I am a computer scientist who conducts research in the areas of *computer vision* and *machine learning*. My goal is to make fundamental contributions towards building AI systems with rich visual reasoning capabilities. Light is abundant and easy to measure making it a versatile sensing modality. Moreover, much of what we create for communication is visual in nature. Thus, techniques for analyzing visual data can not only enable numerous applications, but also provide insights to various physical, cultural, social, economic phenomenon that profoundly affect our lives. For example, data being collected by sensor networks such as satellites and weather radars provide an unprecedented opportunity to understand planet-scale phenomena such as the impact of climate change on biodiversity. I argue that a primary barrier to realizing the full potential of these emerging data resources is *computational*. In particular, current AI systems are unable to *extract rich semantic information from visual data* and *solve novel tasks with limited supervision* to support reasoning and decision making.

Application areas. My research focuses on improving the *robustness*, *efficiency*, *generalization* and *interpretability* of visual recognition architectures in the context of the following application areas:

- *Texture understanding.* Models and theories of texture perception date back to the early work of Bela Julesz, 1981, who proposed a representation of texture as a collection of elementary units called "textons". The seminal work of Leung and Malik, 2001, provided a computational framework for textons. Its variants, and other texture representations such as the "bag-of-words" model of Csurka et al., 2004, were arguably the most influential image representation in the previous decade. My work on "deep filter banks" [CVPR'15] unified decades of early work on texture perception with recent advances in deep representations leading to new insights and models for texture analysis. In subsequent work we proposed an architecture called "bilinear networks" [ICCV'15] that provided a framework for combining multiple views of data in a texture-like representation. These models and their extensions have been shown to be effective at fine-grained recognition (e.g., identifying species of birds), scene understanding, and visual question answering tasks.
- *3D shape understanding.* Another line of my research focuses on models for 3D shape understanding, a critical requirement for autonomous navigation, grasping, manipulation, and other tasks. Together with collaborators we proposed one of the first deep architectures for 3D shape classification called "multi-view CNN" [ICCV'15]. Subsequently, we have developed several innovative architectures for 3D shape classification and semantic segmentation, as well as for estimating shapes from a single image, image collections, or even sketches. One of these techniques called SPLATNET was awarded a *best paper honorable mention* at CVPR 2018.
- *Ecology and astronomy.* With recent improvements in computational infrastructure and the ability to access large-scale datasets, computer vision is poised to provide fundamental scientific insights in many disciplines. I am part of an ongoing collaborative research effort to develop automatic tools to extract quantitative measurements of bird migration from the entire archive of US weather radar data (more than 200 million scans). Another collaboration with astronomers aims to develop algorithms for fine-grained classification of star clusters in high-resolution images of galaxies to shed light on the physics of star formation.

Research themes. In the context of the aforementioned applications my research follows two major themes:

- *Recognition with humans in-the-loop.* My research has developed a framework to discover semantic parts and attributes associated with a category to learn *interpretable* visual representations [CVPR'12,'13], user-interfaces to efficiently annotate images [AISTATS'14], and techniques to incorporate human knowledge during inference [CVPR'14, WACV'15]. These makes it possible to deploy vision systems for collaborative tasks.
- *Efficient learning and inference.* My PhD thesis made then widely-used image classifiers exponentially faster [CVPR'08], which was influential in the development of region-based object detection systems, a precursor to the modern R-CNN detectors. My recent research has focused on learning from *limited* supervision, a capability that can enable many practical applications. To this end my research has explored the role of datasets, representations, architectures, and algorithms for few-shot learning [e.g., ECCV'18, CVPR'19, ICCV'19].

Research impact. Since my PhD, I have published 27 articles in highly selective CV, ML and AI conferences, six journal articles, and two book chapters [7796 citations, h-index 32, according to Google scholar]. Papers that I have co-authored have received the *best paper honorable mention* at CVPR 2018, and *best student paper* at WACV 2015. Besides these, I have contributed several widely-used open-source software, datasets, and benchmarks to the AI community; and I am a long-term organizer of the fine-grained recognition workshops ($FGVC^3$— $FGVC^6$) that bring together researchers across diverse domains such as biology, neuroscience, psychology, and ecology.