# Statement of Purpose

Yash Sanjay Bhalgat

### Applying to the PhD program in Computer Science at MIT

My primary interest lies in the intersection of Computer Vision, Natural Language Understanding and Machine Learning. As a Computer Science graduate student at the University of Michigan, Ann Arbor, I have identified three key problem areas that I want to address in the field of Artificial Intelligence - (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 Medical Imaging, so that I can convince doctors to use 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][2] 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 (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 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[3].

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

[2] https://yashbhalgat.github.io/CV_YashBhalgat.pdf

[3] https://ieeexplore.ieee.org/abstract/document/8462088

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[4] 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**.

My gradual exposure to Machine Learning has made me aware of the tremendous application space it has. As I pursue my PhD, I not only want to delve deeper into my area of expertise but also to thoroughly understand the limitations of my field and build on what can be done. I was glad to know that **Jacob Andreas** will be joining MIT in fall 2019! His idea of using modular deep learning architectures that can dynamically adapt to solve different tasks is simply amazing. He is also targeting the problem of interpretability of AI models head on - the idea of using belief states to develop a translation model for the messages in a multi-agent environment struck me the most. I would like to be a part of this research to explore the next steps in the evolution of deep learning algorithms. On a similar note, I am also excited about **Prof. Aude Oliva**'s work on building saliency models and deciphering the black box of deep networks. I believe this is one of the most crucial problems in AI today and this is something I would be eager to pursue in my PhD. An incredible application of computer vision is in the area of Medical Imaging. **Prof. Polina Golland**'s work on statistical shape analysis and using shape-changes for predictive modeling of disease processes is awe-inspiring. I feel that my current research in image segmentation closely aligns with her interests and I would love to work in this direction under her guidance during my PhD.

Apart from research, I also find great joy in teaching. I am one of the top nominees (final results still pending) for the **Towner Prize for Outstanding GSIs** for my work as a Graduate Student Instructor previous semester. For a class of 139 students, I was in-charge of the lab and worked on the theory part of the course alongside a PhD student. I was highly commended by Prof. Matthew Smith on my work, especially, revamping the lab on Finite State Machines. At IIT Bombay, I worked as a TA for three semesters and I am currently a GSI for the course Computational Data Science with Prof. Raj Nadikuditi. I recently got an opportunity to deliver a special topics lecture on ConvNets in lieu of Prof. Raj. I have always been thankful to have had these opportunities to create impact through my teaching and would like to continue to do so.

[4] `https://yashbhalgat.github.io/papers/smsnet_camera_ready.pdf`