NJIT
New Jersey Institute of Technology

SUMMER RESEARCH PROGRAMS

URI SUMMER RESEARCH & INNOVATION SYMPOSIUM

SUMMER 2025

RONALD E. MCNAIR POSTBACCALAUREATE ACHIEVEMENT PROGRAM

NSF - REU OPTICS AND PHOTONICS: TECHNOLOGIES, SYSTEMS, AND DEVICES

HERITAGE INSTITUTE OF TECHNOLOGY

JULY 23 – 24, 2025
CAMPUS CENTER
• BALLROOM
A & B

“Research is formalized curiosity. It is poking and prying with a purpose. It is a seeking that he who wishes may know the cosmic secrets of the world and they that dwell therein.”

-Zora Neale Hurston

Brochure Creation & Publication
Ronald E. McNair Postbaccalaureate Achievement Program
Designer/Editor
Jeremy De La Rosa
Yamir Perez

Acknowledgements

On behalf of all the participating Summer Research Programs, McNair, Heritage Institute of Technology Research Programs and NSF REU on Optics and Photonics: Technologies, Systems, and Devices we are extremely appreciative of your support and efforts in making this summer program successful. We would like to give special thanks to the following individuals: Marieta Chemishanova, Vaughn C. Rogers, Denis Dyshko, Luis A. Guillen, Mike Dabrowski, Sanjeannetta Worley, Archana Srivastava, Lisa Easton, John Gruppo, Sandra Masibayi, Monica O'donnell, Lea Ronchi, Nicole Bosca, Jessica Simon, Djadji Sylla-Samassa, Soham Sen, Subhodeep Ghosh, Stephen Eck, Nellone Reid, Grisele Gonzalez-Ledezma, Johanna Deane, Shivon Boodhoo, Darnell Simon, Jose Rodriguez, Sumbel Yaqoob, Scott Sire, Vickram Ramoutar, Keren Johnson, Brain Hart, Zoe Mooneyhan and Michelle Kudelka . In addition, we are deeply indebted to the faculty mentors, who have such an important role in our students 'academic, research and social development.

Lastly, we must acknowledge our gratitude to the students for their hard work and perseverance. We encourage you to continue on your path of scholarly inquiry. As you have accomplished a great deal in such a short time with us, we know that you will be successful in your future academic and research endeavors.

NJIT
New Jersey's Science & Technology University

New Jersey Institute of Technology
UNIVERSITY HEIGHTS, NEWARK, NJ 07102-1982
973.596.5590
McNair@njit.edu

Ronald E. McNair
Postbaccalaureate Achievement Program

July 24, 2025

Welcome to New Jersey Institute of Technology's 2025 International URI Undergraduate Summer Research and Innovation Symposium. It is an honor and privilege to be part of this year's Symposium. Ten participants of Ronald E. McNair Postbaccalaureate Achievement Program started their summer research program from May 20, 2025. Five students from Heritage Institute of Technology (HIT), India started their research from June 16, 2025. Ten students of 2025 NSF REU Site: Optics and Photonics: Technologies, Systems, and Devices also participated in this year's summer research program. The students worked very hard to carry out their research projects with literature review, data analysis and simulation in addition to lab experiments.

The successful student exchange program during summer for research between NJIT and HIT has entered the 18th year this year. The purpose of this exchange is to promote international understanding, scholarly collaboration, cultural interaction, and friendship by supporting educational professional and cultural activities among faculty and students of the two institutions.

The NSF REU Site on Optics and Photonics: Technologies, Systems, and Devices was renewed for 2025 after a one-year gap. Students from all over the United States are participating in this program. All the 25 students participated in research, attend many seminars and workshops throughout the summer program.

This year's success comes because of Ms. Zara Williams, Assistant Director, Prof. Ashish Borgaonkar, the Faculty Coordinator of the Ronald E. McNair Program and Prof. Abdallah Khreishah, Director of NSF REU Site for their efforts in coordinating the overall program. Staff members of Research Office, Admissions, Office of Global Initiative, Budget Office, Housing, Facility Systems, Photo ID & Parking Services, and ECE Department contributed significantly to the success of the Symposium. Efforts of Mr. Jeremy De La Rosa and Mr. Yamir Perez of McNair office is recognized for the valuable input in producing this program's brochure. In addition, we have an excellent group of experts presented seminars and workshops to the students' success.

The students in the Ronald E. McNair Program, HIT Program and NSF REU Program have the opportunity for presentation of their research accomplishments that was completed under the supervision of dedicated NJIT faculty. Without the time and effort of NJIT faculty and graduate student mentors the outstanding achievement of the students would not have been possible. We are extremely proud of the research efforts of all the students, the quality of the research presentations and grateful for the support of the NJIT administration, faculty, and staff in contributing to the success of the symposium.

Sincerely,
Durgamahab Misra, PhD

Symposium Co-Chair, McNair Program Director, HIT Program Coordinator
Professor and Chair, Department of Electrical and Computer Engineering

Ronald E McNair Postbaccalaureate Achievement Program


Jada Byfield


Juana Perez


Mauricio Huapaya


Gisselle Ambrosio


Camila Martinez Castillo


Chadley Gede


Naketa Williams


Angelo Bustamante


Andrea Pardo


Mario Urla


Adwaita Laha


Suhrid Behari Paul


Piyush Jain


Ankit Kumar


Arnesh Banerjee


Seth Weaver


Wonki Chae


Simone Nelson


Sachi Rele


Matthew Oliveria


Hala Kanaan


Emma Lim


Kaylin Koen


Assaf Izhar


Abdulrahman Aljoudi

Early-stage Endometrial Uterine Cancer Detection using an Electrochemical Sensor that utilizes a Shear-Enhanced, flow through, nanoporous, Capacitive Electrode

Gisselle Ambrosio, Advisor: Nellone Reid, Mentor: Niranjan Haridas Menon
Department of Chemistry and Environmental Science
Department of Chemical and Materials Engineering
New Jersey Institute of Technology

Every woman deserves a fighting chance to battle endometrial uterine cancer. However, far too many women from underrepresented communities are being diagnosed at a later stage, where survival outcomes are significantly lower. Endometrial uterine cancer, often treatable if found at an early stage, continues to illustrate unusually high mortality rates in African American women due to aggressive tumor biology and delayed diagnosis. Strikingly, p53 gene abnormalities are linked to the more aggressive type of endometrial cancer, and is more commonly found in African American women. Machine learning techniques are shown to be excellent at diagnosing cancer through the analysis of data and images. An Electrochemical Sensor that utilizes a Shear-Enhanced, flow-through, nanoporous, Capacitive Electrode (ESSENCE) is proposed to integrate machine learning and enable high sensitivity and rapid early detection of endometrial uterine cancer. The proposed microfluidic platform was designed to detect cancer biomarkers such as P53 and PTEN proteins by functionalizing the capture material with the relevant antibodies. Electrochemical Impedance Spectroscopy (EIS) was used to measure the biomarker binding and find the concentration of the target analyte (cancer biomarkers) in the sample of interest. Machine learning was then applied to analyze signal patterns and enhance detection accuracy. The combination of microfluidics and machine learning has the potential to detect endometrial uterine cancer early and provide life-saving diagnostics to women who have previously been overlooked.

Beyond Usability: How Aesthetic and Functional Design Influence User Preference in Web Interfaces

Angelo Bustamante, Advisor: Dr. Salem Daher
Department of Informatics, New Jersey Institute of Technology

While usability plays a crucial role in effective web design, the relationship between aesthetic appeal and functional efficiency in shaping user preferences remains unexplored. This study will investigate how visual aesthetics and functional design elements collectively influence user engagement and satisfaction in web interfaces. The significance of this research lies in its potential to shift modern design patterns, ensuring interfaces are not only usable but also visually compelling, enhancing user retention and interaction. The study employs a mixed methods approach, combining quantitative user testing with qualitative feedback. Participants interact with different web interface prototypes, one varying in aesthetic and the other functional designs. Metrics such as task completion time, number of mouse clicks, and self reported satisfaction are analyzed to assess performance and preference. Anticipated outcomes include identifying design configurations that optimize both aesthetic appeal and usability, revealing user trade offs between visual attractiveness and functionality. Past findings suggest that while highly aesthetic designs initially attract users, functional efficiency significantly impacts long-term preference. Future work will expand the sample size and incorporate tasks to find throughout each interface, and calculate the amount of time it takes them in order to further analyze visual attention patterns. This research contributes to an understanding of web design, informing best practices for developers and designers seeking to balance form and function.

Evaluating Usability and Accuracy of Different Methods to Align and Place Digital 3D Wounds

Jada Byfield; Advisor: Dr. Salam Daher
Department of Informatics
New Jersey Institute of Technology

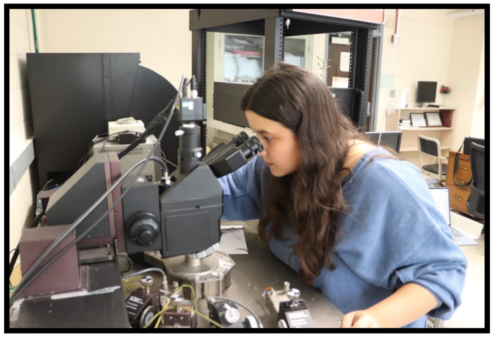
With 10.5 million Medicare beneficiaries affected by chronic wounds, there is an ongoing demand for more precise and consistent tools for wound measurement. In current clinical practice, the most common methods of wound measurement are 2D imaging and the ruler and Q-tip; these methods have limitations due to visual bias and low inter-rater reliability. By incorporating 3D imaging wound measurement software with a manual alignment feature, clinicians can enhance accuracy by reducing visual estimation errors and improving consistency across different users. Accurate wound tracking is critical for patient care, as clinicians rely on consistent measurements to determine whether a wound is healing properly. Improving the reliability of these assessments directly affects patient outcomes and reduces the risk of complications. Addressing these problems can improve the accuracy and efficiency of wound assessment, which is especially important for patients with diabetic ulcers, pressure ulcers, and other chronic wounds or related conditions. 3D wound measurement software provides clinicians with greater control during the wound placement process. This functionality allows users to adjust measurements directly, helping to reduce visual bias and improve consistency in wound tracking. This research aims to examine the usability and accuracy of a wound placement feature in a 3D wound measurement system. This study aims to evaluate how effectively this interface supports ease of use, user control, and overall satisfaction. The study will focus on a small pilot of 10 college students from non-technical backgrounds to understand how users interact with the system. Participants will observe or interact with the software, then complete a structured usability survey to provide feedback on clarity, control, and interface design. Both quantitative data and qualitative insights will be collected to assess how intuitive and flexible the system feels to users. It is expected that most participants will find the interface user-friendly, with the manual adjustment feature offering valuable control and flexibility. This research contributes to the improvement of user-centered medical simulation tools by promoting accessibility and usability, ensuring the software is usable for all users, regardless of their technical experience.

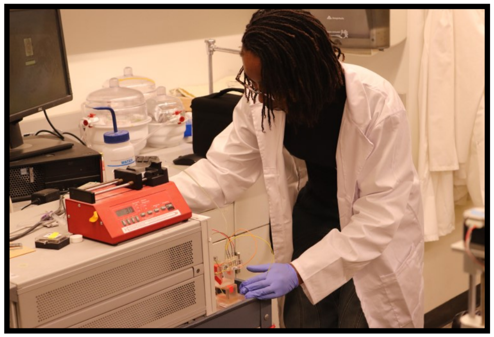
McNair Orientation

Seath Weaver
Physics

Introduction to Statistics for STEM Research

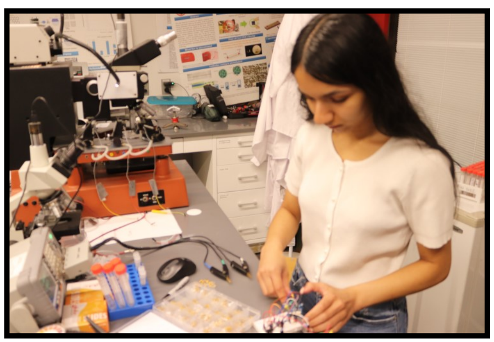
Emma Lim
Physics





Simone Nelson
Chemical Engineering

Matthew Olivera
Electrical Engineering



Sachi Rele
Mechanical Engineering

Sensor for PFAS Real-Time Detection in Water

Chadley Chris Gede
Mechanical Engineering
New Jersey Institute of Technology

PFAS (per/poly-fluoroalkyl substances) are linked to infertility, accelerated puberty, and increased risk of cancer. A 2015 study found that approximately 97% of Americans have PFAS in their blood. PFAS may mimic fatty acids in the body and are strongly correlated to dyslipidemia; dyslipidemia is associated with cardiovascular disease. PFAS are resilient and pervasive environmental contaminants. They are most commonly consumed from groundwater-connected sources such as well water and tap water. Once ingested, they linger in the body from days to decades. PFAS encompasses a large number of molecules that require highly specific processes. Therefore, cost-effective PFAS water treatment is difficult to implement. Lab testing is the conventional method to reliably test for PFAS. Sending water samples to a lab can cost hundreds of dollars, and results are prepared after multiple days. Testing for PFAS is important for individual consumers of water to judge the quality of their drinking water. However, high cost and long waiting times are inconvenient for stakeholders. As an aid to the issue, a capacitive sensor comprised of a gold electrode coated in a polyaniline-chitosan composite detects varying concentrations of PFOA (perfluorooctanoic acid) in deionized water. The polyaniline-chitosan composite facilitates the capture of PFOA molecules and the transfer of charge to the electrode to be read on the capacitance meter. Electrostatic forces between chitosan and PFOA facilitate the capture and sensing of PFOA. Chitosan is known as a high-quality, environmentally friendly adsorbent. Polyaniline facilitates the movement/transfer of electrons via the benzene rings present in emeraldine salt (the specific iteration of polyaniline used in this composite). The sensor is expected to detect PFOA in deionized water within the part per billion range. Successful selective detection of PFOA implies cheaper, more accessible PFAS testing for stakeholders. Future research involves modifying the sensor for increased selectivity and sensitivity. The ideal sensor can ignore other substances and ions in water to detect PFOA alone. The maximum contamination level of PFOA in water is 4 parts per trillion; thus, future research involves developing a sensor capable of detecting lower concentrations of PFOA.

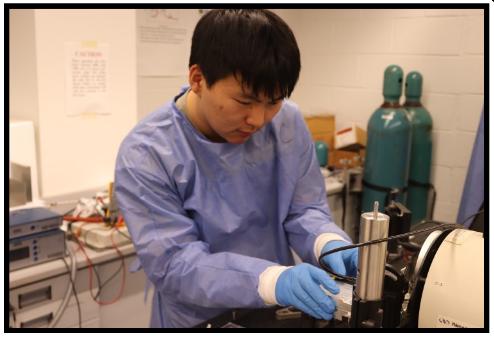
Library Literature Review

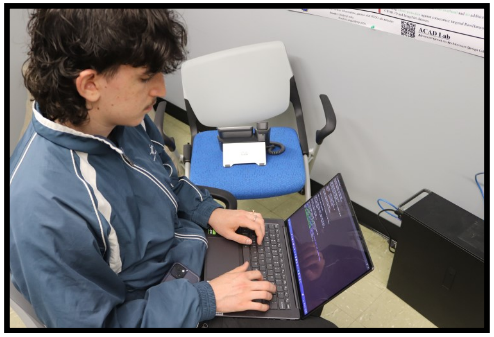
Modeling in-vitro Releases of Crystalline API from HPMC Matrix Formulations

Mauricio Huapaya, Christopher Kossor, Dr. Rajesh Davé
Department of Chemical and Materials Engineering
New Jersey Institute of Technology

Controlled drug release systems are designed to sustain stable therapeutic drug concentrations to ensure proper dosage ensuring the mitigation of adverse effects associated with inconsistent treatment profiles, including nausea, dizziness and toxicity. Hydroxypropyl methylcellulose (HPMC) is a common hydrophilic polymer utilized in controlled-lease drug delivery systems valued for its ability to modulate different drug release profiles with a market projected to reach USD 2.25 billion by 2026, growing at 4.4% annually (Industry Arc, 2021-2026). HPMC matrix formulations are a popular method for compressed tablet solid dosage forms since the Active Pharmaceutical Ingredients (API) release profile can be fine-tuned by adjusting key material attributes including molecular weight and polymer concentration. The HPMC swells dramatically upon contact with the dissolution media and forms a thick gel layer that delays API diffusion. This research addresses these challenges by creating a database of crystalline API in-vitro release profiles from HPMC matrices to further develop a release model for future predictive capabilities. Techniques include assessing the compaction behavior of the desired direct compressed tablet formulations for a range of crystalline API solubilities, API loading, and HPMC molecular weight grade. The experimental API release profiles will be obtained by performing USPI dissolution and UV-VIS analysis to quantify the API release rates across formulations, revealing differences in release kinetics. In addition, water uptake experiments on the HPMC tablets will quantify matrix swelling rates, providing insight to their impact on diffusion pathways. Anticipated results from this study are expected to demonstrate that HPMC formulations with higher viscosity grades, such as K100M, will exhibit slower and more controlled drug release profiles compared to lower viscosity grades like K4M, due to K100M's greater gel layer viscosity and reduced water uptake rate. Several API release quantitative models from literature will be implemented and nonlinear regression will be applied to determine the model fitting parameters to characterize the various formulations of interest to reveal difference in diffusion mechanisms. This systematic approach aims to provide efficient HPMC formulation development guidelines by identifying key material attributes by characterizing a diverse range of controlled release formulations with several predictive models to reduce costs and enhance therapeutic efficacy, significantly improving patient outcomes by ensuring consistent drug delivery in an expanding pharmaceutical landscape.

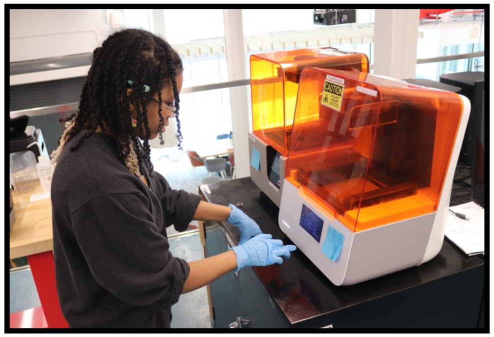
Wonki Chae
Physics





Assaf Izhar
Computer Engineering

Hala Kanaan
Biomedical Engineering

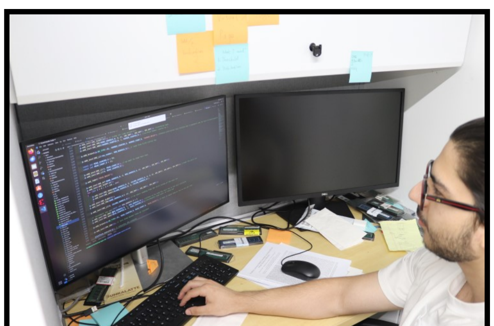


Kaylin Koen
Chemical Engineering

Piyush Jain
Computer Science and Engineering (Data Science)

Ankit Kumar
Computer Science and Engineering

Adwaita Laha
Chemical Engineering



Abdulrahman Aljoudi
Computer Engineering

Developing Biodegradable Mulch for Sustainable Agriculture

Camila Martinez, Advisor: Dr. Lisa Axe,
Mentor: Nikki Rodriguez, PhD Student
Otto H. York Department of Chemical and Materials Engineering
New Jersey Institute of Technology

The practice of mulching in agriculture has been an integral part of human cultivation techniques since well before 500 BC. Since 1948, poly(ethylene) (PE)-based mulch films have been widely used globally. PE mulch is a significant source of microplastic pollution in soil fauna, subsequently integrating into our food chain. Microplastics have been found in human tissues such as the placenta and arteries. The accelerating growth of the human population will exacerbate this issue as a result of higher food demand. For this reason, the need for durable, biodegradable mulch films, which can be directly tilled and degraded in the field, has emerged. However, biodegradable options currently available fail to meet the durability that PE offers; furthermore, the degradability of these films requires further analysis. The overarching objective of this research is to develop an optimized composition of a biodegradable film comprised of microcrystalline cellulose (MCC) and hydroxyapatite (HA) particles intercalated in a polymer matrix of poly(lactic acid) and poly(butylene adipate terephthalate) that exhibits properties similar to PE. Films were melt-mixed in a double-arm mixing unit fitted with roller mixing blades and then hot-melt pressed to produce homogeneous, reproducible samples. Films were then subjected to mechanical testing, ultraviolet (UV) accelerated aging, and water absorption simulations to identify the optimal composition for enhanced durability and environmental performance. Uniaxial mechanical tensile testing was performed to characterize stress at break and elongation at break. Biodegradable mulch films must exhibit a high tensile strength (≥ 18 MPa in machine direction, ≥ 16 MPa in transverse direction) and a high percentage elongation ($\geq 200\%$ in machine direction, $\geq 350\%$ in transverse direction) according to commercial standards. The UV accelerated aging setup uses high-intensity fluorescent lamps to simulate 3 months of sunlight over 9.7 days. Preliminary results show that MCC concentrations strongly affect water absorption, which is important in biodegradation after tillage. However, added MCC and HA particles can compromise the ductility of the film. Therefore, we hypothesize that there will be a composition where properties are balanced for meeting agriculture requirements while undergoing degradation once tilled into the soils within a 6-month period. Future work seeks to identify the optimal composition, compare its properties with PE and biodegradable films currently on the market, and evaluate the films in field studies to ensure commercial viability.

Comparisons using simulations of copula graphic estimators of survival functions based on dependent censored survival data

Andrea Camila Pardo, Advisor: Dr. Antai Wang
Department of Mathematical Sciences, New Jersey Institute of Technology

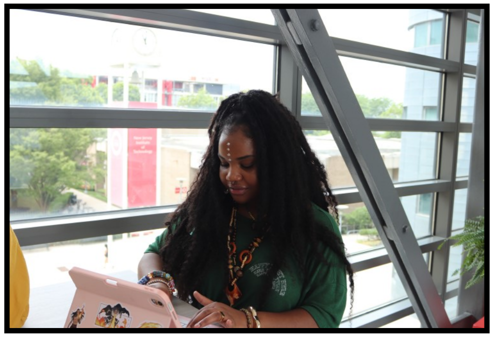
In traditional survival analysis, the widely used Kaplan–Meier estimator assumes that the event of interest and censoring are independent. However, this assumption often fails in real-world settings, such as clinical trials where patients may drop out due to health-related factors. When censoring is dependent on survival time, this can introduce significant bias into survival estimates. To address this, a simulation framework using Archimedean copula models (Clayton, Hougaard, and Frank) was implemented to estimate survival functions under dependent censoring. This project applies the complete set of derivations and estimations formulas from the novel work of my advisor to simulate dependent survival data and apply copula graphic estimators in this context. Dependent censored survival data is generated using statistical software (R) and evaluates the performance of each copula model by comparing their estimated survival curves. The goal is to assess which copula model offers the most reliable and unbiased survival function under different dependence structures. Anticipated outcomes include identifying models that perform robustly even when the true dependence is unknown or mis specified. This research provides a foundation for more accurate survival analysis in biomedical, engineering, and social science applications. Future steps include applying these techniques to real-world datasets and expanding the methodology to more complex copula families.

Fast Antibody Characterization via Microdroplet Digestion with Novel Enzymes and Mass Spectrometry

Juana Perez, Yongqing Yang, Timothy Yaroshuk, Md Tanim-Al Hassa,
Advisor: Hao Chen, Department of Chemistry and Environmental Science
New Jersey Institute of Technology

Monoclonal antibodies (mAbs) are popular biotherapeutics used for the treatment of diseases, such as cancer, infections, and neurological disorders. Exploiting the target specificity of antibodies, cytotoxic small-molecule drugs are sometimes attached to antibodies to form a class of targeted therapeutics known as antibody-drug conjugates (ADCs). To evaluate the safety and efficacy of these therapeutics, bottom-up or middle-up proteomic approaches are generally employed. Typically these approaches use enzymes to digest/break the proteins into smaller fragments. Mass spectrometers (MS) are then used to analyze the fragments to obtain information about the protein. Enzymatic digestions can take hours, however, it has previously been demonstrated that enzymatic reactions in microdroplets can reduce the reaction time to microseconds. This capability was capitalized previously by using Agilent's Jet Stream Source (AJS), which is an ionization source for Agilent's mass spectrometers. This study will therefore use AJS coupled with a quadrupole time-of-flight mass spectrometer to characterize ADCs. Additionally, this study will evaluate various enzymes, such as GingisKHAN and IdeS Xtra from Genovis, that can be used for these bottom-up and middle-up approaches. To assess digestion efficiency, the MS signals of the antibody digest produced by both microdroplet and traditional in-solution reactions will be compared. Next, to ensure clean and simplified MS signals, enzymatic sugar removal or deglycosylation of the ADCs will be assessed with the AJS source. Moving forward, this work can facilitate quantifying the amount of drugs attached to the antibody, otherwise known as the antibody-to-drug ratio (DAR).

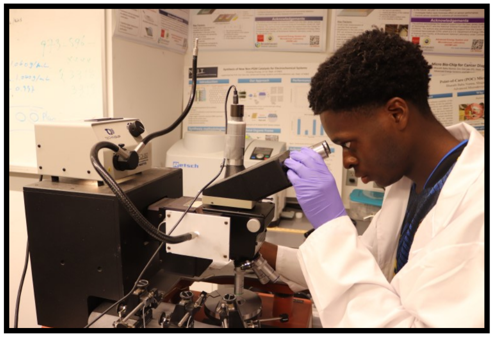
Mario Urla
Computer Engineering



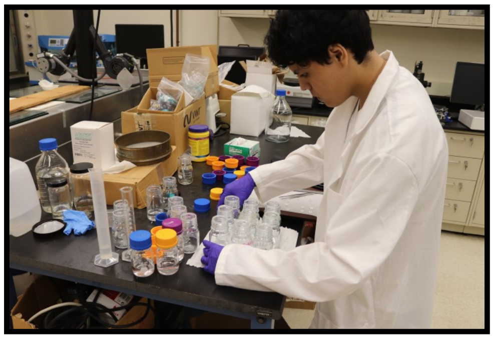
Naketa Williams
Cyberpsychology
Minor: Philosophy & Applied Ethics

Arnesh Banerjee
Computer Science and Engineering (Data Science)

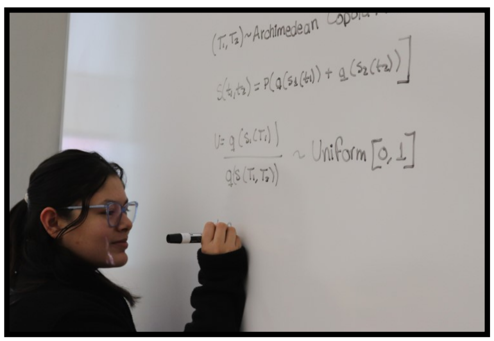
Suhrid Behari Paul
Information Technology



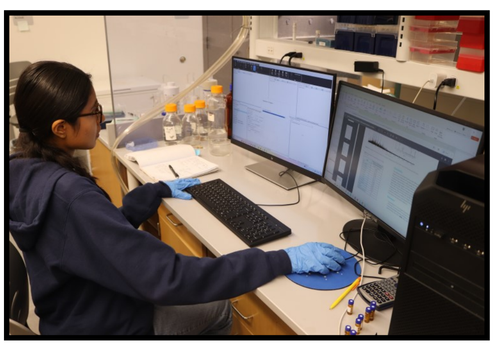
Chadley Chris Gede
Mechanical Engineering



Mauricio Huapaya
Chemical Engineering
Minor: Chemistry



Andrea Pardo
Data Science
Minor: Computational Mathematics



Juana Perez
Biochemistry and Forensic Science

Advancing the Usage of Edge Intelligence in Autonomous Driving

Mario Urla, Advisor: Dr.Tao Han
Department of Electrical and Computer Engineering
New Jersey Institute of Technology

The current network structure is known as the cloud infrastructure. The system has worked as a whole, but as real-life applications, such as autonomous driving, continue to advance, the demands placed on cloud infrastructure grow. These demands are massive volumes of data to make real-time decisions, decisions required by the sensors, such as LIDAR, cameras, GPS, radar, and onboard computing systems. In these strenuous demands on the cloud, the real-life applications become ineffective as the cloud cannot provide the needed support. The solution proposed to this issue is Mobile Edge Computing (MEC) is a crucial part of the modern network architecture as it plays a key role in enabling real-time, low-latency applications. By bringing together the data processor with the source of data itself, Mobile Edge Computing would reduce latency and ease the load from the centralized cloud systems. This research aims to answer how federated machine learning can be utilized to automate and optimize network management in mobile edge computing systems. Through conducting this research, the simulator software CARLA, alongside the real-world RoboRacer AI platform, will be utilized to evaluate the performance, reliability, and responsiveness of edge intelligence within autonomous vehicles compared to cloud infrastructures. In CARLA, federated machine learning is utilized as artificial intelligence in the simulation. CARLA also allows for a cloud-based infrastructure simulation. The CARLA simulator will be leveraged as a testing field to create a comparison between both infrastructures. Both experiments will simulate urban driving with network latency ranging between 10ms - 200ms, bandwidth constraints, and sensor configurations. There are also environmental conditions that are systematically varied to test robustness; these conditions range from dark skies, fog, rain, low visibility scenarios, and lightning storms. The anticipated results would demonstrate that the edge-intelligence structure as a whole allows for autonomous driving to perform better when compared to a cloud structure. In the simulation, tasks increase in complexity from basic lane-following to dynamic obstacle avoidance and pedestrian interaction. Each task is completed in both edge-local and cloud-based infrastructure. To piggyback off these simulations, Roboracer AI serves as the physical manifestation of edge-only AI that must make instantaneous decisions under real-world constraints. These constraints include sensor interference, onboard hardware limits, and track noise. This experiment validates the simulation findings while also highlighting the viability of edge intelligence placed in operational AV systems. Through both forms of experimentation, system performances are compared in terms of decision latency, task success rate, and safety-critical event response. Anticipated results are to demonstrate that edge intelligence deployment into operational AVs will prove superior to cloud-based deployment.

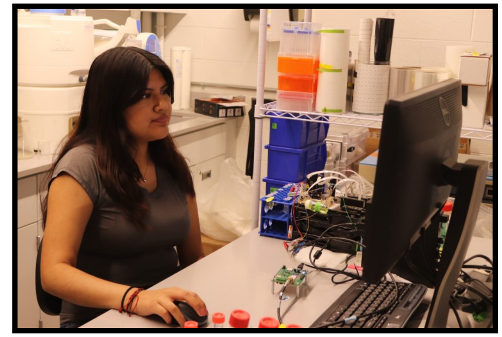
Mycelium to Machine: Exploring Hybrid Intelligence Through Cyberpsychology

KETA Williams
Department of Humanities & Social Sciences
New Jersey Institute of Technology

Current AI development centers the belief that intelligence is individual, rational, and human-like. However, this view limits our ability to envision AI systems whose intelligence and ethics emerge through ongoing interaction with people, environments, and other systems—what is referred to as co-evolving AI. Rooted in methodological individualism, this prevailing paradigm treats intelligence as the property of isolated agents. It often leads to anthropocentric design, where human cognition becomes the default template for machine reasoning and ethics. Inspired by mycelium, this project explores an alternative model for AI grounded in decentralized, cooperative intelligence. Mycelium exhibits a form of non-hierarchical intelligence, as it shares nutrients, adapts to local conditions, and sustains life through decentralized cooperation. Rather than following top-down commands, it coordinates activity through responsiveness and connectivity. These networks prioritize coherence over correctness and relational feedback over rigid control. To explore these properties, this project involves a behavioral experiment that observes how mycelium responds to environmental changes. Petri dishes will be used to cultivate fungal networks under varying conditions (e.g., contamination, nutrient richness), and the researcher will document adaptive behaviors such as rerouting, clustering, or stalling. Observations will be analyzed both qualitatively (e.g., patterns of behavior) and quantitatively (e.g., rates and directions of growth) to develop a prototype of a distributed model of decision-making. These insights will be compared with emerging models of hybrid intelligence that challenge isolated models of cognition, instead viewing intelligence as distributed across human, machine, and environmental relations. Emerging from the mycelial metaphor, the project proposes a framework for AI systems that evolve through interaction rather than isolation.

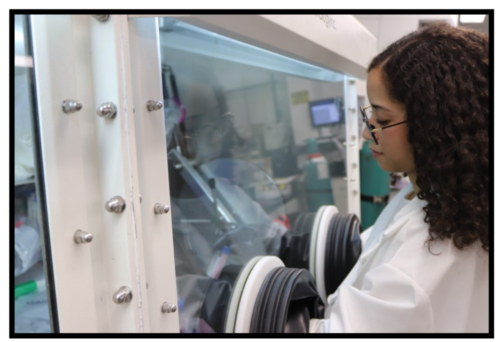
What is Research Seminar

Gisselle Ambrosio
Biochemistry



Jada Byfield
Information Technology

Angelo Bustamente
Web and Information Systems



Camila Martinez Castillo
Chemical Engineering

Portable Readout System for Microfluidic-Based Multiplex Biosensor for Alzheimer's Disease

Sachi Rele, Research Advisor: Dr. Eon Soo Lee
Department of Mechanical and Industrial Engineering
New Jersey Institute of Technology, Newark NJ 07102

Early diagnosis of Alzheimer's disease (AD) is crucial for starting treatments and improving patient outcomes. With more than six million Americans living with AD, the economic impact of AD treatment costs is projected to exceed $1 trillion in the US by 2050, placing an immense financial and emotional burden on families and caregivers. However, many current diagnostic methods, such as MRI or PET imaging, are expensive (ranging from $1300 to $3000), time-consuming, and not accessible to everyone. This gap can be filled by the creation of a point-of-care (POC) diagnostic device for early detection of AD biomarkers such as amyloid beta-42. A microfluidic-based multiplex biosensor that identifies various AD biomarkers by detecting the capacitance changes of the functionalized interdigitated electrodes is proposed. The objectives are to 1) develop and build a portable readout system capable of measuring capacitance changes through the use of an Arduino microcontroller that replaces traditional laboratory equipment and supports POC use; and 2) design the circuit to enhance detection accuracy of the Arduino system so that it is comparable to the traditional LCR (inductance, capacitance, and resistance) meter. This was done by building a circuit combined with the Arduino, programming the system to detect capacitance of the biosensor alongside integrating a display, and building a 3D-printed prototype to miniaturize the system and for portability. A compact, 3D-printed prototype that incorporates the Arduino circuit system and allows for real-time capacitance measuring was developed. Additionally, improvements to the circuit design enhanced the sensitivity of the biosensor's capacitance sensing, bringing it closer to the performance of traditional LCR meters. The system also includes a display module for the instant visualization of capacitance data. A low-cost Arduino circuit system that can function as an effective readout system for capacitive biosensors and can be used as a key alternative to traditional laboratory equipment is possible. The combination of hardware, software, and enclosure design indicates the feasibility of real-time, portable biomarker detection in POC environments. Future study will focus on improving the system's sensitivity to detect capacitance values less than 10 pF which is crucial for early-stage biomarker detection. Additional efforts will focus on full integration of the display interface. These recommendations will improve the accessibility and applicability of diagnostic tools, allowing for early intervention in Alzheimer's disease and broadening their usage to other diseases with detectable biomarkers.

Entomological Photonic Sensing

Seth C. Weaver, Topu Saha, Benjamin P. Thomas
Wyant College of Optical Sciences, University of Arizona. Tucson, AZ 85721,
Department of Physics, New Jersey Institute of Technology, Newark, NJ

Mosquitoes are a uniquely problematic insect for human health as they serve as vectors for disease and cause an estimated 700,000 deaths annually. Current methods utilize physical traps to monitor mosquito populations but may have implicit biases and are costly and time-consuming. The implementation of photonic methods is likely to enhance responsivity and improve the effectiveness of population monitoring and control efforts. However, optical sensing alone generally has a low taxonomic accuracy; notably, it has been shown that photonic identification of mosquito species, sex, and gravidity significantly improves when wingbeat and depolarization ratios are collected. This study assesses the application of near-infrared polarimetric measurements to identify and characterize insects in a transmission configuration under laboratory conditions. Polarization sensitive detectors record transmission intensity to obtain the wingbeat frequency and depolarization ratios of the wings and body of an insect. This data is then processed using an unsupervised machine learning classifier to evaluate mosquito species, sex, and gravidity.

Heritage Institute of Technology

Development and Optimization of microfluidic membrane contactor for gas sensing application

Adwaita Laha, Sreerag Kaaliveetil, Sagnik Basuray
Department of Chemical & Materials Engineering
New Jersey Institute of Technology

Microfluidic gas sensors offer a promising pathway toward compact, sensitive, and selective environmental sensing platforms. In this work, we explore the benefits of using a microfluidic membrane contactor as an electrochemical gas sensor. The architecture consists of two microchannels, one for gas flow and another for confining the sensing liquid, separated by a hydrophobic membrane. These microchannels are sandwiched between two glass substrates, with the bottom glass substrate comprising interdigitated microelectrodes in contact with the sensing liquid. We developed a theoretical model to analyze the mass transfer behavior of this system and found that above a certain threshold gas flow rate, mass transfer is no longer rate-limiting. Experimental validation using CO_2 as the model gas and [EMIM][2-CnPyr] ionic liquid as the sensing element confirmed these findings. We further examined the benefits of using microchannels for both gas transport and liquid confinement. The model demonstrated that directing the gas through a microchannel significantly enhances mass transfer of the target analyte compared to macrochannel configurations, a result corroborated by experimental data. Lastly, we demonstrated that confining the ionic liquid—the sensing medium—within a microchannel enhances the sensor's sensitivity and response time for a fixed volume of analyte, compared to macrochannel-based designs. Together, we showed that a microfluidic membrane contactor architecture for gas-liquid significantly enhances the performance of a gas sensor. These studies lay the groundwork for next-generation microfluidic gas sensors.

Towards achieving safe LLM through rectified penalty using RLHF

Arnesh Banerjee, Kartik Pandit, Sourav Ganguly, Arnob Ghosh
Department of Electrical and Computer Engineering
New Jersey Institute of Technology

Large Language Models (LLMs) have revolutionized the way we approach tasks such as writing, programming, and general knowledge generation. Their ability to understand and generate human-like text has enabled a wide range of applications across industries, education, and research. However, as these models become more integrated into everyday tools and decision-making systems, ensuring their safety and reliability becomes critically important. One of the most pressing concerns is the susceptibility of LLMs to jailbreaking prompts—carefully crafted adversarial inputs that can manipulate the model into producing unsafe, biased, or otherwise harmful outputs. These vulnerabilities not only pose ethical and societal risks but also hinder the adoption of LLMs in sensitive domains such as healthcare, law, and education. In this project, we aim to build a safer and more robust LLM by leveraging Reinforcement Learning from Human Feedback (RLHF) within a constrained reinforcement learning framework. Unlike traditional approaches that focus solely on optimizing helpfulness or informativeness, our method explicitly incorporates safety constraints by treating harmful behavior as a measurable cost. To guide the model's learning process, we introduce a rectified penalty-based mechanism that penalizes unsafe outputs while still encouraging useful, high-quality responses. Anticipated Goals: The primary goal is to design and evaluate a fine-tuned LLM that consistently produces helpful and safe responses, even when faced with jailbreaking prompts. Our method aims to outperform current state-of-the-art safety-tuning techniques.

Spectroscopic characterization of a microfluidic membrane mimic system under dynamic conditions

Simone Nelson, Najamuddin Naveed Khaja, Sagnik Basuray
Department of Chemical Engineering,Hampton University, Hampton, VA
Department of Chemical and Materials Engineering
New Jersey Institute of Technology

Target DNA sensing is the detection of the presence, concentration, and/or identification of a specific DNA sequence. This process is important for disease diagnostics, as it is the process that identifies the genetic mutation or pathogen pertinent to the disease. This project focuses on the development of an electrochemical biosensing platform for the detection of single-stranded DNA (ssDNA) using a microfluidic membrane mimic system (MMM). Studying the binding of ssDNA helps us to better understand how DNA replicates and repairs, specifically how diseases caused by mutations start. A biosensing platform can be applied to lab-on-a-chip technologies and biomolecular diagnostics in tissue engineering. The sensor operates through a pair of microelectrodes that measure impedance changes caused by biomolecular interactions within a microchannel under electrochemical impedance spectroscopy (EIS). The exothermic binding of complementary ssDNA strands generates detectable energy shifts. These shifts can be captured using EIS, allowing for a direct correlation between ssDNA concentration and impedance response. The core innovation lies in modifying a two-electrode, flow-through MMM into a three-electrode MMM utilizing a gold-coated polyester membrane. The membrane serves a dual purpose: acting as both a filter and a working electrode that when optically transparent, enables integration with surface-enhanced Raman spectroscopy (SERS). SERS, which signals are enhanced by the gold surface, provides molecular specificity similar to NMR and IR techniques. When this three-electrode MMM and SERS are combined, the application of the system broadens to many applications, like the detection of biofouling on the membrane.

Continuous Sensing and Machine Learning Analysis of Lead Occurrence in Drinking Water

Matthew Oliveira, Dr. William Pennock, Maryam Talebidaleouei
Department of Electrical and Computer Engineering
Princeton University, Princeton, NJ 08544
Department of Civil and Environmental Engineering
New Jersey Institute of Technology, Newark, NJ 07102

The release of lead and other heavy metals into drinking water supplies poses a significant health risk, with neurological defects especially present in children and infants. It is estimated that at least 170 million Americans were exposed to harmful lead levels in early childhood. In buildings built before 1986, corrosion of lead infrastructure is still an ongoing challenge, as seen in Newark and other cities, and brass fixtures made before 2014 also contain lead. Difficult to detect without off-site lab equipment, it presents a need for onsite, cost-effective lead monitoring. Engineers at the University of Michigan developed a simple four-electrode sensor containing only inert platinum electrodes for the real-time, continuous detection of lead and other heavy metals, such as zinc, copper, and iron, that are commonly present in contaminated drinking water. The sensors can be embedded in water service lines for long-time use until lead or other heavy metals are detected, operating on two 1.5 V batteries and drawing a small current on the order of microamps. The sensors can be fabricated at a low cost (about $.10/sensor) and are ideal for long-term use. We also developed a PCB (printed circuit board) module for transmission of sensor data, which integrates into existing infrastructure for the detection of water pH, temperature, and electrical conductivity. This work provides a path to developing machine learning techniques to correlate with water quality indicators on a water distribution system-wide level, helping further understanding of lead pollution and promote environmental sustainability.

Optimization and Characterization of Low Power In-Memory Computing Devices

Emma Lim, Advisor: Dr. Durga Misra and PhD Mentor: Aseel Zeinati
Department of Electrical and Computer Engineering
New Jersey Institute of Technology

Modern computing systems demand memory technologies that are both energy efficient and scalable. Resistive Random Access Memory (RRAM) and Metal-Insulator-Metal (MIM) devices present an alternative to the conventional von Neumann architecture where the memory and central processing units are separate. In MIM devices, both of these units exist and communicate in the same space, causing a potential for high-speed, low-power, and non-volatile operation. The significance of this work lies in its potential to improve device stability and reliability in next-generation memory systems. Our research focuses on these MIM RRAM devices utilizing zirconium dioxide (ZrO_2) as the active switching dielectric layer due to its simple structure and CMOS compatibility and titanium nitride (TiN) as the metal electrodes because it is a good oxygen reservoir. We aim to evaluate and improve the characteristics of these devices by exploring structural variations and pulse testing protocols. As we conduct this project, we assess different ZrO_2 based configurations, such as 3.5 nm ZrO_2 / H-plasma treatment / 4 nm ZrO_2, and evaluate how these different configurations affect endurance performance under repeated switching conditions. Each device undergoes an initial forming process, wherein an electric field is applied across the MIM device to induce resistive switching. This is applied slowly in order to prevent a hard breakdown and encourage a soft breakdown. The applied electric field drives oxygen ions toward the top electrode while the bottom electrode remains grounded, resulting in the formation of oxygen vacancies within the dielectric layer. These vacancies create conductive paths that enable the device to switch between high and low resistance states. Following successful forming, we conduct endurance testing via a series of pulse operations. We apply repeated voltage pulses of varying pulse widths (i.e. 2 μs, 10 μs) across the device. The resulting resistance state transitions between set and reset tests are monitored across ~10,000 cycles to evaluate device degradation or device failure. Ideally the resulting resistance states form an increasing staircase pattern in graphical outputs, reflecting stable switching in a device. We expect that certain structure modifications, specifically those involving oxide thickness, will yield improved endurance and switching consistency. We also believe that by changing the parameters for the pulse operations, specifically the pulse widths and the applied voltage, we will be able to see what constraints push the limits for the different devices. This research will inform the design of more robust RRAM devices and guide future fabrication and application strategies for in-memory computing.

Query Planning with Agentic AI

Suhrid Behari Paul, Subhodeep Ghosh and Senjuti Basu Roy
Department of Computer Science
New Jersey Institute of Technology

In many real-world applications, we need to find the best way to rank a set of items - like choosing the top products, hotels, or job candidates. Quite often, we rely on LLMs to provide us with such top-k ranking results. But as uncertain as LLMs are, those results are not deterministic. Instead they are probabilistic. With the growing number of items, the number of possible rankings increases exponentially, making exhaustive search using LLMs extremely expensive. This project explores a smart and interactive system that uses AI agents (LLMs) to help break down the user's request into a series of atomic queries that can eventually lead to extracting the top-k result in a cost-effective manner. The system in this project includes a main AI agent (called the Supervisor) and several smaller helper AIs (called Tool LLMs). Each LLM tool can look at a small number of items at a time and give useful feedback about the probable ranking of these items based on a user's request. This can be used to compute the probability of a particular overall ranking encompassing all the items. The Supervisor decides which subset of items to show to which Tool, and uses their responses to update its belief about the best overall ranking. This process repeats until one possible ranking becomes highly probable. The queries to the tools are made in a cost-effective manner as each LLM call gets more expensive depending on the prompt. By doing this, we not only take into account an LLM's inherent non-deterministic nature but also avoid checking every possible ranking, and instead use smart planning and learning to reach the right answer in a much cheaper and more explainable manner.

Search Functionality for Virtual Reality Ontology Editor

Piyush Jain, Advisors, Prof. Margarita Vinnikov, Prof. James Geller
Department of Data Science
New Jersey Institute of Technology

This project extends the functionality of the VROOM (Virtual Reality Ontology Object Manipulation) system. An ontology is a linked dictionary of terms from a specific domain. VROOM was designed for medical ontologies. Medical ontologies are large systems, often containing over 100,000 medical terms. Ontologies are structured like a network of nodes and links. The nodes represent the medical terms, and the links represent relationships between pairs of terms. The most important links are called subclass or IS-A links. They express generalization — for example, pneumonia is more general than viral pneumonia. Due to the size of a medical ontology, a graphics display of its network is overwhelming. The VROOM system addresses this problem by spreading out the network display in THREE-DIMENSIONAL space surrounding the viewer. This is achieved by using an Oculus Headset for VR. The user interacts with the display using a game controller in each hand. Currently, VROOM has no search functionality. In this project, the search functionality will be implemented from scratch. This includes supporting substring search — for example, typing "Can" will return "Lung Cancer," "Skin Cancer," and other related terms. A search bar and keyboard in 3D space will be implemented. The user can type parts of a medical term and the system shows all matching terms. Then, the user can choose the desired term with the game controller. As a result, the user can be teleported in VR space to float in front of the desired term. He can then edit the term or delete it. If time permits, an existing voice search functionality will be re-integrated.

Design and Implementation of a Secure Document Processing System Using Intel SGX

Ankit Kumar
Heritage Institute of Technology, Kolkata, India

The recent leak of over 16 billion compromised credentials, including passwords from major platforms like Google, Facebook, and Apple, highlights the urgent need for robust, privacy-preserving solutions that protect sensitive data from unauthorized access—even in untrusted cloud environments. These incidents reveal a critical vulnerability in conventional software-only security models and demand trusted, hardware-backed frameworks for data protection. This research presents the design and implementation of a secure document processing system that leverages Java for application logic, Amazon Web Services (AWS) for scalable cloud deployment, and Intel Software Guard Extensions (SGX) to create hardware-based trusted execution environments (TEEs). The system ensures end-to-end confidentiality and integrity of sensitive documents during processing, even in the presence of privileged threats. The methodology involves developing a Java-based backend integrated with SGX enclaves via the Java Native Interface (JNI), isolating critical operations such as encryption, decryption, and redaction within secure enclaves. A simulated SGX environment is used for prototyping, followed by deployment on AWS to evaluate performance and real-world applicability. The system architecture (Fig. 1) incorporates secure data transmission (e.g., AWS KMS) and enclave-attested processing to mitigate cloud-level risks. Anticipated outcomes include a working prototype that validates secure enclave-based processing of documents and data, with performance benchmarks under SGX constraints. The system will demonstrate practical integration of TEEs in real-world backend workflows using Java and cloud services. Future work will explore extending the framework to secure processing in IoT and blockchain-integrated systems, enabling trusted automation in areas such as self-driving cars, industrial robots, and critical infrastructure. This convergence of SGX, cloud computing, and decentralized trust aims to build resilient systems that can operate securely even in adversarial environments.

Best Classroom Practices

Miniature Peptide Synthesizer

Kaylin Koen, Advisor: Dr. Nellone Reid, Mentor: Maryom Rahman
Department of Chemical Engineering
New Jersey Institute of Technology, Newark NJ 07102
Georgia Institute of Technology, Atlanta GA 30332

Insulin, made of 51 amino acids, is the first peptide ever synthesized, and its synthesis is an extremely monumental innovation in drug discovery. Since then, it has been able to increase the production of the drug to treat people suffering with certain types of diabetes. Therapeutic peptides can act as hormones, growth factors, ion channels, and other biological structures[1], therefore making their synthesis valuable in the pharmaceutical industry as they have the opportunity to treat several different diseases, such as cancer. Solid phase peptide synthesis (SPPS) is a method of peptide synthesis that builds peptides connected to a resin. 9-Fluorenylmethyloxycarbonyl (FMOC) chemistry is a widely used method of SPPS. It works by using an FMOC group to protect an R-group of an amino acid to orient it correctly when bonding with another amino acid. This FMOC group is then washed away, leaving another R -group available to bond with another amino acid. This process is repeated until the desired peptide is created. Peptide synthesizers like the Liberty Blue uses FMOC chemistry, along with microwave technology, to synthesize peptides.

However, the Liberty Blue can cost between $25,000 to $50,000 even disregarding maintenance costs. Furthermore, its installation requires a certified technician and mandatory training resources from the manufacturer. It also creates very toxic waste that must be handled under a fume hood with the user being required to wear suitable protective equipment. The objective of this research is to make a miniature peptide synthesizer that can address these concerns. The synthesizer will have 32 different slots for 32 possible amino acids, it should be small enough for transport and point-of-use, and it will require no special training to install and operate. To create the prototype of the synthesizer, it was divided into 4 sections: the first-selector valve, the reactor, the second-selector valve, and the output container, which contains the waste, the product, and recycled material like dimethylformamide (DMF) and dichloromethane (DCM). The first selector valve has 4 separate areas with 8 spots each that house the amino acids, the selection of which is controlled by an Arduino-powered stepper motor and software. The reactor is a hollow half-cylinder with multiple shelves inside to increase surface area for the purpose of increasing the productivity of the reaction. The second selector valve is where the products are separated, with 3 areas for the finished peptide, the waste. It is expected that the miniature peptide synthesizer is able to separate waste and product effectively without leaking or wearing down the material after continued use. It is also expected to be fully automated using stepper motors. This device should be used in areas where therapeutic peptides would be needed, but there is no space for or access to a traditional peptide synthesizer, like on spaceships or areas far from hospitals.

Deep Learning for the Study of Particle-Cell Interactions

Hala Kanaan, Advisor: Dr. Xuan Liu
Department of Biomedical Engineering
Duke University, Durham NC 27708
Department of Electrical and Computer Engineering
New Jersey Institute of Technology, Newark NJ 07102

Understanding how particles and cells interact with each other is critical for advancing targeted drug delivery and improving treatments. Phase imaging allows for label-free visualization of these biological interactions and enables precise tracking of particle movement. The U-Net model, a deep learning method used for image segmentation, assists with developing automatic, accurate predictions from labeled data. Prior deep learning methods applied to phase imaging have focused exclusively on cell segmentation, overlooking particle detection. This research aims to manually label particles in phase images, train a U-Net segmentation model using phase augmentation techniques, and evaluate the model's performance in identifying particle-cell interactions. Phase images have been acquired from a novel imaging technology previously developed in this lab, called Modulated Optically Computer Phase Microscopy (M-OCPM). Particles in these images will be manually annotated to produce training data. U-Net will then be trained using phase-augmented images, in which global phase shifts are introduced to account for phase wrapping artifacts. This augmentation strategy is essential because it teaches the model to distinguish real particle features from artificial edges caused by phase wrapping. These methods should result in improved accuracy in identifying particles from wrapped phase images and in successfully distinguishing them from surrounding cells and background. It is expected that the results will demonstrate that deep learning segmentation techniques developed for cell analysis can be effectively extended to particle detection, with the goal of enabling accurate identification from wrapped phase images while avoiding the need for phase unwrapping. This work highlights how deep learning approaches can simplify particle analysis and support future research on particle-cell interactions.

How to Make Effective Poster and Oral Presentations

NSF REU Site: Optics and Photonics

Automating Silicon Photonic Circuit Design Using Large Language Models

Abdulrahman Aljoudi
New Jersey Institute of Technology, Newark NJ 07102

The recent advancements in Large Language Models (LLMs) have enabled significant breakthroughs in the automated generation of text, code, images, and even electronic circuit design. Despite these successes, the application of LLMs in the generation of Photonic Integrated Circuit (PIC) design remains largely unexplored. The manual design of PICs is labor-intensive, repetitive, time-consuming, and highly susceptible to human error, thus, highlighting a strong need for automation in this domain. For this reason, our research aims at leveraging the promising capabilities of LLMs to automate the generation of PIC netlists, thereby reducing design time, improving reliability, and streamlining scalability. In order to automate PIC netlist generation, we are developing a framework that integrates three key components: (i) the construction of a diverse dataset of PIC netlists for LLM training and testing, (ii) the use of photonic circuit simulators to test the syntactic validity and functional performance of generated designs, and (iii) the implementation of an optimization pipeline to refine circuit outputs. We are investigating dataset generation via prompt engineering and Retrieval-Augmented Generation (RAG) techniques, while simulators such as SAX (based on the JAX framework) and PhotonTorch (based on PyTorch) are being explored for their compatibility with automated testing and optimization workflows. We anticipate that this approach will demonstrate the potential of LLMs to produce syntactically correct and functionally viable and optimized photonic circuit designs. Future work will involve scaling the dataset, refining prompt strategies, and benchmarking performance across different LLM architectures to evaluate robustness, efficiency, and creativity in circuit generation.

Infrared Photoconductive Photodetectors Based-on Colloidal Semiconductor Nanocrystals

Wonki Chae, Advisor: Dr. Dong Kyun Ko
The Institute of Optics
University of Rochester, Rochester NY 14627
Electrical and Computer Engineering Department
New Jersey Institute of Technology, Newark NJ 07102

Low-cost, uncooled detectors operating in the mid-wavelength infrared (MWIR) region are in high demand, and their scarcity in the infrared detector market presents a technology gap. Imaging applications such as night driving assistance and missile guidance may not be feasible without such detectors. Lead selenide (PbSe) has the potential to bridge this gap due to its ability to operate at room temperature without the need for large, expensive cooling systems. A cost-effective method for fabricating PbSe thin films involves the use of colloidal quantum dots, which can further reduce manufacturing costs compared to conventional techniques. Regardless of the fabrication method, it is well-documented that a process known as sensitization (oxidation and iodization annealing) is required to achieve high MWIR responsivity in PbSe films. However, its mechanism is still debated, which has hindered the optimization and standardization of process parameters. The goal of this research is to investigate how variables such as annealing temperature, grain size, ligands, and metal contacts influence sensitization, which could lead to more effective film fabrication. A better understanding of this process is expected to enhance MWIR responsivity and enable PbSe films to compete with state-of-the-art technologies.

Deep Neural Network Adaptation For Photonic Hardware

Assaf Izhar Dr. Shaahin Angizi
Ying Wu College of Computer Science
New Jersey Institute of Technology, Newark NJ 07102
Department of Electrical and Computer Engineering
New Jersey Institute of Technology, Newark NJ 07102

Transformer architectures, neural networks that excel at understanding relationships in data, underpin the modern AI revolution and dominate computer vision benchmarks through Vision Transformers (ViTs). Unfortunately, their intensive compute requirements confine real-time inference to energy-hungry data centers. Photonic acceleration offers a promising solution for edge deployment, using light for low-power, faster computation of the extensive mathematical operations required in deep learning models. Yet current optical ViT prototypes forfeit much of this advantage due to complex nonlinear operations that cannot be performed optically, forcing costly conversions between optical and electronic domains with each conversion consuming 10-100× more energy than the optical computation itself. Recent studies show that replacing these nonlinear operations with simpler operators and retraining using knowledge distillation (KD) , where a simpler model learns from a complex teacher model, can recover accuracy with minimal performance loss. Existing optical implementations, however, only offload linear operations to optics, still routing each layer through extensive conversions to achieve nonlinearity.

To address these limitations, we introduce AffineViT: The first ViT architecture whose entire prediction pipeline consists exclusively of simple linear transformations that map directly onto photonic hardware as efficient operations. AffineViT eliminates conversion bottlenecks through four innovations: i) Linearized self-attention that processes and compares image parts without requiring softmax normalization, ii) Simplified feed-forward networks where multi-layer structures collapse into single linear transformations, iii) Minimal conversion architecture that confines nonlinear operations to electronic boundaries, preprocessing before optical entry and classifying after optical exit, requiring only single conversions while keeping intensive operations in the photonic domain, and iv) Multi-objective KD that transfers understanding from teacher to student models through logit matching, attention transfer, and feature alignment.

We anticipate that AffineViT will reduce electronic-optic conversions from ~200 to just 2 per inference, cut total operations and parameters through architectural simplification, slash energy and latency substantially relative to reported optical ViTs, and maintain accuracy within 5 percentage points of traditional baselines. AffineViT marks a decisive step toward truly end-to-end optical vision models, enabling low-power, high-throughput AI inference on edge devices like smartphones, drones, and IoT systems.