Personal Statement

Abdul Fatir Ansari, IIT Roorkee *Ph.D. Applicant* at NUS

With the amalgamation of computing with life sciences, various new doors of possibilities have opened up. Be it medical image analysis or developing ways to study and predict deadly diseases like cancer or sequence analysis to uncover the secrets of the genome, these fields have always captivated me and have inspired me to further investigate these areas by means of research.

Currently a senior year undergraduate in civil engineering at IIT Roorkee, I wish to pursue my studies as a doctoral student in the field of Medical Image Analysis & Bioinformatics. Although civil engineering is a great subject in itself, I realized that my heart was in advancing healthcare through technology. My inclination towards the application of computer science to real-life problems in biology motivated me to undertake various research projects and internships in these fields which imparted me with relevant experience and further stimulated me to pursue a Ph.D.

With a fair understanding of object oriented programming (in Java) from high school, I decided to develop an obstacle avoidance Android game in my free time during my freshman year. During the development period, I acquired a lot of skills ranging from core Android development to memory management and implementation of the laws of Physics in games. In the subsequent summer break, I read an article about fractals. Immediately fascinated by the beauty and complexity of fractals, I started tinkering with code to render fractals. During this phase, I learned the basics of how images are stored and pixels are manipulated. I started reading about image filtering techniques and also started blogging about fractals and basic image filtering. Little did I know that this would have such huge impact on my future work; or as I read about the Butterfly Effect in chaos theory while reading about fractals; that small incidents can lead to great outcomes.

In my sophomore year, I applied to Google Summer of Code (**GSoC**) for a project titled *Mobile Based Blood Analysis* which appealed the most to me because of the skill set it required (image analysis and Android development) and its real world application to life sciences. I realized the areas which needed improvement to enhance my proficiency in medical image analysis after my proposal was rejected in GSoC. Not losing heart, I approached **Dr. Roy** seeking a research project in the field of computer vision with applications in life sciences. I started working on the problem of segmentation and posture classification in the videos of Hammersmith Infant Neurological Examinations (HINE). I learned about and experimented with different classifiers such Support Vector Machine (SVM) and Hidden Markov Model (HMM) and eventually used an HMM-based sequential classifier which gave us the best results with an accuracy of more than 78%. We presented our findings at the International Conference on Computer Vision and Image Processing (**CVIP**), 2016 at IIT Roorkee. By working on this project, I obtained first-hand experience in research, of how existing literature is studied, how novel techniques are developed, and how a research publication is written and published (I have co-authored 2 publications). Besides the improvement of my knowledge in image analysis and machine learning, working on a complex dataset in this project taught me how to be resilient in the face of difficulty.

Since my freshman year, I have dedicated most of my time to learning about computer science, taking online courses and working on projects related to computer science. My research, internships, and projects involving image processing and machine learning and their application to biological data are a reflection of how well I will fare in a research setting in this field. I continued working on other problems in the automation of HINE, such as full-length sequence classification and event segmentation during my junior year with **Dr. Roy** and **Dr. Dogra**. After trying various methods of feature extraction such as Histogram of Oriented Gradients (HOG) and Convolutional Neural Networks (CNN), I developed a Scale Invarient Feature Transform (SIFT) bag-of-words based approach of feature extraction and the accuracy improved to 84% on a larger dataset. We have submitted our results in the Journal of Machine Vision & Applications and our article is currently under review. Through all the work, I realized how much I loved what I was doing, and wanted to contribute to Bioinformatics and medical image analysis.

With my knowledge of medical image analysis and machine learning significantly improved I applied to the 2016's edition of GSoC. Last year's project titled **Mobile Based Blood Analysis** was still open as the algorithm which was developed previously did not give good results. A task related to the project was given to test the applicants' knowledge on which my error rate averaged 7%, far better than any other applicant, while the acceptable limit was 20%. My proposal was subsequently selected and I worked for a computational biology research group at the University of Nebraska-Lincoln under the supervision of **Dr. Helikar**. Specifically, I designed a mobile application to enable a researcher to take the picture of a card with a blood sample and several reference points and using pixel intensities of the samples, the application calculates a standard curve, which is then used to estimate the concentration of a given molecule in the blood sample. After trying techniques such as color based segmentation and SIFT pattern matching, which failed on poorly lit and blurry images, I developed a technique based on adaptive thresholding algorithm with some heuristics which performed exceptionally well even on bad quality images. I also worked with **Dr. Helikar** on the full-stack development of an Android health application, which lets users track various health indicators such cholesterol, glucose, calories etc.

In the summer of 2016, I also did an internship at **MapMyIndia** in the image processing and machine learning team where I developed an algorithm for automatic license plate blurring for privacy protection in street view images. My algorithm was up to 5 times faster than the algorithm previously used and achieved a recall of more than 90%. It was subsequently used to process a dataset of more than a million 30-megapixel images. Having a strong interest in reading and writing poetry, I also wrote a Markov chain based and another RNN based poetry generator.

The research work carried out by **Prof. Wing-Kin Sung** interests me the most, especially the work in DNA Sequence Analysis. I want to apply my data analysis and machine learning skills to sequence analysis. With the relevant background in machine learning (HMM, neural networks, SVM, etc.), I think I will fit perfectly in Prof. Wing-Kin Sung's research group. **Prof. Wynne Hsu**'s research projects *Retina Cloud* and *GeoVisualization of Spatio-Temporal Disease Spread* also interest me. I intend to apply my knowledge and experience in computer vision and medical image analysis to develop techniques for health assessment and disease recognition if given a chance to work with her.

In terms of career, I see myself as an academic and a researcher in the long run. I want to develop algorithms and techniques which will lead to the advancement of biological research and ultimately to the betterment of human lives. The exuberance of discovery that comes after rigorous hours of research, and realizing that my work is advancing healthcare for people is what makes me persevere through my failures. The doctoral program at NUS being one of the finest with world-renowned faculty will provide me an excellent platform for the realization of my goal. I will be able to continue my research with reinforced vigor and renewed outlook in the advancement of healthcare if given a chance to work on my Ph.D. in the intellectually stimulating environment of NUS.