Research Statement

The overall focus of my research is on computer vision, in particular fine-grained image classification and object instance segmentation. This area of computer vision is important, not only in developing intelligent robotic systems with broad societal impact, but also in accelerating interdisciplinary scientific research. Fine-grained image understanding can serve as a useful module in many applications such as navigation and collision avoidance in self-driving cars that segment pedestrians and other nearby cars in videos captured by on-board cameras, household robots that assist the elderly in daily activities, satellites and drones that perform environmental monitoring, etc. Furthermore, automated image-analysis techniques can enable experts to realize high-throughput studies in cross-disciplinary research, e.g. segmenting and counting *C. elegans* in microscope images for studying animal lifespan, and recognizing pollen species in phytology to study plant diversity and climate change over thousands of years. Solving such fine-grained computer vision problems requires great attention to fine details of the objects in an image, posing new challenges and difficulties whose solution will advance scientific understanding of vision algorithms. Therefore, I've been developing and publishing innovative machine learning approaches to attack these problems, and applying them to emerging inter-disciplinary problems.

The following elaborates my two main areas of research which I term "ubiquitous fine-grained computer vision".

Fine-grained classification is the problem of classifying images into subordinate categories within an entry level category, e.g. recognizing car model or identifying the bird species from images. This is more challenging than the generic classification counterpart, which, for example, aims to categorize images as containing a car or a bird. As objects in images have various poses and sizes, in order to use machine learning tools, great attention should be paid to the image details and meaningful bird parts (e.g. in birds the shape of beak or color texture on the wing). To automatically combine the part information efficiently and effectively, I am developing novel approaches that require only category labels without any part-level annotations, reducing the cost of collecting and annotating datasets to train machine learning models. These methods can be widely used not only in consumer applications like recognizing different styles of items sold online, but also in other areas of scientific research. I have demonstrated the application of these methods to identifying species from images of fossilized pollen grains to estimate the plant diversity in a particular area and particular time.

Fine-grained segmentation aims at segmenting semantic object instances present in an image, i.e. identify which pixels belonging to a specific object instance. This combines two traditional tasks in computer vision: object detection and semantic segmentation, which identify the location of target object instance (e.g. by drawing a bounding box around it) and segment regions in terms of category labels, respectively. Fine-grained instance segmentation attempts to unify these tasks, yet more challenging in terms of the goal to not only detect object instances, but also segment them by carefully refining regions around the object boundary. I have used a unified convolutional neural network to segment object masks and predict instance bounding boxes which can be combined as features to segment the object instances by grouping image pixels. I am also implementing top-down mechanisms in the network to encode cognitive process for better segmentation performance. A general framework for performing this task would have a wide range of applications, such as the self-driving cars which need to precisely segment pedestrians and other cars nearby, ground surveillance through satellites and drones, high-throughput studies over microscope images on lifespan, cell aging, gene mutation, drug, etc.