian Yang, Leticia Pinto-Alva, Franck Dernoncourt, Vicente Ordonez. Findings of the Association for Computational Linguistics: EMNLP 2020.

Double-Hard Debias: Tailoring Word Embeddings for Gender Bias Mitigation.
Tianlu Wang, Xi Victoria Lin, Nazneen Fatema Rajani, Bryan McCann, Vicente Ordonez, Caiming Xiong. Association for Computational Linguistics. ACL 2020.

Generative-discriminative Feature Representations for Open-set Recognition.
Pramuditha Perera, Vlad I. Morariu, Rajiv Jain, Varun Manjunatha, Curtis Wigington, Vicente Ordonez, and Vishal M. Patel. Conf. on Computer Vision and Pattern Recognition. CVPR 2020.

Testing DNN Image Classifiers for Confusion & Bias Errors.
Yuchi Tian, Ziyuan Zhong, Vicente Ordonez, Gail Kaiser, Baishakhi Ray. International Conference on Software Engineering. ICSE 2020.

Drill-down: Interactive Retrieval of Complex Scenes using Natural Language Queries.
Fuwen Tan, Paola Cascante-Bonilla, Hui Wu, Xiaoxiao Guo, Song Feng, Vicente Ordonez. Conf. on Neural Information Processing Systems. NeurIPS 2019. Vancouver, Canada.

Balanced Datasets Are Not Enough: Estimating and Mitigating Gender Bias in Deep Image Representations.
Tianlu Wang, Jieyu Zhao, Mark Yatskar, Kai-Wei Chang, Vicente Ordonez. International Conference on Computer Vision. ICCV 2019. Seoul, South Korea.

Text2Scene: Generating Compositional Scenes from Textual Descriptions.
Fuwen Tan, Song Feng, Vicente Ordonez. Conf. on Computer Vision and Pattern Recognition. CVPR 2019. Long Beach, California. (Oral Presentation – Best Paper Finalist – top ~ 1% of submissions)
Chat-crowd: A Dialog-based Platform for Visual Layout Composition.
Paola Cascante-Bonilla, Xuwang Yin, Vicente Ordonez, Song Feng.
North American Chapter of the Association for Computational Linguistics. NAACL 2019. System Demonstrations Track. Minneapolis, Minnesota.

Gender Bias in Contextualized Word Embeddings.
Jieyu Zhao, Tianlu Wang, Mark Yatskar, Ryan Cotterell, Vicente Ordonez, Kai-Wei Chang.
North American Chapter of the Association for Computational Linguistics. NAACL 2019. short. Minneapolis, Minnesota. (Oral Presentation)

Deep Feature Aggregation and Image Re-ranking with Heat Diffusion for Image Retrieval.
Shanmin Pang, Jin Ma, Jianru Xue, Jihua Zhu, Vicente Ordonez.
IEEE Transactions on Multimedia 2019. (Journal Paper)

Feedback-prop: Convolutional Neural Network Inference under Partial Evidence.
Tianlu Wang, Kota Yamaguchi, Vicente Ordonez. Conf. on Computer Vision and Pattern Recognition. CVPR 2018. Salt Lake City, Utah.

Gender Bias in Coreference Resolution: Evaluation and Debiasing Methods.
Jieyu Zhao, Tianlu Wang, Mark Yatskar, Vicente Ordonez, Kai-Wei Chang.
North American Chapter of the Association for Computational Linguistics. NAACL 2018. short. New Orleans, Louisiana.

Building Discriminative CNN Image Representations for Object Retrieval using the Replicator Equation. Shanmin Pang, Jihua Zhu, Jiaxing Wang, Vicente Ordonez, Jianru Xue.
Pattern Recognition 2018. Volume 83. Pages 150-160. Accepted April 2018. (Journal Paper)

Where and Who? Automatic Semantic-Aware Person Composition.
Fuwen Tan, Crispin Bernier, Benjamin Cohen, Vicente Ordonez, Connelly Barnes.
Winter Conference on Applications of Computer Vision WACV 2018. Lake Tahoe, NV.

Men Also Like Shopping: Reducing Gender Bias Amplification using Corpus-level Constraints. Jieyu Zhao, Tianlu Wang, Mark Yatskar, Vicente Ordonez, Kai-Wei Chang.
Empirical Methods in Natural Language Processing. EMNLP 2017. Copenhagen, Denmark. (Oral Presentation) (Best Paper Award)

Obj2Text: Generating Visually Descriptive Language from Object Layouts.
Xuwang Yin, Vicente Ordonez.
Empirical Methods in Natural Language Processing. EMNLP 2017. Copenhagen, Denmark. (Oral Presentation)

Commonly Uncommon: Semantic Sparsity in Situation Recognition.
Mark Yatskar, Vicente Ordonez, Luke Zettlemoyer, Ali Farhadi.
Int. Conf. on Computer Vision and Pattern Recognition. CVPR 2017. Honolulu, Hawaii.

XNOR-Net: ImageNet Classification Using Binary Convolutional Neural Networks.
Mohammad Rastegari, Vicente Ordonez, Joseph Redmon, Ali Farhadi.
European Conference on Computer Vision. ECCV 2016. Amsterdam, Netherlands. (Oral presentation)
Stating the Obvious: Extracting Visual Common Sense Knowledge.  
Mark Yatskar, Vicente Ordonez, Ali Farhadi. North American Chapter of the Association of 
Computational Linguistics. NAACL 2016. short. San Diego, CA  
(Oral presentation)

Learning to Name Objects. Vicente Ordonez, Wei Liu, Jia Deng, Yejin Choi, Alexander C. Berg, Tamara L. Berg. Communications of the ACM. March 2016. Vol 59, No. 3. CACM 2016. (Research Highlight, Invited Paper)

Large Scale Retrieval and Generation of Image Descriptions.  
V. Ordonez, X. Han, P. Kuznetsova, G. Kulkarni, M. Mitchell, K. Yamaguchi, K. Stratos, A. Goyal, J. Dodge, A. Mensch, H. Daume III, A.C. Berg, Y. Choi, T.L. Berg. International Journal of Computer Vision. Special Issue on Big Data. IJCV 2016.  
(Journal Paper)

Predicting Entry-Level Categories. Vicente Ordonez, Wei Liu, Jia Deng, Yejin Choi, Alexander C. Berg, Tamara L. Berg. International Journal of Computer Vision - Marr Prize Special Issue. IJCV 2015. (Journal Paper)

ReferitGame: Referring to Objects in Photographs of Natural Scenes.  
Sahar Kazemzadeh, Vicente Ordonez, Mark Matten, Tamara L. Berg  
Empirical Methods in Natural Language Processing. EMNLP 2014. Doha, Qatar.  
(Oral presentation)

Learning High-level Judgments of Urban Perception.  
Vicente Ordonez, Tamara L. Berg  
European Conference on Computer Vision. ECCV 2014. Zurich, Switzerland.

TreeTalk: Composition and Compression of Trees for Image Descriptions.  
Polina Kuznetsova, Vicente Ordonez, Tamara L. Berg, Yejin Choi.  
Transactions of the Association of Computational Linguistics. TACL 2014  
Presented at EMNLP 2014. Doha, Qatar. (Oral Presentation, Journal Paper)

FurnitureGeek: Understanding Fine-Grained Furniture Attributes from Freely Associated Text and Tags. Vicente Ordonez, Vignesh Jagadeesh, Wei Di, Anurag Bhardwaj, Robinson Piramuthu. IEEE Winter Conference on Applications of Computer Vision. WACV 2014. Steamboat Springs, CO

From Large Scale Image Categorization to Entry Level Categories.  
Vicente Ordonez, Jia Deng, Yejin Choi, Alexander C. Berg, Tamara L. Berg.  
IEEE International Conference on Computer Vision. ICCV 2013. Sidney, Australia.  
(Oral presentation) (Best Paper Award - Marr Prize)  
(Selected for publication in the Research Highlights of the Communications of the ACM Magazine)

Generalizing Image Captions for Image-Text Parallel Corpus.  
Polina Kuznetsova, Vicente Ordonez, Alexander C. Berg, Tamara L. Berg, Yejin Choi. Association for Computational Linguistics. ACL 2013. short. Sofia, Bulgaria.

Babytalk: Understanding and Generating Image Descriptions.  
G. Kulkarni, V. Premraj, V. Ordonez, S. Dhar, S. Li, Y. Choi, A. C. Berg, T. L. Berg. IEEE Transactions on Pattern Analysis and Machine Intelligence. TPAMI 2013. (Journal paper)
Collective Generation of Natural Image Descriptions.
Polina Kuznetsova, Vicente Ordonez, Alexander C. Berg, Tamara L. Berg, Yejin Choi.
Association for Computational Linguistics. ACL 2012. Jeju, South Korea.
(Oral presentation)

Im2Text: Describing Images Using 1 Million Captioned Photographs.
Vicente Ordonez, Girish Kulkarni, Tamara L. Berg. Neural Information Processing Systems.
NeurIPS 2011. Granada, Spain. (Spotlight presentation)

High Level Describable Attributes for Predicting Aesthetics and Interestingness.
Sagnik Dhar, Vicente Ordonez, Tamara L. Berg.
IEEE Computer Vision and Pattern Recognition. CVPR 2011. Colorado Springs, CO.

The Ariadne Infrastructure for Managing and Storing Metadata.
S. Ternier, G. Parra, B. Vandeputte, K. Verbert, J. Klerkx, E. Duval, V. Ordonez, X. Ochoa.
Emerging Internet Technologies and Applications for E-learning.
IEEE Internet Computing 2009. (Journal paper)

**Patents**

Explanatory visualizations for object detection. Adobe Inc.
Rajiv Jain, Vlad Morariu, Vitali Petsiuk, Varun Manjunatha, Ashutosh Mehra, Vicente Ordonez. US Patent No. US11227159B2. (2020).

Techniques for automatic photo album generation. Google – Android.
Vicente Ordonez, Wei Hua, Rodrigo L. Carceroni, Jennifer Gillenwater, Amarnag Subramanany. US Patent No. 8983193. (2015).

Correlating image annotations with foreground features. eBay Inc.
Anurag Bhardwaj, Robinson Piramuthu, Vicente Ordonez, Vignesh Jagadeesh, Wei Di.
US Patent Application No. 20150067471. (2015).

**External Funding**

Total: $1,839,740 ($1,520,015 for our group)

Amazon AWS Cloud Credits for Research.
Project: Large Scale Models for Image-to-Image Matching.
$100,000 Cloud Credits as sole-PI 2023-2024.

Google Award for Inclusion Research.
Project: Measuring and Mitigating Societal Biases in Vision and Language Models for Open Recognition. $60,000 total as sole-PI 2022-2023.

Google’s exploreCSR Program.
For organizing: Regional Workshop on Applications and Research in Data Science (REWARDS). $18,000 + $7,000 total as co-PI. 2022-2023 (with Tilda Oré).

NSF Faculty Early Career Development Program (CAREER) Award.
Teaching Machines to Recognize Complex Visual Concepts in Images through Compositionality. $499,760 total as sole-PI. 2021-2025.

NSF Program on Fairness in Artificial Intelligence in Collaboration with Amazon (FAI).
Measuring and Mitigating Societal Biases in Generic Image Representations. $600,000 total as PI ($340,275 for our group), with Co-PI Baishakhi Ray. 2021-2024.
NSF amount: $375,000 under award #2040961, Amazon amount: $225,000.
REU Supplement: $14,000.
Salesforce AI Research Grant 2021  
Project: Addressing Biases on Generic Pretrained Models  
$50,000 as PI (Unrestricted Gift Funding)

Facebook Research Award 2020  
Project: On-Device Efficient Neural Networks  
$75,000 as PI (Unrestricted Gift Funding)

eBay Research Gift Funding 2020  
Project: Object-specific Image Retrieval  
$48,843 as PI (Unrestricted Gift Funding)

Adobe Research Gift Funding  
Project: Vision, Language and Multi-lingual Reasoning  
$5,000 in 2019 + $30,000 in 2020 + $15,500 in 2021 + $22,000 in 2022 as PI (Gift Funding)

Google Cloud Credits for Research 2020.  
$5,000 Cloud Credits, as PI.

Leidos Gift Funding, 2019.  
Project: Component-based Reasoning for Image Recognition  
$50,000 as PI.

SAP Gift Funding 2018  
$23,000, as PI (Unrestricted Gift Funding)

SAP Research Contract 2018.  
Project: Efficient Deep Learning through Compact Representations.  
$50,000, as PI.

SAP Research Contract 2018.  
Project: Spatial Reasoning for Visually Grounded Dialogs.  
$50,000 as PI.

Google Cloud Credits for Research 2019.  
$16,500 Cloud Credits, as PI.

Google Faculty Research Award 2017.  
Project: Mitigating Biases in Visual Recognition  
$49,700 + $10,437 Cloud Credits, as PI (Unrestricted Gift Funding).

IBM Faculty Award 2017.  
Project: Interactive Dialog for Image Retrieval and Synthesis  
$40,000, as PI (Unrestricted Gift Funding)

**INTERNAL FUNDING**

Total: $90,000 ($25,000 for our group).

4-VA Research Program 2019 (co-PI with Jia-Bin Huang – Virginia Tech).  
Mitigating Bias for Interpretable Human Activity Understanding.  
$25,000 ($5,000 for our group).

3 Cavaliers Program 2019 (co-PI with Paul Humphreys and Chip Levy).  
Comparing and Contrasting Artificial Neural Networks with Biological Neural Networks for Improved Representation Learning.  
$60,000 ($20,000 for our group).
La era de la Inteligencia Artificial: ¿Tenemos cabida en ella?. Telemundo Houston. 11/08/2023. https://www.telemundohouston.com/videos/videos-destacados/la-era-de-la-inteligencia-artificial-tenemos-cabida-en-ella/2367745/

If you can type it, you can see it: Text2image AI turns words into pictures. The Houston Chronicle. 10/20/2022. https://www.houstonchronicle.com/business/tech/article/If-you-can-type-it-you-can-see-it-Text2image-AI-17519217.php

How NSF and Amazon Are Collectively Tackling Artificial Intelligence-Based Bias. NextGov. 02/18/2021. https://www.nextgov.com/emerging-tech/2021/02/how-nsf-and-amazon-are-collectively-tackling-artificial-intelligence-based-bias/172126/

NSF Makes 11 Research Project Grants Under Amazon-Backed Fairness in AI Program. Tech Register UK. 02/12/2021. https://www.techregister.co.uk/nsf-makes-11-research-project-grants-under-amazon-backed-fairness-in-ai-program/

Biometrics experts call for creation of FDA-style government body to regulate facial recognition. Biometric Update. 06/08/2020. https://www.biometricupdate.com/202006/biometrics-experts-call-for-creation-of-fda-style-government-body-to-regulate-facial-recognition

Researchers call for new federal authority to regulate facial recognition tech. TechXplore. 06/05/2020. https://techxplore.com/news/2020-06-federal-authority-facial-recognition-tech.html

Racist facial recognition technology is being used by police at anti-racism protests. Verdict UK (VE). 06/05/2020. https://www.verdict.co.uk/facial-recognition-technology-racist-police-protests

Investigating the Best Features for Predicting a Movie’s Genre and Estimated Budget. TechXplore. 30/08/2019. https://techxplore.com/news/2019-08-features-movie-genre.html

La equidad de género en los momentos de la IA. Forbes México. 01/10/2019. https://www.forbes.com.mx/la-equidad-de-genero-en-los-momentos-de-la-ia/

Researchers Combat Gender and Racial Bias in Artificial Intelligence. Bloomberg News. 12/04/2017. https://www.bloomberg.com/news/articles/2017-12-04/researchers-combat-gender-and-racial-bias-in-artificial-intelligence

Home robots will turn into crude sexists, experts warn. The Times of London. 08/23/2017. https://www.thetimes.co.uk/article/home-robots-will-turn-into-crude-sexists-experts-warn-gnmj09rgq

Machines Taught by Photos Learn a Sexist View of Women. WIRED. 08/21/2017. https://www.wired.com/story/machines-taught-by-photos-learn-a-sexist-view-of-women/

Beyond Silicon: Squeezing More Out of Chips. The New York Times. 10/30/2016. http://www.nytimes.com/2016/10/31/technology/beyond-silicon-squeezing-more-out-of-chips.html

Artificial Intelligence at Your Fingertips. University of Washington CSE News. 10/30/2016. https://news.cs.washington.edu/2016/10/31/uw-cse-and-ai2-in-the-new-york-times-artificial-intelligence-at-your-fingertips/
Invited Talks

Invited Talk: Challenges and Opportunities with Multimodal LLMs
University of British Columbia, Centre for AI, Decision-making and Action (CAIDA) April 15, 2024
Centro de Investigación en Matemáticas (CIMAT), Guanajuato, MX (online) May 13, 2024

Keynote: Envisioning the Next Generation of Vision and Language Models
IEEE International Conference on Pattern Recognition Systems (ICPRS) 2023.
Guaquila, Ecuador.

Invited Talk: Vision and Language: Opportunities and Limitations
Computer Science Seminar Series, University of Houston April 5, 2023

Invited Talk: Vision and Language: Opportunities and Limitations
Google Brain Montréal, Alphabet Inc. (online) March 2, 2023

Invited Speaker: Searching for Images: Granularity, Compositionality, Interpretability and Multimodality
Instance-level Recognition (ILR) Workshop at ECCV 2022, Tel-aviv, Israel (online) October 24, 2022

Invited Talk: On the Success of Large Scale Vision and Language Models
Forum on Artificial Intelligence at the University of Texas at Austin October 14, 2022

Keynote Speaker: Reunión Internacional de Inteligencia Artificial y sus Aplicaciones (RIIAA)
Intl. Conference on Artificial Intelligence and Applications. Universidad San Francisco de Quito (USFQ), Quito, Ecuador September 29, 2022

Keynote Speaker: Past, Present and Future of Vision and Language Models
LatinX in Computer Vision Workshop at CVPR 2022, New Orleans, LA. June 19, 2022

Invited Speaker: Vision and Language: Compositionality and Fairness
Johns Hopkins University, Human Language Technology Center of Excellence March 2022
Texas A&M University, Department of Computer Science & Engineering
Rice University, Department of Computational and Applied Math

Invited Speaker: CODECON 256 – ITSJBA - Daule, Ecuador September 2021
Title: Teachings from Computer Science to Become a Better Programmer. (online)

Invited Speaker: CVPR Workshop on Responsible Computer Vision (RCV) June 2021
Title: Toward Compositional Models that Mitigate Bias and Responsible Model Deployment

Invited Speaker: Compositional Representations for Visual Recognition
University of Rochester, Computer Science Vision Seminar (online) April 2021
Arizona State University, Vision and Language Seminar (online) February 2021
Carnegie Mellon University, VASC Seminar (online) September 2020

Invited Speaker: Measuring and Mitigating Biases in Computer Vision and Beyond
University of Florida, Dept. of Computer & Information Science & Engineering March 2021
Rice University, Department of Computer Science Colloquium October 2020

Invited Speaker: Compositionality, Robustness, and Flexibility in Vision and Language
Amazon Alexa AI, Amazon Inc April 2021
Samsung Research, Cambridge, UK October 2020

Invited Speaker: Visual recognition on the edge through platform-aware model optimization
Facebook AI Systems Faculty Summit October 2020
Invited Speaker: Reunión Internacional de Inteligencia Artificial y sus Aplicaciones (RIIAA)  
Intl. Conference on Artificial Intelligence and Applications  
August 2020

Invited Speaker: Fair and Compositional Representations for Vision and Language  
ICCV Workshop on Linguistics Meets Image and Video Retrieval  
October 2019

Invited Speaker: Building Compositional, Interpretable and Robust Visual Recognition  
Georgetown University - Dept. of Computer Science. Washington DC  
October 2019

Invited Speaker: Human-guided Visual Recognition with Language and Interaction  
SAP Leonardo Machine Learning Research Retreat. Berlin, Germany  
October 2019

Invited Speaker: Building Fair and Robust Representations for Vision and Language  
Oak Ridge National Laboratory AI Workshop. Oak Ridge, TN.  
September 2019

Seminar Speaker: Building Fair Representations for Images and Text  
University of Maryland College-Park, CLIP Seminar Series  
August 2019

Invited Speaker: Building the Next Generation of Representations for Vision and Language  
Adobe Research, San José, California  
March 2019

Invited Speaker: Building the Next Generation of Representations for Vision and Language  
Escuela Superior Politécnica del Litoral, Guayaquil, Ecuador  
January 2019

Invited Speaker: Challenges in Vision and Language Research  
Workshop on Shortcomings in Vision and Language (SiVL)  
European Conference on Computer Vision (ECCV), Munich, Germany  
September 2018

Invited Speaker: Feedback Propagation in Deep Neural Networks  
SAP Leonardo Machine Learning Research Retreat. Munich, Germany  
September 2018

Invited Speaker: Overcoming the Next Challenges in Vision and Language Research  
Microsoft Research Montreal. Montréal, Canada  
August 2018

Mini-Plenary Speaker: Reducing Gender Bias in Machine Learning Systems  
National Center for Women & IT, NCWIT Summit. Grapevine, TX  
May 2018

Invited Speaker: Understanding the Visual World through Language  
Applied Machine Learning Conference. Charlottesville, VA  
April 2018

Keynote Speaker: Integrating Vision and Language through Feedback-based Neural Inference  
Conference on Integrating Vision and Language Processing. Tartu, Estonia  
March 2018

Invited Speaker, Investigación en Reconocimiento Visual Artificial  
Universidad Técnica de Machala (UTM), Machala, Ecuador.  
May 2017

Invited Speaker, Language and Perceptual Categorization in Computer Vision.  
Toyota Technological Institute at Chicago TTI-C  
April 2015

Allen Institute for Artificial Intelligence AI2, Seattle, Washington  
March 2015

California Institute of Technology (Caltech), Pasadena, California  
March 2015
Carnegie Mellon University, The Robotics Institute VASC Seminar Series February 2015
Disney Research Pittsburgh, The Walt Disney Company. January 2015
Stanford University, Department of Computer Science, Vision Group. December 2014

Seminar Speaker, Integrating Vision and Language.
University of Virginia, Computer Science Dept. Charlottesville, Virginia March 2015
Virginia Tech, Computer Science Dept., Blacksburg, Virginia March 2015
Drexel University, Computer Science Dept. Philadelphia, Pennsylvania February 2015

Invited Presentation, Learning High-level Judgments of Urban Perception. September 2014
ECCV 2014 Workshop on Storytelling with Images and Videos. Zurich
ECCV 2014 Workshop on Human-Machine Communication for Visual Recognition. Zurich

Invited Speaker, Understanding Image Descriptions in the Wild. July 2013
Yahoo! Labs, Sunnyvale, California.

Invited Student Speaker, Data-driven Generation of Image Descriptions. June 2013
NAACL Workshop on Vision and Language (WVL) 2013. Atlanta, GA

---

**RESEARCH GROUP**

Jefferson Hernandez, PhD Student (Rice) Fall 2022 - present
Catherine He, PhD Student (Rice) Fall 2023 - present
Moayed Haji Ali, PhD Student (Rice) Fall 2023 - present
Jaywon Koo, PhD Student (Rice) Fall 2023 - present
Zilin Xiao, PhD Student (Rice) Fall 2023 - present
Aman Shrivastava, PhD Student (UVA) (advised now by Tom Fletcher) Fall 2020 - present
Micah Kepe, Undergrad (Rice) Spring 2024 - present
Zachary Kepe, Undergrad (Rice) Summer 2024 - present

**RESEARCH GROUP ALUMNI**

Paola Cascante-Bonilla, PhD – Rice University
PhD Thesis: *More from Less: Learning with Limited Annotated Data in Vision and Language.* May 2024
Next: Assistant Professor at Stony Brook University (SUNY)
Academic homepage: [https://paolacascante.com/](https://paolacascante.com/)

Ziyan Yang, PhD – Rice University
PhD Thesis: *Vision and Language: Information Integration and Transformation.* November 2023
Next: Research Scientist at Tiktok – San Jose, CA
Academic homepage: [https://ziyanyang.github.io/](https://ziyanyang.github.io/)

Tianlu Wang, PhD – University of Virginia
PhD Thesis: *Measuring and Mitigating Biases in Vision and Language Models.* June 2021
Next: Research Scientist at Facebook AI Research – Menlo Park, CA
Academic homepage: [https://tianlu-wang.github.io/](https://tianlu-wang.github.io/)

Fuwen Tan, PhD – University of Virginia
PhD Thesis: *Learning Local Representations of Images and Text.* April 2021
Next: Researcher at Samsung AI Centre – Cambridge, UK
Academic homepage: [https://fwtan.github.io/](https://fwtan.github.io/)

Ayush Sachdeva, BS/MS Student Summer 2023 - Spring 2024
Daniel Wang, MS Student (next ML Engineer at Tiktok) Spring 2023 - Fall 2023
Jaspreet Ranjit, Research Assistant (next PhD Student at USC) Spring 2021 - Spring 2022
Hannah Lei, Undergrad/MS Student (next at Uber) Fall 2021
Leticia Pinto-Alva, RA / Visiting Student (next PhD Student at USC) Fall 2019 - Fall 2020
Fuxiao Liu, MS Student (next PhD Student at UMD) Fall 2019 - Fall 2020
Lindsey Shavers, Ugrad/MS Student Fall 2018, Fall 2019 - Spring 2020
Xuwang Yin, Research Assistant (next PhD at UVA CPE program) Fall 2016 - Summer 2018
Fengyang Zhang, MS - Independent Studies (next at Facebook) Spring 2018
Abhimanyu Banerjee, MS - Independent Studies (next at Cvent Inc)  Spring 2017 - Spring 2018
Anudeep Konda, MS - Thesis (next at VoxelCloud)  Fall 2017 - Spring 2018
KS Sivaraman, MS - Independent Studies (next at Microsoft)  Spring 2017 - Fall 2017
Alex Biyanov, Undergrad Student  Spring 2024
Diego (Letao) Wang, Undergrad Student  Spring 2021 - Spring 2022
Nikash Sethi, Undergrad Student (next at Scale.AI)  Spring 2021
Katherine Weinschenk, Undergrad Student  Fall 2020 - Spring 2021
Jeffrey Tan, Undergrad Student  Summer 2018 - Fall 2020
Brandon Peck, Undergrad Student - (next at Microsoft)  Summer 2019
Mengjia Luo, Undergrad Student - (next at Microsoft)  Fall 2018 - Spring 2019
Arun Kannan, Undergrad Capstone Project - (next UVA MS in Math)  Fall 2017 - Spring 2018
Vijay Edupuganti, Undergrad Capstone Project (next at OpenDoor Inc)  Fall 2017 - Spring 2018
Shijia Wang, Undergrad Student - (next Georgia Tech MS)  Summer 2017 - Fall 2017
Nova Zhang, Undergrad Student  Spring 2017
Divya Bhaskhara, Undergrad Capstone Project - (next Johns Hopkins MS)  Fall 2016
Ernesto Cruz-Esquivel, Visiting Student – UDLAP Mexico  Summer 2023
Veronica Flores, CRA-DREU Visiting Student – Santa Clara University  Summers 2022, 2023
Sindhu Kothe, CRA-DREU Visiting Student – UC San Diego  Summer 2023
Lucas Resck, Visiting Student – Fundaçao Getulio Vargas, Brazil  Fall 2022
Shanmin Pang, Visiting Scholar – Asst. Prof. at Xi’an Jiaotong University  Fall 2017 - Fall 2018
Eleanor Lin, CRA-DREU Visiting Student – Columbia University  Summer 2022
Jonathan Rodriguez, CRA-DREU Visting Student – Tufts University  Summer 2018
Rosangel Garcia, CRA-DREU Visting Student – Le Moyne College  Summers 2017, 2018
Ian K. Torres, CRA-DREU Visting Student – UMass Amherst  Summer 2017

Teaching Experience

COMP 646: Deep Learning for Vision and Language. Rice University  
Course website: https://www.cs.rice.edu/~vo9/deep-vislang  Spring 2024
Course website: https://www.cs.rice.edu/~vo9/deep-vislang/2023  Spring 2023
Course website: https://www.cs.rice.edu/~vo9/deep-vislang/2022  Spring 2022

COMP 648: Computer Vision Seminar. Rice University  
Course website: https://www.cs.rice.edu/~vo9/cv-seminar  Fall 2023
Course website: https://www.cs.rice.edu/~vo9/cv-seminar/2022  Fall 2022

CS4501: Introduction to Computer Vision. University of Virginia  
Course website: http://www.cs.rice.edu/~vo9/vision  Spring 2021
Course website: http://www.cs.rice.edu/~vo9/vision/2019  Fall 2019
Course website: http://www.cs.rice.edu/~vo9/vision/2018  Spring 2018

CS6501: Vision and Language. University of Virginia  
Course website: http://www.cs.rice.edu/~vo9/vislang  Fall 2020
Course website: http://www.cs.rice.edu/~vo9/vislang  Spring 2017

CS6501: Deep Learning for Visual Recognition. University of Virginia  
Course website: http://www.cs.rice.edu/~vo9/deeplearning  Spring 2020
Course website: http://www.cs.rice.edu/~vo9/deeplearning/2019  Spring 2019

CS6501: Computational Visual Recognition. University of Virginia  
Course website: http://www.cs.rice.edu/~vo9/recognition  Fall 2017
Course website: http://www.cs.rice.edu/~vo9/recognition/2016  Fall 2016
Teaching Assistant.
Foundations of Computer Science. Stony Brook University. 2009-2010
Object Oriented Programming. Escuela Superior Politécnica del Litoral, Ecuador. 2005
Introduction to Programming. Escuela Superior Politécnica del Litoral, Ecuador. 2004
Physics for Engineering. Escuela Superior Politécnica del Litoral, Ecuador. 2004

I have also participated as one time guest lecturer for the following courses:
Generative Artificial Intelligence and Humanity, Glasscock School of Continuing Studies at Rice University (Fall 2023)
Computer Vision. ECE Department – Rice University (Spring 2022)
Vision and Language. Computer Science – University of British Columbia (Spring 2021)
Big Data and Marketing Analytics. McIntire School of Commerce – UVA (Fall 2019)
Introduction to Computer Vision. University of Virginia (Spring 2017)
Language and Vision. University of North Carolina at Chapel Hill (Spring 2015)
Computer Vision. University of North Carolina at Chapel Hill (Fall 2014)
Artificial Intelligence. University of North Carolina at Chapel Hill (Spring 2014)
Advanced Multimedia. Stony Brook University (Spring 2013)
Computational Photography. Stony Brook University (Spring 2013)

Volunteer Mentor:
• Doctoral Symposium, Conf. on Computer Vision and Pattern Recognition (CVPR) 2023.

Member of Organizing Committee and Host:
• Workshop on Frontiers in Visual Language Reasoning: Compositionality, Prompts and Causality at the Conf. on Computer Vision and Pattern Recognition (CVPR) 2023.

Outreach Workshop Co-organizer:
• Regional Workshop on Applications and Research in Data Science (ReWARDS) 2023.
  Sponsored by Google at Universidad Técnica de Machala (UTMACH), Machala, Ecuador.
• Regional Workshop on Applications and Research in Data Science (ReWARDS) 2022.
  Sponsored by Google at Universidad de Piura (UDEP), Piura, Perú.

Member of Organizing Committee:
• Responsible Computer Vision Workshop (RCV) at ECCV 2022.

Member of Diversity, Equity and Inclusion Committee:
• Conference on Computer Vision and Pattern Recognition, 2022.

Tutorials Chair / Member of Organizing Committee for the main conference:
• International Conference on Computer Vision (ICCV) 2021 – Montréal (Virtual).

Member of Proposing Team:
• 1st LatinX in Computer Vision Workshop at CVPR 2021.

Member of Organizing Committee:
• Toward Human-level Video Story Understanding Workshop (VTT) at ECCV 2020.

Tutorial on Bias and Fairness in Natural Language Processing, at the conference:
• Empirical Methods in Natural Language Processing (EMNLP) 2019 – Hong Kong.

Member of Organizing Team of Panel on Bias in Machine Learning at the
• ACM Richard Tapia Celebration of Diversity in Computing 2018.

BigVision Workshop on Large Scale Visual Recognition and Retrieval at the
• Conference on Computer Vision and Pattern Recognition (CVPR) 2016.

Co-Director with Prof. Paul Humphreys (Dept. of Philosophy):
• University of Virginia’s Human and Machine Intelligence Seminar, 2017 - 2021.

Co-Organizer with my PhD Students
• University of Virginia’s Computer Vision Seminar, 2016 - 2021

Co-Organizer as part of the Computer Science Graduate Research Council
• Stony Brook University Computer Science Graduate Research Conference 2010
Area Chair / Program Committee / Meta-Reviewer.
Conf. on Neural Information Processing Systems (NeurIPS) 2023, 2024
European Conference on Computer Vision (ECCV) 2020, 2024
Winter Conference on Applications of Computer Vision (WACV) 2024
International Conference on Computer Vision (ICCV) 2019, 2023
Conf. on Computer Vision and Pattern Recognition (CVPR) 2020, 2021, 2022, 2023
Association for Computational Linguistics (ACL) Rolling Review 2022.
Neural Information Processing Systems (NeurIPS) 2022 – Benchmarks and Datasets
AAAI Conference on Artificial Intelligence (AAAI-22) as Senior Area Chair
British Machine Vision Conference (BMVC), 2021
Annual Meeting of the Association for Computational Linguistics (ACL) 2020
North American Chapter of the Assoc. for Computational Linguistics (NAACL) 2018, 2019
Discussion Panels.
Science of Deep Learning Workshop at ICLR 2021 (as moderator)
Escape Velocity 2019 - Ethics in AI Panel. Washington DC’s National Harbor
Workshop Program Committee / Reviewer.
NeurIPS Black in AI Workshop 2018, 2019, 2020, 2021
ACM Richard Tapia Celebration of Diversity in Computing: Panels and Workshops, 2019
NAACL Workshop on New Forms of Generalization in Deep Learning 2018
ICCV Workshop on Closing the Loop Between Vision and Language 2015, 2017
ICCV Workshop on Web-scale Vision and Social Media 2015, 2017
NeurIPS Workshop on Efficient Methods for Deep Neural Networks 2016
ECCV Workshop on Web-scale Vision and Social Media 2016
EMNLP Workshop on Vision and Language 2015
ECCV Workshop on Storytelling with Images and Videos - VisStory 2014
ECCV Workshop on Human-Machine Communication for Visual Recognition 2014
Panelist for Reviewing Research Proposals.
Swiss National Science Foundation (SNSF), 2020, 2021, 2024
Natural Sciences and Engineering Research Council of Canada (NSERC), 2023
US National Science Foundation (NSF), 2018, 2019, 2021, 2022
Dutch Research Council (NWO), The Netherlands, 2022
Flanders Research Foundation (FWO), Belgium, 2017
Reviewer / Program Committee
Conference on Computer Vision and Pattern Recognition (CVPR) 2015 - 2019, 2024
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 2014 - 2023
Neural Information Processing Systems (NeurIPS) 2016, 2020, 2021, 2022
ACM Conference on Fairness, Accountability and Transparency (FAccT), 2022
Artificial Intelligence Journal (AIJ) 2021
Winter Conference on Applications of Computer Vision (WACV) 2021
AAAI Conference on Artificial Intelligence (AAAI-21)
International Conference on Machine Learning (ICML) 2020 (as “Expert Reviewer”)
Journal of Artificial Intelligence Research (JAIR) 2020
Empirical Methods in Natural Language Processing (EMNLP) 2015, 2017, 2018, 2020
ACM Richard Tapia Celebration for Diversity in Computing – Panels and Workshops, 2019
European Conference on Computer Vision (ECCV) 2016, 2018
Association for Computational Linguistics (ACL) 2014, 2016 - 2018
International Conference on Computer Vision (ICCV) 2015, 2017
International Joint Conference in Artificial Intelligence (IJCAI) 2016
International Journal of Computer Vision (IJCV) 2014 - 2016
Asian Conference on Computer Vision (ACCV) 2016
Int’l Conference on Computer Graphics and Interactive Techniques (SIGGRAPH) 2016
North American Chapter of the Association for Computational Linguistics (NAACL) 2016
IEEE Transactions on Multimedia (TM) 2013, 2016
Elsevier Computer Vision and Image Understanding (CVIU) 2014, 2015
Elsevier Information Processing Letters (IPL) 2014
IEEE Transactions on Neural Networks and Learning Systems (TNNLS) 2014
IEEE Transactions on Image Processing (TIP) 2013
ACM Multimedia (MM) 2010. International Multimedia Conference.