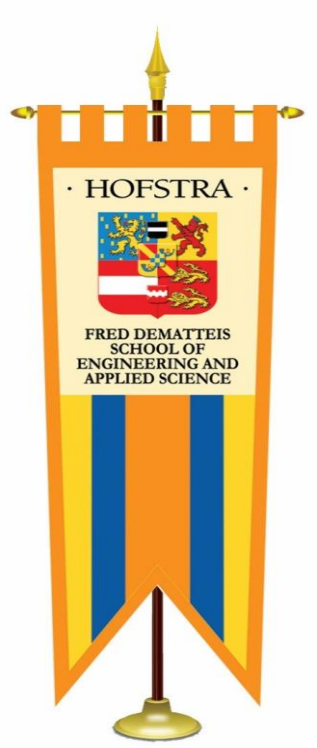
Fred DeMatteis School of Engineering and Applied Science

ASPiRe

Advanced Summer Program in Research



8th Annual Symposium

Hofstra University
Science & Innovation Center, Room 124

August 27, 2024
9:20 a.m. – 3:40 p.m.

This is the eighth annual Advanced Summer Program in Research (ASPiRe) Symposium to be hosted by the Fred DeMatteis School of Engineering and Applied Science and I am pleased that twenty-three students have participated. A key component in its success and growth is the most generous financial support it receives from distinguished benefactors of the DeMatteis School.

Donor support of these exceptional researchers is of immeasurable value to the students themselves and adds luster to all our educational endeavors.

Sina Y. Rabbany, PhD
Dean
Fred DeMatteis School of Engineering and Applied Science

TABLE OF CONTENTS

PRESENTATION SCHEDULE

	TIME	STUDENT/ ADVISOR	PROJECT TITLE
	9:20 AM	**Dean Rabbany Welcome**	
1	9:30 AM	Mason Malone Dr. Simona Doboli	Analyzing Ideas Using Knowledge Graphs and Linguistic Methods
2	9:45 AM	Gianluca Tredici Krishan Rathee Dr. Edward Segal	Anchorage Attachment Using Drone Technology
3	10:03 AM	Shannon O'Brien Dr. Lynn Albers	Utilizing Maker Space Resources to Create New Pedagogy for First Year Courses; Systematic Review on Seatbelt Safety
4	10:18 AM	Raymond Spinelli Dr. Lynn Albers Joseph Heaney III PE, Principal Walden Environmental Engineering	CLCPA, NYSERDA, Walden and Hofstra: Getting to Net-Zero One Building at a Time
5	10:33 AM	Matthew Ryan Dr. Manuel Miranda	Moisture Diffusion and Carbonation of Concrete
Break	10:48 AM	**BREAK**	
6	10:55 AM	Caleb Williams-Anderson Dr. Minjeong Suh	Sunlight Driven Transformation of Pesticide Isoxaben
7	11:10 AM	Aurisha Rahman Dr. Margaret Hunter Dr. David Weissman	Harmful Algal Bloom Detection Using Remote Sensing
8	11:25 AM	Miko Amican Dr. Elliott Williamstyer	Design of a Spherical Antenna Positioner
9	11:40 AM	Raquel Brown Daniel Doyon Dr. Mauro Caputi Dr. Yimin Zhao	Wound Image Segmentation using Deep CNN Architectures
10	11:58 AM	Lidia Lojewski Dr. Edward Currie	Mojo & Modular Programming
Break	12:13 PM	**LUNCH BREAK**	
11	12:55 PM	Gianna Imeidopf Dr. Nicholas Mema	Optimization of Plant-Derived Extracellular Matrix Structure and Mechanics for Vascular Repair
12	1:10 PM	James Bacheller Dr. Sina Rabbany	Optimization of Pressure-Flow Relations on Vascularization
13	1:25 PM	Evan Carroll Dr. Roche de Guzman	Anti-fibrotic Effects of Atorvastatin Delivered via Keratin-based Hydrogels
14	1:40 PM	Kaita Odani Kyle Raymond Dr. David Rooney Dr. John Vaccaro	The Effect of Freestream Turbulence Level on the Drag & Lift Coefficients of Rotating Baseballs and Spheres
Break	2:00 PM	**BREAK**	
15	2:05 PM	Artem Pugach Dr. Xiang Fu	An Approximation Framework for Making Regex Malware Signatures Zero Knowledge Proof Friendly
16	2:20 PM	Aisha Ahmad Dr. Gerda Kamberova	The Nature of Generalizing Same-Different Visual Relations in Deep Learning
17	2:35 PM	Cade Ferguson Dr. Simon Shamoun	Comparing tournament rankings by distance and back edges
18	2:50 PM	Mani Tofigh Dr. Jianchen Shan	Optimizing Memory Allocation in Cloud VMs with Accurate Cache Abstraction
19	3:05 PM	Banmeet Kaur Dr. Krishnan Pillaipakkamnatt	Optimizing Documents for Retrieval Augment Generation
	3:25 PM	**ASPiRe Group Photo**	

1.

Analyzing Ideas Using Knowledge Graphs and Linguistic Methods

Researcher: Mason Malone
Advisor: Dr. Simona Doboli

Characterizing a group's brainstorming conversation helps flag relevant events such as early fixation, or lack of exploration of different ideas or semantic regions. Our research goal is to extract information on the ideas expressed by a group, such as, semantic connections between ideas (e.g. add details to a previous idea, combine ideas), their relative novelty (e.g. a novel idea in the conversation), their absolute novelty (e.g. their novelty in the domain), their relevance to the subject of the conversation, and their level of elaboration.

Our previous attempts at doing so involve using AI models to embed the ideas and assigning a novelty value based on the distance between embeddings. This approach assigns a novelty score to each idea, but does not offer an explanation: what is similar and/or different between ideas, what is being added or changed between ideas.

In this work we aimed at constructing graphical representations of ideas using linguistic methods, such as dependency parsing and part of speech labeling. Thus, we can find direct connections between ideas (e.g. common relevant words), and compute the distance between ideas as a shortest path in a graph. We explored a model trained to extract semantic relationships between words in a text (REBEL), extracting ontological and semantic relationships between words (wordnet), and utilized SpaCY and Stanford Core NLP to examine linguistic relationships in sentences. Our results both visual and numerical show that the approach finds sensible semantic connections between ideas in an explainable way.

2.

Anchorage Attachment Using Drone Technology

Research Team: Gianluca Tredici, Krishan Rathee
Advisor: Dr. Edward M. Segal

Today, drones are providing assistance with rescue operations in areas such as thermal imaging and damage assessment. But, is it possible to expand the impact of drones in rescue operations by using them to make anchorage attachments? In this scenario, drones can be piloted from one side of a chasm, move quickly across long distances, and make precise connections to anchor points. The primary objective of this research was to develop a system that would allow the drone to attach a carabiner, with a pilot line affixed, to an anchorage point on the other side of a divide. The research also included understanding how precisely the drone could be flown and which anchor type would be most suitable for the connection sequence. Early in the design process it became clear that having a purely mechanical release system would be ideal as there could be a wider margin for error with remotely activated systems. The resulting design consisted of a frog carabiner, 3D printed parts, and fastening screws. The unique characteristic of the frog carabiner is that it is fixed in the open position until a sizable force hits its center which forces it to lock, enclosing the anchor point and creating a stable connection. The 3D printed housing and prongs hold the carabiner in a particular orientation, resisting translation and rotation. Once the carabiner makes its connection with the anchor point, the drone pulls away from it allowing the prongs to bend and

(continued)

2. *(continued)*

Anchorage Attachment Using Drone Technology

release the carabiner. In total, the release system weighed approximately 0.2 pounds. The drone utilized in this research was a 2019 DJI Phantom 4 Pro V2.0. The drone's flight parameters were adjusted in the DJI Go 4 mobile app, allowing it to drift no further than 4 inches in any direction with average wind speeds when flying at a height of 5 feet above the ground. After the addition of a carbon fiber drone cage, it became clear the release system could weigh no more than 0.3 pounds to maintain proper mobility. The anchor type was selected after constructing multiple prototypes. It was a grill made up of twelve closely spaced dowels resting at a height of approximately 5 feet above the ground. This anchor type allowed for an easier connection based on the drone's flying tolerances. When flying from approximately 20 ft away, this method of connection proved to be successful around 83% of the time. Moreover, when flying from roughly 33 ft away, connection was successful 33% of the time, despite heavy winds during these test flights.

3.

Utilizing Maker Space Resources to Create New Pedagogy for First Year Courses; Systematic Review on Seatbelt Safety

Researcher: Shannon O'Brien
Advisor: Dr. Lynn Albers

With the creation of the new Maker Space in the First-Year Engineering Lab of the Science and Innovation Center, there are many opportunities for student learning and creating. New pedagogy utilizing equipment in the Maker Space was designed for students to gain experience using and safely operating the equipment while learning STEM concepts for ENGG 016, a first-year design course. The design specification for an active learning environment led to the development of hands-on activities combining one piece of equipment and one STEM concept. This process started with three set projects: laser cutter & buoyancy, ShopBot® & anatomy, and CNC mill & energy conversion. While ideating, the scope organically evolved into 3D printing coral out of natural and sustainable materials as a third project, and electric skateboards for the second project.

A systematic review of vehicle impact testing and seatbelt safety led to the understanding that the female impact test dummy is not modeled on the female body but is a sized down version of the male impact test dummy. As the female body is not accurately represented, women are more likely to be severely injured in car accidents and are particularly more prone to whiplash injuries from the constriction of their seatbelts during a collision. Certain safety features in vehicles benefit the male body more than the female body due to the lack of representation of women in impact testing.

4.

CLCPA, NYSERDA, Walden and Hofstra: Getting to Net-Zero One Building at a Time

Researcher: Raymond Spinelli
Advisor: Dr. Lynn Albers, Hofstra University
Mentor: Joseph Heaney III PE, Principal Walden Environmental Engineering

Getting to Net-Zero is a combination of understanding where we are at, where we need to be, and engineering a plan to achieve this goal. Laws such as the Climate Leadership and Community Protection Act (2019) provide direction, performing energy audits helps create a baseline understanding of the current energy usage, and NYSERDA provides funding opportunities to reduce energy usage and begin the road to net-zero emissions. Data from the HOBO data logger in Weed Hall Room 106 was downloaded and analyzed using MATLAB. Conservatively, an estimate of $0.18 per kWh was used in electric energy calculations to account for both electric usage and demand charges. A total of 14,525 kWh costing $2,615 was spent from May 1, 2023 to April 30, 2024 supporting the activities in Weed Hall Room 106. A lighting audit was conducted on Hofstra University's Physical Plant and a recommendation to retrofit T8 bulbs and respective fixtures with LED lighting was proposed. The existing fixtures use an estimated 68,860 kWh costing $12,395 per year assuming a 10-hour day, 250 workday operation. Retrofitting the T8 fixtures to LED lighting with the same amount of usage will save an estimated 36,327 kWh and $6,539 per year. Plans for Hofstra University's Sustainability Outreach Program were discussed giving faculty and local companies an opportunity to bring energy efficiency awareness to the community at monthly events.

5.

Moisture Diffusion and Carbonation of Concrete

Researcher: Matthew Ryan
Advisor: Dr. Manuel Miranda

This research project had two main objectives, both related to the durability of concrete. First, to measure the evolution of the relative humidity within the pore structure of three high-strength concrete samples over the course of approximately three months. Two of our samples were sealed with wax, having one open face, and the third was completely sealed. The data collected was used to track the diffusion of moisture through the samples and out through the open face, as well as study the process of self-desiccation in the fully sealed sample. Self-desiccation is an important aspect of the investigation because the high-strength concrete of the samples has a very low water-to-cement ratio and was not previously cured, resulting in an incomplete hydration of the cement paste. During the experiments, the self-desiccation caused a significant portion of the evaporable water to be consumed in hydrating the cement particles, especially in the early stages of the experiments.

The second objective was to measure the carbonation depth of concrete samples that were exposed to different environmental conditions. Carbonation is a process in which atmospheric carbon dioxide diffuses into the concrete. The carbon dioxide reacts with calcium hydroxide in hydrated cement paste to form calcium carbonate, lowering the pH of the concrete. Due to the lower pH, the passivating film protecting the concrete reinforcement will be removed, which may cause it to corrode. In our experiments, the samples used were split using a hammer and chisel, and then the depth of carbonation was measured. The carbonation constant was then calculated to determine the rate at which the reaction took place in each sample.

6.

Sunlight Driven Transformation of Pesticide Isoxaben

Researcher: Caleb Williams-Anderson
Advisor: Dr. Minjeong Suh

Around 1 billion pounds of pesticides are used annually in the U.S. alone (U.S. Environmental Protection Agency). Herbicides, a class of pesticides, are used to control or inhibit the growth of unwanted vegetation. Isoxaben is a primary active ingredient in numerous commercially sold pre-emergent herbicides used specifically for broadleaf weeds. While isoxaben is hazardous to humans through inhalation and contact with eyes, the fate of isoxaben and its transformation products remains largely unexplored.

The primary objective of this study was to investigate the transformations isoxaben undergoes in the environment under solar irradiation. Sunlight can cause compounds to degrade via two major pathways: direct photolysis and indirect photolysis. In direct photolysis, the target compound is transformed by direct absorption of sunlight. During indirect photolysis, other components in the environment (e.g., natural organic matter, nitrate) absorb sunlight to form highly reactive molecules, which can then react with the target compound. We developed a system using the Q-SUN Xenon Test Chamber equipped with a water bath for temperature control of irradiated samples. The xenon lamp inside the Q-SUN has a comparable spectrum as solar irradiation, and we conducted actinometry experiments to ensure that the irradiance inside the reactor was comparable to that of natural sunlight.

We first investigated the degradation of isoxaben via direct photolysis by irradiating aqueous solutions of isoxaben. We then exposed suspensions of isoxaben and fulvic acid, a type of natural organic matter (NOM), to determine the effect of the addition of NOM, and therefore indirect photolysis, on isoxaben degradation. The degradation of isoxaben over time was monitored using liquid chromatography (LC) quadrupole time-of-flight (QTOF) mass spectrometry (MS). Our results demonstrate a 90% decay of isoxaben by direct photolysis in 11 days. Preliminary results from indirect photolysis experiments demonstrate that the addition of fulvic acid results in faster degradation of isoxaben, but further studies are needed to investigate contributions from different reactive species in isoxaben degradation, in addition to changes in the transformation products from indirect photolysis.

7.

Harmful Algal Bloom Detection Using Remote Sensing

Researcher: Aurisha Rahman
Advisors: Dr. Margaret Hunter, Dr. David E. Weissman

Algal blooms are a phenomenon in which algae rapidly overgrows due to excess nitrogen and phosphorus. The algal blooms reduce the oxygen content in the water and blocks sunlight from entering. The resulting dead zones in the water ruin the aquatic life of the water body. Although this alone is an issue, the bigger problem arises when algal blooms transform into harmful algal blooms. Harmful algal blooms (HABs) are when the bacteria in the algae release toxins, typically cyanobacteria, into the water. HABs destroy water quality, preventing the distribution of drinking water due to the stomach, liver, and nerve damage that can occur to people through consumption. The enduring presence of climate change has created conditions in which HABs have been able to become more prevalent and dangerous. We are interested in demonstrating if and how remote sensing data can be utilized to detect and locate harmful algal blooms. NASA's newly launched PACE Satellite was used to get imagery of the New York City and Long Island water bodies throughout a 10-week period. The harmful algal blooms were identified through their chlorophyll-a concentration and phytoplankton carbon content. Additionally, sea surface temperature data from the MODIS-Aqua satellite was used to analyze how warmer water bodies were causing algal blooms to flourish. Future research aims to better refine the reach of satellite data, as well as incorporate this technology at the local water management level.

8.

Design of a Spherical Antenna Positioner

Researcher: Miko Amican
Advisor: Dr. Elliott Williamstyer

A new antenna measurement range is being designed in Adams Hall. Antenna measurements are used to analyze the performance characteristics of an antenna, which are necessary for the reliability and efficiency of wireless systems. In the range, a probe antenna measures the power radiation from an antenna under test (AUT). As the orientation of the AUT is manipulated, a 3D map of the power radiated in each direction can be measured. To achieve this, antenna ranges use a device called a “positioner” to change the orientation and control the motion of the antenna.

The design process began by defining performance specifications informed by research of existing technology. These specifications were refined as alternative concepts and equipment were considered. Since the lab will be used for research purposes, the precision requirements of the positioner were prioritized. One of the challenges of this project concerned the ability to meet performance criteria whilst remaining under budget constraints. Initial CAD models of the positioner were developed to generate rough estimates for the mass and inertia. Afterwards a Simulink model was developed to perform a dynamic analysis of the positioner to determine the speed-torque characteristics required for motor selection.

Currently, multiple rotary stages are being evaluated for selection. Rotary stages are a device typically used in metrology applications and rotate an object about a single axis where precise angular positioning is required. Once selected, custom mounts will be designed to interface the rotary stage with the frame of the positioner. After the final design and simulations are completed, components will be ordered and assembled.

9.

Wound Image Segmentation using Deep CNN Architectures

Research Team: Daniel Doyon, Raquel Brown
Advisors: Dr. Mauro Caputi, Dr, Yimin Zhao

Wound closure is a tedious process that depends on the skill of the individual surgeon, and is prone to non-precision. Robotic Process Automation (RPA) has the potential to overcome inherent limitations of humans and offer to automate and streamline this process faster and more precisely than humans. The segmentation of wound regions from patient images is critical for quickly marking the region of interest (the wound area), which can play a significant role in wound closure, and is the objective of this research. We implemented a wound image segmentation model of high accuracy via Deep Convolutional Neural Networks (CNN), which can be integrated into a gantry robotic wound system that can segment wounds on foreign data. Using this information, a PC controller can then automatically detect the edges of the wound using gray-scale processing, and guide the gantry system apply polymignytes fixtures to close the wound.

The U-Net architecture was used as the main framework of the model, which is the gold standard for medical image segmentation with over 71,000 citations as of august 2024 [1]. The problem is that the base U-net architecture is limited in its accuracy. Tweaks must be made to achieve a high-performance model, especially for specific use cases like wound image segmentation.

The goal of this research was to maximize the dice score output by the model. The dice score is a measure of the overlap between the predicted segmentation mask and the ground

(continued)

9. *(continued)*

Wound Image Segmentation using Deep CNN Architectures

truth mask. Several methods, such as data pre-processing, hyperparameter tuning, architecture tweaks, and model ensembling, can be the difference between a model that performs subpar, and groundbreaking.

So far, we have been able to achieve a dice score of accuracy .937, or 93.7%. With the implementation of better pre-processing and post processing methods, we may be able to increase the accuracy by a wider margin. Even with such high accuracy, the precision of the model is of the utmost importance in use cases like medical imaging. A 100% accurate model is nigh impossible at the risk of overfitting. So, getting close to .95 - .98 is the objective we should strive for in the future.

References

[1] Liu et al., “Ore image segmentation method using U-Net and Res_Unet convolutional networks”, RSC, Issue 16, 2024.

10.

Mojo & Modular Programming

Researcher: Lidia Lojewski
Advisor: Dr. Edward Currie

Mojo is a new programming language bridging the gap between research and production by combining Python syntax and ecosystem with systems programming and metaprogramming features. Despite Mojo being a new standalone language, it is designed to become a superset of Python over time. Mojo is designed with simple and composable abstractions, it abstains from unnecessary punctuation, and it is built with the dynamic metaprogramming features. MAX (Modular Accelerated Xecution) is a unified set of APIs (Application Programming Interface) as well as tools that help build and deploy high-performance AI pipelines. It is built using a first-principles methodology and modern compiler technologies to ensure its programmability and scalability for all future AI models and hardware accelerators.

Mojo has a greater advantage in comparison to Python when it comes to speed, as Mojo can execute code 35,000x faster than Python, while also having the compatibility of several Python libraries. While Python has a single execution thread, Mojo has the ability to use parallel processing across multiple cores. Mojo has a higher performance than Python, as it is near C/C++ performance with Python-like syntax and it is ideal for machine learning and scientific computing. Mojo also allows users to choose between static and dynamic typing based on their specific needs and has built-in memory features to prevent massive leaks and errors.

For my project, I have tested the difference in speed between Mojo and Python as well as executed Llama2 in Mojo, which is a family of pre-trained and fine-tuned large language models (LLMs), to demonstrate the difference in performance compared to the other languages. I have also tested MAX by executing Llama3 (which has better performance than Llama2) for quantization to reduce memory usage, and tokenizer to enable efficient text preprocessing, which helps the generative AI model handle the natural language input with accuracy and high performance.

11.

Optimization of Plant-Derived Extracellular Matrix Structure and Mechanics for Vascular Repair

Researcher: Gianna Imeidopf
Advisor: Dr. Nicholas Merna

With the prevalence of cardiovascular disease, there is an urgent need for the development of novel biomaterials for clinical applications. However, synthetic materials frequently fail when used for small-diameter vascular repair. Common issues include thrombosis, initial hyperplasia, and as with any biomaterial, immune rejection. Fortunately, recent developments in the field of bioengineering have enabled scientists to explore plant-based tissues as a solution. They can be engineered to have minimal plant DNA, and they can be seeded with the proper cells to be implanted in humans. This lowers the chances of rejection. Since plant-based biomaterials are so new, there is limited research on this topic. This summer I was fortunate enough to begin researching methods that will lead to standardization of plant use in biomaterial engineering. I studied different decellularization detergents, enzymes, durations in harmful yet necessary solutions, and locations of sample collection. The differences and impacts were evaluated through tensile testing and histological processing and image analysis. As we reduced decellularization time, there was less damage to the structure and mechanical properties of the plant's extracellular matrix. Through optimization of our protocols, we hope to improve the standardization of these emerging decellularization strategies for vascular repair.

12.

Optimization of Pressure-Flow Relations on Vascularization

Research Team: James Bacheller; Daniel Gosnell
Advisor: Dr. Sina Rabbany

This research aims to optimize the effect of pump flow rates on vascularization within a 3D culture. It has been previously shown that the incorporation of flow significantly increases the expression of certain transcription factors and decreases the permeability of endothelial cells. *In vivo*, endothelial cells are exposed to constant flow, making it essential to understand how the shear forces imparted by blood flow affect vessel lumen diameter, density, and fractal dimensions. The objective of this study is to promote neovascularization in a fibrin gel within a microfluidic platform and examine the role of gravity and varying flow rates on vessel stability. We aim to identify optimal strategies for creating physiologically relevant microenvironments to vascularize organoids *in vitro*. Understanding the influence of flow rate on endothelial cell behavior and subsequent vascularization will allow us to better mimic physiological conditions and recapitulate *in vivo* conditions.

13.

Anti-fibrotic Effects of Atorvastatin Delivered via Keratin-based Hydrogels

Researcher: Evan Carroll
Advisor: Dr. Roche DeGuzman

Implantations of degradable biomaterials for drug delivery or restoration and regenerative medicine still cause fibrosis and acute inflammation which may lead to chronic side effects. Additionally, the fibrous encapsulation can itself inhibit the local cells from proliferating and forming new healthy tissue. Current drug delivery models require a new method to decrease the adverse effects following implantation of biomaterials in patients and inhibit the activation signaling sent to quiescent fibroblasts. Statins are commonly prescribed medications used to decrease cholesterol by inhibiting 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase. Importantly, they have pleiotropic anti-fibrotic effects, possibly through the transforming growth factor beta signaling pathway which either prevents the activation of quiescent fibroblast or transdifferentiation of progenitors into myofibroblast. In this study, the bioactivities of statin, as Atorvastatin, were investigated when delivered using keratin-based hydrogels to use novel human-derived hydrogels as drug-delivery vehicles for disease management. The goal is to control the drug loading and release to produce a therapeutic effect for scar tissue inhibition.

The drug pharmacokinetic properties were measured using long-term absorption and release profiles. Two fibroblastic cell models were exposed to control keratin and alginate hydrogels and those hydrogels conjugated to calcium ions and Atorvastatin (n = 3). The drug-carrying hydrogels were implanted subcutaneously into mice to observe their immune response after four weeks (n = 4). The areas of interest were embedded, sliced, and stained using Masson's Trichrome protocol to detect the thickness of the fibrous tissue encapsulation.

Keratin gels alone have many negative charges to more freely bond to the Atorvastatin calcium ions than keratin calcium gels do. This was observed when keratin gels absorbed and released the drug faster than the keratin gels with calcium ions intertwined in the matrix did by day 7. The mass of drug was normalized in all groups after calculating their timed release, and the keratin-based hydrogel conjugated with calcium ions and Atorvastatin inhibited fibroblast activity. Drug-absorbed hydrogels were observed to have thinner fibrous encapsulation around the implant site than untreated hydrogels, and less immune cells were observed in the area. Further clinical and preclinical studies must be conducted to determine the immune benefits of infusing Atorvastatin in biodegradable hydrogel matrices implanted subcutaneously or elsewhere in the body.

14.

The Effect of Freestream Turbulence Level on the Drag & Lift Coefficients of Rotating Baseballs and Spheres

Research Team: Kaita Odani, Kyle Raymond
Advisors: Dr. David Rooney, Dr. John Vaccaro

An experimental investigation was carried out on the effect of freestream turbulence on the drag and lift coefficients of rotating baseballs and smooth spheres in an Aerolab Open-Circuit Subsonic wind tunnel. An ATI Gamma Load Cell capable of measuring both forces and torques in the x, y, and z directions was used to collect force data on an MLB Official Baseball and a 3D-printed smooth sphere. The baseball and the sphere were connected through a steel shaft to a DC motor capable of rotating the baseball or sphere at rates ranging from 0 to 2000 rpm. The area of the shaft exposed in the wind tunnel was shrouded with a 3D-printed airfoil in order to minimize the interference on drag and lift measurements. The baseline turbulence level of the wind tunnel was measured to be 0.3%, and acrylic turbulence screens were designed and placed upstream of the test section which caused turbulence levels of 1.6%, 4.5%, or 7.8%. For each turbulence level, data were collected at freestream velocities from 10 m/s to 30 m/s in an increment of 1 m/s, and at rotation rates of 0 rpm, 400 rpm, 1600 rpm, and 2000 rpm.

(continued)

14. *(continued)*

The Effect of Freestream Turbulence Level on the Drag & Lift Coefficients of Rotating Baseballs and Spheres

Test conditions covered a Reynolds number range from 25000 to 150000, which meant that the flow field was largely subcritical, although the higher levels of freestream turbulence (FST) between 1.6% and 7.8% can trigger critical flow, where the boundary layer on the sphere becomes turbulent, and therefore separates further downstream, resulting in decreased drag coefficient. This phenomenon was noted for the smooth sphere in the absence of rotation, but the baseball experienced a slightly higher drag with increasing FST. When a low spin rate is introduced, the boundary layer on the side of the sphere or baseball that is moving in the same direction as the oncoming flow moves further back, whereas the opposite occurs on the side rotating into the oncoming flow, producing a positive lift coefficient. However, for the smooth sphere, as the spin rate increases, the direction of the lift force reverses because on the side rotating into the flow the separated flow reattaches and then separates again as a turbulent boundary layer further downstream than on the opposite side. The level of FST is a major factor in achieving this lift reversal, known as the Inverse Magnus Effect. For example at a Reynolds number Re = 100000, and moderate spin rates (less than 2000 rpm), the sphere never experienced negative lift for a background turbulence of 0.3%, but when the FST level increased to 1.6%, the lift was negative for almost all spin rates examined. Importantly this complete reversal of lift force was only recorded for the baseball in the absence of spin, and was weaker than for the sphere under the same conditions, and was also more persistent when the FST level was higher. The rough surface of a seamed baseball prevents the reattachment on the side of the ball rotating into the oncoming flow, minimizing the possibility of a lift reversal.

15.

An Approximation Framework for Making Regex Malware Signatures Zero Knowledge Proof Friendly

Researcher: Artem Pugach
Advisor: Dr. Xiang Fu

This summer project is one part of a research project which aims to prove that the Linux system is malware-free in zero-knowledge. For example, proving that encrypted collections of emails are malware-free without decrypting the emails. However, there is a challenge with this proof. A zero-knowledge prover requires a DFA (Deterministic Finite Automaton), and the translations to DFA are enormous. With this challenge comes the goal of approximating the regex signatures and determining the sufficiency of these approximations.

This project addresses the challenge by developing a framework for translation and a unit testing framework for verification. The work is divided into two parts: verifying that the approximation and translation processes preserve the integrity of the original signatures, and studying the "loss" of precision that may occur during approximation. The translation framework handles backreferences, \w, \W, \s, etc., and the testing framework shows how the approximations differ from the original regular expressions.

Sometimes the approximations cannot be made identical to the original signature and this is when the problems with the false positives occur. Most of the false positives occur when a file is not correctly recognized and is categorized as infected while in reality it is malware-free. The work on this project is still ongoing and is a part of the superset that Dr. Xiang Fu is working on.

16.

The Nature of Generalizing Same-Different Visual Relations in Deep Learning

Researcher: Aisha Ahmad
Advisor: Dr. Gerda Kamberova

Deep neural networks are known to have the capability of achieving human-level performance on various object recognition benchmarks. Despite this, neural networks typically struggle to learn simple abstract relations, such as classifying two objects as the same or different. A recent study “Deep Neural Networks Can Learn Generalizable Same-Different Visual Relations” by Alexa Tartaglini, et al. claims that with the right architecture, deep learning networks can learn this abstract relation. Our study used Tartaglini’s best performing model, a CLIP-pretrained visual transformer, and tested its performance on an adversarial and a double.These are specialized inputs designed to confuse a neural network. While both confuse a deep learning classifier MobileNetV2 model, adversarial images are distinguishable from the original image by humans and doubles are indistinguishable images from the original. Although both images are misclassified by the MobileNetV2 model, the latter poses a great threat to applications that encounter them. We study how same-different abstraction generalizes when doubles enter the scenario. We generate adversarial and double inputs for a class of natural images and test the model’s performance on these specialized inputs to determine if deep neural networks can learn the abstract relation of same-different.

17.

Comparing tournament rankings by distance and back edges

Researcher: Cade Ferguson
Advisor: Dr. Simon Shamoun

Effective tournament ranking strategies which focus on minimizing back edges have become increasingly relevant as they become applicable to a wide range of fields including social network analysis, biological networks, choice optimization and fraud detection. Helpful comparisons of different sorting algorithms and strategies for the problem as well as the impact of tournament design have been carefully considered in research over the past several decades. The GreedyFAS algorithm developed by Eades is determined to be one of the most effective fast algorithms for ranking tournaments. Through our research, we consider several scoring schemes for double round-robin tournaments and compare their rankings by overall points and by Eades’ GreedyFAS algorithm. The rankings are compared by Spearman’s Footrule distance between given rankings, Kendall's Tau distance between given rankings, and total number of back edges. We also account for team strength by assigning a probability of a win and compare the number of back edges in rankings for these tournaments. Because each scoring system and ranking method presents unique benefits and drawbacks, conflict and disagreement surrounding ranking strategies is a notable concern in many circumstances. When selecting a scoring system and ranking method for a tournament ranking problem, it may prove helpful to consider the discrepancy between rankings by different criteria. In circumstances where scoring criteria is subject to change or dispute this may be the best route.Our results aid in identifying a strategy which satisfies as many outcomes as possible. Additionally, our test results demonstrate that the same ranking method does not always satisfy the MFAS for every kind of tournament. The type of tournament and importance of accounting for team strength should also be considered when selecting a ranking method.

18.

Optimizing Memory Allocation in Cloud VMs with Accurate Cache Abstraction

Researcher: Mani Tofigh
Advisor: Dr. Jianchen Shan

The CPU cache is an important hardware resource that is used to accelerate data access, as it provides much lower latency compared to main memory and disk by multiple magnitudes. The Operating System (OS) relies on accurate cache information to store the hottest data in the cache with minimal conflicts for reduced data access latency. However, in the virtualized clouds, the virtual machines (VMs) provisioned using the hypervisor cannot provide an accurate abstraction of the virtualized CPU cache.

Misguided memory allocation by the OS could cause performance degradation for data and memory intensive applications, including AI, ML, and Data Science workloads. To capture such dynamics more accurately and facilitate the VM's OS in better utilizing the cache, we studied existing tools that aim to provide cache information in cloud VMs, alongside their problems and limitations, and we propose to develop a set of micro-benchmarks, named vCache, to reverse engineer the CPU cache information, by probing cache abstraction via cache eviction sets inside a VM in the virtualized Cloud. This work can provide more observability into the dynamic cloud resource and improve data-intensive workloads deployed in cloud environments.

19.

Optimizing Documents for Retrieval Augment Generation

Researcher: Banmeet Kaur
Advisor: Dr. Krishnan Pillaipakkamnatt

Large Language Models (LLMs) are advanced artificial intelligence (AI) programs designed to understand and generate human language by learning from vast amounts of text data. Despite their capabilities, LLMs often encounter issues such as hallucinations–instances where the model generates text that is factually incorrect, misleading, or entirely fabricated, though it presents this information with a high level of confidence.

Retrieval Augment Generation (RAG) attempts to increase the accuracy and reduce the hallucinations of Large Language Models (LLMs) when answering questions pertaining to a specific topic. RAG LLMs improve the quality of their responses by injecting relevant context to user prompts. This context is retrieved from databases that are built from documents supplied by the user herself. It has been anecdotally observed that document formats and layouts have a substantial effect on the speed, accuracy, and quality of RAG LLMs.

We are interested in methodically examining the impact of document formats and layouts on RAG systems and proposing optimal formats for various combinations of embedding and language models. In this process, we aimed to improve the performance of the Ollama LLM specifically for Hofstra-related information using RAG techniques.

Future research focuses on addressing an open question: how can bulletin content be formatted to be more language model-friendly?

LAST PAGE