CCSC NE '24

The Twenty-Eighth Annual
Consortium for Computing Sciences in Colleges
Northeastern Conference

April 12 – 13, 2024

Hosted by The College of Saint Rose
Albany, New York

Undergraduate Student Poster Abstracts

CCSC
CONSORTIUM FOR
COMPUTING SCIENCES
IN COLLEGES
ccsc.org

The College of Saint Rose

CONTENTS

Poster 5001

Using Funds of Knowledge and Funds of Identity to Assess Computational Thinking in K-12 Students

Annamaria Butler, Samantha Huck
Marsal Family School of Education
University of Michigan – Ann Arbor, Ann Arbor, MI

Faculty Advisor
Jennifer Randall
University of Michigan – Ann Arbor, Ann Arbor, MI

The lack of BIPOC representation (and the associated negative consequences) in STEM fields broadly, and in computer science specifically, is well documented. In response, the National Science Foundation's Computer Science for All (CSforAll) program seeks to provide access to computer science to all K-12 students, including those from racially and ethnically minoritized populations. Within this effort, computational thinking has emerged as a promising area of focus, specifically with elementary-aged students. Moreover, other scholars have encouraged the use of culturally responsive computing (CRC) pedagogical practices to facilitate more culturally appropriate CS experiences for BIPOC students. Although culturally appropriate instructional practices should be complemented with culturally appropriate assessment practices, the literature reveals that this is rarely the case. The project described here seeks to address this gap. Specifically, by leveraging students' funds of knowledge (FoK), we sought to uncover early elementary students' computational thinking skills. Using data collected from family interviews (in an urban New England school district) about their daily routines and practices, we developed two formative assessment tasks: the Packing Game and the Library Game. We then analyzed the clinical interviews of students' engagement with the tasks, identifying instances of CT and recognition of FoK.

As a result of the development of these two tasks, teachers possess additional tools for assessing their students' CT skills. The availability of these tools holds significant implications for educators who are looking to create a more inclusive and equitable classroom environment that supports the learning and success of all students, regardless of their sociocultural background. By providing teachers guidance on how to incorporate FoK into their teaching and expose minoritized children to CT and CS from a young age, they are better able to prepare them for future STEM or CS careers.

Poster 5002

Explorations in Cycle Sort; Running time at the cost of memory usage

Carter Kelly
Department of Mathematical Sciences
Eastern Connecticut State University, Windham, CT

Faculty Advisor
Garrett Dancik
Eastern Connecticut State University, Windham, CT

Various sorting algorithms exist; merge, bubble, insertion, and quick sort are generally the ones individuals are most familiar with. Cycle sort is a less known algorithm centered on the idea that for a random finite set, the set can be decomposed into disjoint "cycles" such that successive swaps return each element in the cycle to its correct location. When all cycles are complete the set is sorted. Cycle sort is primarily used to minimize writes, such as in flash memory which has a limited number of erase/write operations. For a random set of size n (no duplicates), cycle sort operates in quadratic time, $O(n^2)$, with auxiliary space complexity of O(1). For a set of size n where all values are [1, n]; time complexity is reduced to O(2n-1). O(2n-1) comes from that once a cycle is complete, there is no way to determine which elements have not been operated on without iterating though the set until a misplaced element is found, thus initiating a new cycle. In trying to improve time complexity at the cost of memory usage, the idea of a "run" was used, a sequence of consecutive sorted elements. By tracking the start and ends of runs, the next untouched element will be q+1 where the run is from zero to q. A hash table was used to store the start and end of runs using index as the hash function input. Statistical analysis was performed for values of n=10k,100k, and 1mil (with 10k replicates). We report various metrics such as the most values stored simultaneously in a hash table of size n, number of collisions for a hash table smaller than n, and the improvement in running time for dual processing (as cycles are disjoint).

Poster 5003

A New Engine to Support Map-Based Visualization of Parallel Algorithms in METAL

Michael Plekan
Computer Science
Siena College, Loudonville, NY

Faculty Advisor
James D. Teresco
Siena College, Loudonville, NY

I present a new addition to the Map-based Educational Tools for Algorithm Learning (METAL) Project. The goal of the project is to help computer science students learn and understand graph algorithms using map-based graphs. Students being able to see what the algorithm is doing can help with learning. During my time working on the project we designed a new engine for visualization of parallel algorithms. This includes a big boost in efficiency but also a platform that can be expanded on in the future. With the end goal of this new engine in mind, we decided to build it from the ground up after realizing the extensive reliance on serial execution in the existing engine. I started with building up the map manager to handle updating and displaying the maps that will be needed. This was done to make it easier for people implementing algorithm visualizations (AVs) in the future. I then worked on threads which handle keeping track of local data and per-thread execution location and a way for them to interface with the map manager. Next I worked on the scheduler which is where the simulated parallelism is done, picking which thread should run next. Lastly I worked on a new AV format that would allow for the parallelism to be easily implemented, along with some example AVs. These examples were basic parallel algorithms to show off the parallel simulation. Aside from the new parallelism, there was also a massive performance boost and a new UI to better suit the multi-map design. The poster will discuss an overview of what I worked on, and the challenges I encountered during the project. The full description and implementation features of the website will be presented in detail.

Poster 5004

Using an Agile Approach to Develop a Small Studio Inventory Management System

Eric Rosenhein, Maximus Streeter, Ryan Fanning, Ephrem Kassahun
Department of Computing Sciences
SUNY Brockport, Brockport, NY

Faculty Advisor
Sandeep Mitra
SUNY Brockport, Brockport, NY

We present our experience participating in a service-learning project in the Computing Sciences department at SUNY Brockport. The project goal is to create a web application to assist local artists in organizing/maintaining small studio spaces. Ultimately, the app may be used to track the nature, location, and quantities of materials available in the studio, determine shopping lists for inventory replenishment at regular intervals, calculate costs of orders to be placed, and display a portfolio of items. Development technologies include C#, Blazor, ADO.net, and SQL Server. A key aspect about this project is its student-led nature, with the faculty supervisor mainly playing the role of a consultant. Over the past academic year, a team of 4 or 5 students have adopted an Agile approach to the development process. These teams have sought to adopt the Agile self-organizing team approach and assumed responsibility for assigning and tracking work and progress. Students decide on the content of each Sprint backlog, manage the regular Scrum meetings, and focus on creating a solid (potentially shippable) prototype at the end of each Sprint. Also, the team focuses on continuous improvement of the system under development and on ensuring quality, through continuous interactions with potential customers. At the time of writing, the team has discovered that adopting techniques for writing extensible and maintainable code, with good modularization, should be a major focus of this project, to adapt to various types of artists and their required needs. Consequently, we have learned that instruction in such techniques is a necessary component in the onboarding process for new team members. Implementation of the initial set of features is expected to be completed by the time we present this work, and we hope to present our experiences in learning how student-run teams can be productive in delivering a real-world product.

Poster 5005

ChemBase: Elevating Course Maintenance with Revolutionary Intelligent Chemical Inventory Solutions

Spencer Wondolowski, Kenneth Yao, Timothy Fraedrich, Rohit Reddy Tangadipalli
Computer Science
University of New Haven, West Haven, CT

Faculty Advisor
Mehdi Mekni
University of New Haven, West Haven, CT

The golf industry is confronted with significant challenges in efficiently managing chemical inventories for course maintenance. Current methods lack sophistication and precision, resulting in inefficiencies, higher costs, and potential environmental impact. These shortcomings hinder optimal chemical application practices and compromise the sustainability of golf course maintenance. To tackle these pressing issues, we propose the development of an intelligent, portable, chemical inventory solution named ChemBase. Our Solution will empower golf course professionals to effectively track, manage, and analyze their chemical inventory. Through cutting-edge technologies, the application will provide real-time inventory tracking, data analytics for informed decision-making, and formulation calculations for precise chemical applications. By optimizing resource utilization, reducing costs, and promoting sustainability, ChemBase will serve as a catalyst for improving maintenance practices in the golf industry. The successful development and deployment of ChemBase will yield tangible benefits for golf course management professionals. Key outcomes include enhanced operational efficiency, supported by streamlined chemical management processes, improved decision-making capabilities through real-time data analytics, and the promotion of sustainable maintenance practices. ChemBase is poised to revolutionize course maintenance in the golf industry, offering a sophisticated yet accessible tool for professionals. Assisted by our sponsor, Jason Dogowitsc, the VP of Product Development at the Plant Food Corporation, ChemBase benefits from his institutional knowledge and expertise in plant health. His invaluable contributions have been pivotal in refining our system, enhancing its effectiveness, and ensuring it meets the evolving needs of golf course professionals. Through collaborative efforts, ChemBase continues to evolve as a sophisticated yet accessible tool, poised to revolutionize maintenance practices in the golf industry.

Poster 5006

Stress Mastery Advancing Responder Training (SMART-VR)

James Mok, Sean Vargas, Matthew Lamour, Kyle Muldoon
Computer Science
University of New Haven, West Haven, CT

Faculty Advisor
Mehdi Mekni
University of New Haven, West Haven, CT

In the demanding field of firefighting, firefighters often find themselves grappling with the profound impacts of high stress on the job. This challenge not only diminishes job efficiency but also takes a toll on their health. Recognizing the importance of firefighter training, it is important to note that it is hindered by its own set of challenges. Real-world training is expensive and lacks effective stress quantification methods. Moving training to virtual reality (VR) platforms could be a cost-effective solution, but integrating stress management features and realistic scenarios present new challenges. We, as part of our senior capstone project at the University of New Haven's computer science department, aim to address this problem. Our approach involves developing stress management features and firefighter officer scenario training within Pleiadian Systems' Fire Rescue VR software. We are utilizing Meta Quest VR headsets, the Unity development environment, and HeartMath hardware while adhering to agile scrum methodologies for efficient project completion. Our project focuses on creating a stress management solution that integrates real-time stress monitoring into the VR environment. Our goal is to demonstrate the efficacy of VR-based firefighter training, enhancing both training frequency and stress identification. The project's success relies on a solid partnership with Pleiadian Systems and our collaboration with the Cleveland Fire Department for validation and verification of our proposed software solution.

Poster 5007

Comparing Touch-Based Coding to Traditional Computer Programming Methods

Jannice Salazar
Computer Science
Caldwell University, Caldwell, NJ

Faculty Advisor
Vladislav D. Veksler
Caldwell University, Caldwell, NJ

This study presents a heuristic evaluation of touch-based (mobile-friendly) coding prototypes, intending to make computer programming accessible and approachable.

We designed two touch-based coding prototypes using Figma, one similar to Scratch [1] and another to Python [2]. We ran heuristic evaluations on these prototypes and Python (run on Google Colab [3]) and Scratch coding environments. All prototype evaluations were performed using an 11" Android tablet and a single coding task that involved variable assignments and output. The evaluations focused on standard UI/UX metrics and employed four evaluators, as suggested by Neilsen [5,6]. Additionally, we recorded time-on-task data and included a brief post-evaluation survey.

Results showed a unanimous consensus in favor of tapping instead of dragging in Scratch or typing in Python, with reviewers expressing that tapping felt more “natural” and “easier”. Reviewers preferred portrait mode, mentioning that compatibility with touch devices, consistency with other mobile apps, and displaying more code on screen were the primary reasons for this preference. Surprisingly, all reviewers preferred Prototype 2, looking more like a text-based programming environment, over Prototype 1, which looked like the supposedly "easier-to-learn" Scratch visual coding environment. Time-on-task data revealed that participants were faster using touch-based than traditional coding methods, with Prototype 2 being the overall winner.

Our research suggests the potential for touch-based programming to offer a developer-friendly, more efficient, and mobile-accessible alternative to traditional computer programming methods. Future work will include developing a full-featured touch-specific programming language and a larger comparison study for significance testing.

[1] Scratch. https://scratch.mit.edu/about
[2] Python. https://www.python.org/
[3] Colab. https://colab.research.google.com
[4] 10 Usability Heuristics for User Interface Design. Nielsen Norman Group. https://www.nngroup.com/articles/ten-usability-heuristics/
[5] Nielsen, J. (1995). How to conduct a heuristic evaluation. Nielsen Norman Group, 1(1), 8.

Poster 5008

Malicious Accounts Detection on Social Media using Machine Learning

Ajay Bhattacharya
Computing Sciences
SUNY Brockport, Brockport, NY

Faculty Advisor
Adita Kulkarni
SUNY Brockport, Brockport, NY

We present our experience using machine learning techniques to detect and categorize malicious accounts on major social media platforms. The goal of this research is to create accurate, efficient machine learning models for detecting and categorizing social media accounts, enabling the general public to scan profiles in real-time through a web application. Over the past several years, the amount of money being lost to social media scams has grown exponentially, due to cybercriminals growing more sophisticated in their creation of bots and other fake accounts on social media platforms. We created a simple web-based user interface, where a user could enter a link to a suspicious profile, and receive insights based on the results of our machine learning models. We obtained a dataset of Twitter users and their affiliated Tweets that had been divided into account categories, including genuine accounts, fake followers, and various types of spambots. We experimented with various data processing and machine learning techniques, with a focus on the text content of the Tweets, until we were satisfied with the accuracy of the results. In the end, we determined that the pre-trained large language model DistilBERT was the most effective way of using text data to classify these accounts. We also created a risk score function by training a logistic regression model on some of the numerical features from the dataset and pulling the underlying probabilities from the predictions. Our poster will discuss the various machine learning techniques we experimented with, and some of the challenges that we encountered. The accuracy results and the cloud architecture that we deployed the project on will be discussed in detail.

Poster 5009

CSP-Style Parallel Composition in Golang

Daniel Goldman
Computer Science
Vassar College, Poughkeepsie, NY

Faculty Advisor
Marc Smith
Vassar College, Poughkeepsie, NY

Golang is a modern, mainstream programming language with support for Hoare's CSP-style concurrency. In particular, Golang provides direct support for processes in the form of asynchronous function calls and synchronous channels for inter-process communication. Golang also provides a select statement to, among other things, avoid deadlock while attempting to read from multiple synchronous channels. The only CSP feature not directly supported by Golang is a parallel composition operator, like the PAR keyword in occam. What Golang does provide is a sync WaitGroup, which is a lower-level abstraction for implementing a barrier. The goal of this project was to implement occam's PAR operator in Golang, utilizing Golang's sync and reflect packages, along with Golang's support for variadic functions.

Golang's motto is: do not communicate by sharing memory, share memory by communicating. Golang processes communicate with each other via synchronous channels. CSP-style processes are composable, i.e., the parallel composition of a group of processes is itself a process. By implementing occam's PAR operator in Golang, we hope to provide a higher level of abstraction for developers than Golang's sync WaitGroup.

We implemented a pair of functions in Golang to support CSP/occam-style parallel composition. PAR is a higher-level function, whose inputs are the processes (asynchronous function calls) being composed. To implement PAR using Golang's sync WaitGroup, we needed to abstract away the sync WaitGroup. To do this, we implemented a second higher level function, PARify, which allows developers to prepare processes for parallel composition without being concerned with sync WaitGroup. To implement PARify, we used Golang's reflect class for reflection, and variadic functions. PARify's input parameters include the process and its parameters, and the function returns a process whose only parameter is a sync WaitGroup. PAR ensures that the same sync WaitGroup is passed into all the PARify'ed functions.

Poster 5010

A Layer-Freezing Approach for Reduced Backpropagation Demand

Nika Gedenidze
Computer Science
Caldwell University, Caldwell, NJ

Faculty Advisor
Vladislav D. Veksler
Caldwell University, Caldwell, NJ

Backpropagation in traditional Convolutional Neural Networks (CNNs)[1] faces challenges due to its intensive computational demands. Despite these challenges, CNNs employ backpropagation across all layers, utilizing filter layers to improve learning. This study investigates a way to make CNN training more efficient by reassessing the necessity of backpropagation for the network's filter layers. It introduces a novel approach, inspired from the human brain, that enhances training efficiency through a selective layer-freezing strategy. This strategy adopts a cyclical training regimen, systematically alternating between freezing and activating individual layers for training across epochs to reduce the resource demands typically associated with backpropagation in CNNs.

We conducted simulations to train neural networks on the HasyV2[2] dataset and comparison focused on two models: a traditional CNN model and an innovative model employing a freezing schedule. Both models shared a similar architecture, featuring convolutional, dense, and dropout layers, and were trained until they each achieved a training accuracy limit of 80%. Figure 1. Accuracy and Efficiency of Traditional and Scheduled CNNs. (Figure elided). The average accuracy and efficiency results of two models throughout 10 simulations are shown in Figure 1. The traditional CNN model met the accuracy threshold within 6 epochs, taking 543.61 seconds to train and achieving a validation accuracy of 76.86% (training accuracy = 80.60%). In contrast, the CNN model with a layer freezing schedule also reached the threshold in 6 epochs but was faster, at 517.24 seconds, and exhibited a higher validation accuracy of 78.02% (training accuracy = 80.10%). These findings highlight the potential efficiency and accuracy benefits of using a layer freezing schedule in CNNs, particularly for larger and more complex datasets. Future studies will focus on further exploring the advantages of scheduled layer freezing in CNNs on larger datasets and comparing the performance of these models with traditional CNNs, including the application of transfer learning to assess efficiency and accuracy improvements.

[1] doi:10.1162/NECO_a_00990
[2] https://www.kaggle.com/datasets/guru001/hasyv2/metadata

Poster 5011

Forensic Investigative Realistic Environment (F.I.R.E)

Aliyah Jones, Tony Khamphouy, Kaylie Neal, Julia Roberts
Computer Science
University of New Haven, West Haven, CT

Faculty Advisor
Mehdi Mekni
University of New Haven, West Haven, CT

As education transitions to online platforms, challenges arise for classes necessitating in-person training, such as fire investigation. Educators traditionally rely on physical burn rooms for skill practice, but these are costly to build and maintain, and is limited in student capacity. In response to the growing trend of online education, the Forensic Investigative Realistic Environment (F.I.R.E) project proposes utilzing Virtual Reality (VR) to effectively support arson investigation, providing an immersive experience with engaging interactions while mitigating associated costs, risks, and scalability. The primary objective involves creating virtual arson scenes for user interaction, ensuring educational value by incorporating proper investigative procedures, and eliminating the constraints linked to physical burn rooms. Our approach includes recreating an existing burn room in VR, complete with detailed 3D models and functionalities allowing students to interact with the instructor and each other along with items and evidence within the scene. The expected outcomes of this project include a virtual replication of a real-world burn room, empowering students to enhance their investigative skills in a virtual setting with the same authenticity as a physical one. To develop an immersive learning platform this project is in cooperation with the Henry C. Lee College of Criminal Justice and Forensic Sciences at the University of New Haven. Using this resource, we will conduct experiments and engage university students to gather data and validate the user platform. The main contribution of this paper is to propose the first project for online arson investigation training. Challenges anticipated throughout the project include integrating our project into our sponsor's current environment, refining 3D models for accuracy, developing a notification system to give feedback, adding an investgation-based toolset into the scene, and providing students with all the necceasry features needed to replicate a real-world burn room investigation.

Poster 5012

Navigating the Iterative Process Between Machine Learning and Linguistics Research

Cameron Jester, Madeleine Guettler
Math, Statistics and Computer Science
Simmons University, Boston, MA

Faculty Advisor
Nanette Veilleux
Simmons University, Boston, MA

This work is part of an ongoing research project funded by the National Science Foundation on mapping prosody (how something is said) to pragmatic meaning (what is implied). As Data Science students, we were responsible for developing machine learning models for the project. This project emphasized data cleaning and validation before and during modeling which led to uncovering areas to improve upstream annotation and collection processes. Making such discoveries ultimately allowed for data recovery, which prompted principal researchers to ask new questions about the data.

To begin the iterative process, we student researchers analyzed a robust data science pipeline created by a Principal Investigator to investigate the data extraction and manipulation processes implemented. We then constructed a random forest model to identify critical variables in the dataset, which yielded an accuracy of 86%. After investigating variables identified as important in the model, anomalous f0 (pitch) slope data was discovered in the data. Such a discovery could point to issues with the linguistic hypotheses or with the data representation; careful examination revealed a subtle data collection inaccuracy. After removing instances with spurious f0 values, our model's accuracy decreased by 1.2%. However, while performance suffered slightly, our model had a more credible performance.

Reporting this discovery back to principal investigators led to a change in the annotation system, which was instituted to recover data that would otherwise be excluded from analysis. The random forest model also uncovered what parts of the utterance contributed most to perception of meaning, leading to further fruitful linguistic inquiry.

This iterative process is particularly useful in this work with prosody-meaning mappings where linguistic theory does not provide a clear framework. Having the privilege to work at the intersection of data science and linguistics has provided us with a critical understanding of the back-and-forth nature of research.

Poster 5013

Spotify Emotions Analyzer: Predicting Song Lyric Emotions with BERT LLM

Alexander Elguezabal, Derek Costello
Computational Sciences
Merrimack College, North Andover, MA

Faculty Advisors
Christopher Stuetzle, Zachary Kissel
Merrimack College, North Andover, MA

This research project presents Spotify Emotions Analyzer, a large language model-based approach to analyzing the emotional content of song lyrics from a user's Spotify history. Leveraging web scraping techniques, Spotify Emotions Analyzer collects lyrical data directly from the user's listening history, enabling a personalized emotional analysis. The core methodology revolves around employing a BERT (Bidirectional Encoder Representations from Transformers) Large Language Model (LLM) trained on the Google GoEmotions dataset. The GoEmotions dataset consists of 58,000 human-annotated Reddit comments extracted from popular English-language subreddits data matched to one of the following emotional categories: admiration, amusement, anger, annoyance, approval, caring, confusion, curiosity, desire, disappointment, disapproval, disgust, embarrassment, excitement, fear, gratitude, grief, joy, love, nervousness, optimism, pride, realization, relief, remorse, sadness, surprise, and neutral.

Through the utilization of this sophisticated language model, Spotify Emotions Analyzer predicts emotional weights for sentences extracted from song lyrics. The emotional weight predictions are then aggregated to classify sentences of lyrics into one emotional category.

To enhance user experience and comprehension, Spotify Emotions Analyzer presents the emotional analysis results using interactive infographics within a React-based web application. The web application and server are cloud hosted. This integration of infographics not only enhances the accessibility of emotional analysis but also provides a user-friendly platform for exploring and understanding the emotional dimensions of music consumption. The combination of advanced analytical techniques with user-centric visualization methods positions Spotify Emotions Analyzer at the forefront of personalized emotional analysis tools for music enthusiasts.

Reference: https://github.com/Frostfire25/Spotify_NLP_Service

Poster 5014

Cheap Monocular Depth Estimation for Small Drones

James Biondi
Computer Science
Saint Anselm College, Manchester, NH

Faculty Advisor
Michael Huelsman
Saint Anselm College, Manchester, NH

We present our research into cheap monocular depth estimation for small drones. Our goal was to take a video feed from a small single camera drone and pass the images through a convolutional neural network (CNN) on a remote device to create a depth map of that image in real time. These images would then be sent back to the drone where it would be used to assist in autonomous flying and avoiding obstacles by giving an estimate of how far objects in view are away from the drone. Depth perception is typically achieved by using two or more cameras, or by using technology such as radar or infrared. We chose to not use these methods as they are both heavy and expensive, and we were creating a cheap lightweight solution for single-camera drones. To accomplish this, we trained a CNN on a dataset of indoor images and their corresponding depth maps to be able to estimate the depth of an image that is passed through it. One of the biggest constraints with our research was the time it took to train the model. We did not have access to any large datacenters that can quickly train models, so we instead used relatively cheap hardware. Because of this evaluating a particular architecture often took days or weeks, extending the training time and limiting the scope of the project. While progress was made towards finding a solution to the problem, we did not find any model architectures that met our needs in the time we were given. Given further development and a better understanding of what makes a good architecture, we may find a solution to the problem.

Poster 5015

Doc-Talk, Apology & Disclosure Simulator

Antonio Craveiro, Michael Daly, Danny Farhart, William Hancock, Gabriella Zola
Computer and Data Sciences Department
Merrimack College, North Andover, MA

Faculty Advisors
Christopher Stuetzle, Julie LeMoine (iCELS)
Merrimack College, North Andover, MA

The project is a collaborative project between Merrimack College's Computer and Data Sciences Department and the UMass Chan Interprofessional Center for Experiential Learning and Simulation (iCELS) aimed at addressing the critical need for immersive and accessible training in healthcare communication. Recognizing the importance of effective Disclosure and Apology (D&A) skills in healthcare, the project's objective is to develop a Virtual Reality (VR) application that facilitates the practice and refinement of these skills among healthcare professionals and students. The VR application offers three immersive scenarios simulating real-world situations where healthcare providers engage in D&A conversations with patients and their families. These scenarios include: 1) Medication Error - Low Acuity (Not life threatening), 2) Medication Error - High Acuity (Life threatening), and 3) Delayed Diagnosis (Diagnosis could have been made sooner). Each scenario is carefully designed to provide learners with opportunities to practice D&A skills and receive personalized feedback to enhance their proficiency. Through immersive VR experiences, users can navigate challenging conversations with empathy, integrity, and professionalism, ultimately contributing to improved patient care and safety. The project requires the innovative use of VR technology to create a dynamic and interactive learning environment for D&A skill development. By leveraging VR technology, learners can engage in realistic scenarios that closely resemble clinical practice, bridging the gap between theory and real-world application. Feedback and surveys will be conducted to assess the effectiveness of the VR application in enhancing D&A proficiency and learner confidence. This collaborative effort between academia and healthcare practitioners underscores the importance of experiential learning in healthcare communication and the potential of VR technology to revolutionize training methodologies in the field.

Poster 5016

Revolutionizing Chess: The Development of a Smart Chessboard for Enhanced Game Analysis and Digital Integration

Joshua Kenna, Richard Kaden, Akashdeep Damrath
Computer Science
Widener University, Chester, PA

Faculty Advisor
Suk-Chung Yoon
Widener University, Chester, PA

In our project within the Computer Science Department at Widener University, we embarked an endeavor inspired by our personal engagement with chess and a vision to integrate it with the digital analysis capabilities of today. The core innovation of our project involves the integration of magnetic sensors within the chessboard and magnets in the pieces, coupled with an Arduino to track piece positions. This data is then processed by a Raspberry Pi, converting the game states into standard chess notation and uploading it to Lichess for post-game analysis. The selection of appropriate sensors and the integration with the Lichess API were pivotal moments in our project. Additionally, the development of a mobile application to interface with the Raspberry Pi and manage Lichess accounts further exemplifies our project's aim to bridge traditional chess play with modern technology. Our endeavor highlights our technical prowess and innovative approach, underscoring our dedication to enhancing the chess-playing experience through technology. Our poster will detail the conceptualization, challenges, and development process of our smart chessboard, showcasing the integration of traditional chess with digital technology. This presentation aims to share our journey, the solutions we've implemented, and the unique contribution our project makes to the intersection of technology and traditional gaming.

Poster 5017

Storytelling with Natural Language Processing

Stacey Do
Computer Science and Mathematics
SUNY Brockport, Brockport, NY

Faculty Advisor
Ning Yu
SUNY Brockport, Brockport, NY

The project endeavors to develop an interactive storytelling chatbot through the integration of Google Vision API and OpenAI API using natural language processing and deep learning algorithms. This collaboration aims to create a machine-learning model capable of generating storytelling narratives based on given images. Leveraging Google Vision API, the system extracts diverse features from images, including landmarks, objects, colors, text, and logos, using optical character recognition. Subsequently, OpenAI API processes the extracted information and produces concise image descriptions through natural language processing. Big companies such as Google and OpenAI have already made huge success in AI technology but there is still much continuous growth that can be done. The key significance of this project lies in the seamless fusion of visual understanding and language generation, resulting in a compelling and personalized storytelling chatbot. The goal of this project is to explore various methods to synchronize image analysis and natural language processing, offering users an immersive and engaging narrative experience. Ultimately, the chatbot has the potential to create a diverse range of user needs, enhancing education, entertainment, creativity, and accessibility while displaying the possibilities of cutting-edge API technologies in meaningful ways. The poster will display the processes of integrating the APIs together along with the natural language processing tactics to produce meaningful storytelling experiences.

Poster 5018

Deeper Understanding of Depth Sensing Technology

James Petkin
Data Science
Southern Connecticut State University, New Haven, CT

Faculty Advisor
Winnie Yu
Southern Connecticut State University, New Haven, CT

This research focuses on depth perception and how the current depth-sensing technology can improve image recognition and biometric authentication. Against the backdrop of evolving AI technologies and deepfake practices, this study aims to explore and evaluate the vulnerability in authentication using 3D (with depth) versus 2D images. We employed the StereoBM algorithm from OpenCV to compute disparity maps from pairs of stereo images, mimicking human depth perception. The algorithm introduces a function, extract_2d_features, which utilizes grayscale conversion, histogram analysis, and the Canny edge detector to extract comprehensive 2D features from images. Similarly, extract_3d_features are defined to analyze depth maps through Gaussian filtering, gradient computation, and depth magnitude assessment, forming a multidimensional feature vector. Real-time image processing is demonstrated through continuous frame capture from a camera feed, applying Laplacian filters and Canny edge detection to enhance and visualize edges. Finally, we explored authentication between 2D and 3D images using ORB feature detection and brute force matching in OpenCV, establishing a secure feature comparison and verification methodology. The findings will enhance the understanding of depth sensing technology and inform its secure application in facial recognition and broader implications for the field's future trajectory.

Poster 5019

Proximity Mapping and Spatial Analysis: Optimizing Ithaca's Transit Network

Caitlyn O'Dell, Cole Wissink
Computer Science
Ithaca College, Ithaca, NY

Faculty Advisor
Ali Erkan
Ithaca College, Ithaca, NY

We extend our gratitude to the Community Foundation of Tompkins County and Ithaca College for their grant funding. We also thank the Computer Science Department of Ithaca College for their support. The goal of the project was to identify gaps in efficiency in TCAT routes and propose a system to determine a measure we refer to as average walk/distance; for all households on a bus route, what's the average distance to the closest bus stop? The project utilizes qGIS to determine the distance every address in Ithaca has to walk to access the closest bus stop and taking an average distance for the entire city. We aimed to create a system where when changes to routes or new routes are proposed, this system can be utilized to determine whether the average distance to a bus stop in the city has changed, making it easier to make adjustments to a transit system that is heavily utilized. The Tompkins Consolidated Area Transit System, known commonly as TCAT, is the transit infrastructure utilized by the Ithaca/Tomkins region. This transportation system is vital to the population in the area. Therefore, making sure this system operates at the benefit of the population is crucial to better public transportation as well as reducing the impact individual vehicles have on the environment. This project has created a mapping system that calculates the average distance each household has to walk to the nearest bus stop so that new proposed routes can be layered over it to determine if the new routes will decrease the average distance to a bus stop for residents.

Poster 5020

Analyzing Engagement and Growth in YouTube's Creator Economy

Samprada Pradhan
Computing Sciences
SUNY Brockport, Brockport, NY

Faculty Advisor
Derya Ipek Eroglu
SUNY Brockport, Brockport, NY

Today creators have a significant influence in the digital sphere, with their creativity and skills that help them reach and monetize social capital across various platforms. According to a 2022 survey, there are more than 300 million creators across nine large nations, including more than 85 million Americans, which accounts for a market of $104.2 billion. In this dynamic ecosystem with four main stakeholders, (i.e. creators, viewers, businesses, and the platforms), we focus on the interplay between the business decisions of creators and their viewers. We study how YouTubers build their brand, approach other businesses, and how they influence user perceptions. Our plans include analyzing YouTube video data using video descriptions, comments, and community engagement to further build a framework for gaining insights which can ultimately inform data-driven strategies for content creators.

Poster 5021

Mitigating the Unmanaged Attack Surface in Cloud-Based Architectures

Belier Javier
Departments of Computer Science & Cybersecurity
Utica University, Utica, NY

Faculty Advisor
Unnati Shah
Utica University, Utica, NY

This research aims to improve comprehensive knowledge on the unmanaged attack surface in cloud computing, and investigate current mitigations. Cloud computing has become a renowned technology implemented by many enterprises for complex solutions due to its vast benefits. On the contrary, cloud computing solutions are generally utilized for handling large loads of data; where cloud computing services are processed via the internet opening vulnerabilities on several aspects. In this poster, we categorize cloud computing architectures into a taxonomy of layers based on their prominence in the research community. Each layer represents a delegated component within cloud-based systems which we then analyze and determine their respective functions in a system, alongside their occupying vulnerabilities and threat priority level. From this, we observed that the network and data storage layers produce a higher threat priority level compared to the others, based on their subject frequency in research contributions with a higher regard to refinement. Additionally, we provided a demonstration to test the unmanaged attack surface. Using AWS, we developed a web hosting service which has been misconfigured to replicate data leakage; a common fault in cloud solutions. This is followed by the exploitation flow a threat-actor would embark to take advantage of the unmanaged attack surface. We found that threat exposure on complex cloud solutions is fairly difficult to track and manage, leading to exploitation. Furthermore, we evaluated several approaches previously conducted by other researchers that aim to secure the unmanaged attack surface. These mitigations were selected by their efficiency and specified topics which they targeted including ML anomaly detection, DDoS prevention, privacy inference mitigations, and data leakage prevention. Our evaluation shows that the most effective solutions use some combination of ML algorithms to manage the unknown.

Poster 5022

Generative Model Upscaling with Deep Learning

Tye Leckinger
Computer Science
SUNY Brockport, Brockport, NY

Faculty Advisor
Zakariya Qawaqneh
SUNY Brockport, Brockport, NY

I present my experience from the Computer Science department at SUNY Brockport. Equipment such as 3D scanners work by measuring the distance between an object and the sensors. This results in a point cloud representation of the object. However, without any data augmentation, a 3D mesh cannot be generated from this data since there is no information that tells how points should be connected to make edges and faces. For this reason, algorithms such as The Marching Cubes algorithm exist to help generate this data. The objective of this project was to try to improve the Marching Cubes algorithm with Deep Learning. Since meshes with a high vertex count look smoother, a low-resolution mesh could be fed into a Deep Learning network and compared with a High-resolution mesh to learn the difference between the two. The loss function would then be used to get the similarity between the two meshes. This poster will present the results obtained on our investigations into this project.

Poster 5023

Qwirkle Self-Learning AI

Nicholas Chaves
Computer Science
Widener University, Chester, PA

Faculty Advisor
Edwin Dauber
Widener University, Chester, PA

I present my experience participating in the Summer Undergraduate Research and Creative Activities Program at Widener University, with help from my advisor, Dr. Edwin Dauber. My project incorporates artificial intelligence and machine learning principals into a tabletop game named Qwirkle. Qwirkle is played similarly to Scrabble, but instead of tiles with letters, there are 108 tiles with different colored shapes, and similairly, players must add onto previous lines of tiles with combinations of their own tiles to score points. I chose Qwirkle because it was one of my favorite childhood games, and nobody ever attempted to create a self-learning AI for the game. Along with the self-learning AI, my Project incorporates computer vision techniques to recognize all tiles within a photo. To do this, the AI finds different contours in a given photo, and measures the angles of them to classify shapes, then measures the pixel hues to classify the color. The AI correctly positions all the tiles onto a virtual board, where the self-learning AI can then determine what move should be played. The AI selects the best move by running a Monte Carlo tree search, where the heuristic value of each gamestate is decided by the neural networks. This process is also used to create new training data for the AI. The AI becomes better by playing a series of games against a copy of itself trained on the new data, and if the newer model wins it becomes the current best AI model. More training data is created and the process repeats. I was able to accomplish my goals of creating a gamestate image detector and a self-learning AI. My poster will offer full in-detail descriptions of the image detection process and the self-learning AI.

Poster 5024

Unicon Debug Adapter Development

Maximus Streeter, Austin Barr
Department of Computing Sciences
SUNY Brockport, Brockport, NY

Faculty Advisor
Zakariya Qawaqneh
SUNY Brockport, Brockport, NY

We present our experience participating in the Unicon debug adapter development project at SUNY Brockport. The goal of the project is to further development of the Unicon debug adapter and to gain experience building real world systems as a team. The Unicon debug adapter development project was conceived upon a previous semesters work, the Unicon language server project. Unicon is a high-level programming language built upon the Icon Programming Language. Unicon adds new features such as objects, networking, databases, threads, and more to the base Icon programming language. Unicon's built-in debugger, UDB, allows for debugging of Unicon programs within a terminal environment. This project enables UDB to be used in an IDE or editor instead. We wanted to challenge ourselves for this project, to learn as much as we could about modern day, team-oriented software development, and how to overcome the difficulties associated with it. For the first semester of this project, we based the system around a specific program architecture. During the second semester, we decided to do an architecture overhaul, in the hopes to improve system performance. This experience was a great introduction on how to handle big difficult tasks, like system overhauls, as a team. We learned about how important it is to stay coordinated, divvy up tasks appropriately, and stay in contact with our fellow team members. Our goal for the project is to have a fully functioning, commercial ready product, available online, for anyone to use. Our poster will discuss the background, technical details, and results of the project.

Poster 5025

BrockportGPT: A Chatbot for SUNY Brockport

Matthew Saad
Computer Science
SUNY Brockport, Brockport, NY

Faculty Advisor
Zakariya Qawaqneh
SUNY Brockport, Brockport, NY

We develop BrockportGPT, an advanced chatbot designed to serve the SUNY Brockport community. BrockportGPT leverages Large Language Models (LLMs) and approaches such as Retrieval Augmented Generation (RAG), fine-tuning techniques, and the creation of a from-scratch transformer model. Emphasizing the enhancement of text retrieval and the integration of question-topic classification, this project systematically explores various retrieval strategies to improve the accuracy and relevance of responses to queries related to campus life, academics, and more. The project underpins its methodology on comprehensive data collection, including web scraping and the generation of a synthetic question-answer dataset using GPT-4 to train and refine the models. Preliminary results demonstrate BrockportGPT's efficacy in streamlining information access, showcasing the potential of conversational AI in educational settings, and highlighting ethical considerations in AI deployment. Through this work, BrockportGPT emerges as an example of applying LLMs to enhance university-level information dissemination, paving the way for future research in the field.

Poster 5026

ML for Plagiarism Detection in Coding Assignment

Khagendra Khatri, Shadip Khadka
Computer Science
Caldwell University, Caldwell, NJ

Faculty Advisors
Vladislav D. Veksler
Caldwell University, Caldwell, NJ

Recent advances in Artificial Intelligence (e.g., Chat-GPT) have exacerbated plagiarism in student assignments. The goal of this research is to evaluate the effectiveness of machine learning algorithms for automatically detecting plagiarism. We examined cursor position and document length changes over time, recorded from student Python programming assignments in a first-year Computer Science course (recordings of student work were gathered via the LectureAssign VSCode extension [1]). The collected data was assessed for plagiarism by human graders, and then fed into machine learning models, including logistic regression, decision tree, and neural network models to assess whether these algorithms may be trained to identify plagiarism.

Our dataset comprised 133 assignments, with 21 assignments containing plagiarized content. We used Synthetic Minority Over-sampling Technique to increase the size of our dataset, resulting in 224 samples, with 112 of each kind. We partitioned the dataset into 70% for training and 30% for testing purposes. The neural network model was trained for 20 epochs, comprised 2 hidden layers (each comprising 64 nodes with 'relu' activation), and was optimized using the 'adam' optimizer, with sparse categorical cross entropy as the loss function.

Results indicate that the neural network outperforms other classifiers, achieving 95.6% overall test accuracy. The neural net was able to correctly identify non-plagiarized work 100% of the time, and correctly identify plagiarized work 92% of the time on the test dataset (see Figure 1). (Figure elided) Figure 1. Plagiarism prediction accuracy results after training.

Future work will involve gathering a larger training sample (which should aid in improving neural network classification performance), checking minimum edit distance between assignments and completion time of the two closest assignments, using different machine learning algorithms (xgboost, convolutional nets, transformers), and further research into what constraints might enhance detection of plagiarism in student assignments.

[1] https://lecturedot.github.io/LectureAssign-vscode/

Poster 5027

Adapting Sensor Transplantation from Autonomous Rovers to Autonomous Drones

Peyton Hamilton
Math, Computer Science and Statistics Department
St. Lawrence University, Canton, NY

Faculty Advisors
Kevin Angstadt
St. Lawrence University, Canton, NY

We present our research on autonomous drone resilience in the Math, Computer Science, and Statistics Department at St. Lawrence University. The goal of the research is to reduce security vulnerabilities and software defects in autonomous vehicles. With the increase in autonomous vehicle use in recent years, the need for trust in autonomous vehicles has grown. Given previous research that forces software to terminate prior to repairs, we introduce a method that allows a mission to be completed while repairs are deployed. We implement redundant computers to enable hardware switchover mid-mission. Using sensor data transplantation, we plan to counter software vulnerabilities by seamlessly switching control between computers while simultaneously making repairs to flight software. This would circumvent the need to pause a mission in the case of a software deficiency or security breach. At this time, we have successfully implemented sensor data transplantation while continuously running ArduPilot software. Data analysis of initialization times will be shown in the poster presentation, with results mirroring the results found with sensor data transplantation in rovers. Further work is needed to transplant sensor data while flying a mission. Our poster will discuss the challenges we faced during the project as well as the improved performance as a result of computer redundancy. The software and hardware decisions will be presented in their entirety.

Poster 5028

‘AccessiMove’: Assistive Technology

Michael Ruocco, Jack Duggan
Computer Science
Quinnipiac University, Hamden, CT

Faculty Advisor
Chetan Jaiswal, Brian O’Neill, Karen Majeski
Quinnipiac University, Hamden, CT

This paper introduces AccessiMove, an innovative software solution designed to empower users with disabilities or limited hand mobility by providing hands-free control of a computer system. Through the combined use of facial tracking and gesture recognition technologies, AccessiMove enables precise and intuitive mouse control through the analysis of facial positioning and eye events (blinks). Users can manipulate the on-screen cursor, facilitating a natural and responsive computing experience without reliance on traditional input devices. Moreover, the incorporation of gesture recognition technology allows users to execute actions and key presses by performing specific gestures, further enhancing the range of hands-free interactions.

Through the utilization of MediaPipe libraries, we are able to generate points that represent the position of facial landmarks in relation to one another and the camera. These points allow us to calculate vectors that can be used to make assumptions on the current ‘state’ of the user's face, further mapping these states to customizable computer inputs. The software will also include an interface with customizable settings, catering to individual preferences and varying degrees of motor abilities. AccessiMove's adaptability extends to diverse applications, encompassing general computing tasks, multimedia consumption, and even simple gaming.

To conclude, AccessiMove stands as a promising advancement in the field of assistive technology, demonstrating the feasibility and effectiveness of facial tracking and gesture recognition for hands-free computer control. The software not only addresses the specific needs of users with limited hand mobility but also contributes to the ongoing efforts to create more inclusive and accessible computing environments.
.

Poster 5029

CLIMB: A Narrative Strategy Game about Time Management

Alaa Ramzi, Connor Paul, Sean Lucibello, Nicolas Perez
Tagliatela College of Engineering
University of New Haven, West Haven, CT

Faculty Advisor
Mehdi Mekni
University of New Haven, West Haven, CT

In order for someone to satisfy all of their life commitments, they must allocate time for them, which many student athletes are unable to find. Pursuing both athletic and academic success, many of them fail to find time for both, leading to them dropping the former. Time management is a special issue that demands special attention, and such attention is rarely accessible or engaging enough, especially before it's too late. We believe that student athletes deserve a proper method to improve time management without jeopardizing themselves even further. Thus, we propose an interactive and dynamic mobile game that challenges the user's time management within a virtual, consequence free environment. We codenamed this project CLIMB. CLIMB can allow you to improve their time-management skills through strategization and storytelling. Such skills should be applicable to the player's real, personal life. To achieve our goals, CLIMB must be able to allow the player to approach time-management within the simulation and its narrative as they would in real life. Analogous time-management between real life and CLIMB demands creativity and clever, logical thinking on behalf of the player. An emphasis on player-choice is necessary to leave a larger impact on the player themselves. Furthermore, the player is responsible for strategizing how they'll allocate time within CLIMB's simulation. CLIMB's' design must reward good strategy and punish harmful strategy to provide real-time feedback of the users' performance. During CLIMB's development, we expect to face challenges involving presentation and content output. In response, we have already begun preparing the prototype for CLIMB.

Poster 5030

Talon Offensive Cybersecurity Framework (TOCF): Empowering Red Team Operations

Chase Leckinger

Computer Science
SUNY Brockport, Brockport, NY

Faculty Advisor

Zakariya Qawaqneh
SUNY Brockport, Brockport, NY

This thesis presented by Chase Leckinger of and supervised by Dr. Qawaqneh of the Computer Science department at SUNY Brockport, Investigates Offensive Cybersecurity techniques that help defend against the ever-evolving threat actors within the Cyber landscape. This study focuses` on the devolvement of a command line Command and Control (C2) framework that emulates the Tactics, Techniques, and Procedures of different threat actors. The development of this framework enhances the operations of Red Team operators by incorporating a reverse Socks5 proxy, an API that communicates with AI-generated phishing templates, and evasion techniques. Through the research, development, and testing of this framework this project aims to offer a sophisticated tool that helps to better aid red team professionals in engagements. The final form will be a deployable open-source application on Github.

Poster 5031

GraySim: a Page Replacement Simulator

Melinda Zhang
Department of EECS
United States Military Academy, West Point, NY

Faculty Advisor
Maria R. Ebling
United States Military Academy, West Point, NY

We aim to address the challenge of effectively teaching students both fundamental and advanced concepts in operating systems, with a particular focus on memory handling and page replacement policies. Page replacement policies, such as Optimal, First-In First-Out (FIFO), Least Recently Used (LRU), and Clock, are crucial for understanding how operating systems manage virtual memory. However, grasping these concepts can be challenging for students without hands-on practice.

We are developing a program called GraySim to support student learning through simulating these various page replacement policies and providing feedback to students about their demonstrated understanding. This program is designed using the Scala programming language and follows the Model-View-Controller (MVC) architectural pattern through test-driven development to ensure accuracy and effectiveness. By providing practical exercises through this simulation tool, students can deepen their understanding of memory handling in operating systems in a more interactive and engaging manner.

To evaluate the simulation, we plan to perform a qualitative study with participants who have completed a college level operating systems course. We will ask them to observe and use the simulation and then complete a short survey that provides feedback about its usability and expected utility for novice students of operating systems. We will then analyze the results and make prioritized recommendations to improve the simulation in preparation for a planned quantitative study to take place in the next academic semester. During this poster presentation, we will demonstrate this simulator and collect feedback from willing attendees (IRB approval pending).

This approach not only aims to enrich the cadet learning experience but also to contribute valuable insights into the effectiveness of practical simulations in computer science education.

Poster 5032

Physical Activity and Cigarette Smoking in Adults Living with HIV

Rachel Kunz, Matthew Gilliland, Caroline Collins
Data Science
Merrimack College, North Andover, MA

Faculty Advisor
Christopher Stuetzle
Merrimack College, North Andover, MA

Our goal for this research is to characterize the relationship between physical activity (PA), depression, and smoking in people living with HIV (SLWH), by using deep machine learning and various clustering techniques, to compare to those living without HIV. We are hopeful to produce an outcome to aid future trials in bettering the quality and length of life among individuals SLWH.

Understanding the dynamic between PA, cigarette smoking, depression, and their effects on the health of SLWH is crucial for developing effective interventions and helping to improve the lives of this population. Prior research has determined that there is a direct correlation between PA as well as cigarette smoking among adults living with HIV. PA has proven to help increase short-term abstinence and mitigate the chances of a nicotine relapse. Increasing your levels of exercise have been shown to correlate with decreased levels of depression, while reducing depression has been associated with lower levels of smoking. However, there have been no studies that have focused on the relationship between all variables within this at-risk population. We will analyze how these behaviors can affect this group's health.

This data is gathered from a multicenter cohort known as the MultiAids Cohort Study (MACS)/Women's Interagency HIV (WIHS) Combined Cohort Study (MWCSS). Our analyses include 1) characterizing aggregate PA behaviors and comparing SLWH to people with HIV non-smokers; 2) exploring relationships between PA and smoking by serostatus; 3) exploring the relationship between smoking and other drug use, HIV biomarkers, by PA and serostatus, and 4) explore differences in PA by descriptive and potentially confounding variables including gender, race, age, HIV status, depressive symptoms, site location, and BMI. These analyses will be done using multi-linear regression, segmentation modeling, hierarchical clustering, ensemble voting, neural networks, and other machine learning and data analysis techniques.

Poster 5033

AI Enhanced Autonomous Navigation: Integrating AI Processing and LiDAR in Robotics

Eric Barton, YongJun Lee, Anthony Puelo
Department of Computer Information Systems
SUNY Buffalo State University, Buffalo, NY

Faculty Advisors
Gang Hu
SUNY Buffalo State University, Buffalo, NY

This research explores autonomous robotics by integrating LiDAR, advanced imaging, and artificial intelligence (AI) for enhanced navigation and interaction in diverse environments. Our innovative approach utilizes both Light Detection and Ranging (LiDAR) sensors and a camera system to collect spatial and visual data, enabling precise obstacle detection and navigation. Central to our methodology is a sophisticated AI framework, developed from the ChatGPT-4 model, designed to process and analyze this multimodal data. By synthesizing insights from LiDAR and image inputs, the AI generates autonomous movement commands, guiding the rover with improved adaptability and awareness.

The project employs the Robot Operating System (ROS) for seamless integration of hardware and software, ensuring efficient real-time decision-making and data exchange among the rover's components. This synergy between LiDAR sensing, AI-driven image processing, and ROS architecture marks a significant advancement in autonomous robotics, demonstrating potential applications in exploration, surveillance, and environmental monitoring.

Our findings contribute to the robotics field by showcasing a novel sensor fusion technique and AI-enabled navigation, setting new standards for autonomous capabilities. This research underscores the importance of integrating diverse data sources and AI in enhancing robotic autonomy, paving the way for future innovations in robotic exploration and monitoring.

Poster 5034

Facial Recognition Privacy Concerns: Pixels, Patterns, and Privacy

Akshar Patel
Department of Computer Science and IT
Western New England University, Springfield, MA

Faculty Advisor
Hanieh Shabanian
Western New England University, Springfield, MA

Facial recognition has become an accepted function in our everyday lives, ranging from the ease of unlocking our smartphones to its application in law enforcement. However, this convenience comes with a price – the storage of biometric data in corporate databases, often without individuals' explicit consent. Privacy concerns are raised by facial recognition technology, which functions by identifying and analyzing distinct facial patterns. This study seeks to examine the ethical considerations associated with facial recognition and how it affects personal privacy. We will explore the societal effects of facial recognition technology, by examining its deployment in public and commercial domains. It will dive into the fine line between the benefits gained by utilizing the technology, and the potential privacy concerns for individuals. Moreover, this study will highlight the inherent inaccuracies and biases within facial recognition systems, with a specific emphasis on the potential occurrence of false positives and negatives, which could result in erroneous identifications. Addressing these biases is crucial to guarantee the fair operation of the technology. Through a thorough investigation of facial recognition technology, this study aims to inform individuals about the challenges and important factors related to its use, revealing how it might affect their lives without them realizing it.

Poster 5035

Quinnipiac Virtual Medical Simulation

Jonathan Mason, Matthew Merritt, Michael Merritt
School of Computing and Engineering
Quinnipiac University, Hamden, CT

Faculty Advisors
Chetan Jaiswal
Quinnipiac University, Hamden, CT

Michael Russo, Thomas Nowicki
Hartford Healthcare, Hartford, CT

Medical training and simulation have always been areas where providing real-world practice is both crucial yet challenging to achieve in safe and accurate environments. Traditionally, students gain experience by observing or performing procedures under supervision, or practicing in simulations that use medically accurate mannequins. Additionally, medical professionals are required to complete training and practice simulations for their ongoing professional practice evaluation (OPPE) requirements, but access to accurate resources is limited when working in large organizations. The Quinnipiac Virtual Medical Simulation is being developed with Hartford HealthCare and the Center for Education, Simulation, and Innovation (CESI) as a collaborator, client, and sponsor to address these issues. By offering a self-regulating virtual reality simulation that can record and play back user inputs, our product seeks to reduce the strain on the schedules of medical professionals in charge of overseeing and managing these trials.

The solution is set up to be flexible, scalable, and adaptable, as it supports the creation of user-generated simulations and custom asset uploading including 3D models, audio, and text resources. With our application, staff at the hospital can take custom 3D models scanned in from real medical environments and then use these objects to build simulations that mirror the environments that medical students and staff operate in during day-to-day practice. Users can select from the saved scenarios created by staff and then generate a VR experience for a selected procedure. Currently, the application has an example of a trauma team leadership scenario, but will also support other specialized training simulations for nurses and other staff. The results of this simulation are then recorded and saved in a database that can be used by staff for verifying completion of required training or for filtering applicants by performance and attentiveness to detail.

Poster 5036

AITracker: A neural network designed for efficient eye tracking

Jacob Hogrefe, Emilio Cruz
School of Computing and Engineering
Quinnipiac University, Hamden, CT

Faculty Advisors
Chetan Jaiswal, Jose Riofrio
Quinnipiac University, Hamden, CT

Eye tracking technology has typically been limited by the necessity for expensive and specialized hardware. However, the AITracker project aims to revolutionize this field by introducing an affordable and accessible alternative. This project presents a vision for an innovative webcam-based eye tracking system that utilizes deep learning algorithms to accurately detect eye movements in real-time. The primary objective of AITracker is to provide users with a seamless and intuitive interface for controlling and interacting with external devices using only their eyes. By leveraging the capabilities of deep learning, the software can precisely track eye movements in eight directions (up, down, left, right, up left, up right, down left, and down right) as well as detect blinks, effectively transforming a standard webcam into a powerful input device.

Despite challenges related to data collection for AI training and optimization of algorithms, the project aims to achieve the highest attainable accuracy, with potential for further improvement. With an emphasis on user-friendliness and accessibility, AITracker is designed to run on standard laptops with minimal system requirements, making it widely available to users. In conclusion, the AITracker project represents a significant advancement in eye tracking technology, offering an affordable and versatile solution with broad applications in various different fields. From determining screen-time to hardware integration, the applications of the project are wide-reaching in both the hardware and software spaces. This project aims to not only be a complete solution for hardware integration, but a solid starting point for developers to build upon the network for their specific applications. In our case, we will use the neural network to control a motorized wheelchair with just our eyes.

Poster 5037

CodeForge: Online Collaborative Code Editor

Joshua Reed, Nicholas Tourony, Ty Hutchison
School of Computing and Engineering
Quinnipiac University, Hamden, CT

Faculty Advisors
Chetan Jaiswal, Christian Duncan, Robin Chataut
Quinnipiac University, Hamden, CT

For beginners who are learning to code, version control tools, such as GitHub, can often seem overwhelming and complicated. These tools are often unnecessary for beginners to use on assignments and small projects. Being able to collaborate on a project online, in real time, would allow for programmers to bypass the issues and complications of version control, while still allowing for development in a team. CodeForge allows users to effectively work in a team, like Google Docs, but for software development. CodeForge has many standard code editor features built into the browser, as well as additional features, such as a whiteboard for brainstorming and annotation, and a classroom system where professors can post announcements, assignments, and share projects with students during lectures. All these features come together to allow students, and new developers to focus more on learning, and their projects, rather than the complications of version control. Experienced programmers also benefit from our project, by gaining an alternative tool to use when working on smaller, short-term projects, or when working in real-time.

CodeForge is an online tool, written in JavaScript using the Node.js library, utilizing Express as a web framework, and MySQL to manage our user project database. All our user's code is run in Docker containers to protect both the users, as well as the server. Our site uses proper web security practices, such as salting and hashing sensitive data. Our site is HTTPS enabled, ensuring that all network traffic is encrypted in transit. To give users a more secure and time friendly alternative, there is a "Sign in with Google" option that uses Google's authorization servers to use our website.

Poster 5038

Using Artificial Intelligence to Predict the Outcome of Batted Balls and Evaluate the Impact of New MLB Rules

Samuel Trumbley
Computer Science
Southern Connecticut State University, New Haven, CT

Faculty Advisor
Winnie Yu
Southern Connecticut State University, New Haven, CT

I present an analysis of the MLB's new rules by training a machine learning model on 2023 batted balls, then predict 2022 batted balls and compare to the actual data from that season. Recently, the MLB has seen certain patterns of change in the sport that have seemingly hurt the viewing experience. The length of games has seen consistent increase, and the sabermetric approach most teams have taken, which values home runs and slugging, has led to less action on the bases and less balls being put into play. The MLB decided to implement several new rules for the 2023 season that would hopefully combat these issues. These rules include a pitch clock, larger bases, and the banning of infield shifts. In the first season with these rule changes, it seems the MLB has been successful in their attempt to increase offense and make the game more engaging. Statistics like stolen bases, hits, and batting average all showed an increase from the 2022 season to the 2023 season. However, I wanted to dig beyond these statistics to see just how impactful these new rules are, specifically the ban on infield shifts. This idea inspired me to utilize artificial intelligence and machine learning to evaluate the impact by comparing the two seasons. The data from the two seasons were collected from Baseball Savant. Each dataset uses six unique features to predict the outcome of batted balls, and three different classifiers are utilized to project the outcomes. The results from comparing the predicted hits from the 2022 season support the claim that the rules were effective in increasing offense. The predicted hits were greater than the actual number of hits from the 2022 season, confirming that if the rules were implemented in that season, it would have led to more offense.

Poster 5039

What Lies in the Listings? A Computational Analysis of House Descriptions in the Greater Boston Area

JinYu Li, Edith Po
Computer Science
Wellesley College, Wellesley, MA

Faculty Advisor
Eni Mustafaraj
Wellesley College, Wellesley, MA

We present our experience participating in an independent study in the Computer Science Department at Wellesley College. The goal is to investigate economic phenomena in data collected from Zillow. We found that descriptions of houses on Zillow provide valuable resources for us to investigate how factors beyond basic housing features, such as the number of bedrooms and bathrooms, can impact housing prices. Although there are already many housing indexes in the Economics literature, factors more ambiguous, like seller's persuasive efforts in housing descriptions, have not been captured. During our first semester in the independent study, we designed an algorithm to automatically scrape information from the web. We collected data from over 130,000 house listings spanning Norfolk (24,638 houses), Suffolk (15,768), Middlesex (58,075), and Essex (32,055) counties. These listings include houses with most recent sold dates starting in 2020.

This semester, we employ a computational mixed-methods design by combining natural language processing techniques with qualitative content analysis to examine how sellers differ in their cultural representations of their listings. Our findings highlight the pervasiveness of rhetoric focusing on lifestyle in higher income neighborhoods in comparison to descriptions of amenities like schools and hospitals in lower income neighborhoods. These findings underscore the reinforcing dynamics between wealth disparities in Boston area, with subsequent implications towards the ongoing development and stereotypical portrayal of longstanding residential communities.

Our poster will discuss our reasoning for training and choosing between different machine learning models, and the challenges we encountered during the process. The full description and implementation feature of the project will be presented in detail.

Poster 5040

Taking Java Obfuscation to the next level by exploiting SSA Form

Shanyu Thibaut Juneja
Manning College of Computer Science
University of Massachusetts, Amherst, MA

Faculty Advisors
Neena Thota
University of Massachusetts, Amherst, MA

I present my research in the obfuscation field in the form of an open-source Java software-protection tool designed to outperform both free and commercial alternatives available on the market. Stemming from a personal frustration with the many underlying problems in both resilience and quality with the products in which I had invested my personal income, I took, improved and formalized theoretical concepts into a framework that allows for developers to build their own Java app security tools with ease, removing a great barrier of entry to the field. I have built my own obfuscation tool designed to obscure and obstruct reverse-engineering of decompiling Java bytecode, by leveraging compiler techniques such as Static Single Assignment (SSA), including opaque predicates, and by introducing novel obfuscation techniques. This research introduces a new obfuscation metric, extending upon the limitations of prior established research, by accounting for the intricacies introduced by modern SSA-based optimizations. This metric focuses on measuring optimized graph representation over raw assembly code, which results in a more rigorous perspective of how 'messy' the control-flow of a program is visually rendered in modern reverse-engineering tools. This work addresses many of the shortcomings of current practices and sets a new standard for evaluating obfuscation effectiveness in Java. The tool developed as part of this research represents an important advancement in the field of software protection, offering both theoretical insights and practical benefits to developers seeking to safeguard their Java applications. Furthermore, results find a 99.6% effectiveness against modern automatic reversing tools. The project currently stands with 500+ stars on GitHub. Further discussions on the methodology, implementation and detailed algorithms of each technique, and comprehensive evaluations of the obfuscation tool's performance are presented, providing valuable contributions to the domain of software security.

Poster 5041

Let's Talk: Bridging the communication gap with A.I.

Peter Zegarek, Bineeth Reddy
Computer Science
Quinnipiac University, Hamden, CT

Faculty Advisor
Chetan Jaiswal
Quinnipiac University, Hamden, CT

Communication is the key for humans to not only express themselves but also to exchange ideas and information. Communication can take the form of spoken or written and each of these forms the basis of most if not all languages. Communication also enables individuals to convey thoughts, emotions, and information with clarity and precision, fostering mutual comprehension and empathy. Using language as a medium for communication, it facilitates understanding, promotes social cohesion, drives economic opportunities, and enables education. However, language could also form a communication barrier for people, particularly communities and students from different linguistic backgrounds. This could not only lead to social exclusion and discrimination but also could hinder cultural preservation and impede effective communication, leading to confusion, frustration, and interpersonal conflicts.

This work aims to help bridge the gap for translation languages in real-time using A.I. It is an A.I. based engine that translates any selected language to any other language. Its workflow simulates that of a phone call in which party-A chooses to speak (verbally) in the language Li of their choice, but party B on the other side receives and hears the translated speech in the language Lj of their choice, and vice-versa. This translation happens seamlessly and effortlessly by using the Let's Talk speech translation engine that we developed. We are launching this work in the form of an Android app that has a simple UI for anyone to use. The applications for this work are limitless from social work to medical practitioners, helping them to help the community. Additionally, the second phase, which is a work in progress addresses ASL (American Sign Language) to be translated in any language in real time and vice-versa. It allows people who can only communicate in ASL to be able to have conversations and connect with billions of people, without the need for an interpreter.

Poster 5042

Humanizing NPCs for Enhanced Video Game Immersion

Adam Liu, Alesandro Gopal
Amity Regional High School
Woodbridge, CT

Faculty Advisor
Chetan Jaiswal
Quinnipiac University, Hamden, CT

Artificial Intelligence (AI) has been rapidly integrated into various aspects of human life, highlighting a challenge: AI lacks human-like qualities. This is why, in the realm of videogames, playing with Non-Player Characters (NPCs), or AI robots, often results in less enjoyable experiences. To address this, researchers and game developers explored using machine learning to artificially mimic human learning. Long Short-Term Memory (LSTM) models are machine learning algorithms known for processing sequential data. The purpose of this project was to design a human-like LSTM-based NPC to enhance gaming experiences. The project aimed to enhance the Minecraft player's experience and potentially extend these techniques to other fields like healthcare and education, fostering more engaging interactions between humans and AIs.

The project was deemed successful if the NPC could exhibit human-like reactions and seamlessly blend with real players in 80% of testing scenarios. Training was split up into different skills required for the game: aiming, moving, and action-timing. Human participants, mainly 14–22-year-olds recruited through convenience sampling, played small games in Minecraft to generate skill-specific data for training the LSTM model. The NPC was programmed using Mineflayer, a JavaScript library for Minecraft NPCs, and trained on Google Cloud servers. Following the Turing test approach, an online survey asked participants to distinguish between NPC and human players based on gameplay videos. Findings thus far confirm that the designed LSTM-based NPC displays convincing and consistent human attributes during multi-stage gameplay, while vastly exceeding a standard programmed bot. As such, integration into Minecraft and multiplayer environments promises more immersive and enjoyable player experiences overall.

Poster 5043

Evolution of January 6 Framing Devices Throughout Digital Media

Samantha Schumacher
Department of Political Science
Emmanuel College, Boston, MA

Faculty Advisors
Adam Silver, Mark Sherman
Emmanuel College, Boston, MA

The insurrection on the US Capitol on January 6, 2021, exposed deep rooted fissures in our democracy that have come to pose a direct threat to the integrity of America's democratic electoral systems. In the past three years, Jan. 6 has continued to impact political discourse amongst political elites and the American public, remaining a major news story amid congressional hearings, court arraignments, and continued attempts to fuel election denialism. Political elites have helped shape public opinion using specific vocabulary, phrasing, and tonal indicators to emphasize their preferred narrative of the driving forces behind Jan 6. Existing literature shows that the framing of Jan. 6 has created differing accounts of the events and impact of the insurrection. Most studies have focused primarily on framing devices used in traditional news media, as well as the social media posts of participating groups or individuals. While this covers a large part of our current information landscape, it cannot be extrapolated to framing effects on social media sites due to their highly curated atmosphere; the selective nature of recommendation algorithms has irreversibly altered the way people consume and digest information. This study seeks to bridge this gap by analyzing the posts of Rachel Maddow and Sean Hannity on X (formerly Twitter), offering insights into the framing of Jan. 6 by two of the most influential and widely followed American political commentators. A critical discourse analysis shows that terminology directly describing Jan. 6 remains consistent between Maddow and Hannity. However, the narratives described by Maddow and Hannity differ drastically. Hannity's posts predominately question the legitimacy of the House select committee and its subsequent hearings, while Maddow's posts highlight the unprecedented and serious nature of Jan. 6. Unlike previous studies focused on traditional media, these posts describe Jan. 6 with high definitional certainty since the insurrection.

Poster 5044

Experiential Learning via Distributed Hardware Access in Computer Organization

Erin Sorbella
Seidenberg School of Computer Science and Information Systems
Pace University, Pleasantville, NY

Faculty Advisor
Thomas Schmidt
Pace University, Pleasantville, NY

We present our findings following the introduction of remote hardware labs to a Computer Organization course at Pace University. Computer Organization is a traditionally theory-heavy course that familiarizes students with the fundamentals of computer hardware. Motivated by a shared goal to enhance interactive learning in computer science courses at Pace, students were invited to participate in FPGA (Field-Programmable Gate Array) lab modules as part of their coursework. We hypothesized that empowering students to create their own hardware designs with FPGAs would improve student engagement and understanding of the course material. Throughout the semester, students progressed from logical circuits in Verilog to developing their own 8-bit computer processor. To facilitate access to the necessary equipment, we collaborated with LabsLand, a provider of remote FPGA labs. Initial student surveys and interviews showed a positive reception to the newly-incorporated lab modules; however, some concerns arose due to the limitations of a remote lab format. The given lab modules, the design of the 8-bit processor, and an overview of student responses will be presented in further detail. We gratefully recognize the Pace University Department of Computer Science for supporting this endeavor.

Poster 5045

Bridging Natural and Programming Languages with a 'Natural Transpiler'

Eshen Brinkert, Sidney Kaeb, Alyssa Rotondo,
Jessica Schafer, Audrey Taylor
Department of Math and Computer Science
Muhlenberg College, Allentown, PA

Faculty Advisor
Hamed Yaghoobian, Jorge Silveyra
Muhlenberg College, Allentown, PA

We present our semester software project participating in the Culminating Undergraduate Experience Computer Science Seminar course in the Math and Computer Science Department at Muhlenberg College. The goal of this course is to develop a significant software project that builds on computer science knowledge acquired throughout our undergraduate coursework. Our project idea stemmed from an attempt to answer the proposed question: if coding was more accessible to non-native English speakers, would there be more diversity in computer science students and professional developers? To answer this question, we decided to create a program that will allow users to code in Python in their preferred language then translate it to English before compiling it into machine code. The use of a translator is necessary to accomplish this goal to bridge the gap between natural and programming languages. Given the complex nature of language, no two languages have direct translations for every word or phrase. Grammar, dialect, and sentence structure play a role in correct pronunciation and syntax. This brought about the idea to use deep learning to train our large language models. If we can use deep learning to train the model, our LLM can help make correct distinctions between the nuances of languages. This "Natural Transpiler" will translate between natural languages as well as the words that make up the program, rather than restructure formatting and syntax. By envisioning what learning computer science looks like to a non-English speaker or to someone whose first language is not English, we can overcome the barrier to programming faced by many non-native speakers.

Poster 5046

Using a GPT with Reinforcement Learning for More Personalized or Targeted Text Generation

Giorgi Gogitidze
Computer Science
Caldwell University, Caldwell, NJ

Faculty Advisor
Vladislav D. Veksler
Caldwell University, Caldwell, NJ

This study explores a methodology for efficiently adapting GPT [1] text-generation models for personalized or context-targeted results by integrating GPT with Reinforcement Learning (RL). Specifically we used GPT-2 [1] with manual reward assignment based on the quality of generated text. This strategy involves directly influencing the model's learning process by providing immediate feedback for each output, thereby enhancing the efficiency of the training phase and tailoring the GPT-2 model's responses to desired criteria without extensive computational overhead. (Figure elided) Figure 1. GPT with RL.

Simulations were conducted to generate text responses based on a synthetic prompt response dataset [2], comparing a standard GPT-2 adaptation process with the proposed optimized approach. Key hyperparameters included a learning rate of 0.0001, a discount factor of 0.9, over 1000 episodes, and a batch size of 5. Text generation parameters were set to a max length of 40 tokens, a repetition penalty of 1.2, and temperature at 0.6, aiming for a single response sequence from a model comprising 149 layers. (Figure elided) Figure 2. Sample responses before (left) and after (right) RL.

Qualitative text analysis of the proposed approach revealed more precise and concise text generation capabilities (Figure 2). Specifically, we observed a reduced number of repetitions in text responses. For original GPT-2 mean lexical diversity was 0.499, while for GPT-2 with RL showed lexical diversity of 0.616. Overall, these findings suggest that the proposed method has the potential to augment flexibility and precision of pre-trained language models across diverse applications.

Future research will involve refining proposed optimization techniques and evaluating its applicability across larger datasets and more complex generation tasks. The potential for incorporating adaptive learning rates and exploring the effects of different batching strategies on model performance and resource utilization will also be explored.

[1] https://huggingface.co/openai-community/gpt2
[2] https://huggingface.co/datasets/Dahoas/sft-gptj-synthetic-prompt-responses