Research Interests and Background: My research interests lie at the intersection of Machine Learning (ML), Computer Vision (CV), and Human Computer Interaction (HCI) with healthcare applications. I am particularly interested in developing robust ML algorithms to diagnose and monitor diseases effectively. To this end, I am excited about mobile healthcare *i.e.* utilizing the data from mobile sensors and using on-device machine learning to democratize healthcare. My prior research involves using machine learning, image processing, and optical physics to tackle key challenges in computer vision and healthcare, including (1) image compositing, (2) self-supervised representation learning, (3) few-shot image segmentation, (4) anomaly detection, and (5) AI-based medical diagnosis. These have resulted in publications at top conferences like **IJCAI, WACV, SIGIR, EMBC**, and **IMWUT/Ubicomp**. I wish to earn a Ph.D. to tackle core machine learning and computer vision problems focusing on healthcare and medical applications.

Applying ML, Vision, HCI to Healthcare: I am currently working at **Microsoft Research** as a **Pre-doctoral Research Fellow** (residency program) on developing AI-based low-cost diagnostic solutions. Over the past year, I have worked with **Dr. Mohit Jain** and **Dr. Nipun Kwatra** on building a smartphone-based corneal topographer to diagnose keratoconus – a severe eye disease that can lead to blindness if not treated in time. Such a device can be highly useful in rural and remote locations where modern medical facilities are inaccessible.

A commercial corneal topographer is an expensive (∼$10,000) device with proprietary hardware and software. It projects a concentric ring pattern on the eye and reconstructs the corneal surface using ray-optics and 3D geometry principles. The generated topography maps highlight any deformities in the cornea. I directed this project from end-to-end, and after months of consistent efforts we built a working prototype. Building a low-cost device with the same fidelity as a commercial topographer required several key design and technical innovations. First, I designed a 3D printed conical attachment to project a concentric ring pattern on the eye. Second, I developed an AI-assisted smartphone app to ensure proper alignment and image quality with a hand-held device. Third, to process the captured images I implemented an analysis pipeline that involved image processing and a 3D reconstruction algorithm to generate the corneal topography. And to automatically detect keratoconus, we trained a CNN-based classifier on the output maps. Further, due to lack of hardware sensors on the smartphone I developed an approximate vision-based algorithm to estimate the working distance (between camera and eye), which is required for reconstruction. We conducted an extensive clinical evaluation (at Sankara Eye Hospital) of our system and found that it performed at par with the gold-standard medical device 'Optikon Keratron'. This project led to a publication at **ACM IMWUT/Ubicomp 2021**. The proposed attachment costs **$33** (**303x** less than Keratron) and weighs just **140 grams**, making it ideal for use by community health workers. We are now working towards deploying the device at Sankara clinics across India for mass screening of keratoconus, through which we envision early interventions and timely treatment.

Automatic detection of diseases can aid in early interventions and prevent delayed treatment. In this context, I worked on the task of **anomaly detection** from auscultation signals. I developed a "ResNet34" based lung sound classification model to identify abnormalities like crackling and wheezing. Interestingly, a simple CNN model worked best for this task compared to Hybrid CNN-LSTM networks and Transformer networks, which had been the focus of prior work. We proposed and combined 'novel' data augmentations and transfer learning techniques to achieve top scores on a small-sized dataset. The work was accepted at **EMBC 2021**. This has motivated me to explore techniques like transfer learning, domain adaptation and self-supervised learning in greater depth.

My experience at MSR has broadened my perspective on the transformative impact technology can have on critical real-world applications like healthcare and medical diagnosis. I also noticed how introducing domain knowledge could drastically improve the performance of ML systems. E.g., based on medical literature and discussions with doctors, we found that keratoconus is usually bilateral (affects both eyes of a patient). Using this knowledge, we forced the model to classify both eyes as keratoconus if only one eye was classified as having keratoconus. This addition helped increase the sensitivity of our diagnostic system by 7% (from 87% to 94%).

Computer Vision and Machine Learning: My journey in research began in my sophomore year, when I started working with **Prof. P. J. Narayanan** at the Center for Visual Information Technology (CVIT) at IIIT Hyderabad. Initially, I explored problems in the space of computational photography and structure-from-motion. For my first research project, I proposed a data-driven approach to perform color-consistent sky replacement in outdoor images. I received the **Dean's Research Award** for this work, and it led to a publication at **MMM 2018**.

While working on these initial projects, I realised the utility of robust image features for visual tasks like image recognition, retrieval, style transfer, etc. This got me interested in **representation learning** to model the style and content in images. Style recognition and similarity are important measures for understanding abstract concepts like art, fashion, and design. However, the definition of style is contextual and vague. Deep neural networks (DNNs) perform well on image understanding tasks but require densely annotated large-scale datasets. Obtaining such annotations is expensive (in terms of time and money) and often infeasible. I developed a DNN based framework for learning a neural embedding that captures the 'look and feel' of an image. Unlike previous supervised learning methods, the proposed framework is **unsupervised** – it does not use categorical labels but uses a gram matrix feature-based clustering to get proxy labels for forming triplets of anchor, similar, and dissimilar images. These triplets are used to train a siamese network with a triplet loss to learn an embedding useful for style-based search and retrieval. Despite being 16 times more compact than traditional representations, the embeddings achieved state-of-the-art results for style-based image retrieval and recognition on six datasets with different notions of style. Thus, the proposed method provides a general framework to learn style representations and can also be combined with other proxy measures for style-aware grouping. This work was published at **IEEE WACV 2020**.

My research work at CVIT culminated in the form of my Master's thesis. Working at CVIT introduced me to the fascinating world of research, where I developed a solid foundation in computer vision and machine learning, and skills to collaborate effectively and conduct principled research. The brainstorming sessions with my advisor and senior graduate students taught me to communicate well, ask questions, and develop methods for answering them. I left CVIT feeling optimistic about research and wished to explore similar problems further.

After completing my thesis, I joined the **Media and Data Science Research** lab at **Adobe** as an intern. Here I worked on building an efficient solution to detect and segment unseen (not seen during training) object classes in images. Image segmentation is crucial for visual understanding and is a common use case for software like Adobe Experience Manager – a digital asset management solution for brands like Audi, Nike, etc., to maintain their product images. However, supervised methods fall short because of limited availability of good quality annotated data. An automated solution to segment 'novel' object classes with minimal supervision can save a lot of time and effort. During my internship, I carried out an extensive literature survey of the domain and narrowed our focus to the **Few-Shot Image Segmentation** problem, *i.e.,* learning to segment a query image given only a few samples (1-5) as support for each unseen object class. I proposed a novel method that improves the similarity propagation between the support and query image features to get accurate segmentation. The method introduces two key components which provide a non-trivial improvement ($> 6\%$ mean-IoU) over prior methods. First, we jointly predict the support mask and the query mask to enforce self-similarity. Secondly, we introduce a novel foreground-background attentive fusion mechanism to utilize similarities in the background regions of the query and support images. The method achieved state-of-the-art performance on the one-shot and five-shot segmentation benchmarks. The work was published at **IJCAI 2020**, and a **US patent** has been filed for the same.

While working on this project, I observed that class-conditional similarity matching could only match pixels with a similar class mix between the query and the support images. Hence, I wish to explore introducing inductive biases via representations that capture the geometric structure and physics of the real-world to improve visual understanding. My internship experience enhanced my skills as an engineer and researcher. I learned to write well-documented and optimized code, and build applications to integrate my research project as a software tool.

Ph.D. and Future Goals: After completing my Ph.D., I wish to continue in academia. I have found learning about new problems, conducting research, collaborating with fellow researchers, and teaching – fulfilling and enriching. A Ph.D. would help me build a strong foundation and depth to achieve this goal. This decision is informed by my research background and positive experience working as a teaching assistant for graduate and undergraduate level courses. My experience as a TA improved my understanding of multiple core subjects, and it was gratifying to see students indulge in critical thinking and ask challenging questions. During my Ph.D. I wish to continue working on building efficient and trustworthy machine learning models to aid in healthcare. Modern healthcare today is expensive, requires sophisticated medical devices, and is inaccessible to a large fraction of the population. Further, trained doctors are required to interpret medical reports which can suffer from subjectivity. The synergy between human and artificial intelligence can help build more accurate, reliable and affordable diagnostic systems. Building end-to-end AI systems that can be deployed in the real-world require several components: efficient data-collection pipelines, robust machine learning algorithms, interpretable outputs, and exhaustive testing protocols to guarantee reliability. I wish to contribute towards building such systems through my research.

At UW, I am interested in working with Prof. Shwetak Patel on his research efforts on building and deploying mobile health systems. His prior works like Biliscreen and Pupilscreen are similar to my work at MSR that was published at IMWUT. Similar in spirit, Prof. Shyam Gollakota's work on using smart speakers to detect cardiac arrests and mobile phones to identify sleep apnea events is fascinating to me. I am also interested in working with Prof. Ranjay Krishna. My interests align well with his research on combining vision models and human-computer interaction to improve ML-based systems, explainability and interpretability. Additionally, I am excited by the work done in Prof. Sun-In Lee's lab on developing robust, precise, and explainable machine learning models, which are essential for critical medical applications. My experience in machine learning and computer vision research to build and deploy healthcare systems will add value to such projects, and I would love to continue my research with her. UW has a strong and diverse faculty body, close collaborations across research groups, and hence, it would be ideal for conducting fundamental and interdisciplinary research. I believe my strong background in research, software engineering experience, and alignment of interests with faculty makes me a good fit for UW's graduate program.