Cover Letter - Kumar Ayush

My interests are centered around Deep Learning and its applications in Computer Vision, Natural Language Processing and their intersection. In particular, I am interested in conglomerating deep learning representations with generative modelling, inference in generative models and fully end-to-end learning with neural network architectures. I am also drawn to tackling the major challenges in deep learning, which are unsupervised learning in the regime of small data, and simulation-based learning and its transferability to real world. I worked on a project titled "Generative Adversarial Learning for Reducing Manual Annotation in Semantic Segmentation on Large Scale Microscopy Images" as my undergraduate thesis, which has been accepted in Computer Vision for Microscopy Image Analysis (CVMI) workshop - CVPR 2017 and also received the best undergraduate thesis award in the graduating batch of the Dept. of Computer Science and Engineering. During the course of my thesis, I also became interested in visual recognition of human-object interactions in images. A key bottleneck to recognition of such interactions is the categories that have very few training examples. In such cases, semantic knowledge can significantly improve recognition. A rare category might piggyback on an co-occurring interaction that has more training data thus resulting in an improved recognition. I propose to use GANs in a semi-supervised setting for recognition of human-object interactions in images, which is an important subset of the larger problem of action recognition. I also intend to delve deeper and extend the work to videos as well. I am also looking for opportunities in AI assisted healthcare - developing machine learning algorithms and software systems to analyze new sensor data and infer disease progression and medication efficacy.

In my fifth semester, as part of the NLP term project under Dr. Pawan Goyal, I worked on the development of OCR++, an open-source framework designed for a variety of information extraction tasks from scholarly articles including metadata (title, author names, affiliation and e-mail), structure (section headings and body text, table and figure headings, URLs and footnotes) and bibliography (citation instances and references). This project won multiple accolades. It was adjudged as the Best Project by Flipkart from a pool of 30 projects. It has also been accepted for poster presentation at COLING 2016 and won the third prize on IBM Day 2016 (system demonstration competition) at IIT Kharagpur. OCR++ was also selected for the prestigious Gandhian Young Technological Innovation (GYTI) Award 2017 by SRISTI (Society for Research and Initiatives for Sustainable Technologies and Institutions). These awards are given every year during the Festival of Innovations (FOIN) at the Rashtrapati Bhawan (Office of the President of India) in the month of March. Later, in the same academic year, I was involved in guiding a team of sophomores for an inter-hostel event, Open-Soft, where the task was to detect graphs from scanned documents and produce the corresponding data in tabular form. These endeavours taught me many valuable lessons — the perennial struggle of getting code to work under time constraints, the importance of being skeptical of initial results, the many perils of not taking a systematic approach, and above all, the value of persistence when everything seems to fail.

In my sophomore year summer, I was awarded a fellowship by the Indian Academy of Sciences to work as a research intern under the guidance of Dr. R. Venkatesh

Babu at Indian Institute of Science, Bangalore. I was involved in developing a deep neural network for Predicting Human-Eye Fixations in Images. The approach we used was inspired from VGG-16 net upon which we added our own modifications like using inception style convolutional modules, to capture semantics at multiple scales, and incorporating a novel Location Biased Convolutional layer which learns to add center-bias to saliency maps in an image–dependent manner within the CNN framework instead of adding a constant center-bias to saliency predictions of all images. DeepFix, as we named it, was declared the winner of Saliency Prediction task in Large Scale Scene Understanding (LSUN) Challenge, 2016 which was organized by Princeton University in conjunction with CVPR 2016. One major lesson from this internship for me was on how to work in a research group, communicate my ideas to more experienced people and be open to their suggestions in order to avoid being stuck on some issue for long. It served as a great motivation for me to pursue graduate studies in future leading to career in research. DeepFIx has also been accepted for publication in the IEEE Transactions on Image Processing journal.

The following summer, I interned at Big Data Experience Labs, Adobe Research, where I worked on Augmented Reality for Enterprise. I worked as part of a close-knit team of researchers from different disciplines. This increased my confidence in communicating my research to those outside my field and improved my interpersonal skills. This was a huge learning curve which helped me appreciate the demands of accuracy and efficiency and gave me the confidence that I could work successfully in a company. The work resulted in three patents being filed in US and one full paper has been submitted to ISMAR 2017.

My foundation in maths, as well as my long-standing interest and aptitude for learning algorithms and computer vision gives me the conviction to pursue research in this challenging area. I've derived a lot of confidence from my exploits so far and hope to keep up my consistency with persistent efforts. Both my summer internships were major building blocks for my inclination towards research.

At MILA, due to the plethora of available opportunities, I believe that I will be able to gain higher exposure and with it, a deeper understanding of various problems in my fields of interest. I would specifically like to focus on machine learning, vision technologies and healthcare (clinical machine learning). Being a premier and esteemed organization, MILA attracts some of the best brains. I feel MILA is a wonderful place to be at, one that gives an individual a lot of scope for expansion. Being part of such an elite group would foster intellectual growth and enable me to work to my potential. It will be a privilege for me to spend a fruitful and rewarding time at MILA.