Cover Letter - Yash S. Bhalgat

M.S., Computer Science and Eng. | *University of Michigan, Ann Arbor*

Applying to the Uber AI Residency program

Why Uber AI Residency

I am deeply motivated to solve problems which will impact the real world and make a positive change. It is just awe-inspiring to see the efforts undertaken at Uber ATG to make self-driving cars a reality. I am eager to be a part of the ATG R&D Labs and work in the areas of Perception and Planning. In addition to autonomous vehicles, Uber AI Labs has now become a pioneer in cutting-edge machine learning research through its recent efforts and has set new heights for researchers all around the world. Overall, I believe that I have developed a strong grasp on the concepts and a skill-set in machine learning and deep learning. And the AI Residency program would provide me a platform to further nurture a strong research aptitude and combine it with my skill-set to become a part of the scientific feats that would eventually prove beneficial to the society.

My Background

My primary interest lies in the intersection of Computer Vision, Natural Language Understanding and Machine Learning. In a short period of ten years, we've taught machines to achieve superhuman capabilities in these areas. While advances in deep learning and powerful computation technology take all the limelight, I have identified three key problem areas that I want to address in the field of AI - (1) understanding the underlying semantics associated with visual data and its interpretation, (2) reducing the limitations posed by data and reliably integrating unlabeled data sources into training and (3) building smaller and interpretable models that can be deployed into the real world (especially in self-driving cars, so that passengers, insurance companies and law enforcement can trust them!).

I got into the Indian Institute of Technology (IIT), Bombay - with an All India Rank of 155 among 0.5 million students taking the IITJEE Advanced entrance examination. Having been an inquisitive student in high school, I entered undergrad with a strong background in Mathematics which helped me explore multiple domains in the first two years. My first decisive experience was when I joined the robotics team, the Mars Society of India, which focused on building a Mars prototype rover. I worked in the Vision and Navigation system and this was my first intro to robotics and planning. It was simply enthralling to see my work go into action. That is one of the reasons most of my research has been in Computer Vision, because you can immediately "see" the results of your algorithms. My first deep learning project was in a Computer Vision course where I implemented a Convolutional Neural Network from "scratch" - writing classes for linear (dense), convolutional, max pooling, batchnorm layers including the forward pass and the backprop functions/equations. This got me interested in deep learning and to build a solid background, I continued to amass knowledge through my courses in advanced Computer Vision, Natural Language Processing, Medical Imaging, Parallel Computing and several Machine Learning courses.

Other than the courses, I pursued several projects[1] and internships to expand and hone my skill-set. My expertise in deep learning got me an internship at IBM India Research Lab in my junior year, where I built a fast catalog search system for large fashion databases. I used an Autoencoder based network called a CorrNet, which is built on the idea of Common Representation Learning. A major challenge in this project was to choose an appropriate combination of cross-reconstruction, self-reconstruction and correlation loss functions in training, which was constrained by the diversity in the dataset. I used Transfer Learning to solve this by selecting the terms that gave the best representations. As an end product, I implemented an end-to-end pipeline in Theano which performed a query-search over the entire database in 2-3 miliseconds.

At IIT Bombay, I got the opportunity to work with Prof. Vikram Gadre for my undergraduate thesis. I worked on the problem of enhancement of latent fingerprints (speficially, raw imprints obtained from forensic documents). This work was in collaboration with the Department of Cyber Security, Maharashtra with an initiative to reform the fingerprint recognition systems. For this work, I used a Local Non-linear Total

[1] `https://github.com/yashbhalgat`

Variation model to extract the texture components from the images and then used a Scattering Wavelet Network (ScatNet) with synthetic Gabor filters for the enhancement. An extended version of this work on Iris classification has been submitted to the International Journal in Biometrics. I was awarded with the **Undergraduate Research Award** (URA 02) by IIT Bombay for this work. One of my other parallel work on ScatNets was through my internship with the Image and Signal Processing group at IFPEN, Paris on seismic sensor images, which was presented at ICASSP 2018 as a full paper[2].

Inspired by the recent feats of deep learning in Medical Vision, I undertook a course on Algorithms in Medical Image Processing. As a part of the course project, I implemented Cellular Neural Networks and U-Nets for segmentation of blood vessels from retinal images. I also started working on a project on segmentation of anatomical structures in chest radiographs. These projects made me realize that, the cost of annotations in medical image segmentation is very high in terms of logistics and time involved. This made me wonder if we could utilize weaker inexpensive forms of annotations for training. I came up with a method to enable suggestive annotation in a mixed-supervision setting by formulating a budget-based cost-minimization problem - via dense segmentations, bounding boxes, and landmarks. This method achieved state-of-the-art performance with significantly reduced cost of annotations. Our work has recently been accepted as a workshop paper[3] in **Medical Imaging meets NeurIPS 2018** (acceptance rate $\sim 23\%$). This has been a continued collaboration since my undergrad and we are currently working on formulating a joint optimization framework to include the cost-minimization LP into the base segmentation architecture.

The desire to free AI from the limitations posed by labeled data sources led me to another fascinating project during my internship this summer at IBM Almaden Research Center. I worked with the Watson Languages group, which at that time was facing a critical issue of label noise in their training data for the sentiment classification pipeline. To mitigate this issue, I implemented several methods - the Noise Adaptation Layer struck me the most. It was elegant how it optimized the EM likelihood function by adding an extra softmax layer to the network. But pure deep learning based approaches failed due to insufficient data. I eventually posed this as a labeled-unlabeled data exploitation problem, wherein I proposed a Teacher Student paradigm (not the same as Mean Teacher, but a bit similar to Curriculum learning) for efficiently importing pseudo-labeled samples with the help of adaptive thresholds. With extensive comparisons on Semeval 2016, I showed that the proposed model outperformed state-of-art in most cases. The success with this work was a result of frequent discussions with my mentors and peers. Working and interacting with my mentors on a daily basis was invaluable in terms of the lessons I learnt about persisting when faced with what seemed like a brick wall and also, seeing the same problem in different ways when stuck. My work is now a part of the Watson Sentiment pipeline and we are submitting this as a short paper to **NAACL-HLT 2019**.

Deep learning has become a very active, fast moving area of research where new ideas and advances are proposed almost everyday. Starting with a papers reading roadmap provided by one of my mentors, I have developed a habit of reading 1-2 papers daily and if possible, implement a few of these to cope up with the most important developments. I keep my ear to the Talking Machines and TWiML podcasts everyday to learn about the critical real world problems around the world, that can be solved with AI. I was also one of the main contributors to FOSSEE group's open source project to develop an Image Processing and Computer Vision toolbox for Scilab. I continue to actively seek open-source projects I can contribute to, most recently I have just started to explore OpenMined[4].

[2] `https://ieeexplore.ieee.org/abstract/document/8462088`
[3] `https://yashbhalgat.github.io/papers/smsnet_camera_ready.pdf`
[4] `https://openmined.org/`