Attending graduate school at Stanford Computer Science embodies, to me, a unique opportunity to dedicate myself to the forefront of today's technological revolution. And in this increasingly data-driven economy, I have the insatiable desire to help build the world of tomorrow by collaborating with the finest minds in computation at Stanford. In particular, artificial intelligence, from detecting cancer cells to automated speech translation, is exponentially advancing humanity's most important functions, so I aim to use the resources within the CS Department to develop both meaningful expertise in AI and lifelong interpersonal connections with research leaders in machine learning. My principal interest here is developing a robust understanding of contemporary methods in graph mining, computer vision, and other machine learning fields, for example extending my current research on knowledge distillation from knowledge graphs. I have implemented various deep learning vision applications in industry and explored cutting-edge research in graph neural networks, so I believe the Master's program at the Stanford Department of Computer Science provides an unrivaled academic mastery and professional network in my interests, and hence the starting point for my journey to help transform the world through computing.

Over the last summer and through this school year, I have been working as a machine learning engineer at Viasat on the Intelligence, Surveillance, and Reconnaissance as a Service (ISRaaS) team, where I focused on action recognition and object tracking and re-identification tasks in the domain of low-resolution footage. My main takeaway from my time at Viasat that I wish to supplement at Stanford is how to create and use the latest novel ideas in the machine learning field and transform them into suitable software needs. For example, I tackled my task of improving the object tracking and re-identification module by dividing it into two pieces: metrics and model development. The evaluation metric was highly tied to business needs, so after perusing the literature and talking to the various project stakeholders I implemented into the evaluation pipeline the minimum per-class HOTA (Luiten et al., 2020) and IDF1 (Ristani et al., 2016) to best balance detection (ability to detect a new object) and association (correct identification of that object) accuracy. On the model-development side, I found that the previous DeepSORT (Wojke et al., 2017) association method failed to generalize between pedestrians and vehicles, which I successfully combated by integrating the ability for the CNN to dynamically include additional filters of different sizes when needed, inspired by OSNet (Zhou et al., 2019).

Moreover, I pioneered a project on action recognition on the low-resolution video domain at Viasat, implementing an altered version of X3D (FAIR, 2020). The algorithm employed different ways to extend a 2D CNN into the 3rd time dimension depending on how specific of an action was desired—real time recognition for broad action groups (e.g. person-car interaction) or more computationally expensive recognition for reanalyzing videos for specific actions (e.g. person opens trunk). The greater complexity model also integrated optical flow, pumping up the top-1 accuracy to 73% in this difficult application space.

So given exposure to these highly technical aspects of contemporary computer vision, I intend to both enhance my comprehension of Deep Learning methods for video and further develop breadth across all computer vision domains at Stanford. Of particular note, I'd like to contribute to the Stanford Vision and

Learning Lab, which has spearheaded the forefront of computer vision, starting from the original ImageNet. Building upon my experience in action recognition, I found the approach in RubiksNet by Fan and Buch et al. under Dr. Li Fei-Fei and Dr. Juan Carlos Niebles especially noteworthy; learning the proper spatio-temporal shift for video frames resembles a more concise way to learn “attention-aware filters” in the video context, and I would love the opportunity to collaborate with these researchers on further projects.

I also head a research project at Brown University under Dr. Stephen Bach on graph neural networks. The main idea that we have been pursuing is exporting the notion of “foundation models,” which have been heavily explored in frameworks such as BERT (Devlin et al., 2018), GPT-3 (OpenAI, 2020), and T0 (HuggingFace, 2021), to graph neural networks using common-sense knowledge graphs. More specifically, we want to emulate how these large natural language models were pre-trained on generally simpler tasks and then fine-tuned to a variety of downstream tasks. The extension of this to graphs is transferring the knowledge gained from a self-supervised task on a large knowledge graph into graph downstream tasks such as social network analysis. One significant form of self-supervision I have explored is knowledge graph embedding through link prediction using attention-based and relation type-based neighborhood aggregation schemes, employing frameworks including CompGCN (Vashishth et al., 2019) and Multi-hop Attention Graph Neural Networks (Wang et al., 2020) that connect link prediction with general graph tasks. The latter paper was under Dr. Jure Leskovec’s SNAP lab at Stanford, which I’d like to contribute to, having referenced them many times previously in my work. Provided the many facets of graph neural networks such as representation learning and zero-shot learning that I’ve researched with Professor Bach, I believe myself to be a great fit—I yearn to expand my previous experience into areas such as question answering for graphs (QA-GNN, 2021) and feature-imputation on graphs (GRAPE, 2020) under the SNAP lab.

Yet despite my experience in the vision and graph domains of machine learning, it would be naive to limit myself to these spaces—Stanford’s Computer Science Department includes world-class research in areas such as NLP and RL that I am interested in investigating as well. The work of Dr. Christopher Manning from Stanford has provided the foundation of today’s advancements in computational linguistics, and his recent research such as contrastive learning of medical images and text continues this trend, so I would be honored to be able to work with Stanford NLP. Furthermore, I am interested in the work by Dr. Emma Brunskill in Reinforcement Learning and the theoretical learning bounds there. In my graduate RL class with Dr. Michael Littman I investigated the notion of regret, missing out on potential rewards, which has been instrumental to algorithms for autonomous agents, and I’d like to further investigate works similar to Tighter Problem-Dependent Regret Bounds in RL by Professor Brunskill.

Artificial Intelligence today presents the potential to transform humanity for the better, and Stanford’s Computer Science Department stands at the forefront of this revolution. To attend the Master’s program at Stanford is to partake in captivating classes and to collaborate with renowned researchers on the world’s most pressing problems. But most of all, it is to join a wider community of machine learning practitioners crafting the next “industrial revolution” for a better future. I look forward to forging my own contribution towards tomorrow.