Statement of Purpose Hmrishav Bandyopadhyay

From cultivating the patience to read complex research papers in my fresher years to publishing my first-author work internationally as a sophomore, I find my journey of undergraduate research both taxing and fulfilling. With a strong inclination towards research in Computer Vision developed by more than 2 years of laboratory work at Jadavpur University and one year of research at CMU, stepping into a Ph.D. program is a crucial step for me to establish myself as a researcher where I can work on the applicability of vision algorithms in the real world and build robust systems which can function in complex environments without further supervision. I find grad school the perfect place where I can explore the depths of this domain, particularly topics like Incremental Learning, Meta-Learning, and Domain Generalization which align heavily with my research interests and my past publications.

Currently a senior undergraduate at Jadavpur University at the Department of Electronics and Telecommunication Engineering, I was introduced to signal processing in my sophomore year through coursework on Signals and Systems which further pushed me to know more about image processing and computer vision as a discipline of research. That year, I developed an autonomous lane traversal robot equipped with simplistic lane segmentation by edge detection to maintain motion on a specified path as a direct practical outcome of my coursework. Most of my years at Jadavpur University have been spent under the supervision of **Dr. M Nasipuri** and **Dr. N Das**, where I worked on multiple publications including my first first-author publication on the analysis of cervical cancer with the help of K-Nearest Neighbours for extraction of shape features from images.

Tinkering with the analysis of document images, I worked on **RectiNet** and **RectiNet-v2**, deep learning based algorithms that deal with reducing warps in document images arising due to bad image capturing practices. While working on implementing state-of-the-art algorithms in document dewarping, I realized that none of the methods provide channel-specific attention when they generate a backward mapping containing x and y channels for dewarping images. To deal with this, I formulated channel-specific attention, with a bifurcated U-Net while focussing on line-level detail with a gated module in RectiNet. RectiNet became my first major work published at the International Conference of Pattern Recognition (**ICPR**) as a conference paper (2021), beating previous state-of-the-art methods. A few months later, I released RectiNet-v2, currently under review at **Pattern Recognition Letters**. Amongst all other things, working on long-term projects like RectiNet and RectiNet-v2 has taught me the importance of patience in research and the attitude of holding on to my work.

Recommended by supervisors at my college, I started working under **Dr. B Banerjeee** at **IIT Bombay** on several topics ranging from medical imaging to domain adaptation in Computer Vision. With a specific focus on Open-Set domain adaptation, I worked on **Source Free Incremental Open Set Domain Adaptation (SFI-OSDA)** which introduces a novel problem statement to help make vision systems more robust and adaptable to real-world surroundings. At its core, the idea involves the concept of Open Set Domain Adaptation clubbed with Incremental Learning (I-OSDA) allowing us to build vision systems that can adapt to new domains without Catastrophic Forgetting. Along with the problem statement, I also propose an algorithm that fits these constraints and beats state-of-the-art methods by a wide margin. The work on SFI-OSDA spanned more than a year with a huge number of experiments that helped me recognize the importance of being methodical with research and improved my organizational skills.

Introduced to medical imaging at both IIT-B and my university, I was excited to explore further possibilities computer vision offered in this domain. Consequently, I joined **CMU** as a research intern at the **Xu Lab of Computational Biology** where I worked on the unsupervised analysis of Cryo-Et structures as a solution to the lack of annotated data. Inspired by previous methods on the improvement in model generalizability by hyperparameter randomization, I developed a trainable randomization framework that randomizes the input subtomograms before they are fed to the model with the help of distortion and noise modules. When the performance on the target dataset was not up to the mark after randomization, I added a multi adversarial domain adaptation algorithm to help the model adapt to real data in a completely unsupervised manner. While being trained on unannotated data only, the algorithm **Cryo-shift** now accepted at **Bioinformatics**, one of the most prestigious journals in Computational Biology outperforms even few-shot algorithms by as much as 7 percent on the target dataset. Working on Cryo-shift kept me the busiest; being part of a big laboratory with limited compute power, I had to manage a huge number of experiments while making sure I did not overuse the GPUs. While this limited the number of experiments I could perform, it helped me segregate the exact experiments I needed, increasing my resourcefulness.

Around the end of my sophomore year, I was funded under the **Mitacs Globalink** research program to pursue short-term research at **Ontario Tech** under **Dr. M Ebrahimi**. Here, I worked on domains like Cross-modal Image to Image translation for day and night Images and inpainting for resolution enhancement. While this work did not culminate in a research paper, I cherish the opportunity I had to explore new domains and interact with talented doctoral students at the institute, enabling me to work and bond with people from different nationalities and cultures which has helped me learn a lot from their ideologies. This has further reinforced my desire to complete my doctorate from a globally acclaimed university that can offer me a truly diverse peer learning experience and help shape my worldviews.

Through my research endeavors at various labs and my encounters with challenges in Computer Vision, I have found a deep-seated love for the subject. I find my passion in pushing the boundaries of computer vision by working in research and academia. As a researcher, I want to explore subtopics of computer vision like Domain Adaptation and Incremental Learning that can help me build large-scale and robust vision systems that perform well in a variety of environments and have a larger success rate in real-world scenarios. With these aspirations, I believe pursuing a Ph.D. program is a natural transition at this point in my career.

My application to the Department of Computer Science at Boston University is driven by my interest in the works of professors like Dr. Kate Saenko, Dr. Bryan Plummer, Dr. Margrit Betke, and Dr. Venkatesh Saligrama. Dr. Saenko's work on Visda helps make my work on SFI-OSDA possible, the paper making use of the DomainNet dataset primarily to demonstrate its efficacy. I am highly interested in her recent works in domain adaptation and generalization and would love to explore these topics in depth under her supervision. Dr. Saligrama's research focussing on learning with limited supervision with meta-learning and few-shot learning aligns is particularly interesting to me, these domains heavily aligning with my research interests. I am also broadly interested in Dr. Plummer's work on vision-language models and Dr. Betke's work involving the extraction of text from scanned documents, which I would like to explore further under their guidance. Overall, I find the research at Boston University's Computer Vision laboratories to be heavily aligned with my research goals and aspirations and would love to pursue a Ph.D. at BU.