

## EDUCATION

<b>Cornell University</b>	<i>Ithaca, NY</i>	<b>Aug 2022 – Jun 2024</b> (ongoing)
<ul style="list-style-type: none"><li>• <b>Master of Science in Computer Science (Thesis Track)</b></li><li>• <b>Graduate courses:</b> Topics in Computation Sustainability, Advanced Topics in ML, Graphics</li></ul>		
<b>Indian Institute of Technology Bombay</b>	<i>Mumbai, India</i>	<b>Aug 2017 – May 2021</b>
<ul style="list-style-type: none"><li>• Bachelor of Technology in Computer Science &amp; Engineering w/ Honors, Minor in Artificial Intelligence &amp; Data Science</li><li>• GPA: 9.68 / 10, Honors GPA: 10 / 10, Minor GPA: 9.4 / 10</li><li>• <b>Courses:</b> Theoretical &amp; Advanced ML, ASR, Concurrent Programming, Functional Programming</li></ul>		

## WORK EXPERIENCE

<b>Software Engineer, Samsung Electronics</b>	<i>Suwon, South Korea</i>	<b>Sep 2021 – Aug 2022</b>
<ul style="list-style-type: none"><li>• Part of the <b>Physical Uplink Shared CHannel</b> team of Samsung's 5G vRAN Development Lab, working on high performance low latency virtualized physical layer for next-gen wireless communication, working on core-cycle optimization</li><li>• Utilized Intel® Intrinsic (<b>AVX-512</b>) for efficient parallel processing of data</li><li>• Reduced bottlenecks in uplink signal processing pipeline to achieve upto 20% speedup</li></ul>		
<b>Software Engineer Intern, Samsung Electronics</b>	<i>remote from India</i>	<b>Jun 2020 – July 2020</b>
<ul style="list-style-type: none"><li>• Built an automated network-load testing framework using <b>Locust</b> hosted on kubernetes for distributed server simulation</li><li>• Tool used to evaluate performance of Samsung's in-production load balancing services</li></ul>		
<b>Summer Research Intern, TU Braunschweig</b>	<i>Braunschweig, Germany</i>	<b>May 2019 – July 2019</b>
<ul style="list-style-type: none"><li>• Created <b>WeLineation</b> - an application using an <b>Expectation Maximization</b> algorithm called <b>STAPLE</b> to generate ground-truth like sclera segmentations from crowdsourced data.</li></ul>		

## PUBLICATIONS

- **Improving low resource code-switched ASR using augmented code-switched TTS** - Y. Sharma, B. Abraham, K. Taneja, P. Jyothi [INTERSPEECH 2020]
- **WeLineation: crowdsourcing delineations for reliable ground truth estimation** - S. Goel<sup>1</sup>, Y. Sharma<sup>1</sup>, M.L. Jauer, T.M. Deserno [SPIE Medical Imaging 2020]
- **STAPLE performance assessed on crowdsourced sclera segmentations** - M.L. Jauer, S. Goel<sup>1</sup>, Y. Sharma<sup>1</sup>, T.M. Deserno, M. Gijs, T.T. Berendshot, C.J. Bertens, R.M. Nuijts [SPIE Medical Imaging 2020]

## TEACHING ASSISTANTSHIPS

<b>CS3410: Computer System Organization &amp; Programming</b>	<i>Cornell University, Ithaca</i>	<b>Fall 2022</b>
Involves grading exams, staff meetings, holding office hours and leading lab discussions		
<b>CS251: Software System Lab</b>	<i>IIT Bombay, India</i>	<b>Fall 2019, Fall 2020</b>
Orchestrating and preparing assignments for the lab course of "SSL" for CS sophomore students. Awarded <b>best TA</b> in 2020.		
<b>MA105: Calculus</b>	<i>IIT Bombay, India</i>	<b>Fall 2018</b>
Took weekly discussions of 50 freshmen students, graded exam papers and volunteered to teach beyond class hours		

## RESEARCH EXPERIENCE

<b>Undergraduate Thesis and R&amp;D Project</b>	<i>IIT Bombay &amp; Microsoft</i>	<b>Dec 2019 – Jun 2020 &amp; Aug 2020 – Jun 2021</b>
<b>Improving code-switched Automatic Speech Recognition<sup>2</sup></b>		
Focused on improving performance of end-to-end ASR models on Gujarati-English speech by conditioning transformer on language of the text. A Temporal Loss is used to train language specific parameters and add explainability		
<b>Improving Low Resource Code-switched ASR using Augmented Code-switched TTS<sup>2</sup></b>		
Used E2E Automatic Speech Recognition models trained on Hindi and English monolingual data and code-switched Text to Speech (TTS) to improve performance in low-resource settings. Ideated a new loss function to target underlying distributions of languages in the data. Used augmentation and encoder freezing to avoid over-fitting on synthetic artefacts		

## LANGUAGES AND SOFTWARES

C/C++, python, bash, Haskell, Javascript,  $\LaTeX$ , SQL, PyTorch, TensorFlow, AVX, Git, Perforce, Jira, Linux, Docker, MATLAB

## RESEARCH COURSE PROJECTS

<b>Few-shot action recognition on egocentric data; Prof. Kilian Weinberger;</b>	<b>Fall 2022 (ongoing)</b>	<b>Cornell University</b>
<b>Learning systems for cocktail party problem on bird calls; Prof. Carla Gomes;</b>	<b>Fall 2022 (ongoing)</b>	<b>Cornell University</b>
<b>Low Resource Morphological Inflection; Evolutionary RL on maze solving; VQA with dynamic neuralnet</b>		<b>IIT Bombay</b>

<sup>1</sup>Equal contribution

<sup>2</sup>Work done as part of collaboration between **Microsoft India Development Center** and **Indian Institute of Technology Bombay**