Design of Phase-Locked Loop

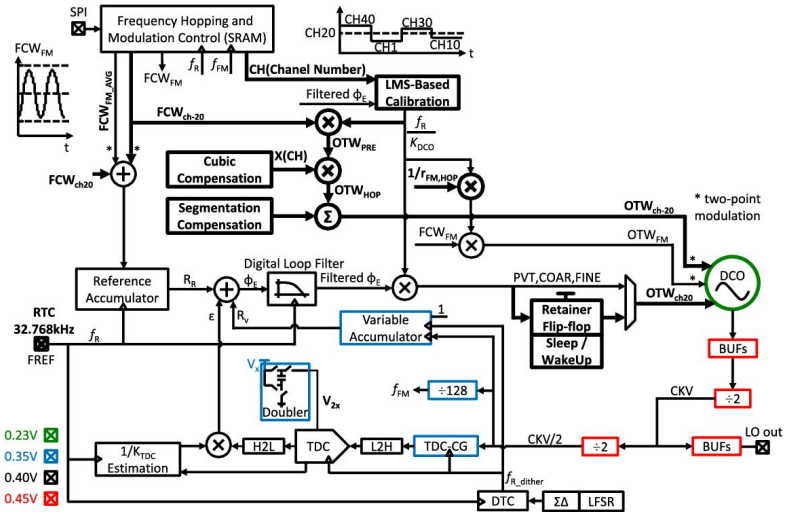
Evan Goldstein, Sophie Jaro, Hadassah Yanofsky

Problem Statement

A Phase Locked Loop (PLL) is a system that generates an output signal that has a constant phase difference from a reference input signal

These are commonly used for clock signal synthesis and can be used to multiply a reference clock by very precise amounts.

Our specific system will be used for bluetooth LNE applications (2.4 GHz)



C. -C. Li, M. -S. Yuan, C. -C. Liao, Y. -T. Lin, C. -H. Chang and R. B. Staszewski, "All-Digital PLL for Bluetooth Low Energy Using 32.768-kHz Reference Clock and ≤0.45-V Supply," in *IEEE Journal of Solid-State Circuits*, vol. 53, no. 12, pp. 3660-3671, Dec. 2018, doi: 10.1109/JSSC.2018.2871632.

Background Research

We used RF Microelectronics 2nd Ed. by Razavi for the bulk of our research into this topic.

There are a bunch of papers that we are using as reference but they have complex circuits aren't useful for industry.

Although other PLLs exist, ours will improve on those by consuming low power while maintaining clean output signal (low noise).

Table 1. PLL Performance Metrics for Bluetooth LE Application in Literature

Ref, Year	Frequency Range	Power Consumption	Phase Noise	Fractional Spur
[4], 2016	2.05-2.55GHz	1.4mW	-92 dBc/Hz @5MHz	-60 dBc
[5], 2017	1.8-2.5GHz	0.67mW	-103 dBc/Hz @1MHz	-56 dBc
[6], 2021	1.8 - 2.7 GHz	2.76mW	-112 dBc/Hz @10MHz	44 dBc

[4]"A Fully Integrated Bluetooth Low-Energy Transmitter in 28 Nm CMOS With 36% System Efficiency at 3 DBm." *IEEE Xplore Full-Text PDF:* https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7476842.
[5]"24.7 a 673µw 1.8-To-2.5GHz Dividerless Fractional-N Digital PLL with an Inherent Frequency-Capture Capability and a Phase-Dithering Spur Mitigation for IOT Applications." *IEEE Xplore*, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7870440
[6]"An Injection-Locked Ring-Oscillator-Based Fractional-N Digital PLL Supporting Ble Frequency Modulation." *IEEE Xplore*, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9737302.

Proposed Solutions

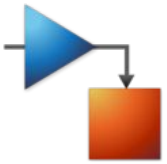
We have various parts that we wanted to work on improving.

We also wanted to work on improving phase noise and power consumption.

We worked on virtuoso, verilog and simulink to build and test our individual circuits.

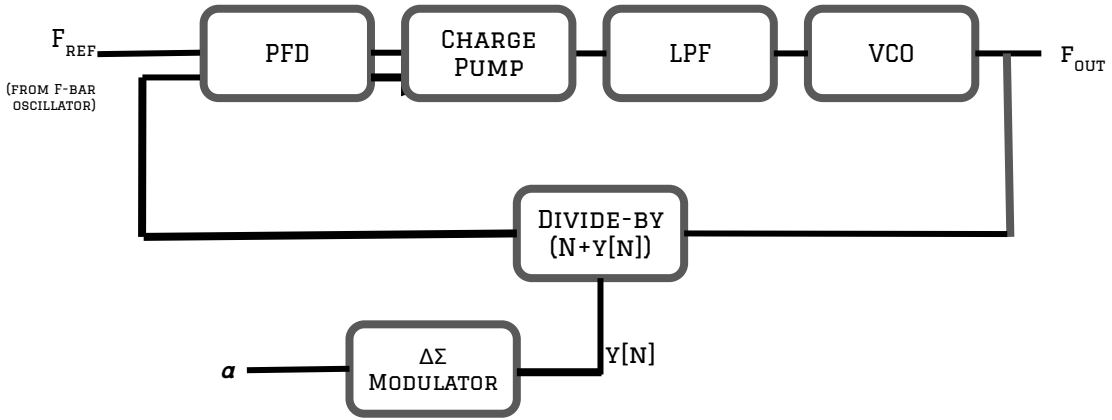
cadence®

Verilog HDL

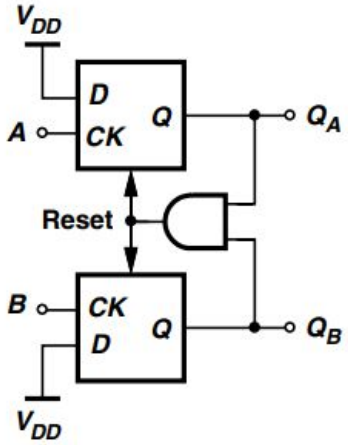


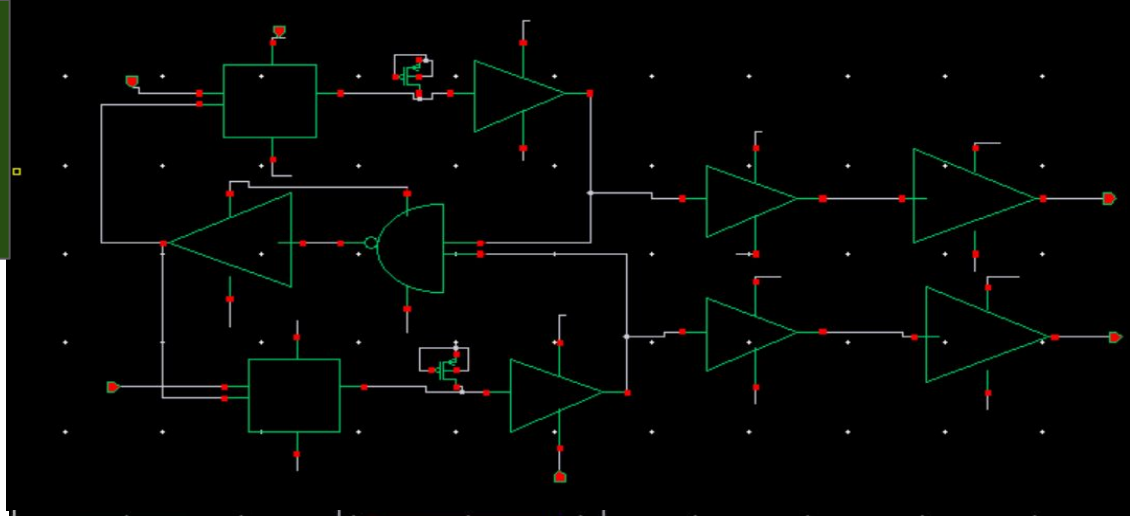
Experimental Solution

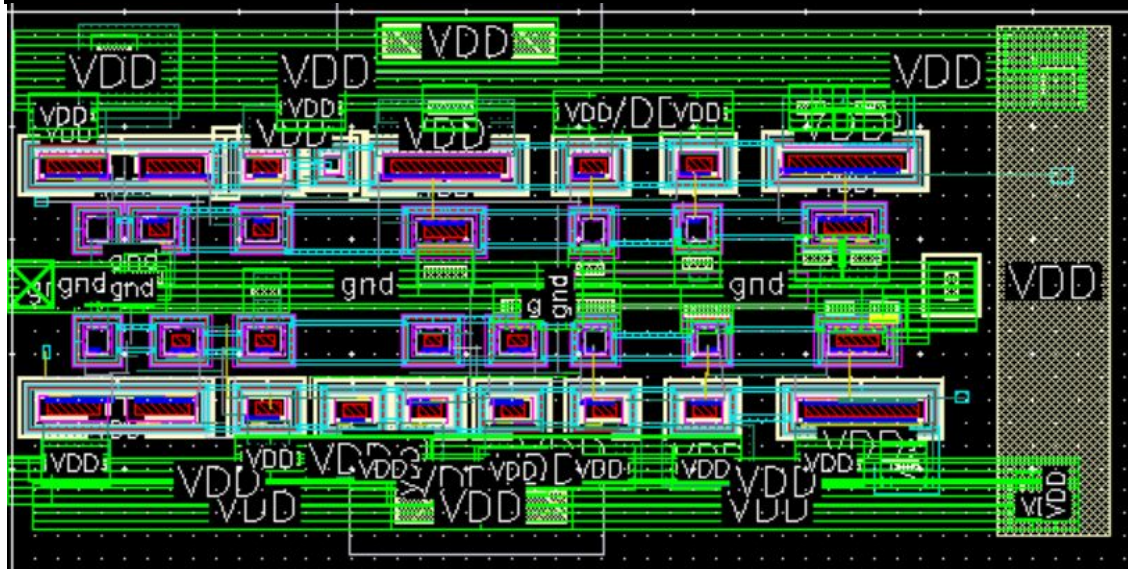
We split up the system into various subblocks, then combined all the individual components into the complete system and manage the interconnection.



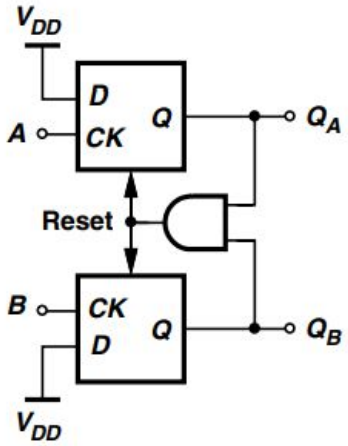
PFD

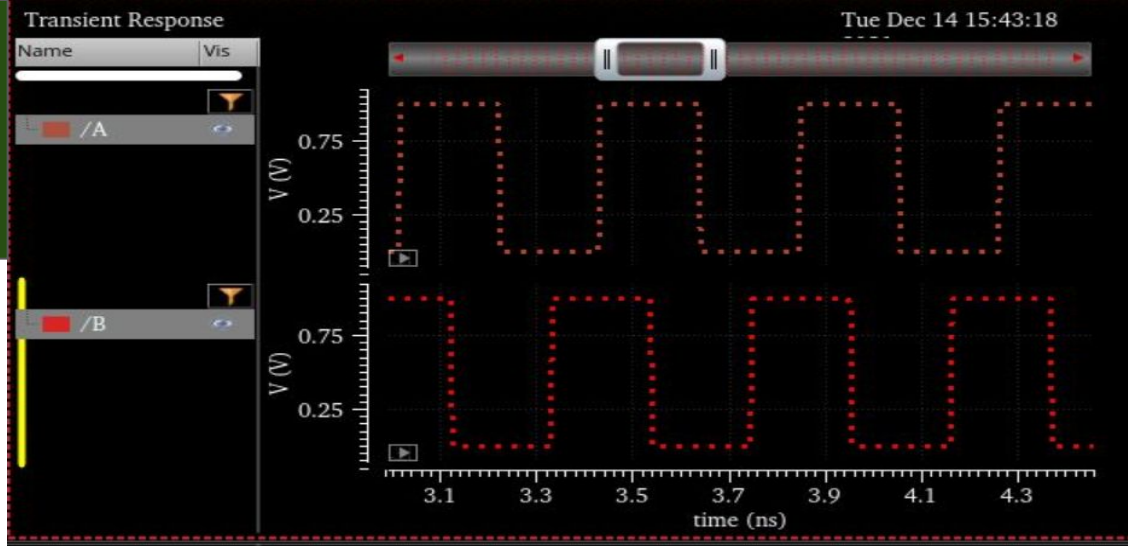


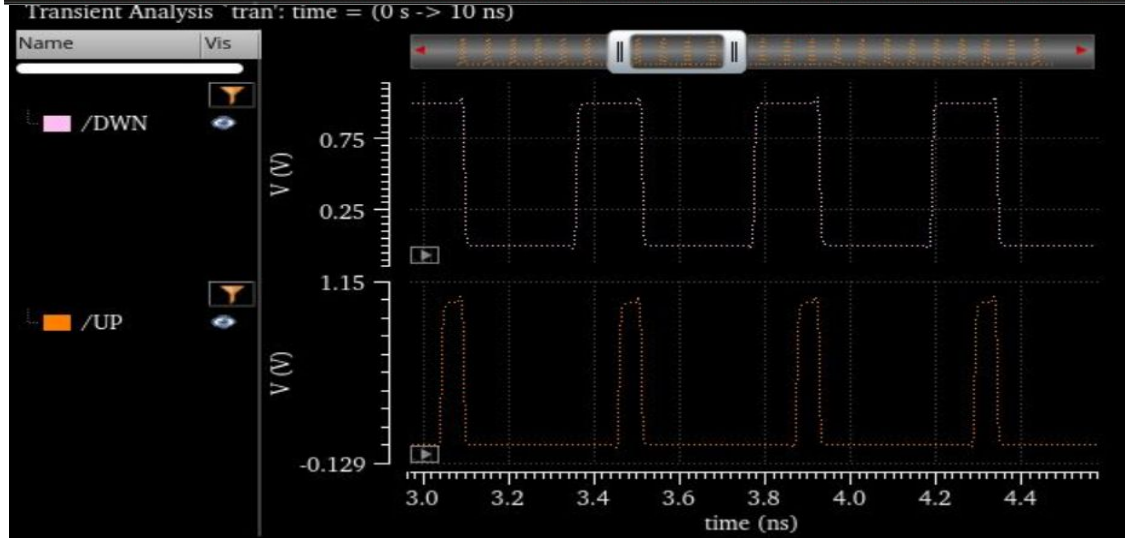




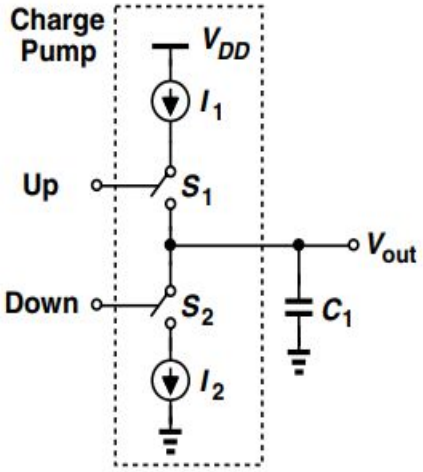
PFD Results

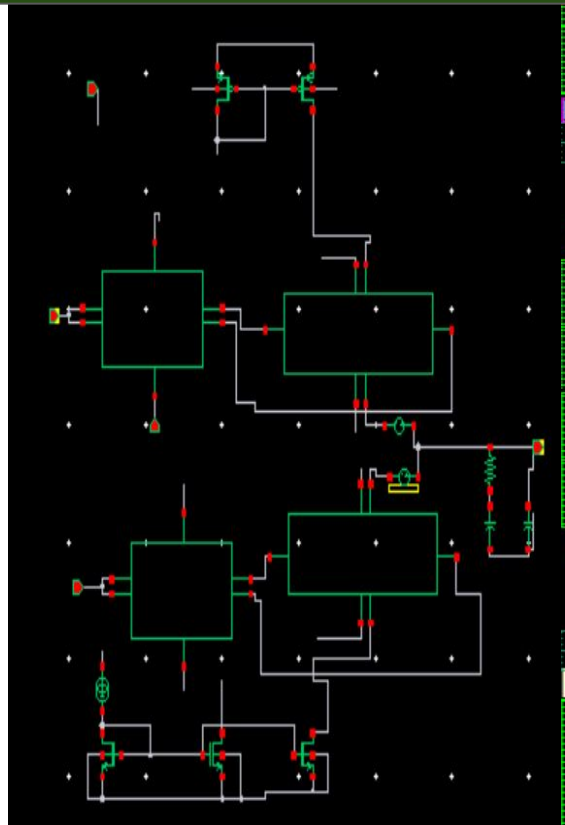

V_{DD}
D
CK
Q
A
Q_A
Reset
B
CK
D
Q
Q_B
V_{DD}

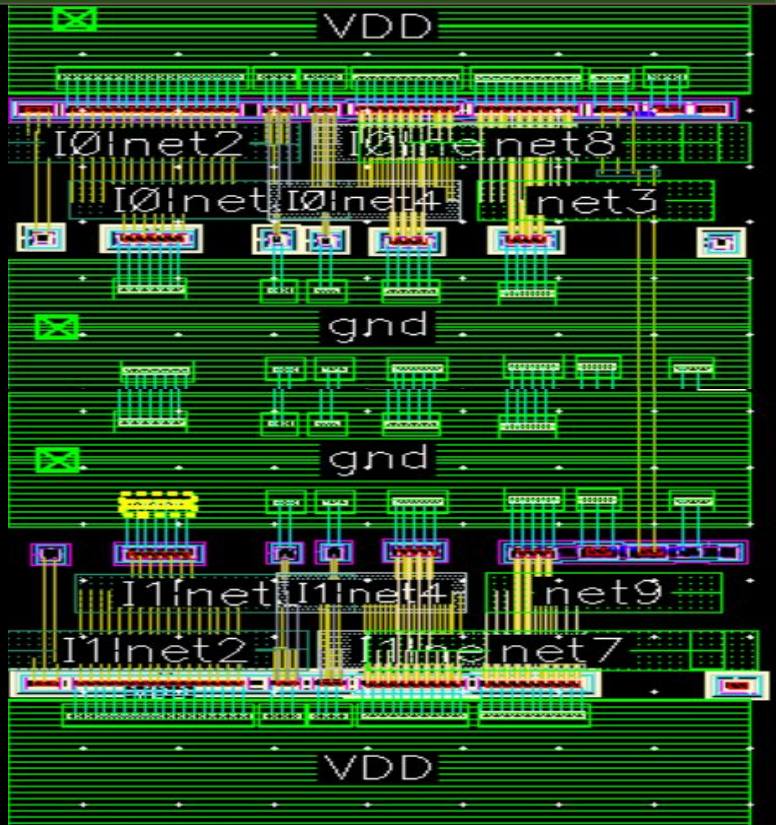

Transient Response
Tue Dec 14 15:43:18
Name
Vis
/A
/B
0.75
0.25
V (V)
0.75
0.25
V (V)
3.1
3.3
3.5
3.7
3.9
4.1
4.3
time (ns)


Transient Analysis `tran': time = (0 s -> 10 ns)
Name
Vis
/DWN
/UP
0.75
0.25
V (V)
1.15
-0.129
V (V)
3.0
3.2
3.4
3.6
3.8
4.0
4.2
4.4
time (ns)

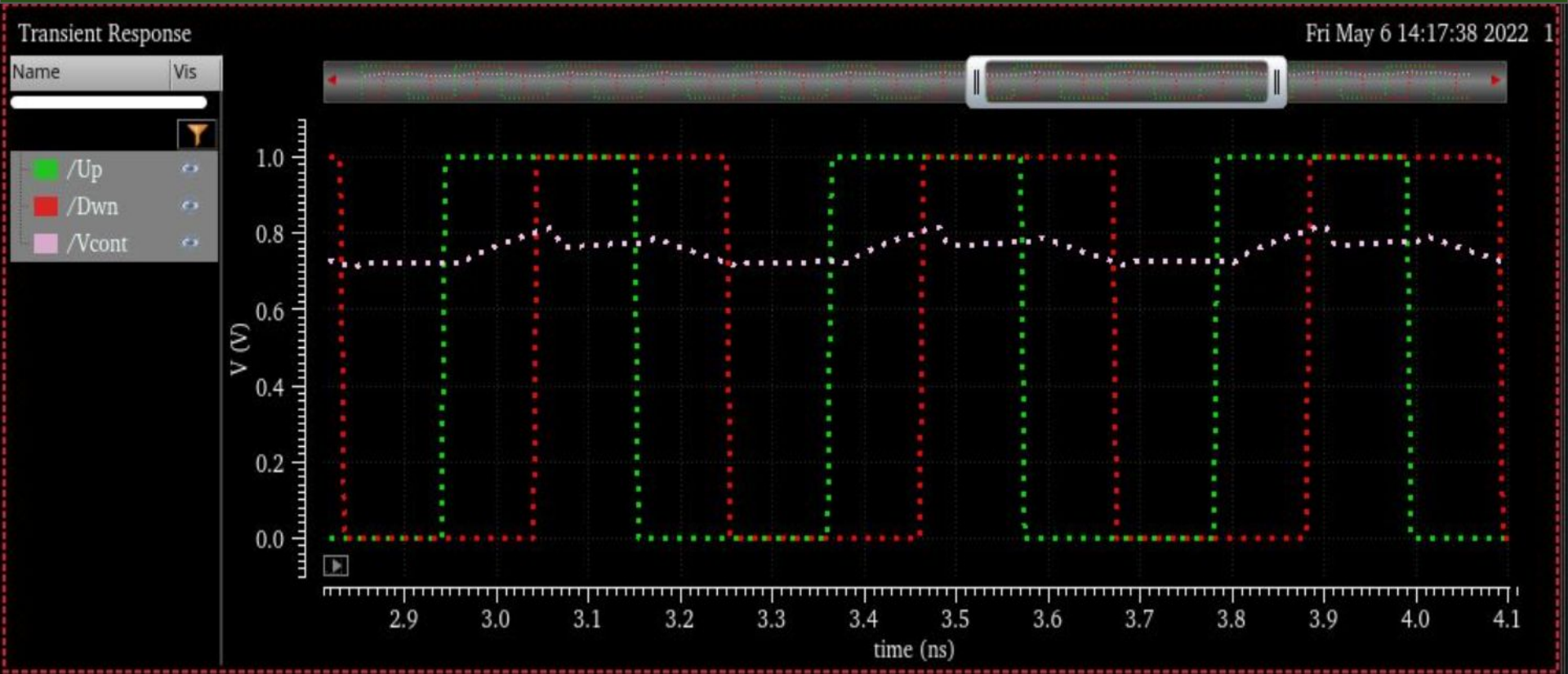
Charge Pump


Charge Pump
V_{DD}
I_1
Up
S_1
V_{out}
Down
S_2
C_1
I_2

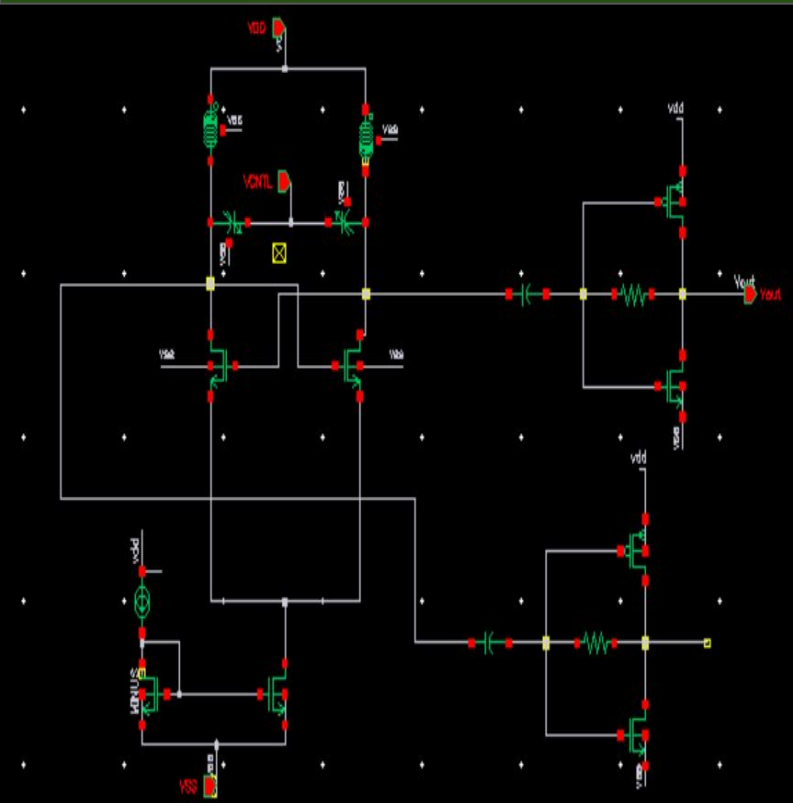


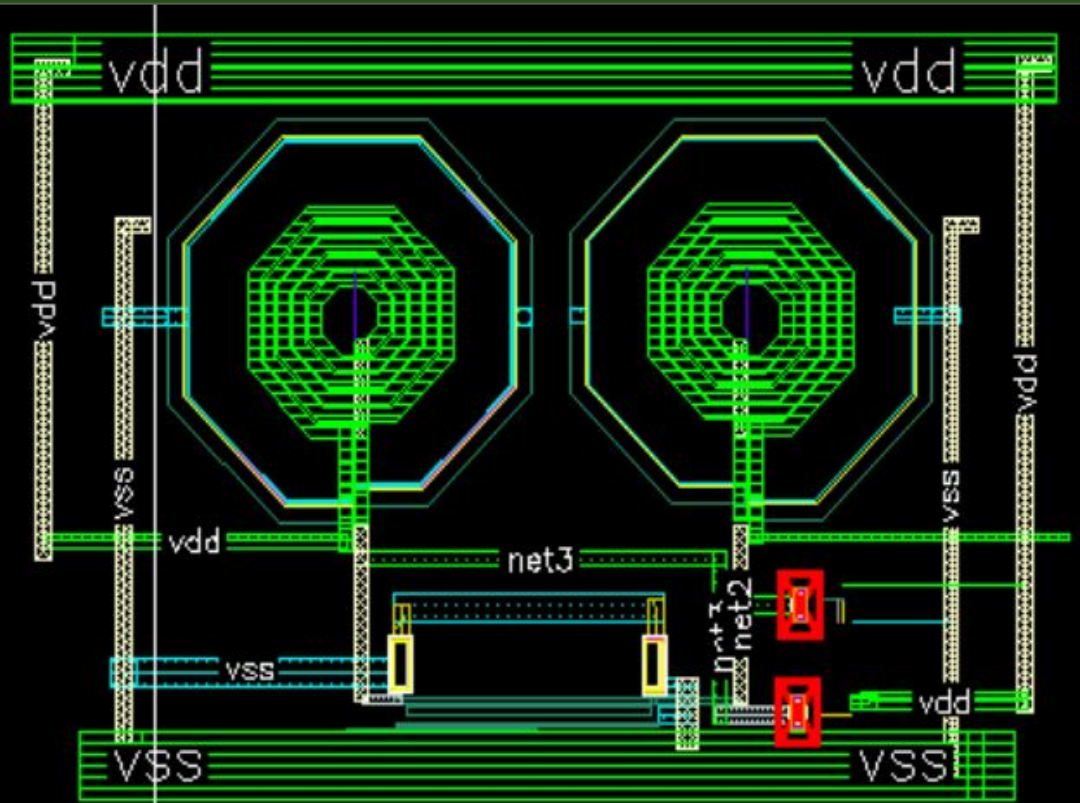

VDD
I0|net2
I0|net8
I0|net
I0|net4
net3
gnd
gnd
I1|net
I1|net4
net9
I1|net2
net7
VDD

Charge Pump Results

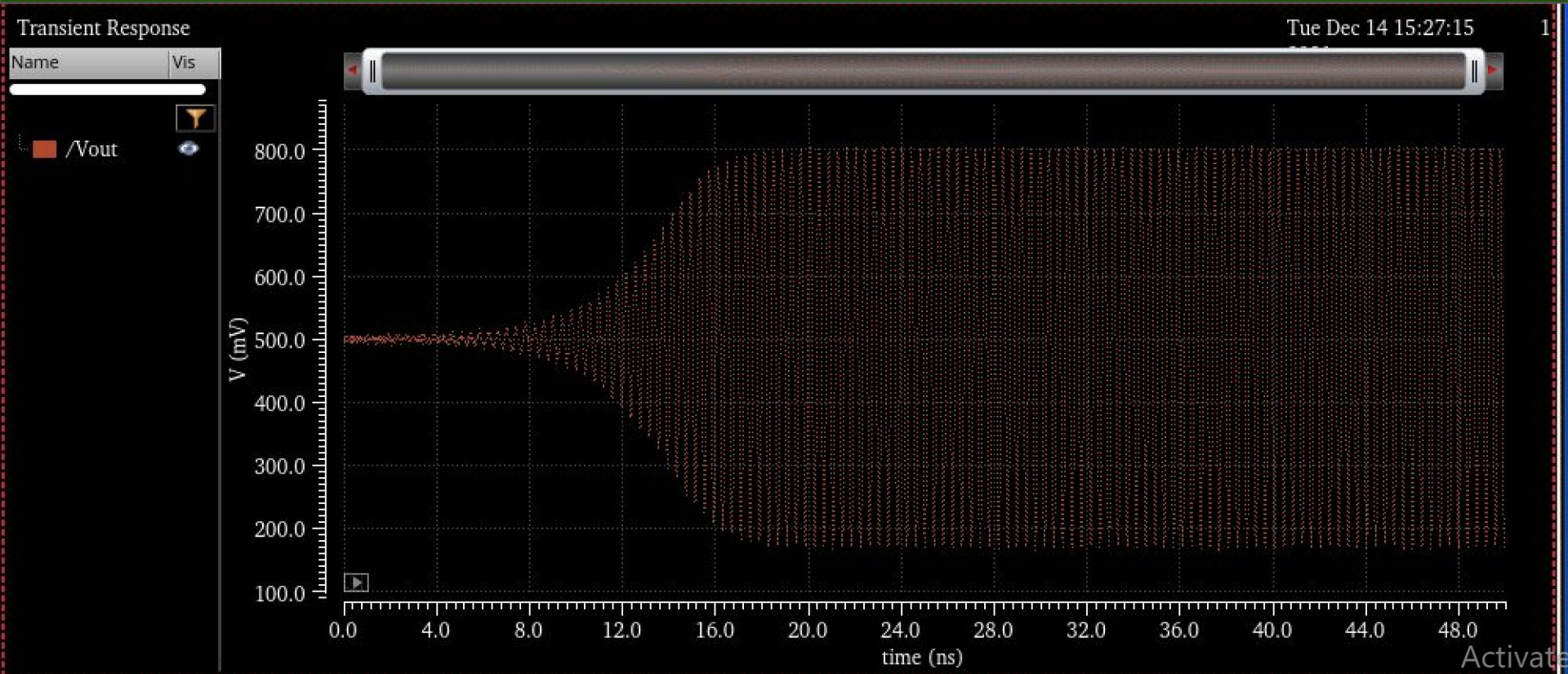


VCO

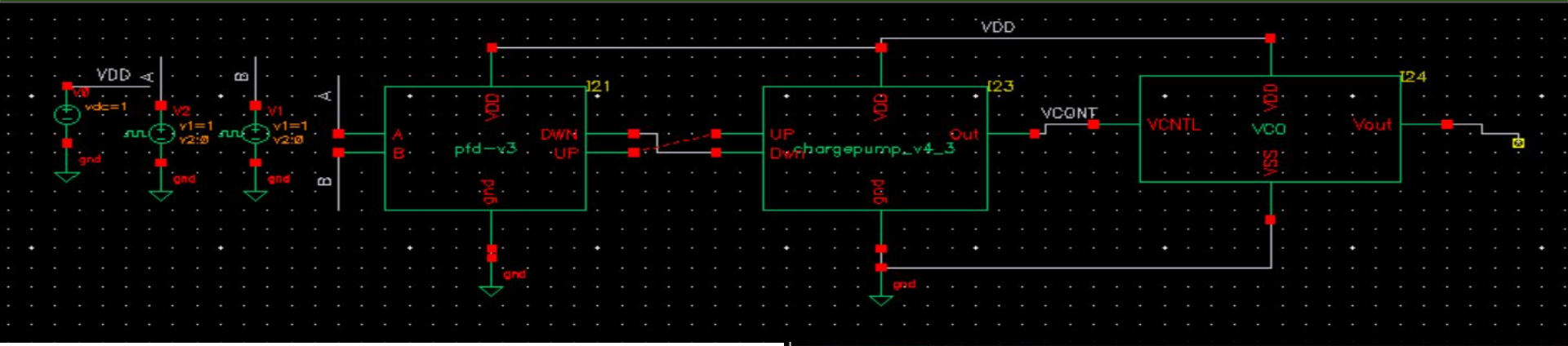

VDD
VCNTL
Vout
vdd
vss


vdd
vdd
vdd
vss
vss
vdd
net3
net2
vss
vdd
VSS
VSS

VCO Results

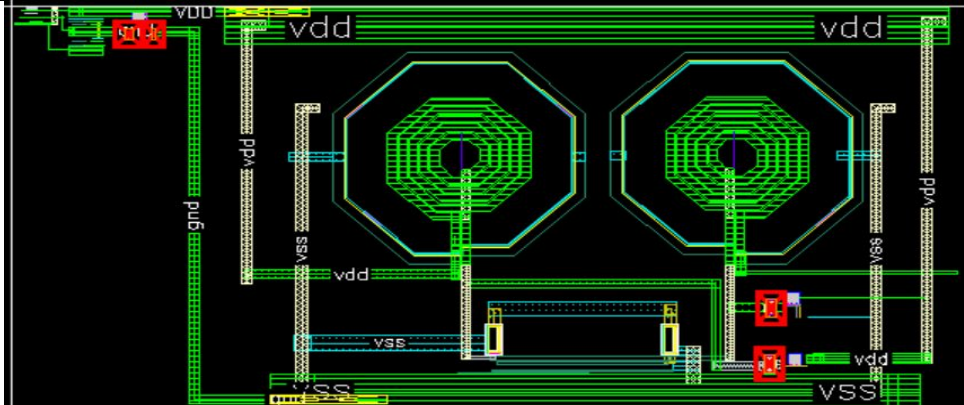

Transient Response
Tue Dec 14 15:27:15
Name
Vis
/Vout
800.0
700.0
600.0
500.0
400.0
300.0
200.0
100.0
V (mV)
0.0
4.0
8.0
12.0
16.0
20.0
24.0
28.0
32.0
36.0
40.0
44.0
48.0
time (ns)

Analog Components

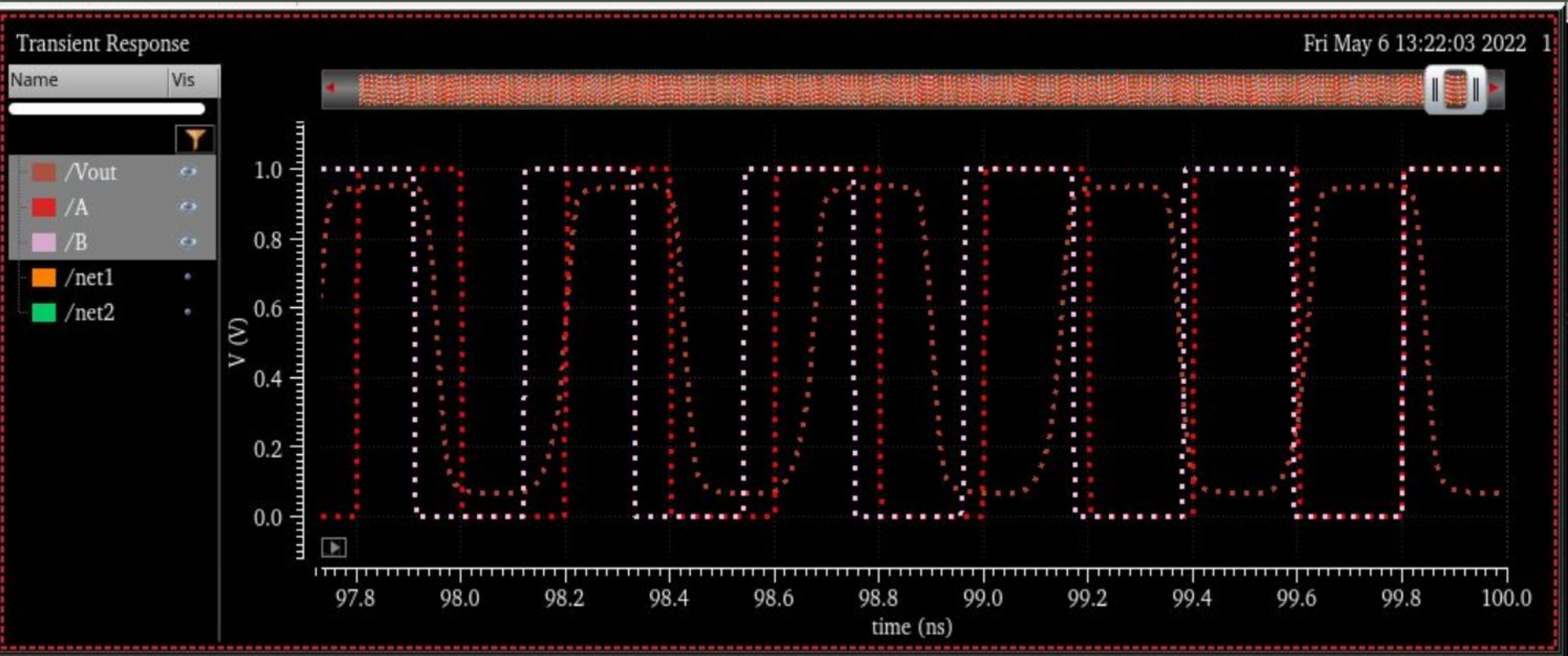


PFD, CP, VCO combined schematic

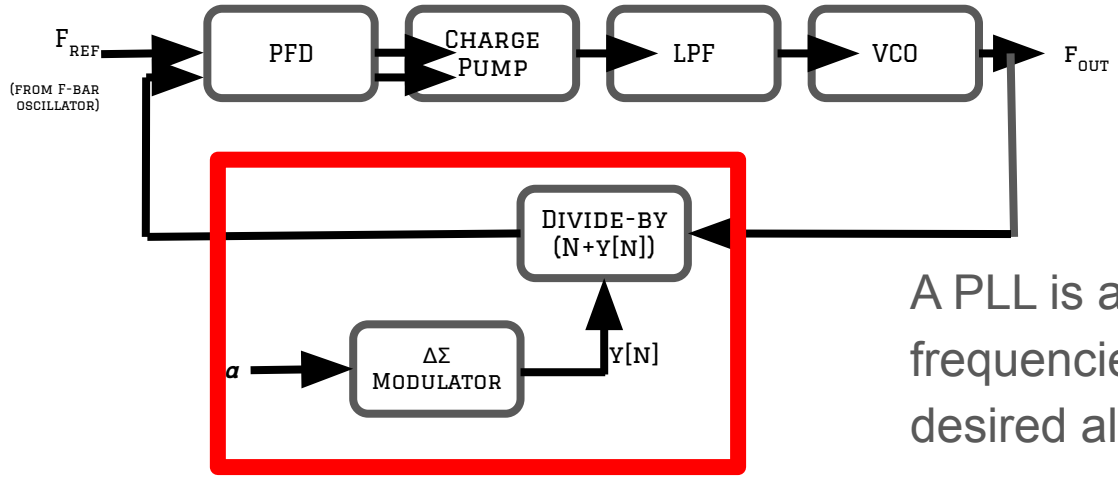
VCO layout dominates due to large inductors



Analog Components Results



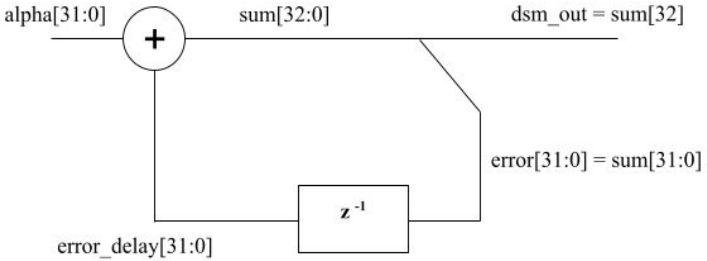
Fractional-N PLL

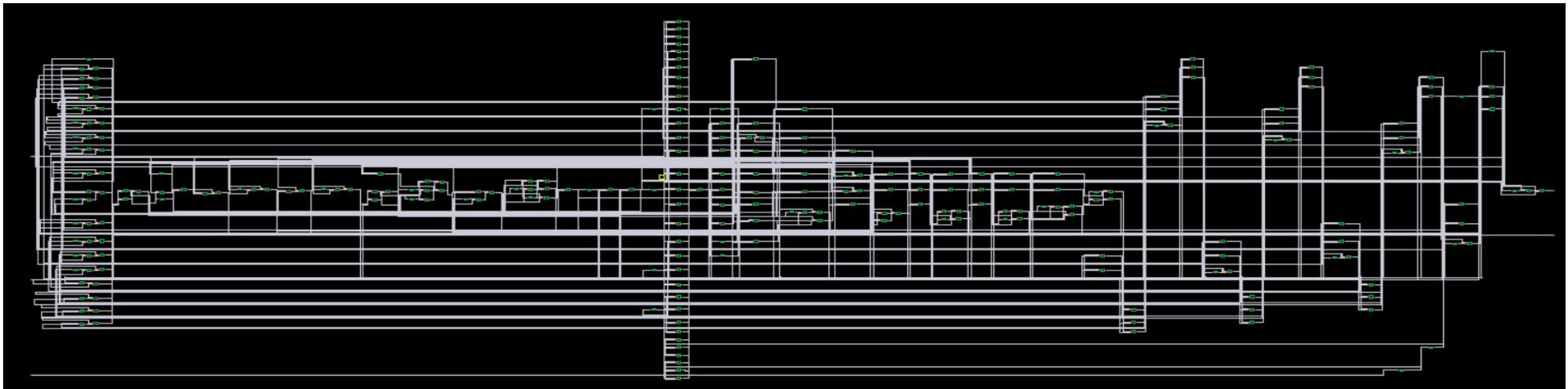


A PLL is able to synthesize fractional frequencies by dividing by N plus some desired alpha.

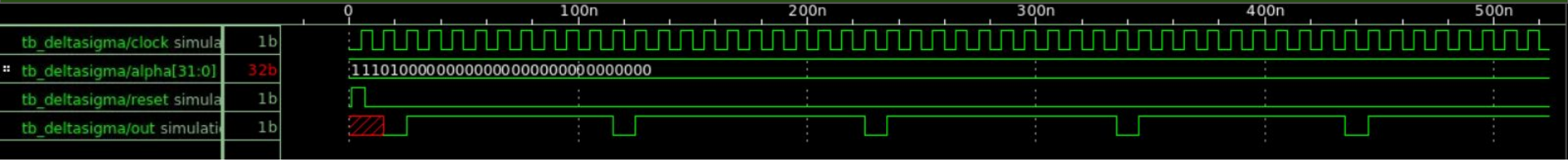
Alpha can be adjusted in tiny increments to achieve fine resolution frequency synthesis.

Delta Sigma Modulator

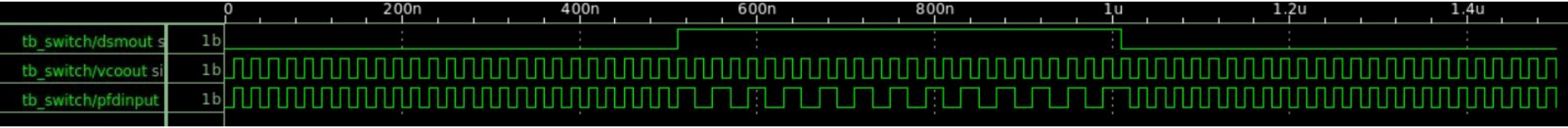




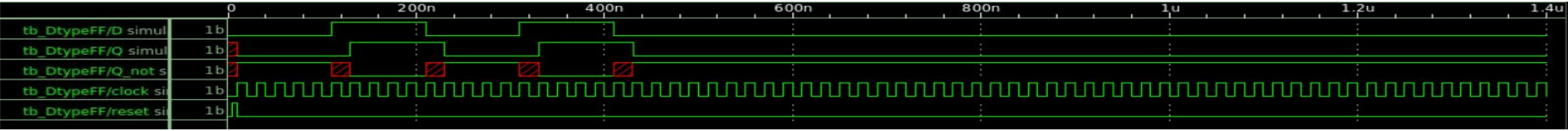
Delta Sigma Modulator Results



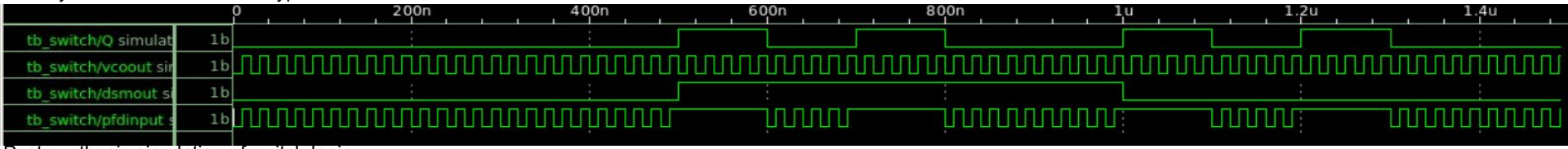
Post-synthesis simulated outputs of DSM module



Simulated PFD input as a result of DSM output signal control and VCO output



Post-synthesis simulation of D type FF



Post-synthesis simulation of switch logic

Next Steps

Next steps would be:

1. Combine the analog and digital components
2. Tapeout
3. Develop package and test procedures for chip

Initial plan was to import digital netlist into Cadence so we can combine with all analog circuits to run top simulation. But, we assumed we had the schematic view for each digital standard cells (we didn't – the TSMC digital library does not have schematic view). This week, we tried to extract the netlist of each analog block (VCO, PFD, and CP) and then build up test bench.

Thank You

Professor Koo

Professor Shlayan

Benchmarking Josephson Junctions for Scalable Quantum Computing

Tamar Bacalu, Alexa Jakob, Mark Koszykowski

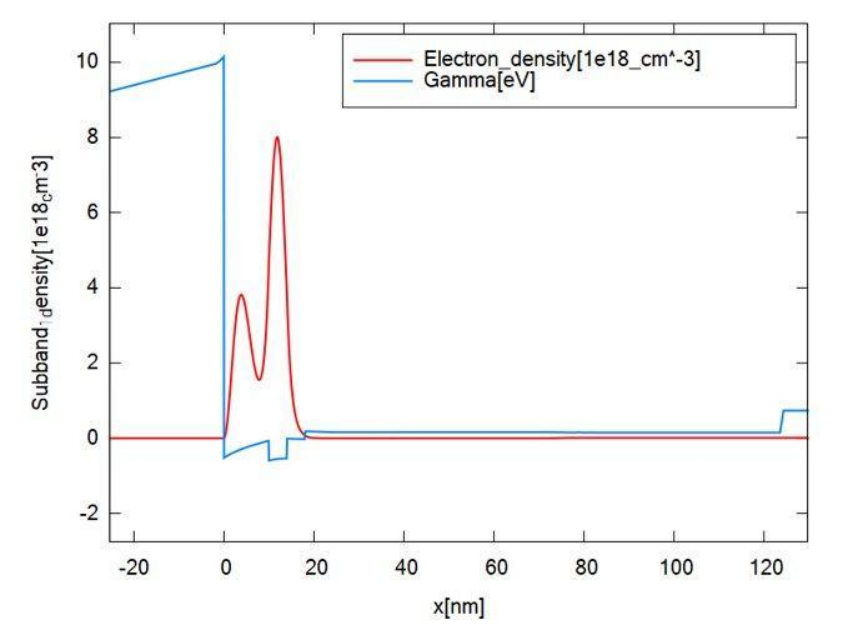
Objective

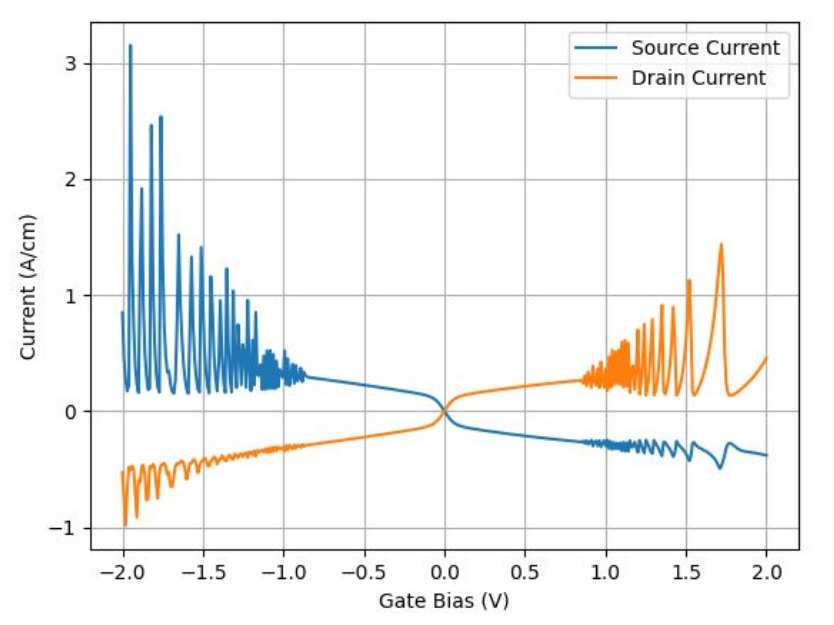
Verification of quantum devices is difficult to scale because of the time constraints associated with measurements. Our goal is to link characteristics of JJFETs at room & cryogenic temperatures to enable faster verification.

We do this by:

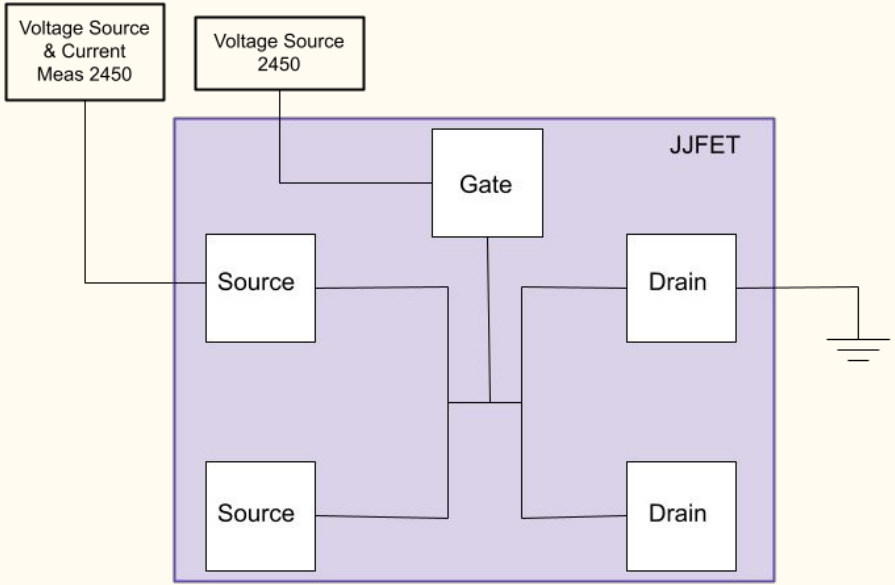
- Simulating devices at room temperature
- Measuring IV curves at room temperature
- Comparing R_N to resistance at room temperature
- Determining a mathematical relationship

Simulation Results

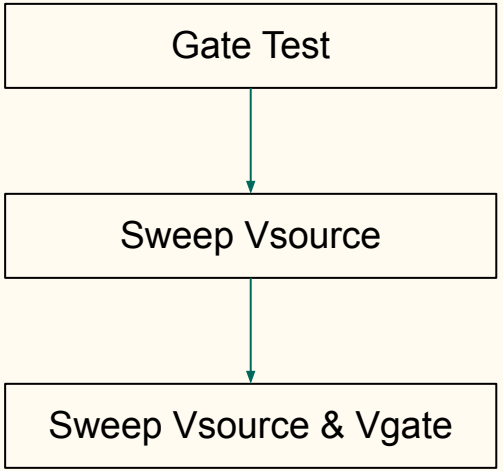




Measurement Setup

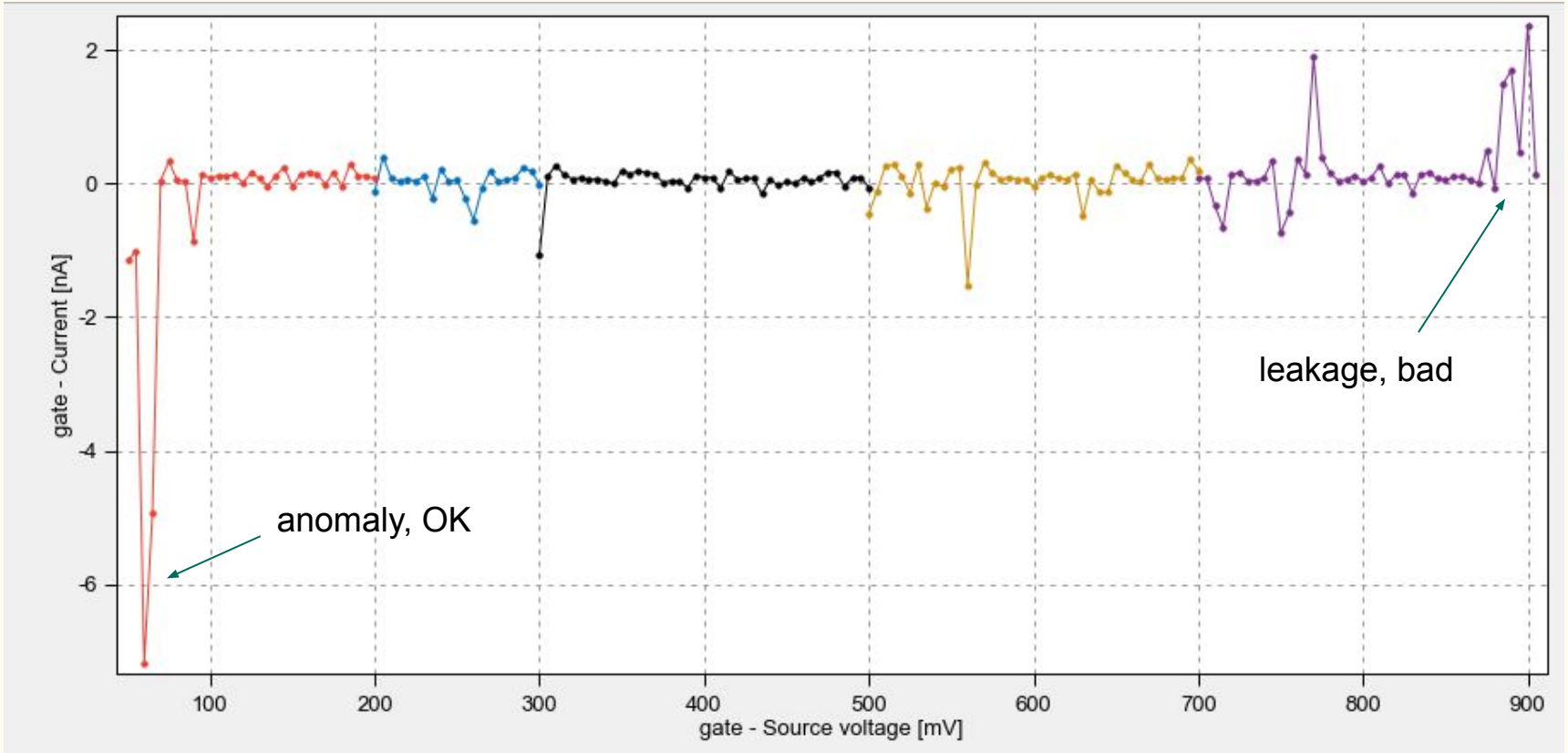


Four-point measurement circuit with voltage bias

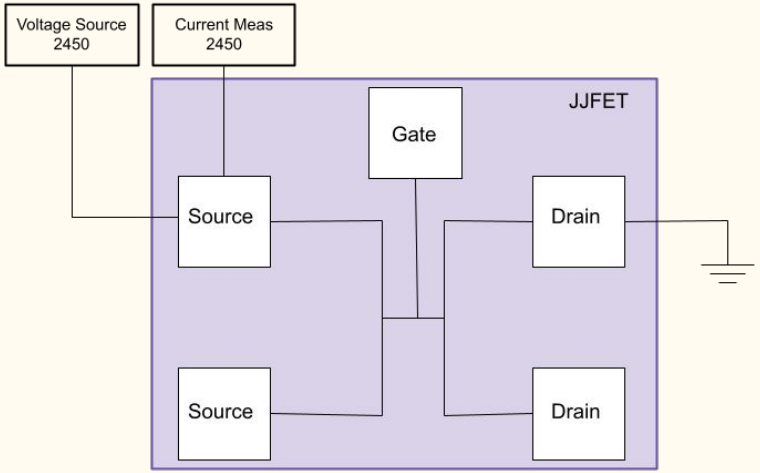


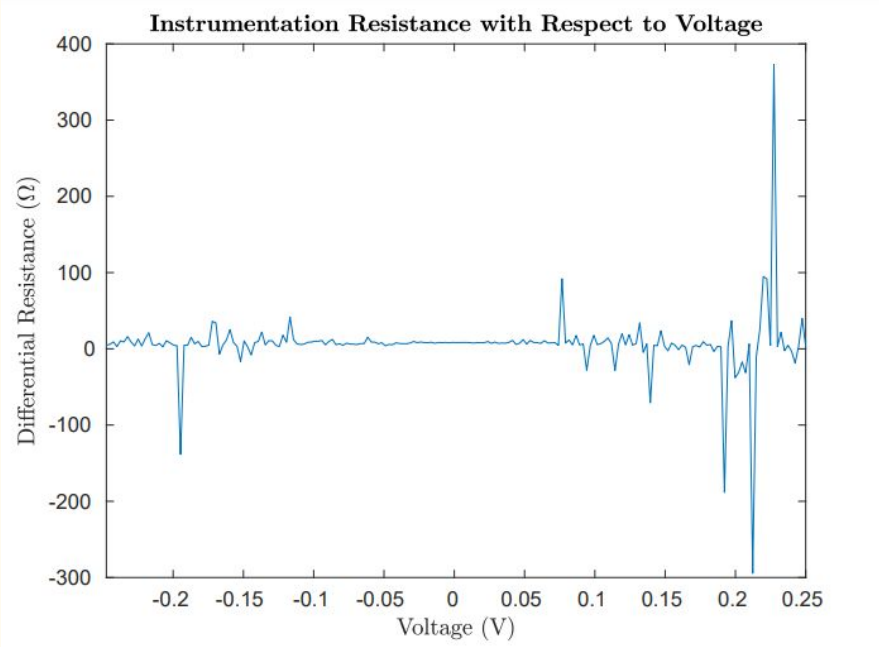
Measurement procedure

Measurement Results - Gate Test



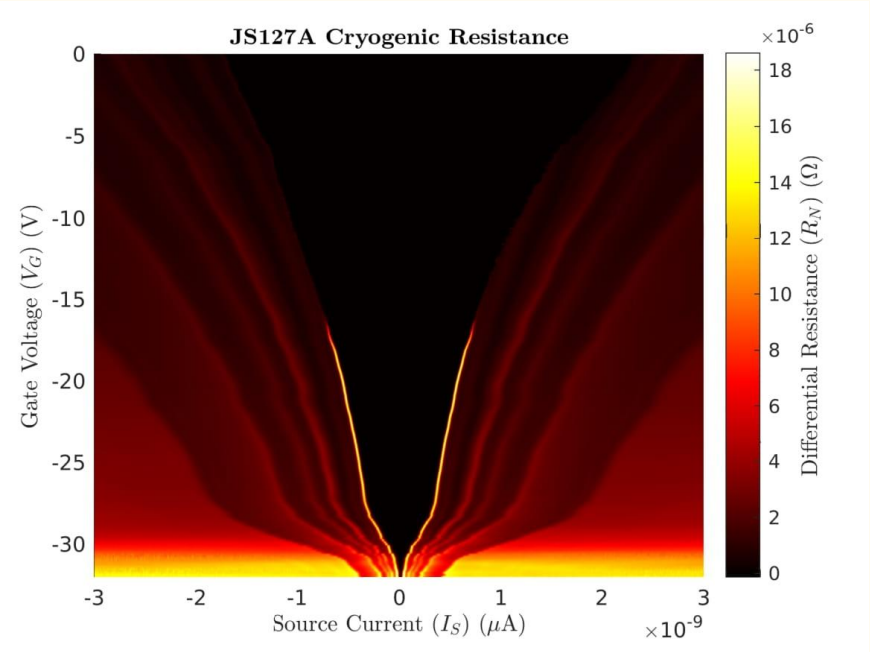
Results & Analysis



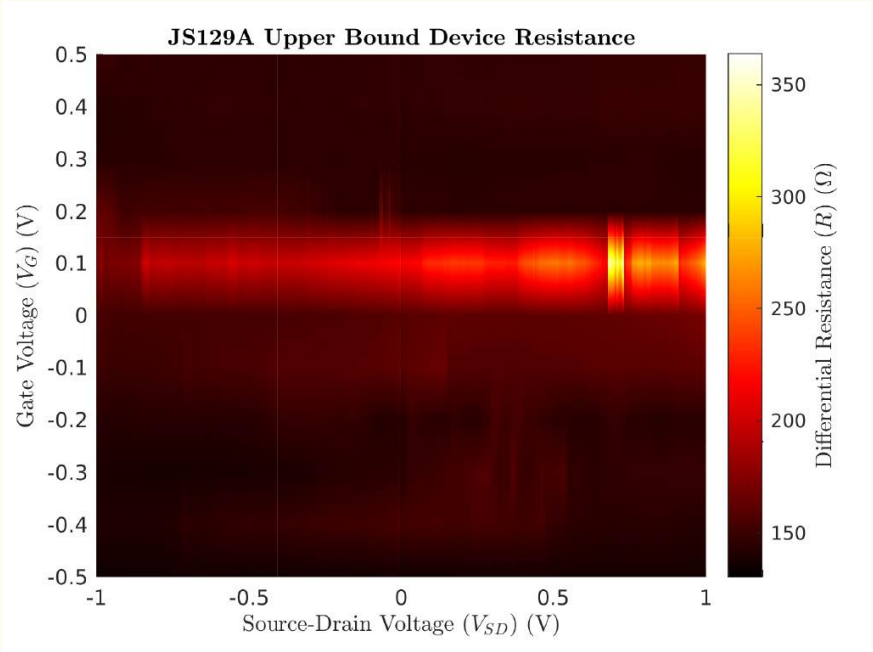


Average contact resistance: 6.3 Ohms

Analysis



Cryogenic Temperature



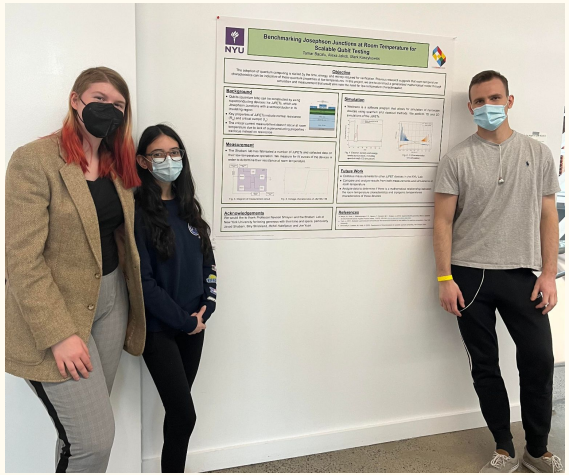
Room Temperature

Conclusion

- Difficult to draw conclusions: lack of data and devices, different measurement setups, yet some evidence to support cryogenic-high temperature relationship
- Provided documented framework for future students interested in project

Many Thanks!

- Cooper Union: Neveen Shlayan
- NYU Shabani Lab: Javad Shabani, Mehdi Hatefipour, Billy Strickland, Mohammed Farzaneh
- NYU Tandon School of Engineering: Zhujun Huang



Cost Effective Gimbal

Anthony Belladonna

Philip Blumin

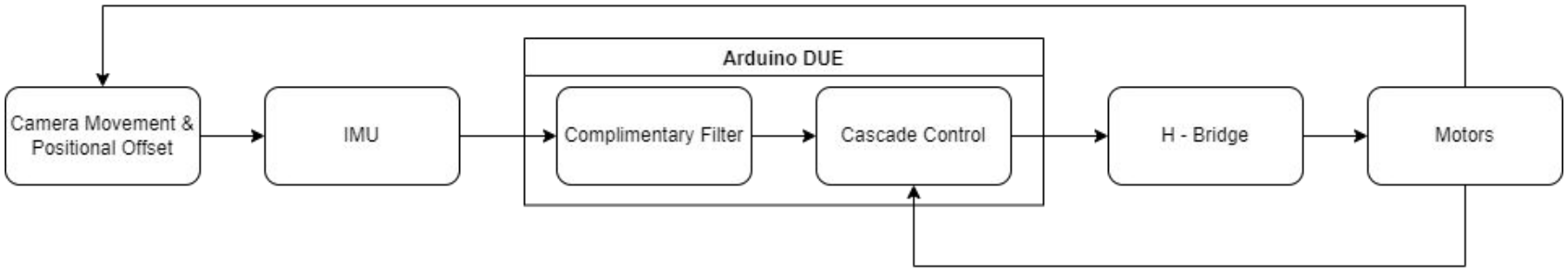
Paul Cucchiara

Daniel Kim

What is a Gimbal?

Features

- Rotational Stability in 2 dimensions
  - (Roll and Pitch)
- Custom MOSFET H - Bridges
- Custom Controller Shield



Gimbal Implementation

Hardware

- Microcontroller
- IMU
- DC Motors
  - Hall Encoders + Planetary Gearboxes
- H-Bridges
- Power Shield
- Power Supply
- Aluminum Frame

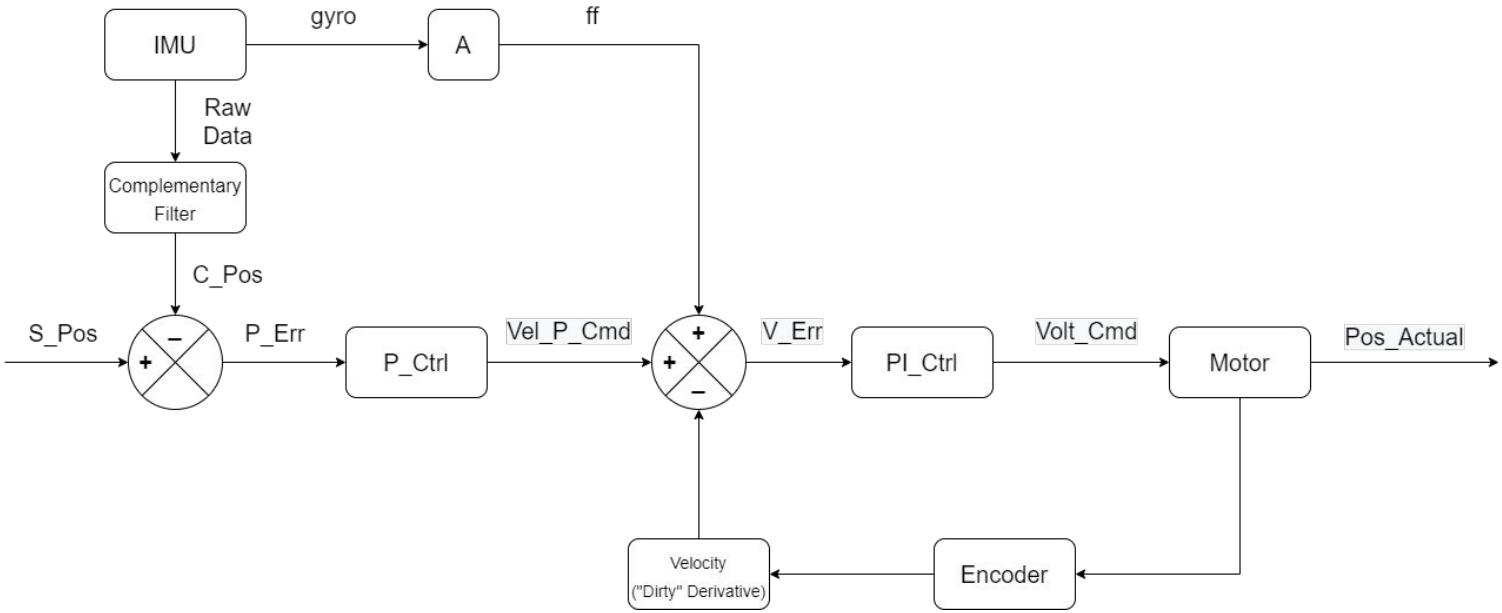
Sensor Fusion Algorithm

- Complementary Filter

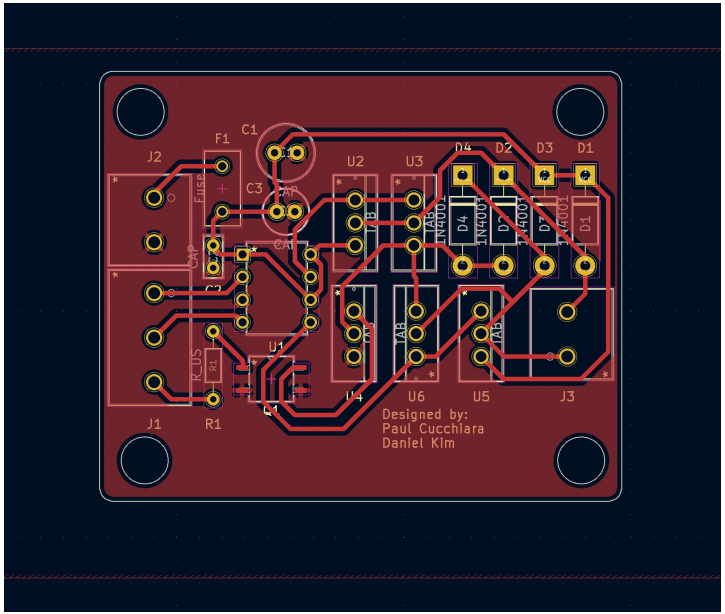
Feedback Controls

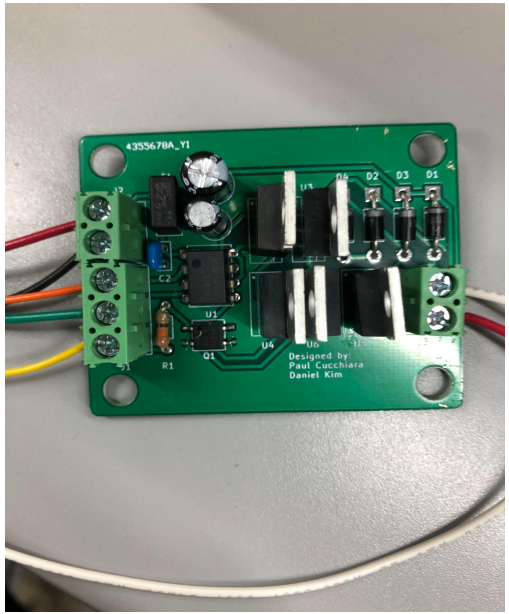
- Cascaded Controller

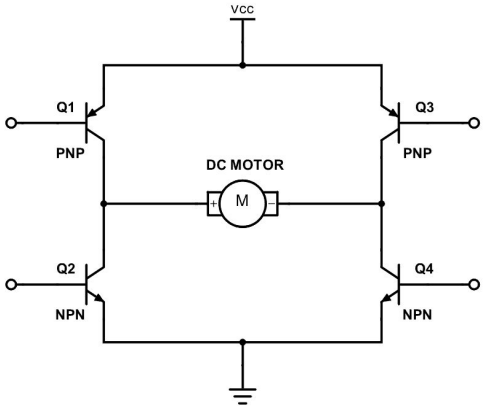
Cascade Control



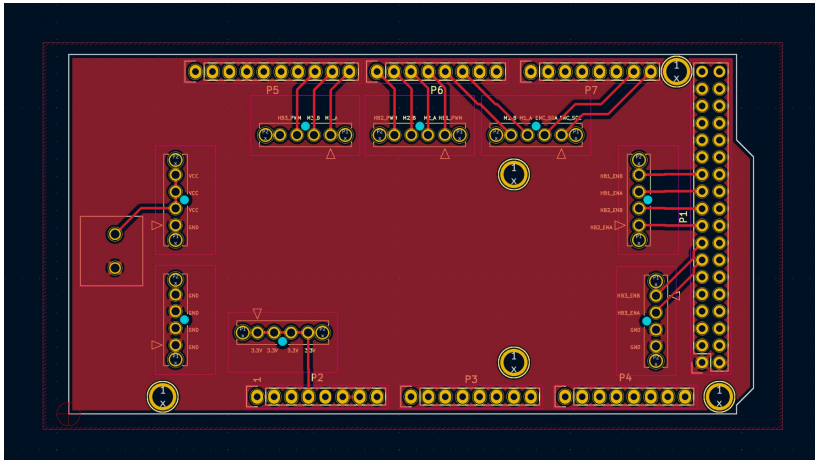
Custom H - Bridges

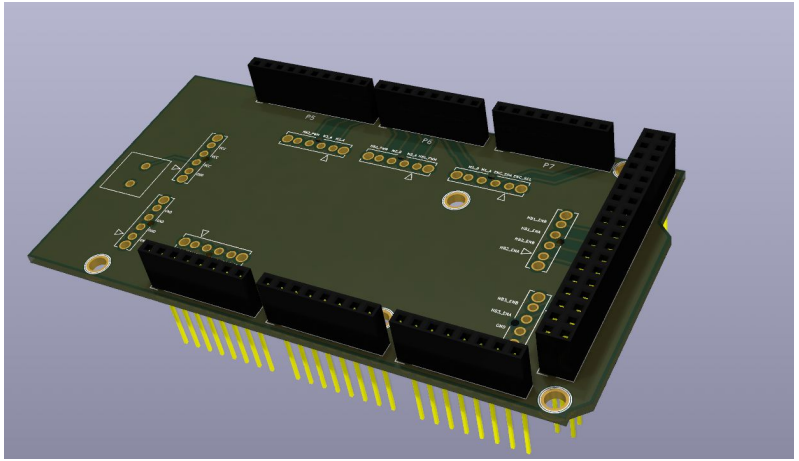




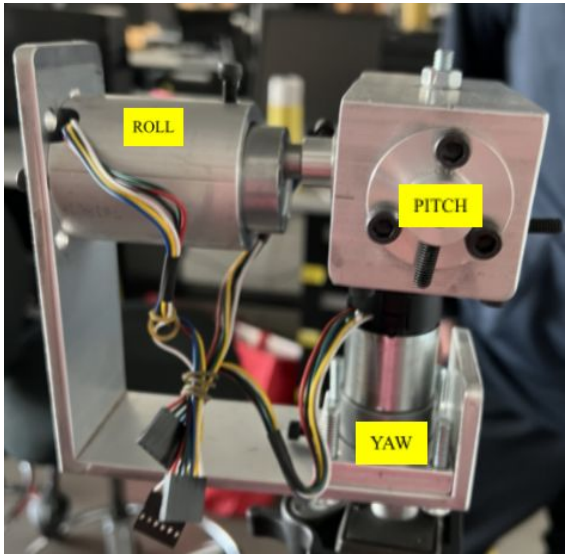


Custom Power Shield





Frame



Final Cost

3 Motors - $72

3 H-bridges - $42

Microcontroller - $40

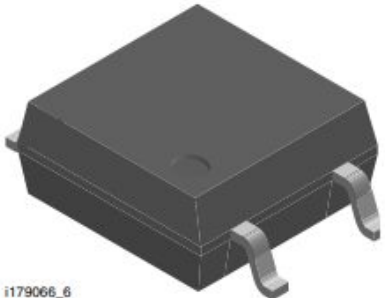
IMU - $3

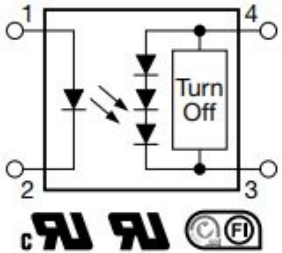
PCB - $2

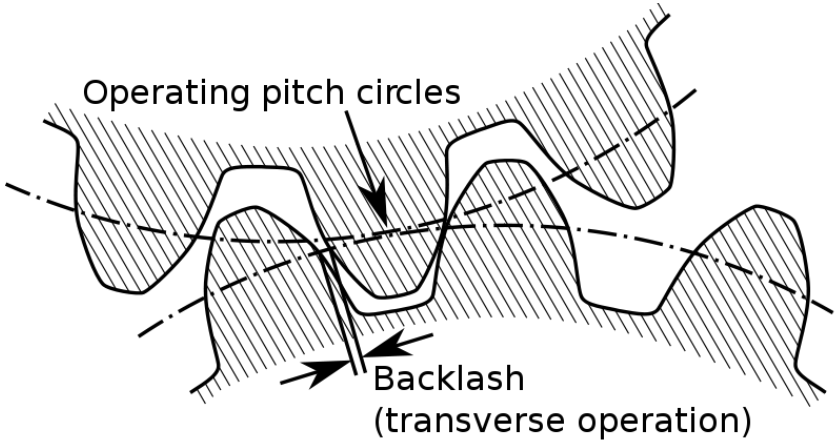
Frame - $5

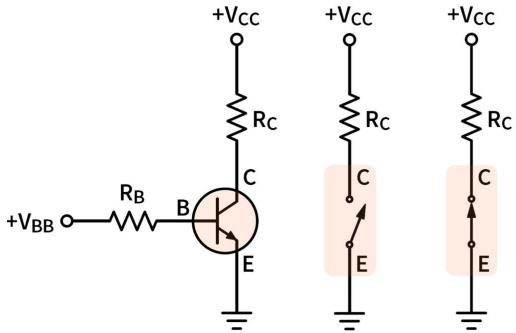
Difficulties & Lessons Learned

- Expensive motors are worth it
- Multiple IMUs are necessary
- Account for microcontroller output









Thank you!


A special thanks to:
Professor Shlayan
Sinisa Janjusevic
Michael Giglia

Look Ahead n Drive (LAnD)

Henry Son

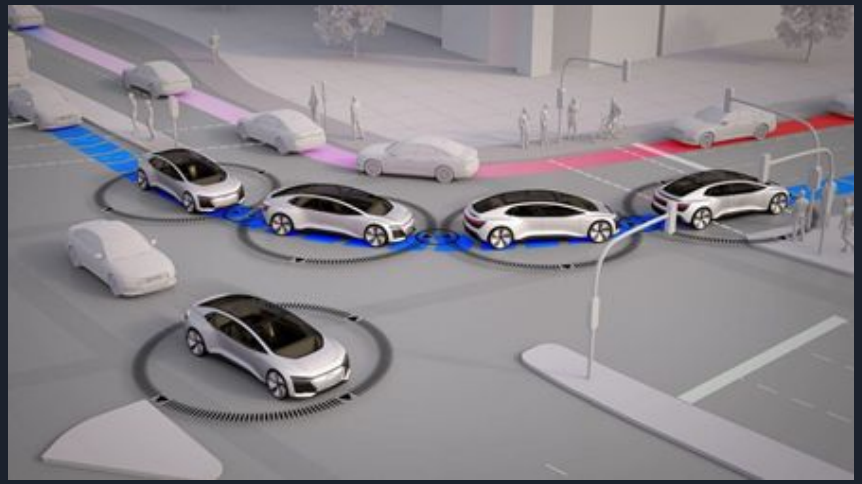
Intelligent Ground Vehicle Competition

IGVC Rules

Sign / Obstacle	Dimensions
"Road Closed"	24" H x 30" W minimum height from ground is 5 feet
"One Way"	12" H x 36" W minimum height from ground is 5 feet
"Stop"	24" H x 24" H minimum height from ground is 5 feet

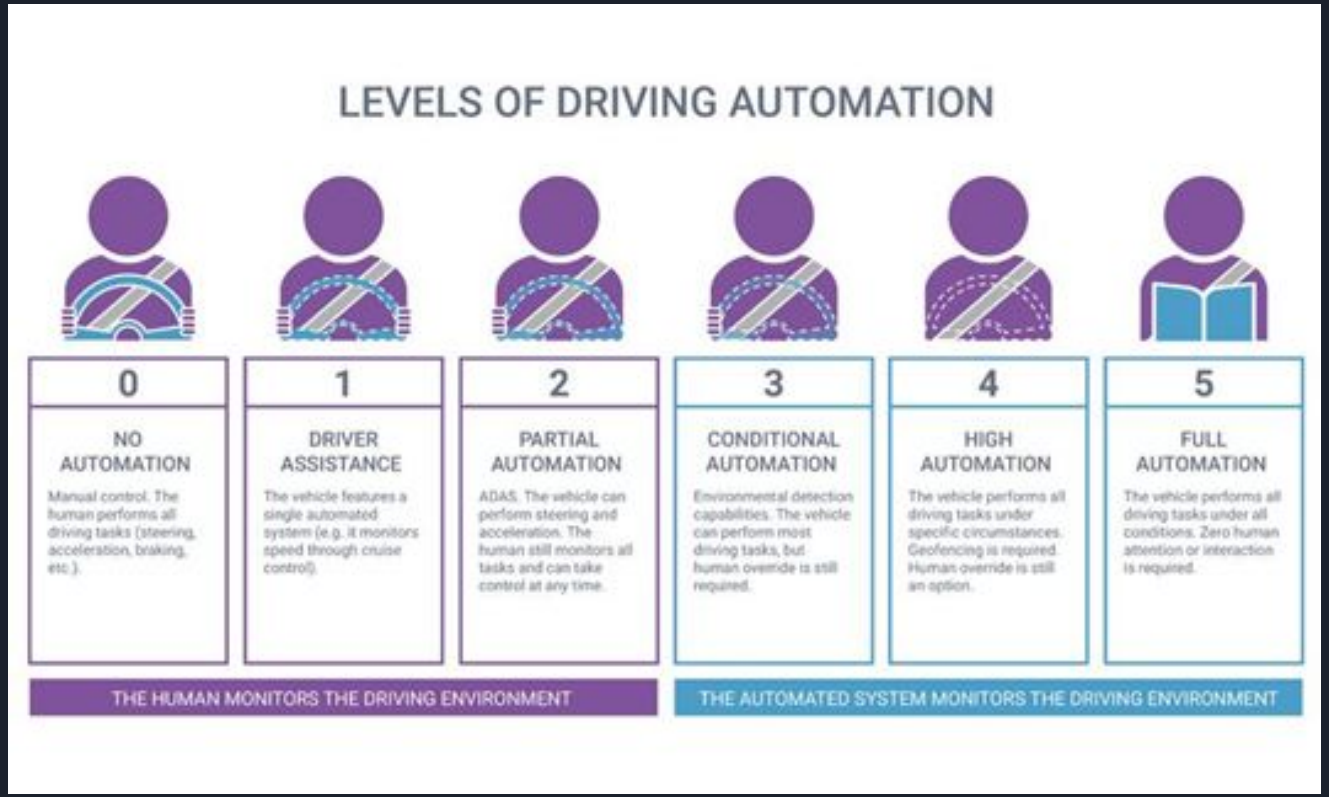
What's the practical application?

- Reliable and quicker transportation
- Reduce deaths from crash fatalities
- Possibly reduce congestion
- Increase productivity from saving time and energy actively driving

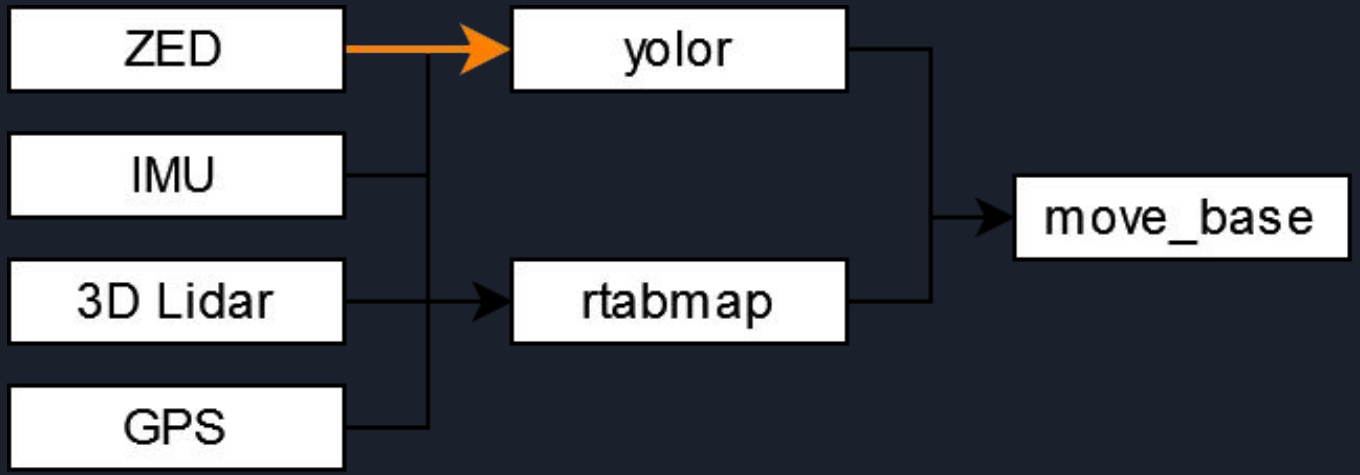


Source: https://drivemag.com/news/pros-and-cons-of-self-driving-cars-will-avs-solve-the-traffic-problem

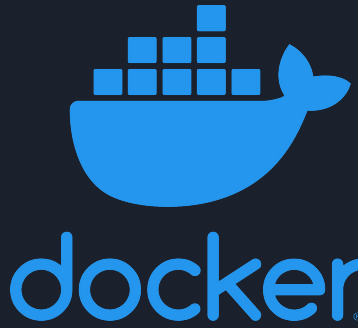
Where are we at?



Technical Stack Block Diagram

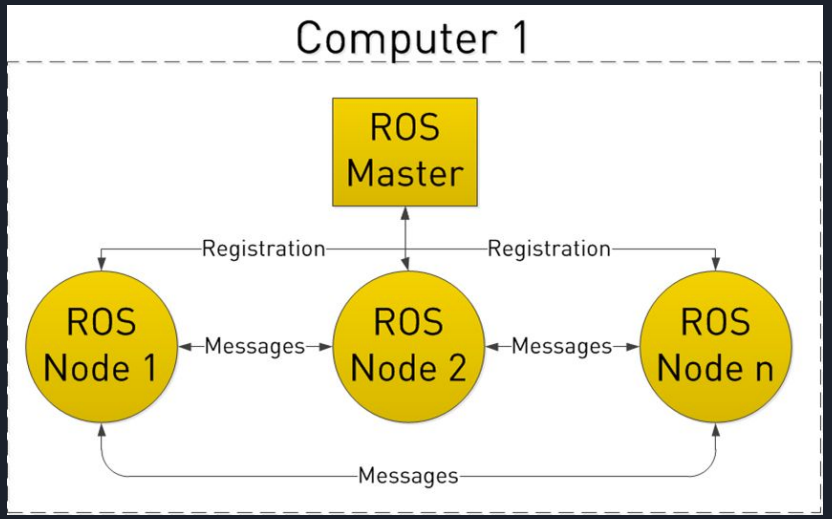


Development Tools

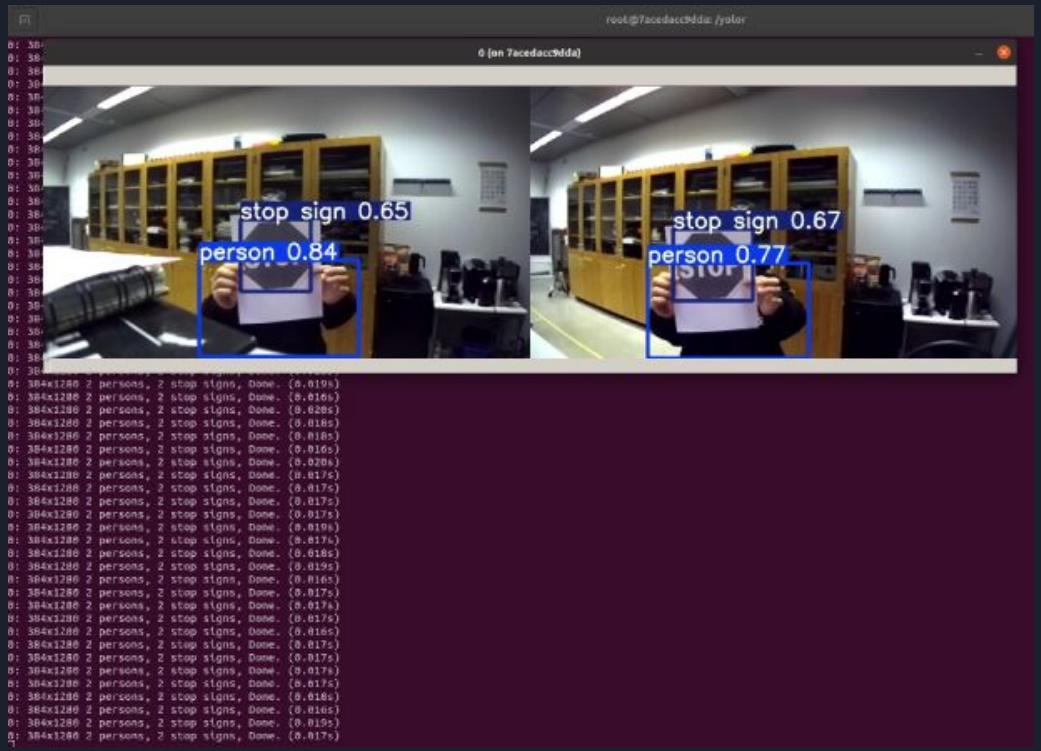


Object Detection

ROS



Object Detection



Future Work

- Create a publisher from the YOLOR node that gives the location of the center of mass of the object with respect to the camera and the object classification
- Optimize the memory usage for the algorithm
  - Shrinking the model
  - Write better code…

Thank you!

- Professor Shlayan
- Mike Giglia
- Dan Mezhiborsky
- Technical Team

Prediction of progression from CKD 4 to CKD 5

A machine learning approach

Presenter: Min Cheng

Group: Min Cheng in collaboration with
Dr. Lili Chan, Dr. Wonsuk Oh
from Mount Sinai

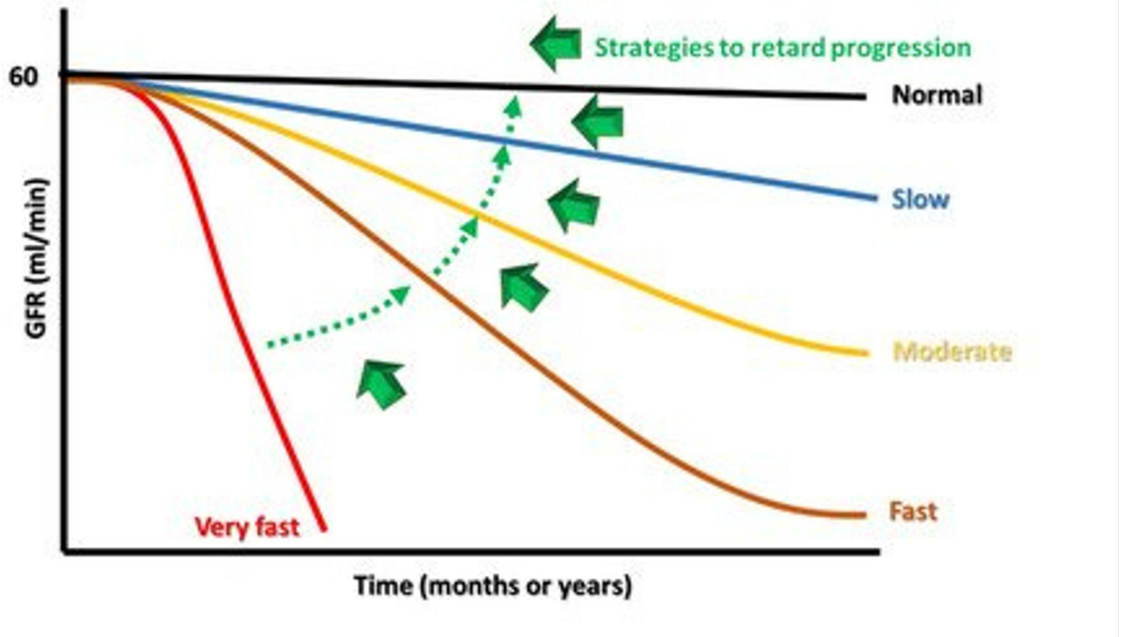
Chronic Kidney Disease stages

CKD stage[a]	GFR (ml/min/1.73m^2 body surface)	Description
1	≥90	Normal or increased GFR, with other evidence of kidney damage
2	60–89	Slight decrease in GFR, with other evidence of kidney damage
3a	45–59	Moderate decrease in GFR, with or without other evidence of kidney damage
3b	30–44	
4	15–29	Severe decrease in GFR, with or without other evidence of kidney damage
5	<15	Established renal failure

CKD, chronic kidney disease; GFR, glomerular filtration rate; NICE, National Institute for Health and Care Excellence.
a. Use the suffix (p) to denote the presence of proteinuria when staging CKD.
Source: National Collaborating Centre for Chronic Conditions (2008).

Patients can progress at different rate

Early identification can allow for dialysis planning.
-- can be referred once eGFR <20



Kidney Failure Risk Calculator - Tangri

Kidney Failure Risk Equations

--Cohort
CKD stages 3 to 5
Mean eGFR: 36 mL/min/1.73 m_2

--Outcome:
Probability of kidney failure at 5 years
Dialysis
Transplantation

Table 2. Hazard Ratios and Goodness of Fit for Sequential Models in the Development Data Set[a]

	Models						
Variable	1	2	3	4	5	6	7
Baseline GFR, per 5 mL/min/1.73 m^2		0.54	0.57	0.58	0.60	0.61	0.64
Age, per 10 y	0.86	0.75	0.80	0.80	0.79	0.82	0.82
Male sex	1.03[b]	1.46	1.26	1.27	1.34	1.16[b]	1.26
Log spot urine ACR[c]			1.60	1.61	1.55	1.42	1.37
Diabetes				0.86[b]			0.88[b]
Hypertension				1.17[b]			0.89[b]
Systolic BP, per 10 mm Hg					1.15		1.14
Diastolic BP, per 10 mm Hg					1.10		1.15
Body weight, per 10 kg					0.91		0.91
Serum albumin, per 0.5 g/dL						0.84	0.83
Serum phosphate, per 1.0 mg/dL						1.27	1.34
Serum bicarbonate, per 1.0 mEq/L						0.92	0.93
Serum calcium, per mg/dL						0.81	0.82
C statistic[d]	0.56	0.89	0.91	0.91	0.92	0.92	0.92
Akaike Information Criterion[d]	5553	4834	4520	4521	4463	4432	4378
P value		<.001	<.001	.40	<.001	<.001	<.001

Abbreviations: ACR, albumin-to-creatinine ratio; BP, blood pressure; GFR, glomerular filtration rate.
[a] Data are presented as hazard ratios unless otherwise specified. Models 2, 3, and 6 columns indicate models based on laboratory data. *P* values are for comparison of C statistics between successive models, except for models 5, 6, and 7, which are compared with model 3.
[b] Hazard ratios with $P > .05$; all other hazard ratios are significant (ie, $P < .05$).
[c] Hazard ratio for ACR represents a 1.0 higher ACR on the natural log scale. For the average patient with 20 mg/g of albuminuria, this represents an increase to 55 mg/g.
[d] Null values for C statistic and Akaike Information Criterion are 0.50 and 5569, respectively. Higher values for C statistic and lower values for Akaike Information Criterion indicate better models.

Tangri, Navdeep & Inker, Lesley & Griffith, John & Tighiouart, Hocine & Djurdjev, Ognjenka & Naimark, David & Levin, Adeera & Levey, Andrew. (2011). A Predictive Model for Progression of Chronic Kidney Disease to Kidney Failure. JAMA : the journal of the American Medical Association. 305. 1553-9. 10.1001/jama.2011.451.

Problem Statement

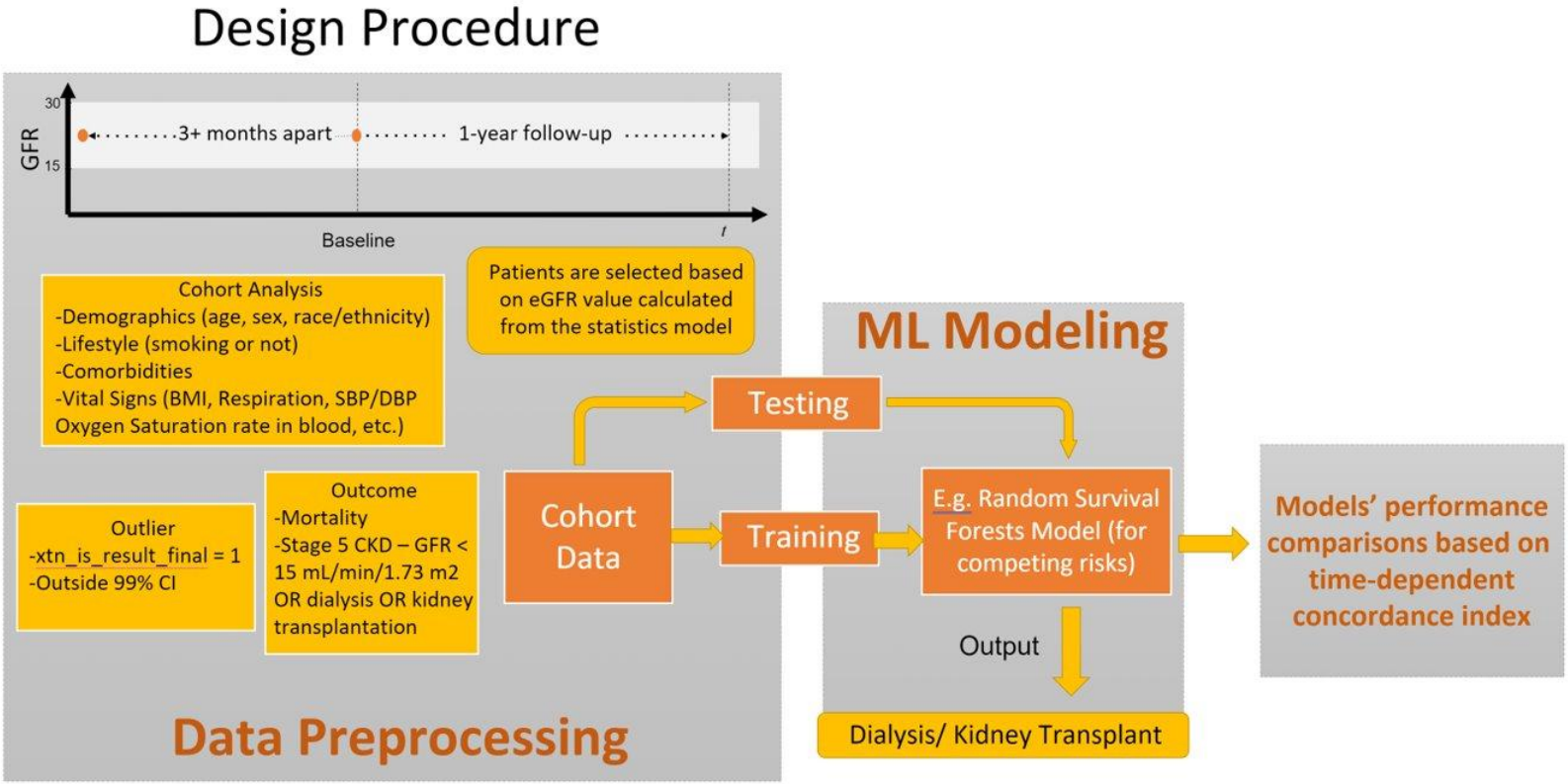
- **Can we build a machine learning model using longitudinal EHR data to generate a risk prediction model for progression from CKD 4 to CKD5**?

- **Hypothesis**
    1. The use of ML models can improve prediction accuracy compared to statistical models.
       --We all know that this hypothesis is no longer a hypothesis; rather, this can be a fact.

    2. The use of competing risk models can improve prediction accuracy than non-competing risk models.

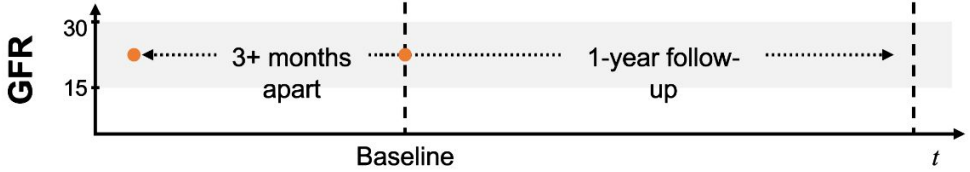
    3. The use of annual laboratory data, e.g., Hb, eGFR, SCr, SCa, etc., can improve prediction accuracy than models using the latest known laboratory data.

Flowchart Overview

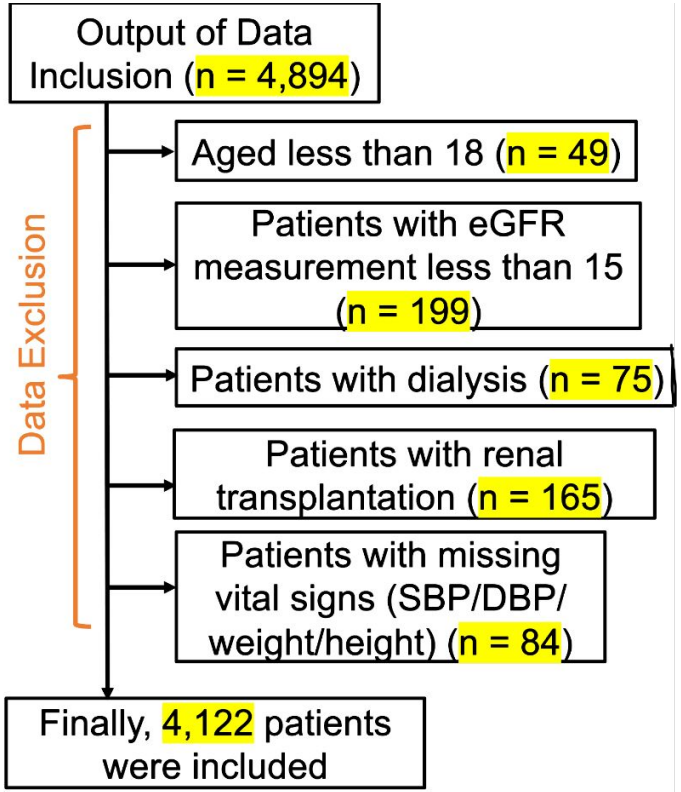


Methods and Results

- **Study design**
  - Mount Sinai Health System
  - Observation: 1/1/2011 – 12/31/2019
  - Inclusion
    - 2 GFR 15-29 with 3 months apart
    - 1+ GFR within 1 year after the 2nd GFR 15-29



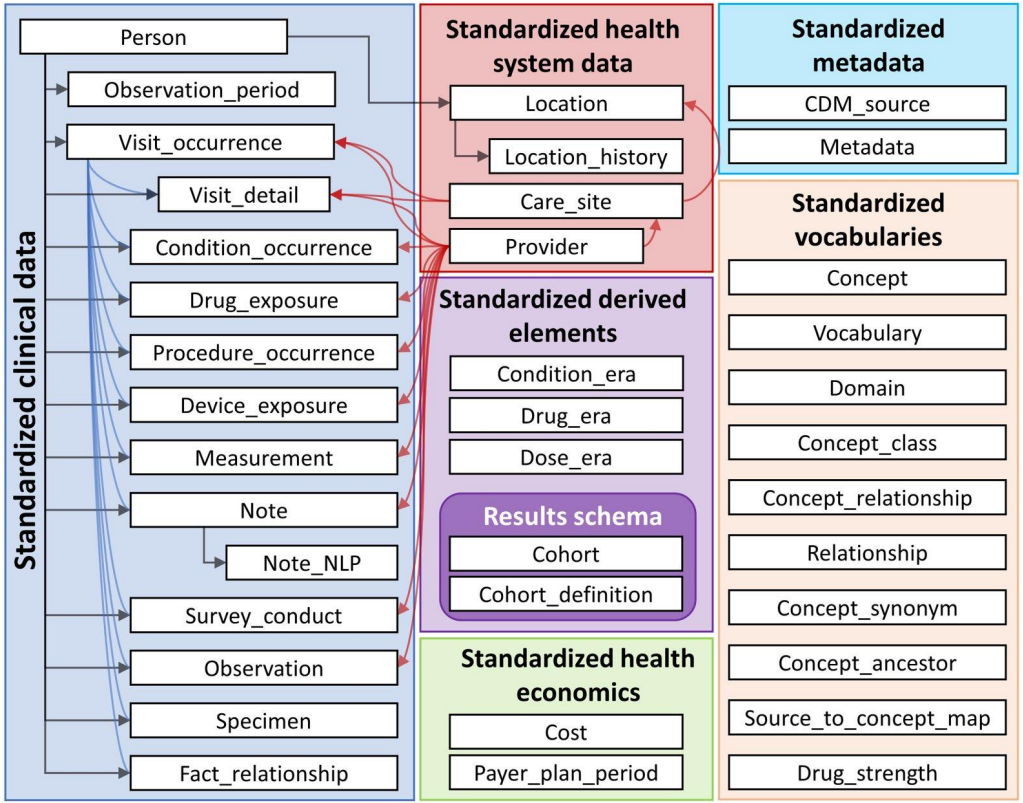
  - Exclusion
    - Aged less than 18
    - Patients with eGFR measurement less than 15
    - Patients with dialysis
    - Patients with renal transplantation
    - Patients with missing vital signs (SBP/DBP/weight/height)



Methods and Results

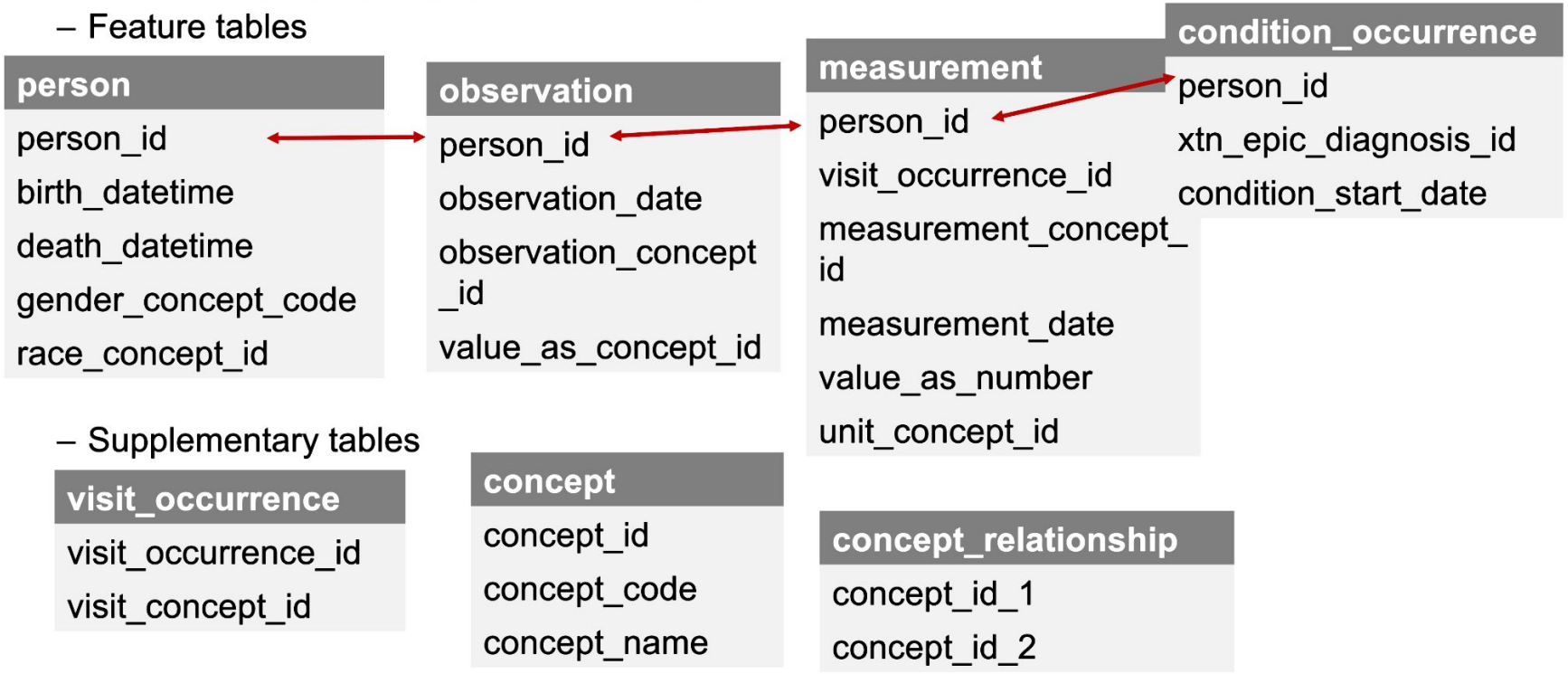
Data collection and Measurements

Mount Sinai Data Warehouse:
-- OMOP Common Data Model



Methods and Results

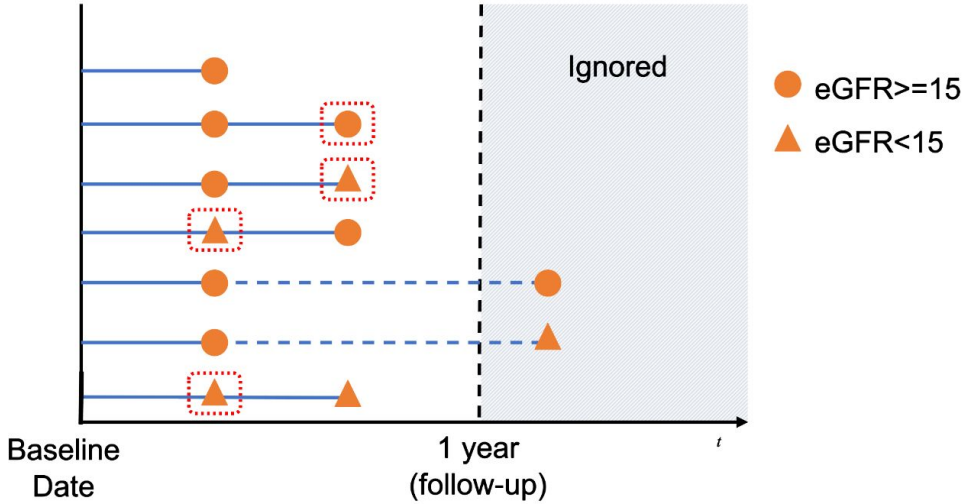
- **Data collection and Measurements**
  - Feature tables



  - Supplementary tables

Methods and Results

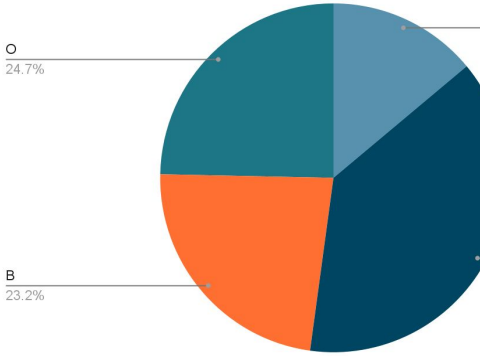
- **Data preprocessing**
  - Outcomes
    - Potential outcomes
      - Mortality
      - Stage5 CKD – GFR < 15 mL/min/1.73m^2 OR dialysis OR kidney transplantation
    - Right censoring (based on eGFR value)



Methods and Results

- **Data preprocessing**

– Demographics
• Age • Sex – male, female • Race/ethnicity – white, black, hispanic, other
– Complications
• Comorbidities used in Elixhauser comorbidity index
– Vital signs (outpatient setting)
• SBP – LOINC 8480-6 • DBP – LOINC 8462-4 • Pulse – SNOMED 78564009 • BMI – body weight (LOINC 8302-2) and body height (LOINC 29463-7) • Body temperature – LOINC 8310-5 • Respiratory rate – LOINC 9279-1 • Oxygen saturation in Blood – LOINC 20564-1
-- **Laboratory data** (GFR, serum creatinine, hemoglobin, serum calcium, serum phosphate, serum albumin, serum bicarbonate, urine albumin, urine creatinine)
-- Treatment



Outcome:

--Survival outcome

→features in the models

--Competing risk outcome

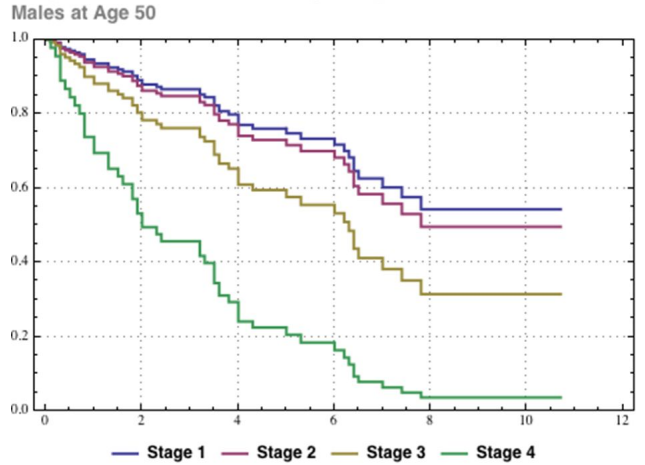
Ongoing and Future Works

- **Models**

	Statistical Models		ML Models	
	Adjusted	Adjusted+ Competing risk	Adjusted	Adjusted+ Competing risk
Semi-parametric	• Cox proportional hazard model	• Fine-Gray model	• Random survival forests • XGBoost • DeepSurv	• Random survival forests for competing risks
Parametric	• Weibull proportional ha zards model		• Deep recurrent survival analysis	• DeepHit

- **Performance metric**
  - Time-dependent concordance index

Expected Survival outcome based on Cox prop-hazard model



Many thanks to

Dr. Lili Chan (Nadkarni Lab, Icahn School of Medicine at Mount Sinai)

Dr. Wonsuk Oh (Nadkarni Lab, Icahn School of Medicine at Mount Sinai)

Prof. Neveen Shlayan (EE Department, Cooper Union)

Prof. Eric Lima (MechE Department, Cooper Union)

Exploring Feedback Control for Autonomous Vehicles

Nathaniel Kingsbury

Custom Motor Drive for Electric Racing Applications

Samuel Shersher

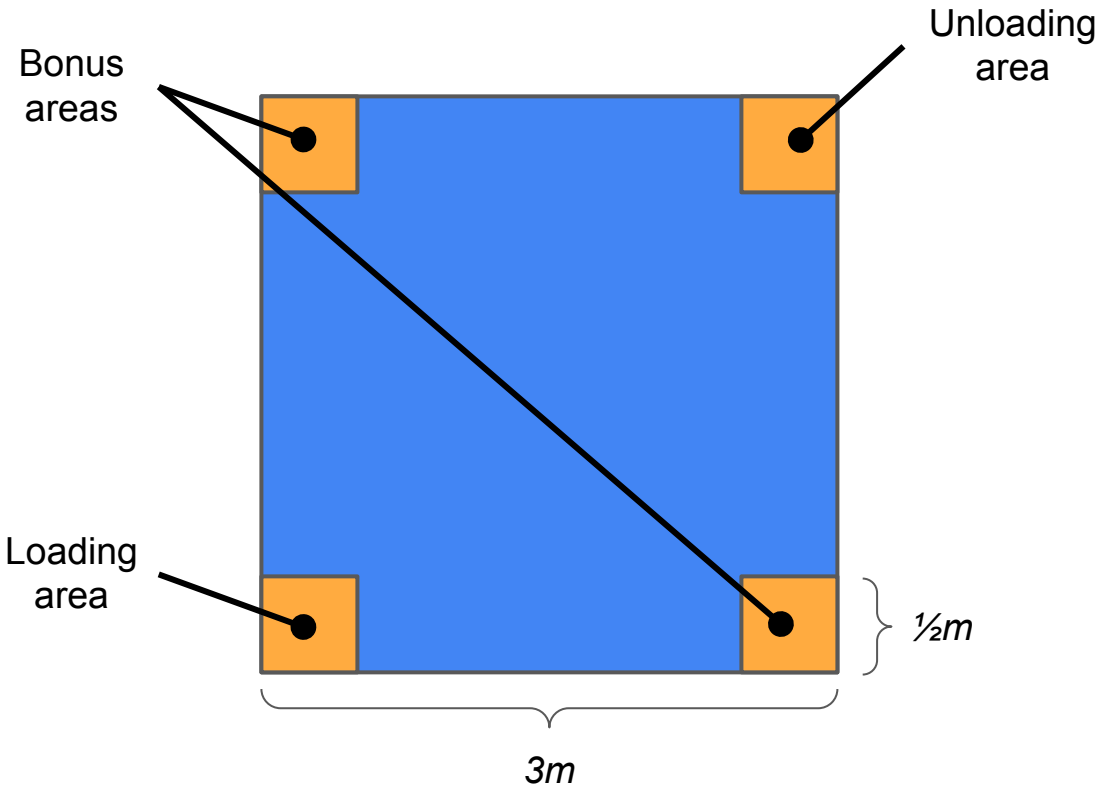
H_2Go: Water-Powered Cart

Brandon Ho (ME'22), Jared Jacobowitz (ME'22), and Josh Yoon (EE'22)

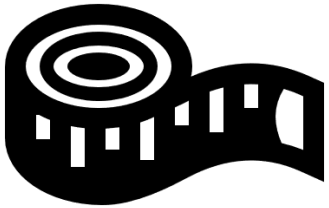
Advisors: Prof. David Wootton and Prof. Neveen Shlayan

Introduction

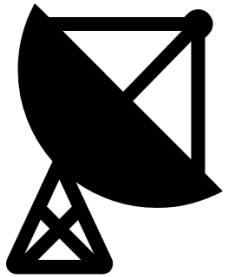
The Competition



H_2Go Competition Requirements



The device must fit within a 50cm x 50cm x 50cm box

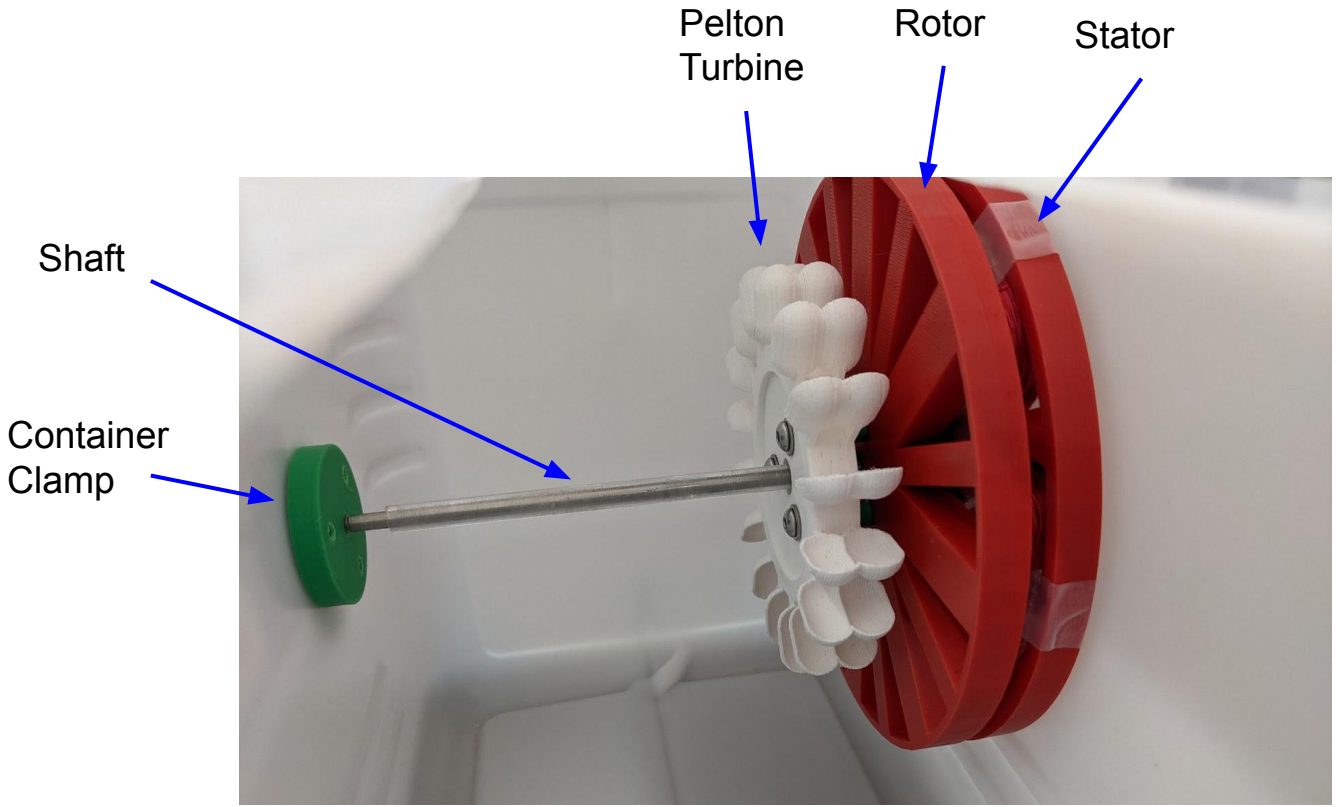


The device must be remote-controlled

The device cannot have any initially stored energy

Design

Assembled Alternator



Rotor

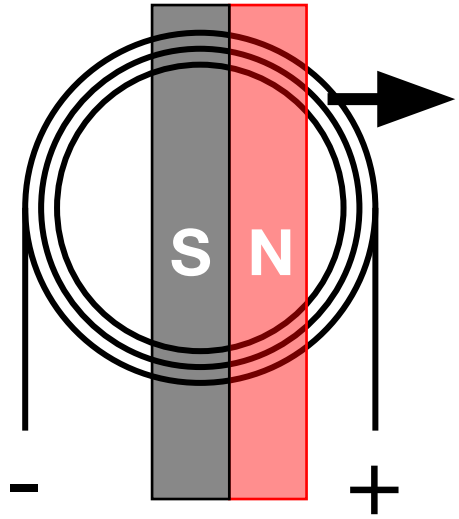
Mounting holes

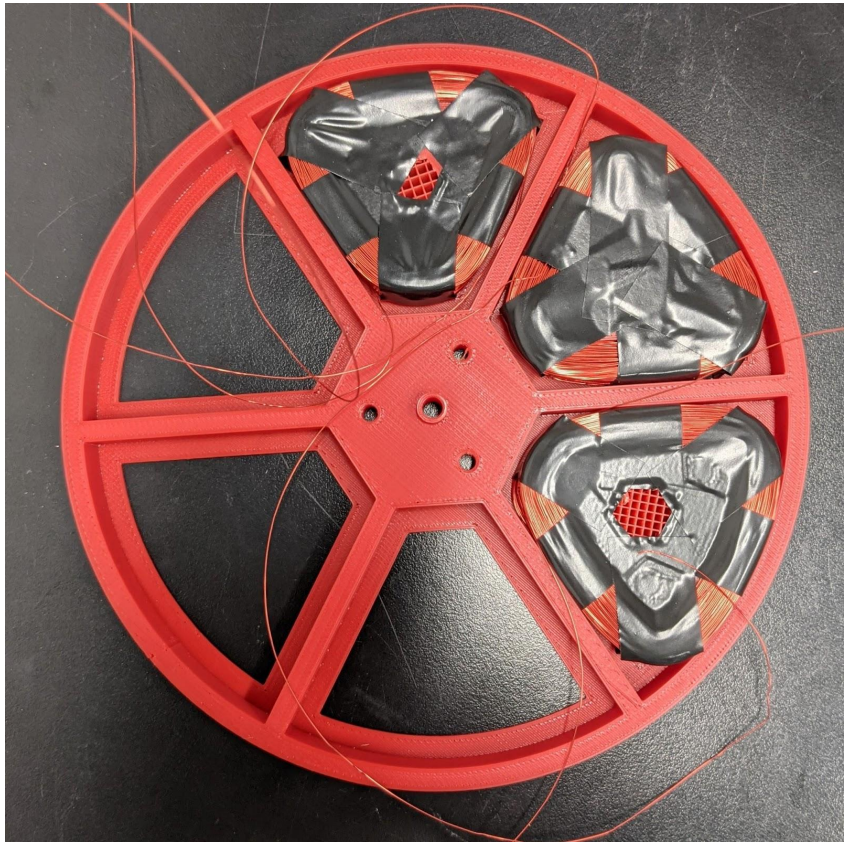
Press-fit bearing

Strong-Field Side

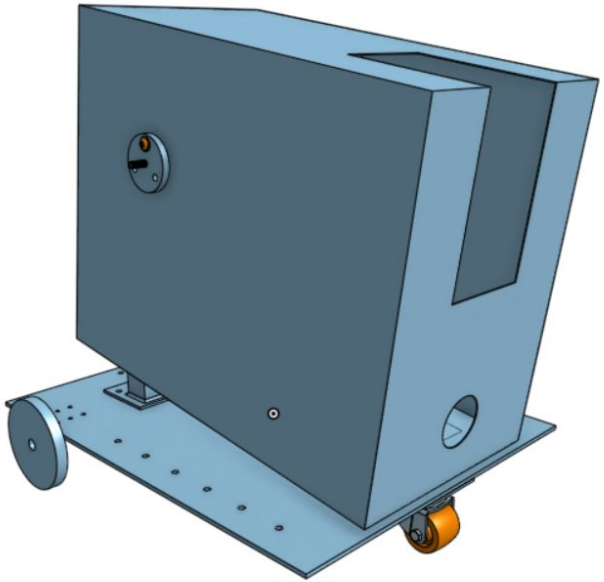
Weak-Field Side

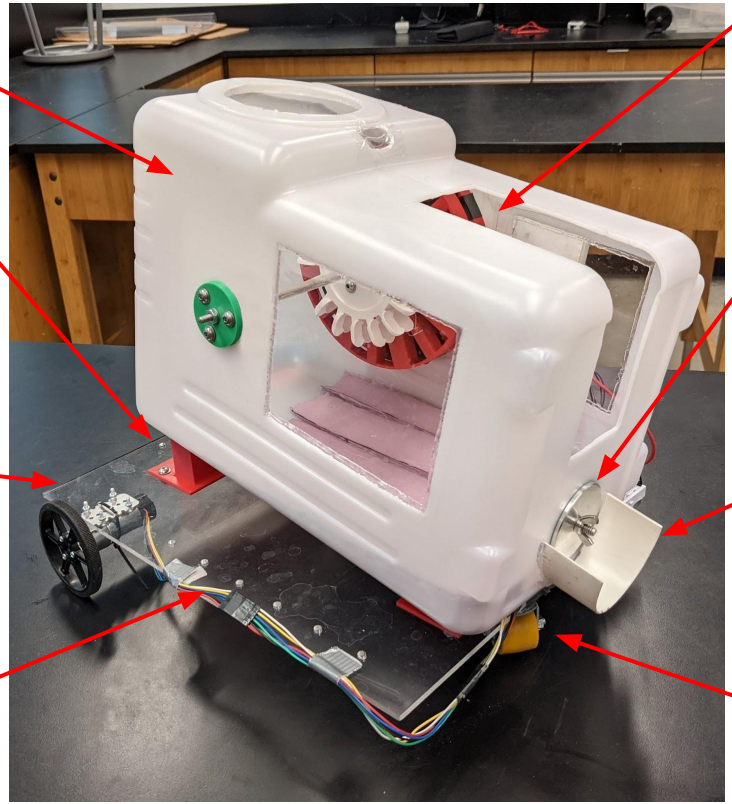
Stator



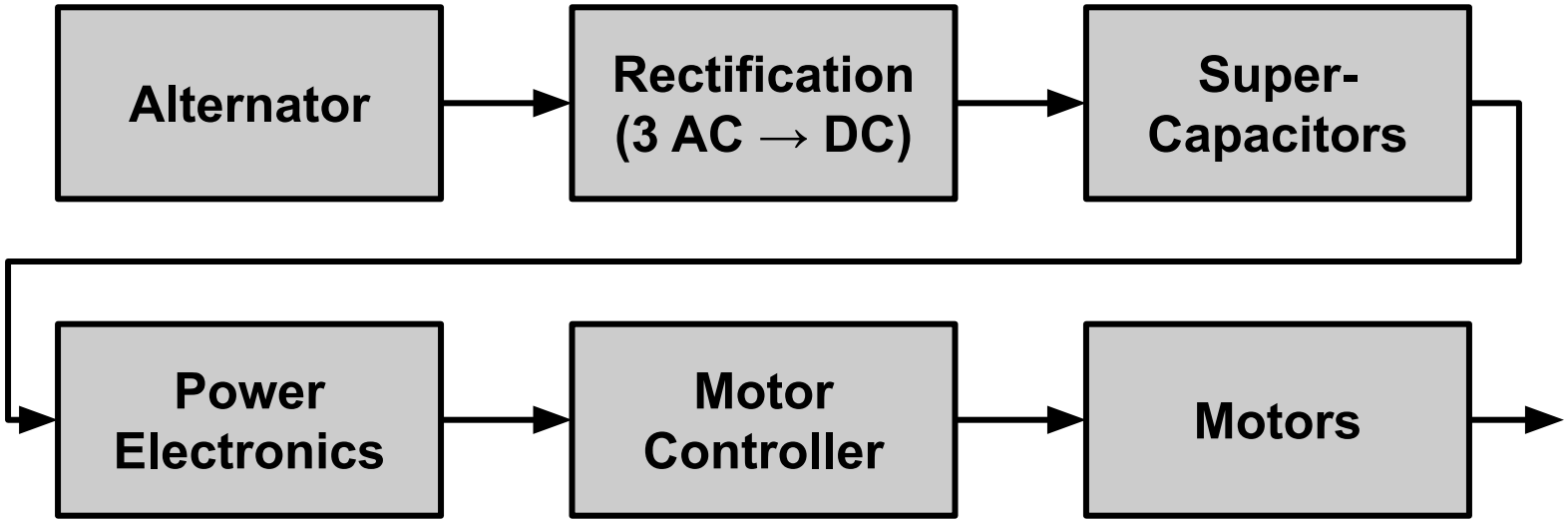


Our Competition Robot



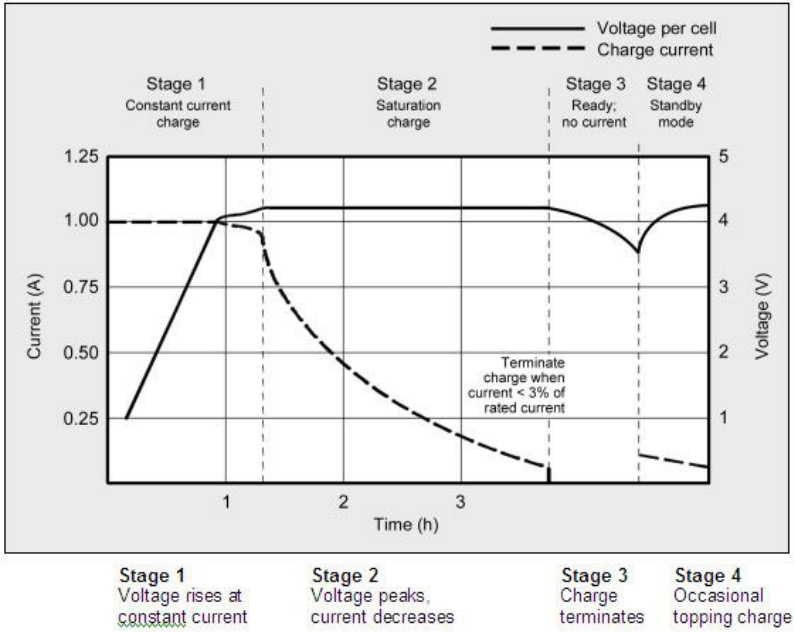


Electronics Overview



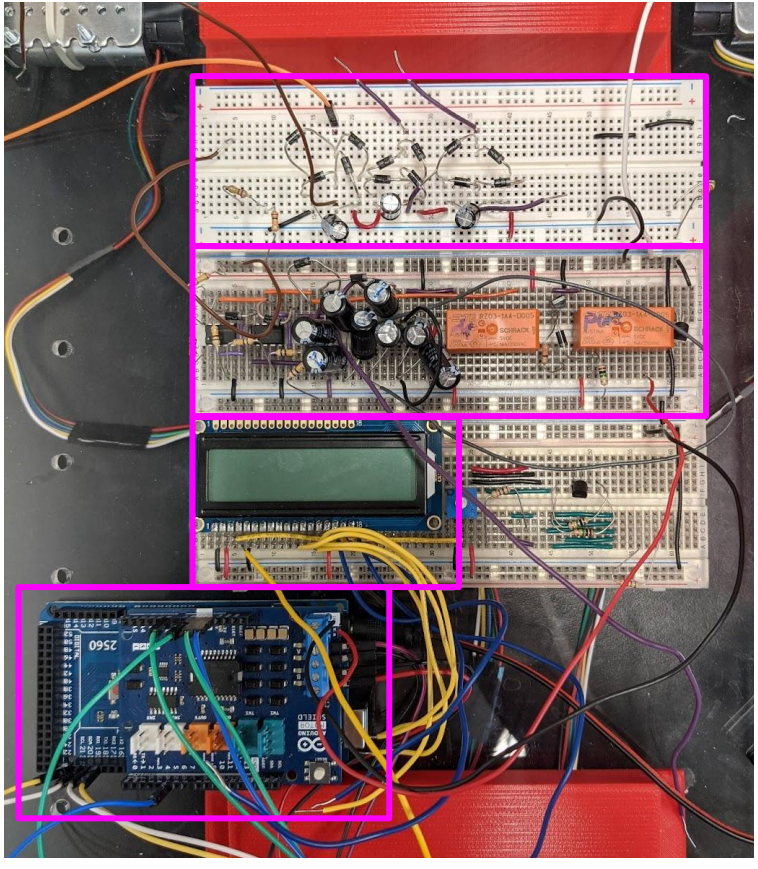
Why supercapacitors?

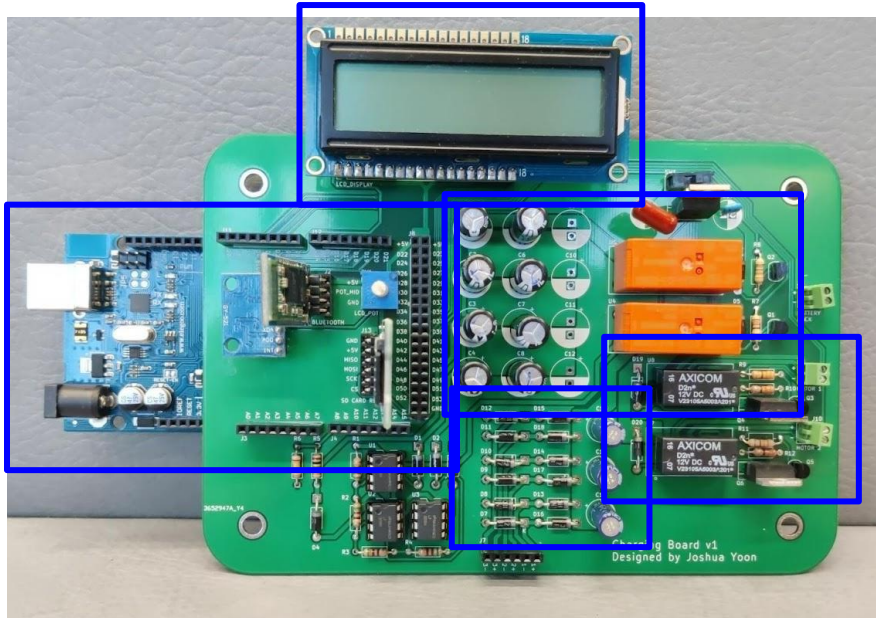
3F 3V Supercap



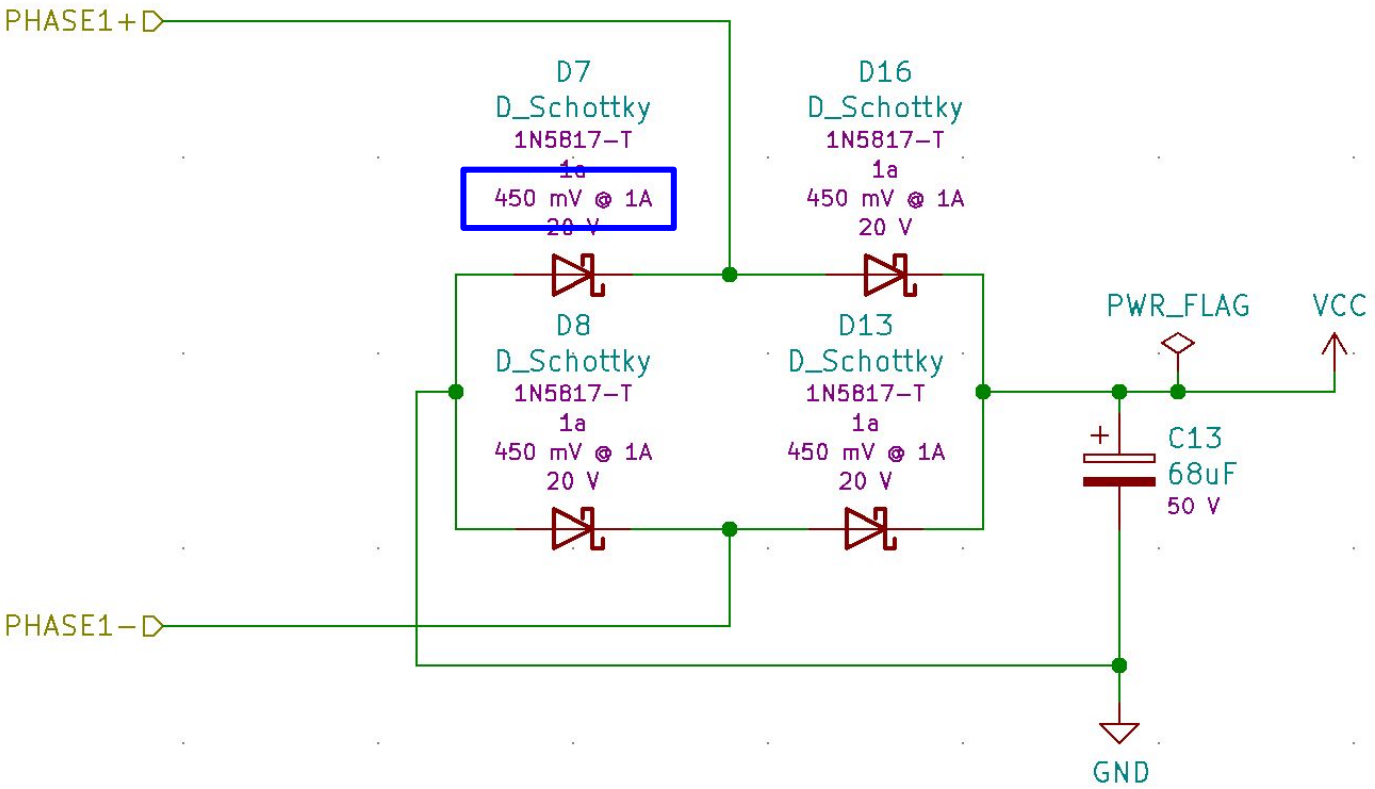
Li-Ion Charging Curve (credit: Battery University)

Electronics Setup

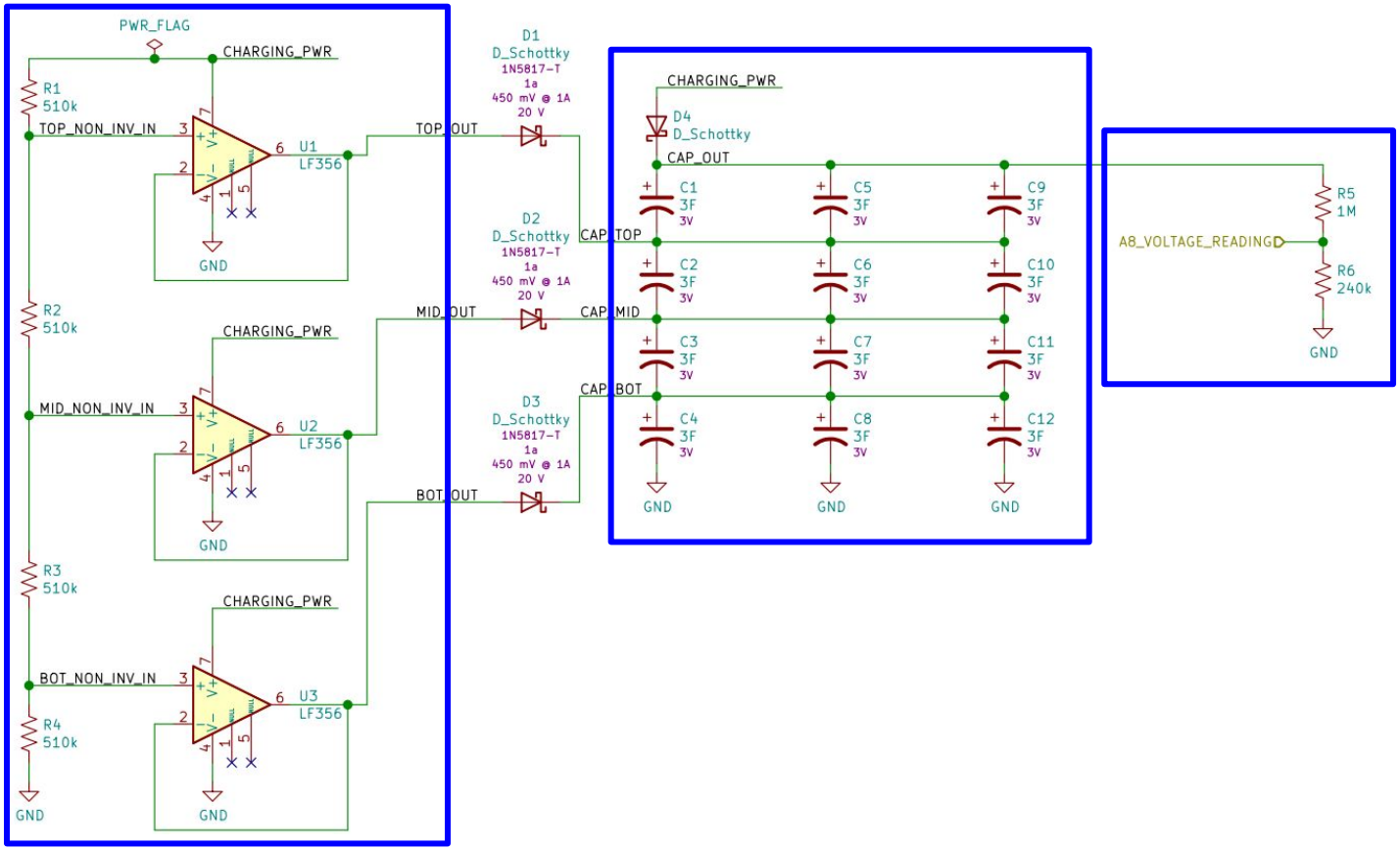




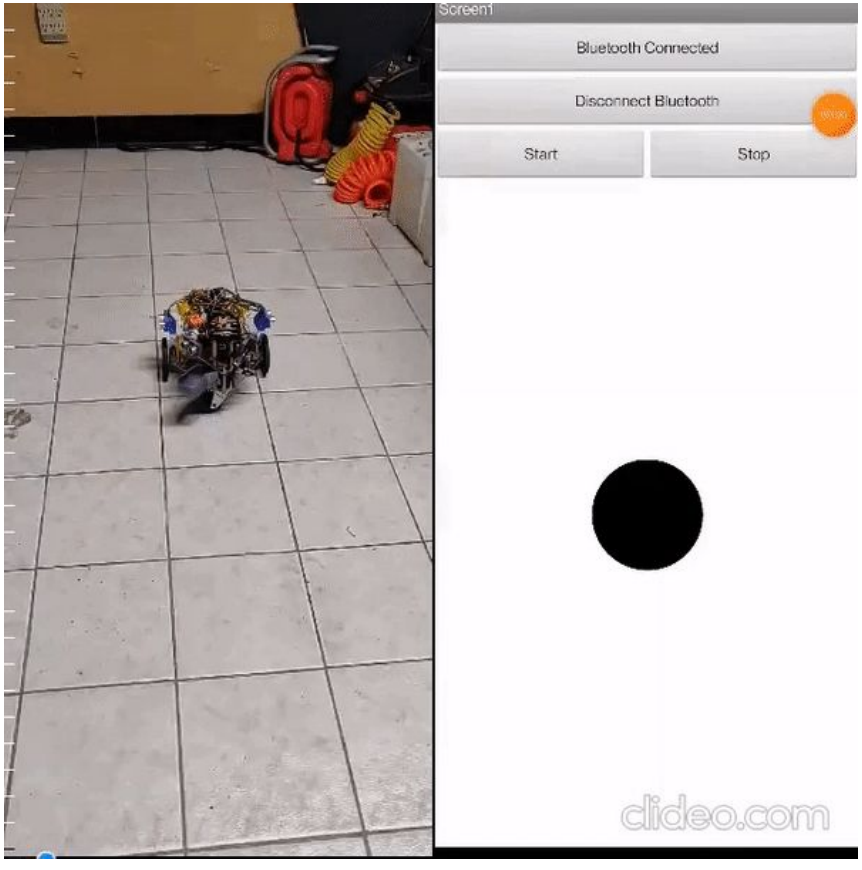
Electronics: Rectifiers



Electronics: Balancing Circuit

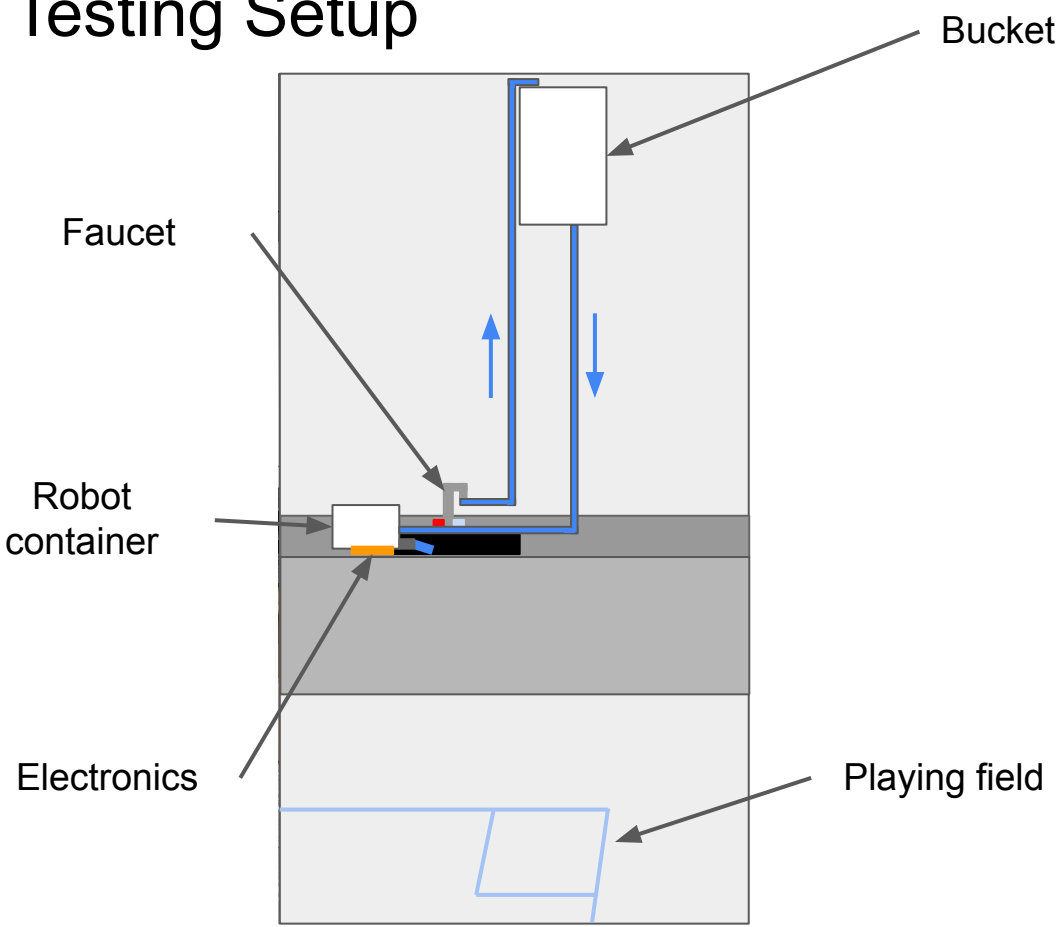


Bluetooth Control

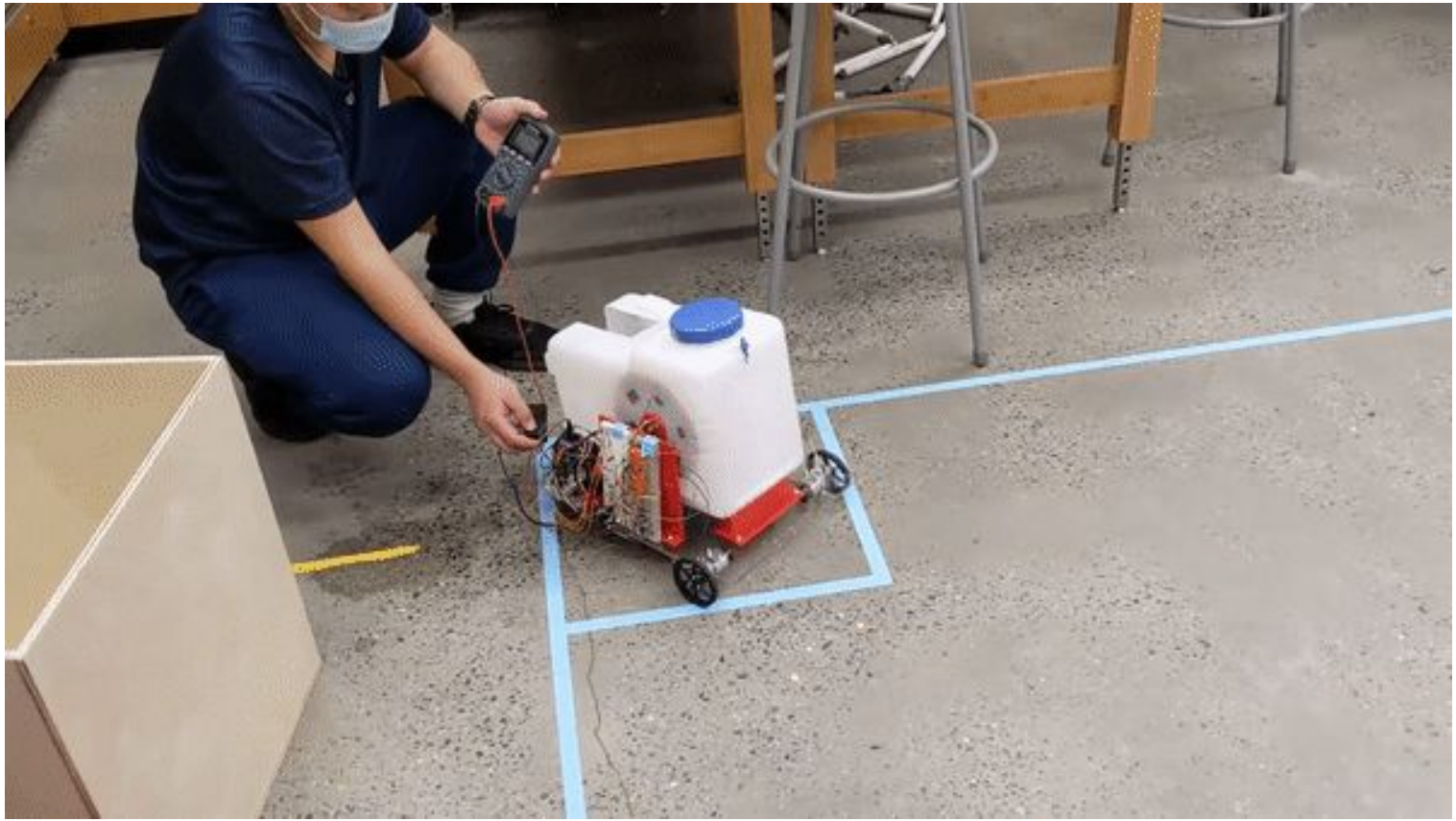


Testing, Redesign, and Competition

Competition Testing Setup

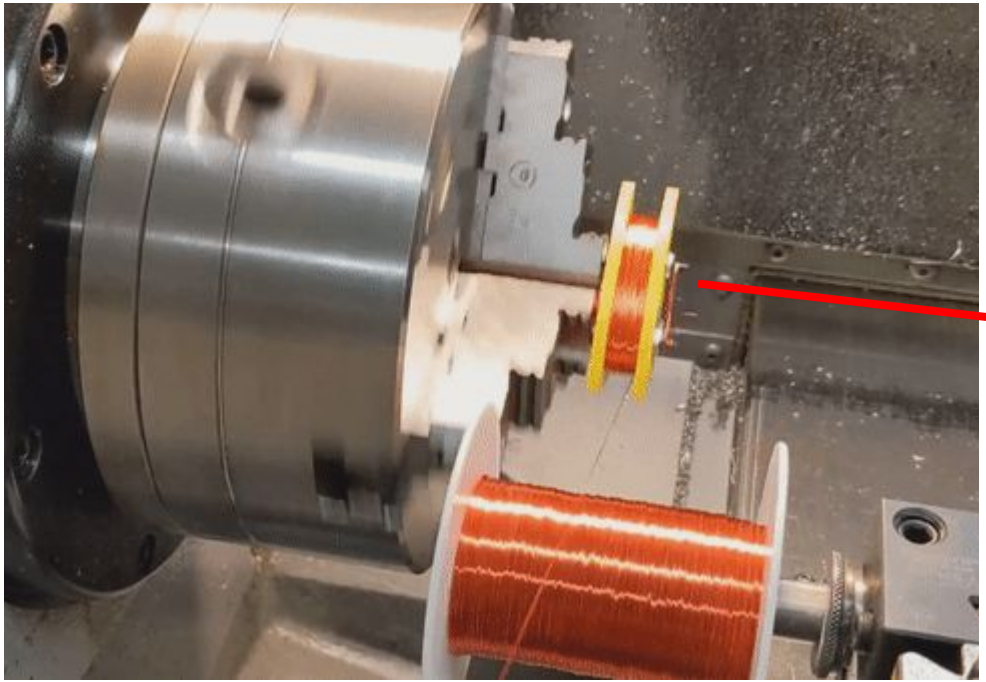


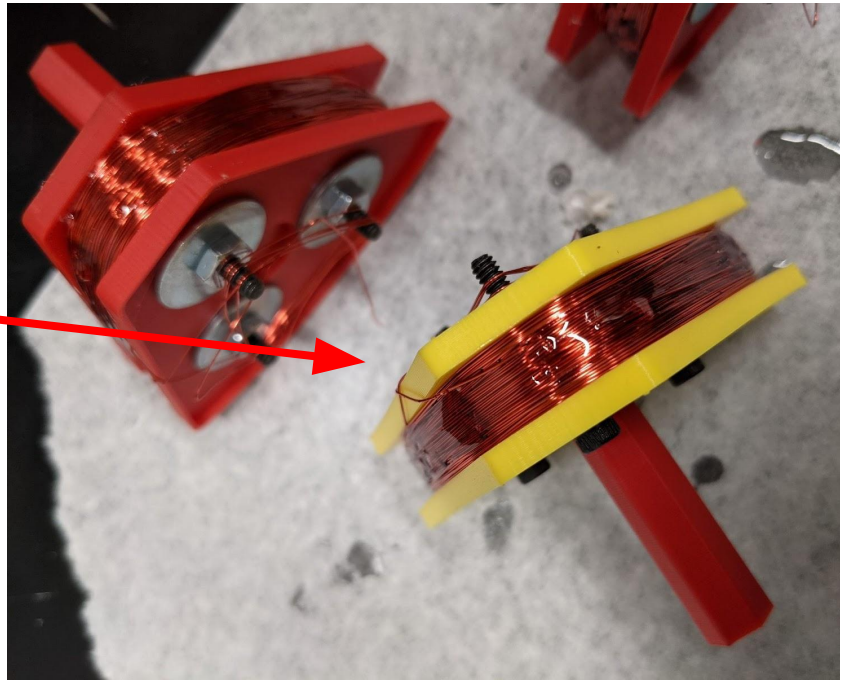
Original Test



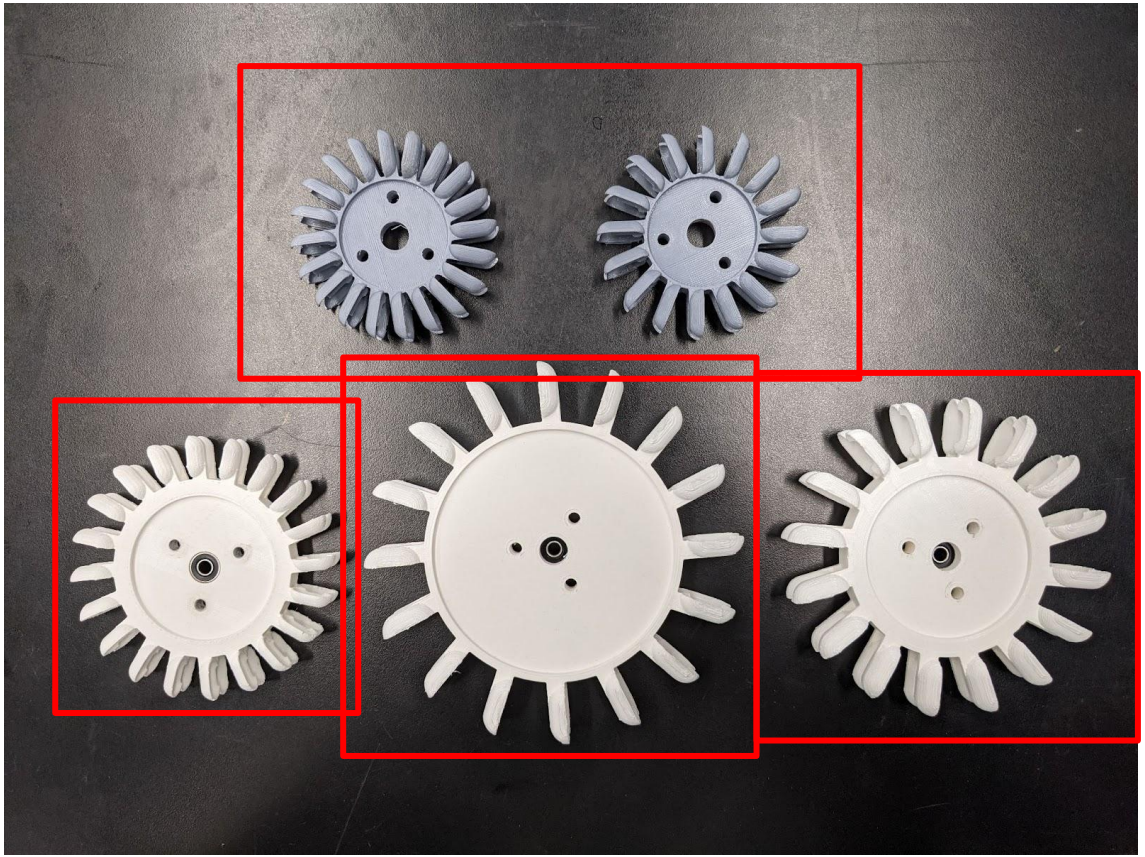
Rotor and Stator Iterations

Stator (cont.)



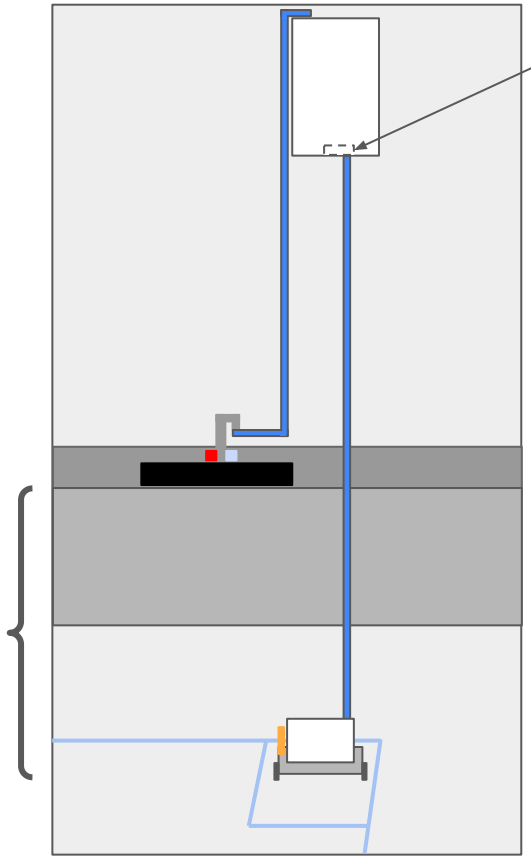


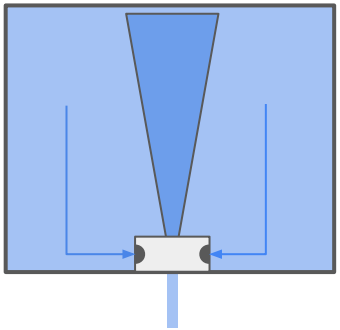
Pelton Turbine



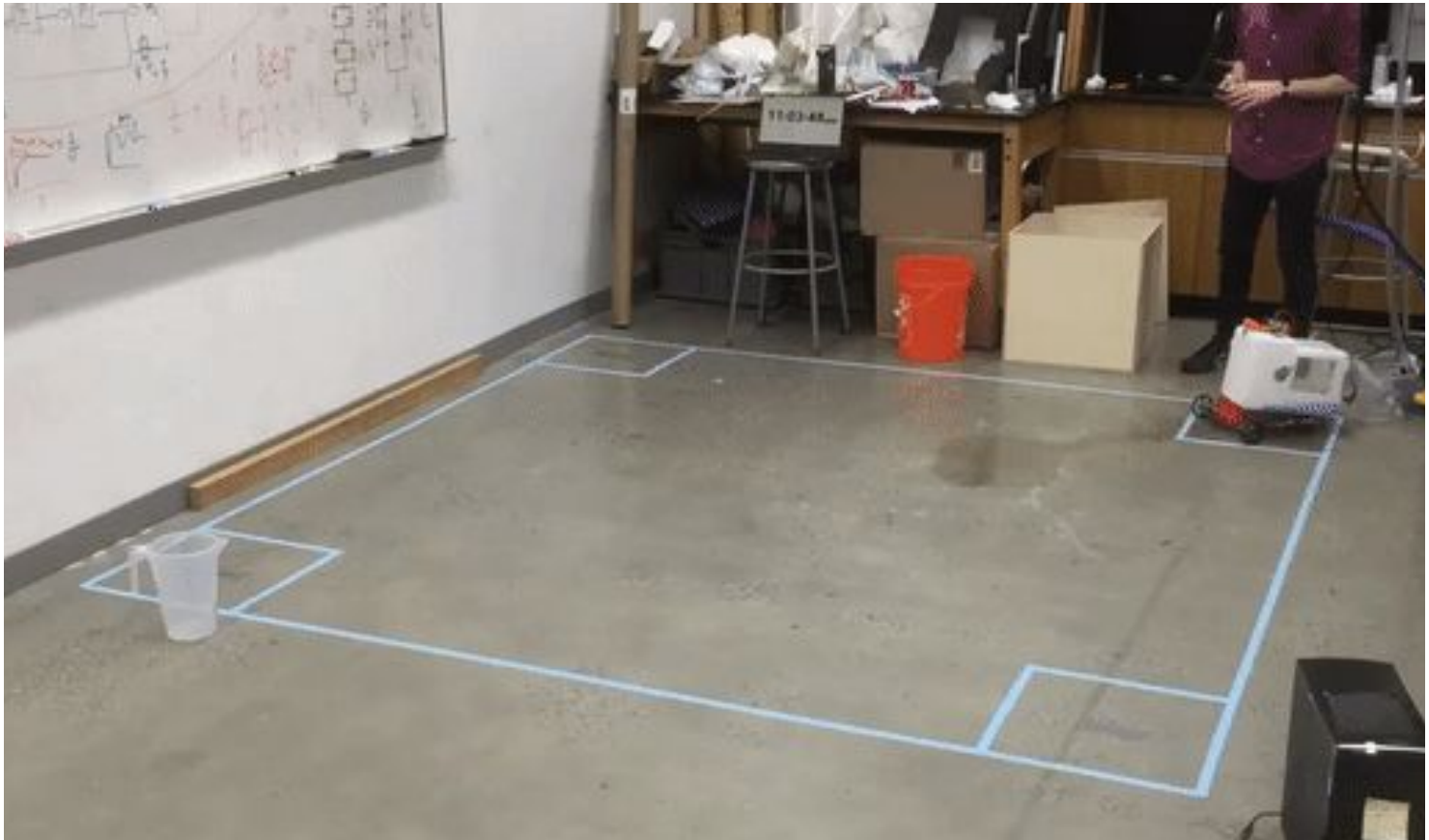
Competition Setup

Result: 48% charging time reduction





The Competition Run



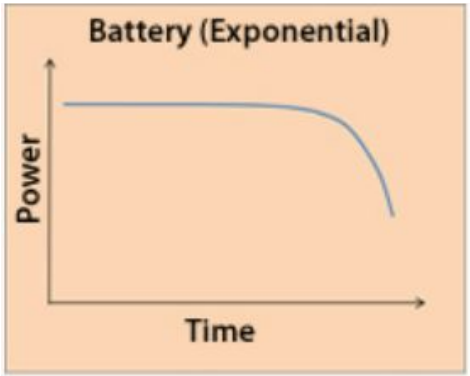
What Went Wrong?

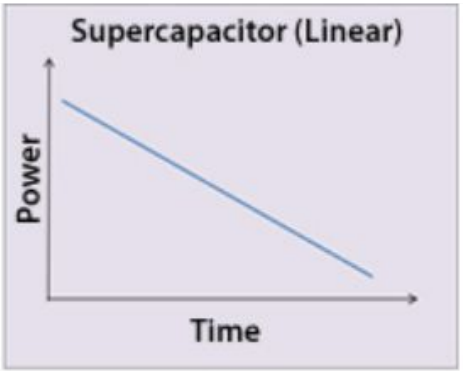
Controllability

Unfortunate Complication

Unequal Weight Distribution

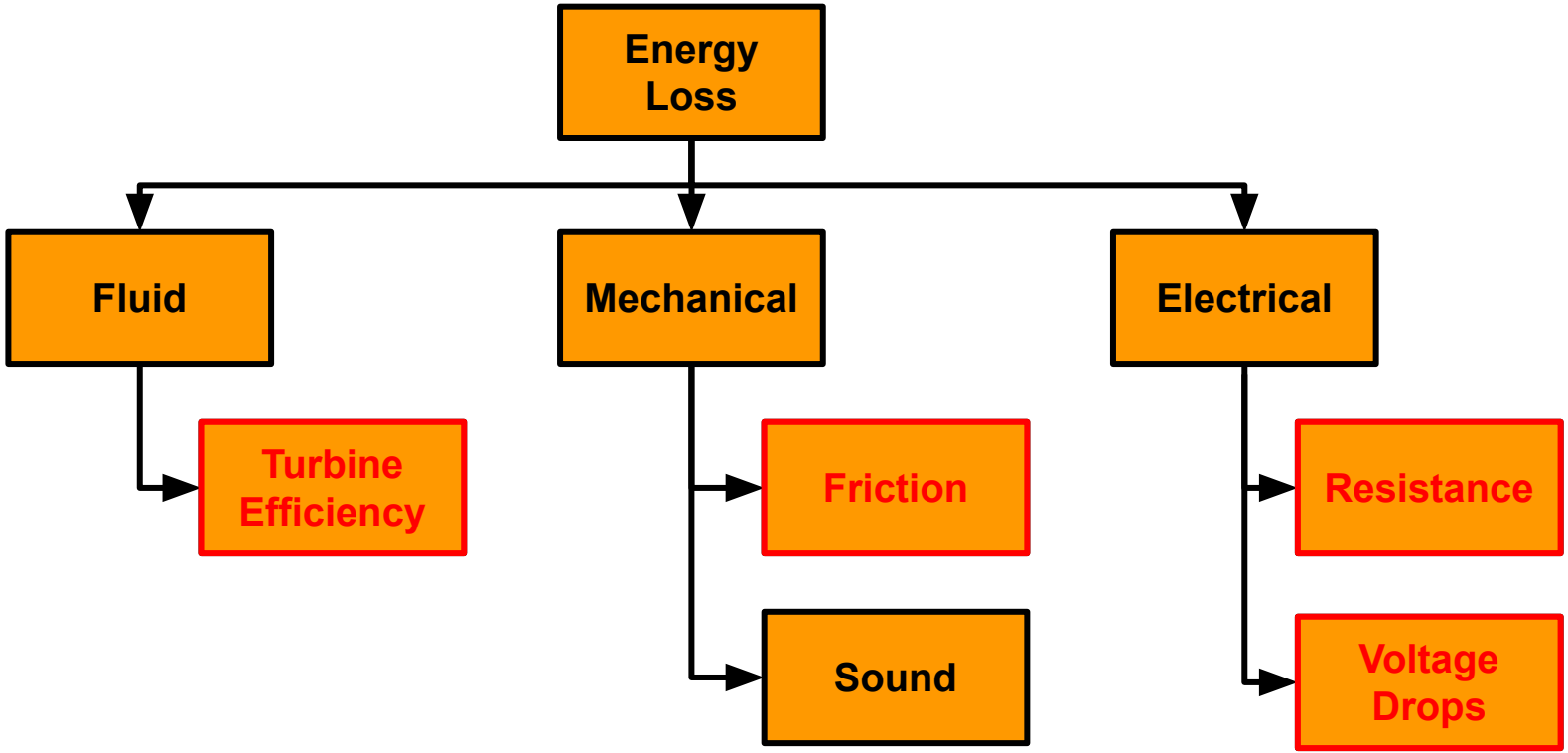
Motor Voltage Threshold



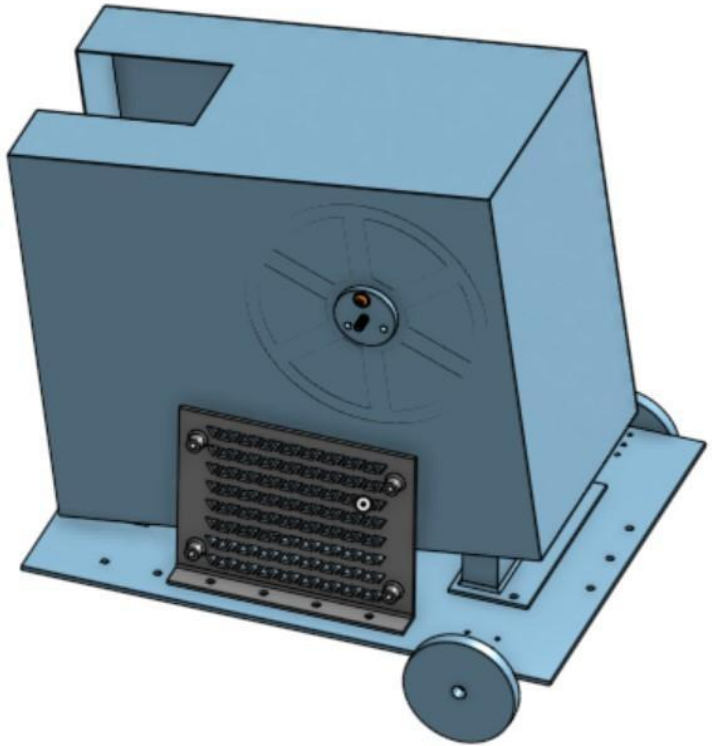


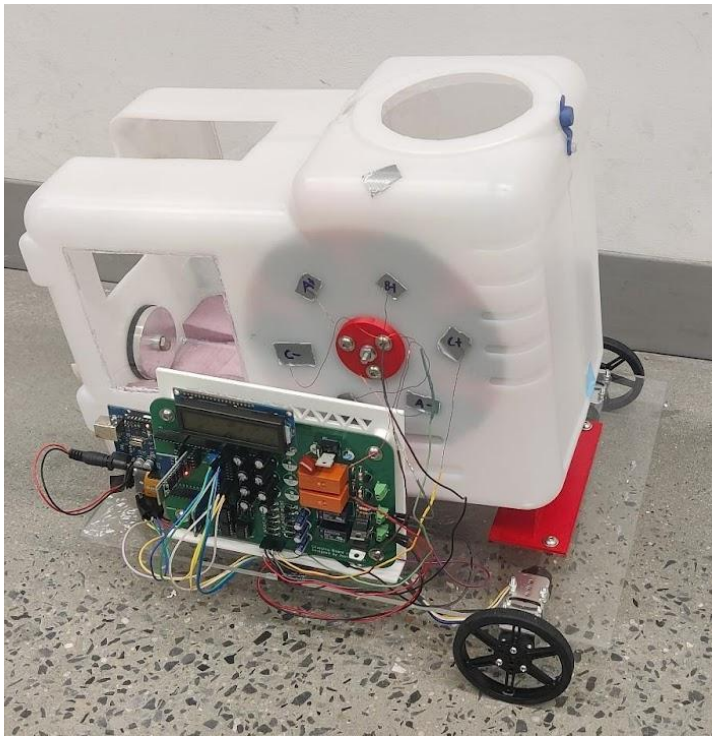
(Credit: Battery University)

Avenues of Energy Loss



Second Iteration of the Cart





Conclusion

Conclusions

Went through **multiple design iterations**

Designed **our own motor controller**

Designed **our own alternator**

Designed **our own PCB**

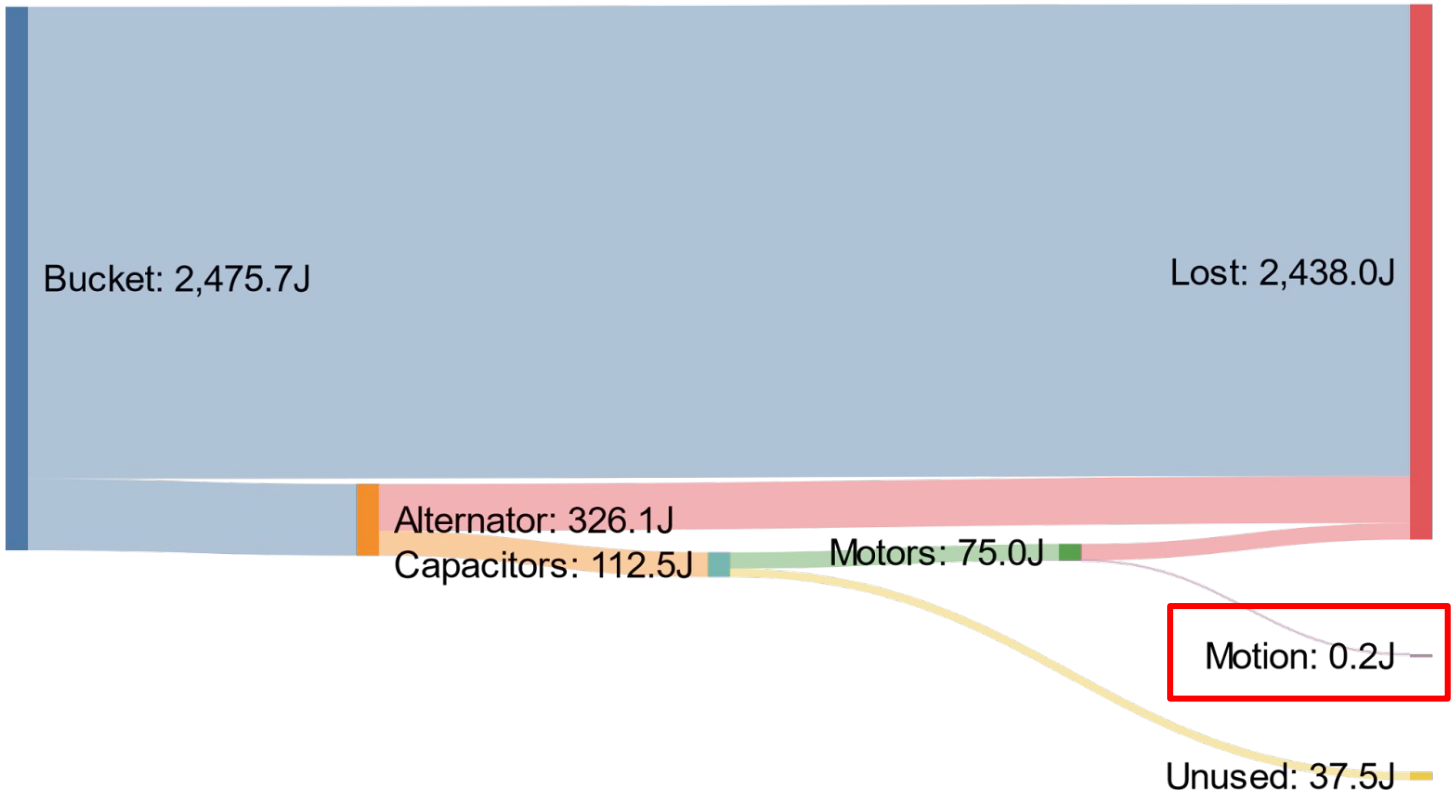
Tripled driving distance

Reduced charging time by **48%**

Tied for **2nd** best design (out of 17 teams)

Thank you for your attention

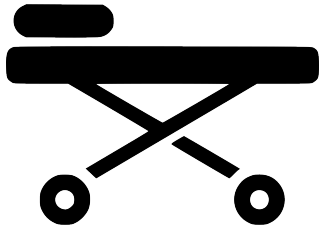
Where'd the Energy Go?



LUNCH

Improving EHR Data Imputation with Distribution-Based Methods

Dan Brody



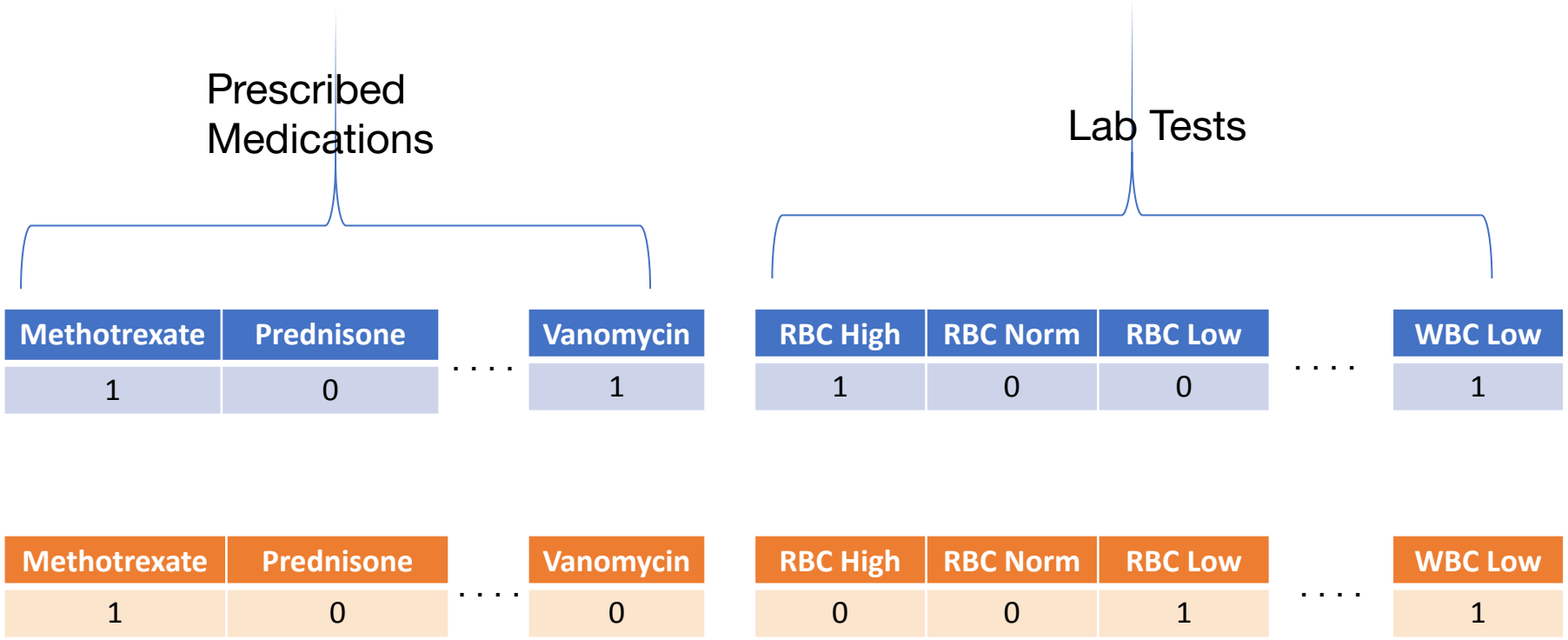
What is an EHR?

Patient_ID	Date	Prescribed Prednisone Medications	Lab Test RBCs Results	Reported Back Pain Symptoms	…
1	2/3/2019	1	50	1	…
⋮	⋮	⋮	⋮	⋮	

Problem Statement

EHR data is challenging to represent and model due to its high dimensionality, noise, heterogeneity, sparseness, incompleteness, random errors, and systematic biases [1,2,3]. This property of EHR data makes it difficult to use in any deep or machine learning algorithms including algorithms in precision medicine such as endpoint prediction.

What is Endpoint Prediction?



Missingness Types

Probability of missing

Prednisone	RBCs
1	1
1	1

- MCAR - unrelated to [illegible] les and unrelated to the variable itself
- MNAR – dependent [illegible] ed variables of a variable Y
- MAR – dependent o[illegible] observed

Methods:

Median Imputation, GAIN

Median Imputation

Median (

Prednisone
1
0
0
0
1
0
0
0

) = 0.5

Fault :
accuracy depletes as the number of samples in a feature decrease.

GAIN Description [4]

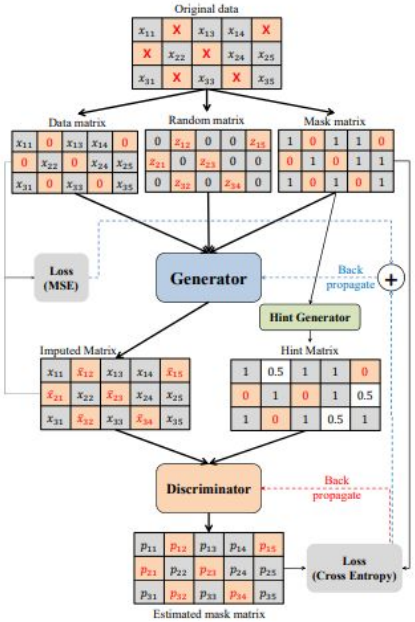


Figure 1. The architecture of GAIN

GAIN Performance [4]

Table 2. Imputation performance in terms of RMSE (Average ± Std of RMSE)

Algorithm	Breast	Spam	Letter	Credit	News
GAIN	**.0546 ± .0006**	**.0513± .0016**	**.1198± .0005**	**.1858 ± .0010**	**.1441 ± .0007**
MICE	.0646 ± .0028	.0699 ± .0010	.1537 ± .0006	.2585 ± .0011	.1763 ± .0007
MissForest	.0608 ± .0013	.0553 ± .0013	.1605 ± .0004	.1976 ± .0015	.1623 ± 0.012
Matrix	.0946 ± .0020	.0542 ± .0006	.1442 ± .0006	.2602 ± .0073	.2282 ± .0005
Auto-encoder	.0697 ± .0018	.0670 ± .0030	.1351 ± .0009	.2388 ± .0005	.1667 ± .0014
EM	.0634 ± .0021	.0712 ± .0012	.1563 ± .0012	.2604 ± .0015	.1912 ± .0011

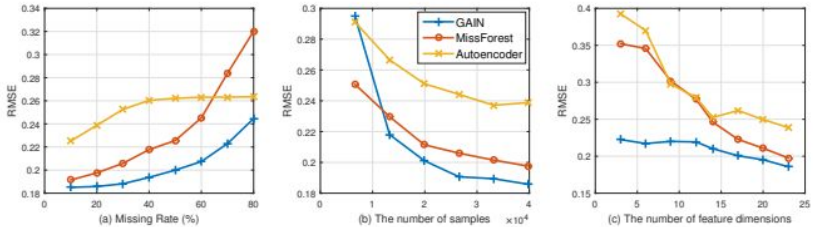


Figure 2. RMSE performance in different settings: (a) Various missing rates, (b) Various number of samples, (c) Various feature dimensions

Table 3. Prediction performance comparison

Algorithm	AUROC (Average ± Std)			
	Breast	Spam	Credit	News
GAIN	**.9930 ± .0073**	**.9529 ± .0023**	**.7527 ± .0031**	**.9711 ± .0027**
MICE	.9914 ± .0034	.9495 ± .0031	.7427 ± .0026	.9451 ± .0037
MissForest	.9860 ± .0112	.9520 ± .0061	.7498 ± .0047	.9597 ± .0043
Matrix	.9897 ± .0042	.8639 ± .0055	.7059 ± .0150	.8578 ± .0125
Auto-encoder	.9916 ± .0059	.9403 ± .0051	.7485 ± .0031	.9321 ± .0058
EM	.9899 ± .0147	.9217 ± .0093	.7390 ± .0079	.8987 ± .0157

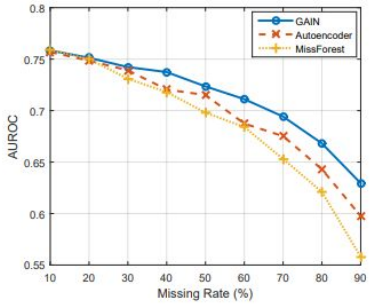


Figure 3. The AUROC performance with various missing rates with Credit dataset

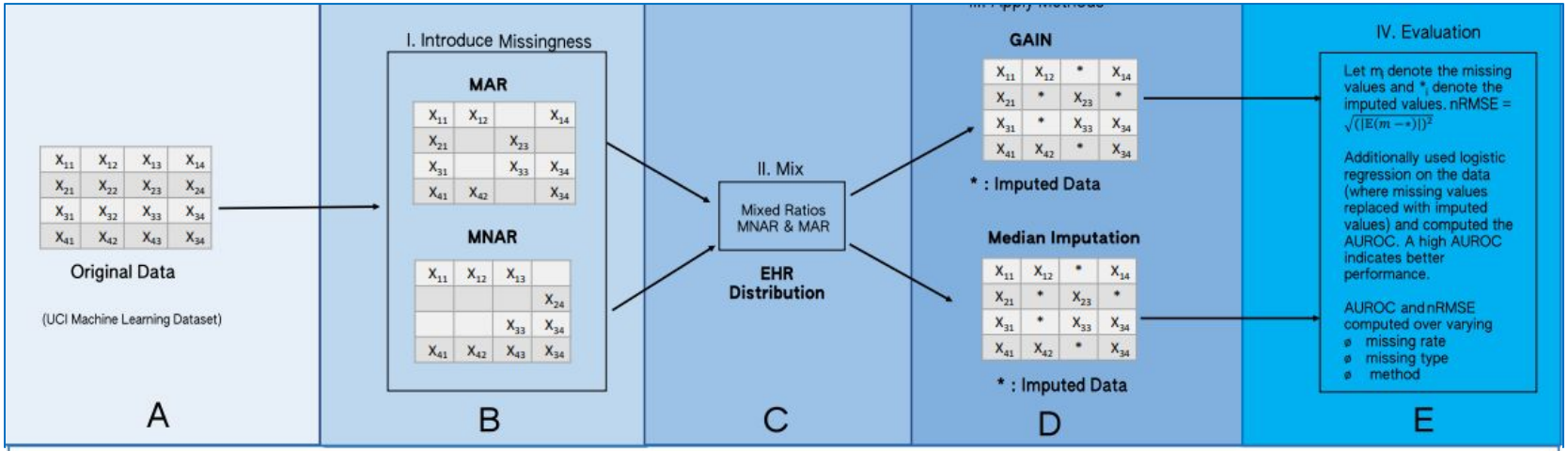
Table 4. Congeniality of imputation models

Algorithm	Mean Bias ($\|\|\mathbf{w} - \hat{\mathbf{w}}\|\|_1$)	MSE ($\|\|\mathbf{w} - \hat{\mathbf{w}}\|\|_2$)
GAIN	**0.3163± 0.0887**	**0.5078± 0.1137**
MICE	0.8315 ± 0.2293	0.9467 ± 0.2083
MissForest	0.6730 ± 0.1937	0.7081 ± 0.1625
Matrix	1.5321 ± 0.0017	1.6660 ± 0.0015
Auto-encoder	0.3500 ± 0.1503	0.5608 ±0.1697
EM	0.8418 ± 0.2675	0.9369 ± 0.2296

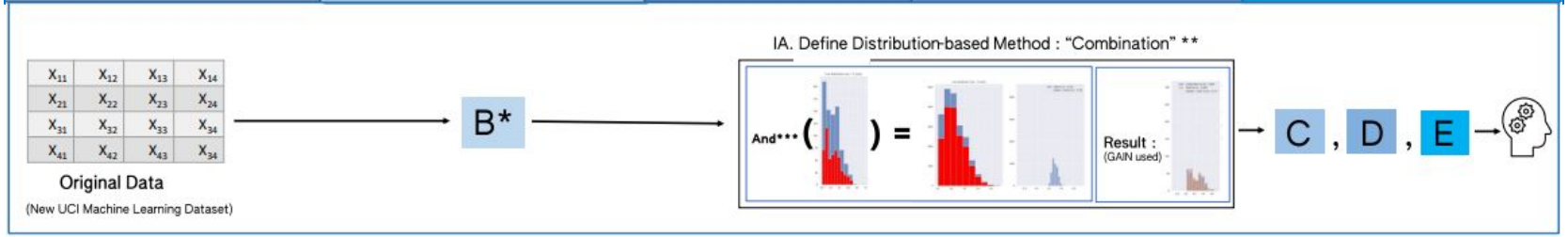
relationship between feature and label. As can be seen in the table, GAIN achieves significantly lower mean bias and mean square error than other state-of-the-art imputation algorithms (from 8.9% to 79.2% performance improvements).

Methodology

1



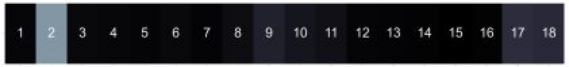
2

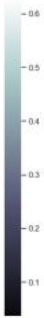


Results

Note : Missing rate of 0.8 and 50% MNAR and MCAR was assessed. Refer to the paper at https://github.com/dbrodyd12 for more examples and information

Annotation : The following graph highlights feature-wise nRMSE as well as AUROC for the respective method and dataset where each block of the heatmaps is feature "x" of the dataset

Method	Messidor Heatmap	Messidor AUROC	Credit Card Heatmap	Credit Card AUROC
Combination		0.694		0.603
GAIN		0.695		0.580
Median		0.687		0.69



Conclusions : The "combination" method is consistent across datasets since never gets outperformed by both methods

Future Work

- Using more methods
- Distribution Search API for UCI Machine Learning
- Applying results to EHR data (MIMIC-III) to improve endpoint prediction of myocardial infarction in []

Works Cited

[1] Giraud T, Dhainaut JF, Vaxelaire JF, Joseph T, Journois D, Bleichner G, Sollet JP, Chevret S, Monsallier JF. Iatrogenic complications in adult intensive care units: a prospective two-center study. *Crit Care Med.* 1993;21:40–51. doi: 10.1097/00003246-199301000-00011.

[2]0 Bracco D, Favre JB, Bissonnette B, Wasserfallen JB, Revelly JP, Ravussin P, Chiolero R. Human errors in a multidisciplinary intensive care unit: a 1-year prospective study. *Intensive Care Med.* 2001;27:137–145. doi: 10.1007/s001340000751.

[3] Abramson NS, Wald KS, Grenvik AN, Robinson D, Snyder JV. Adverse occurrences in intensive care units. *JAMA.* 1980;244:1582–1584. doi: 10.1001/jama.1980.03310140040027.

[4]Jinsung Yoon and James Jordon and Mihaela van der Schaar. GAIN: Missing Data Imputation using Generative Adversarial Nets. 2018. GAIN: [1806.02920] GAIN: Missing Data Imputation using Generative Adversarial Nets (arxiv.org)

Applications of Ultra-wideband RTLS in Architecture (AURA)

Theo Song, Dan Park

Applications of Ultra-wideband RTLS in Architecture (AURA)

Donghyun Park
Theo Song

Problems

1. Inefficient architectural design and no systems to analyze it
    a. How is space being used? Bottlenecks?
    b. Maximize human pathing behavior
2. Lack of non-commercial implementation of RTLS at sub-meter accuracy
    a. BLE is well established but inaccurate
3. Ethical dilemmas behind human data collection

Goals: Engineering

- Hardware: Implement small-scale RTLS network system using barebones UWB hardware
- Software: Integrate a web/mobile based application to store RTLS telemetry data, as well as collect user information

Goals: Interdisciplinary

- Collaborate with Architecture School to create visualizations of the aggregated data
- Perform analysis of how space is utilized, how certain groups have similar behaviors, observe any unique patterns in movement etc.

Goals: Social Science

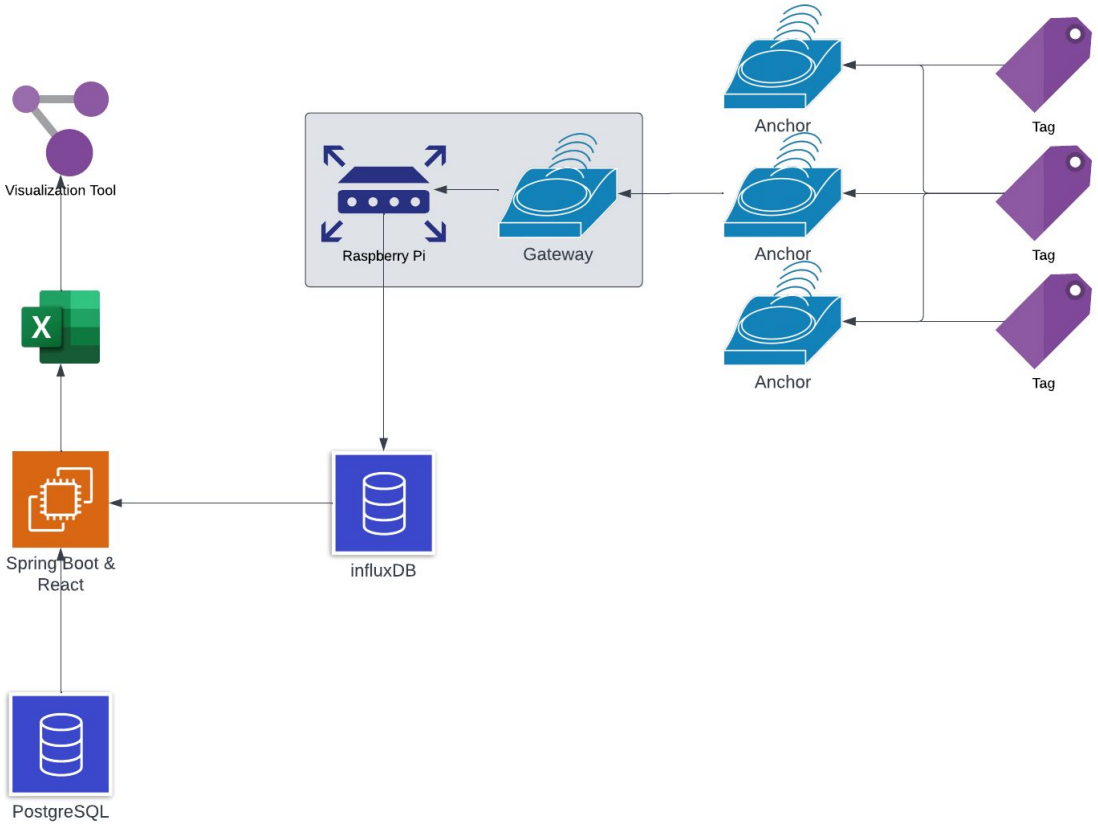
- Discussion of the ethical and moral principles behind tracking technologies applied to humans
- Implicit vs Explicit consent
- Illusion of choice
- Applications of this RTLS systems (beneficial vs detrimental)

Ultra-wideband

+ Great accuracy (10 ~ 20 cm range)
+ Wide Range of 150m
+ Higher data transmission bandwidth
+ Newer generation of localization technology

- Expensive, commercial solutions outside of budget
- Uncommon off-the-shelf parts and commercial implementation

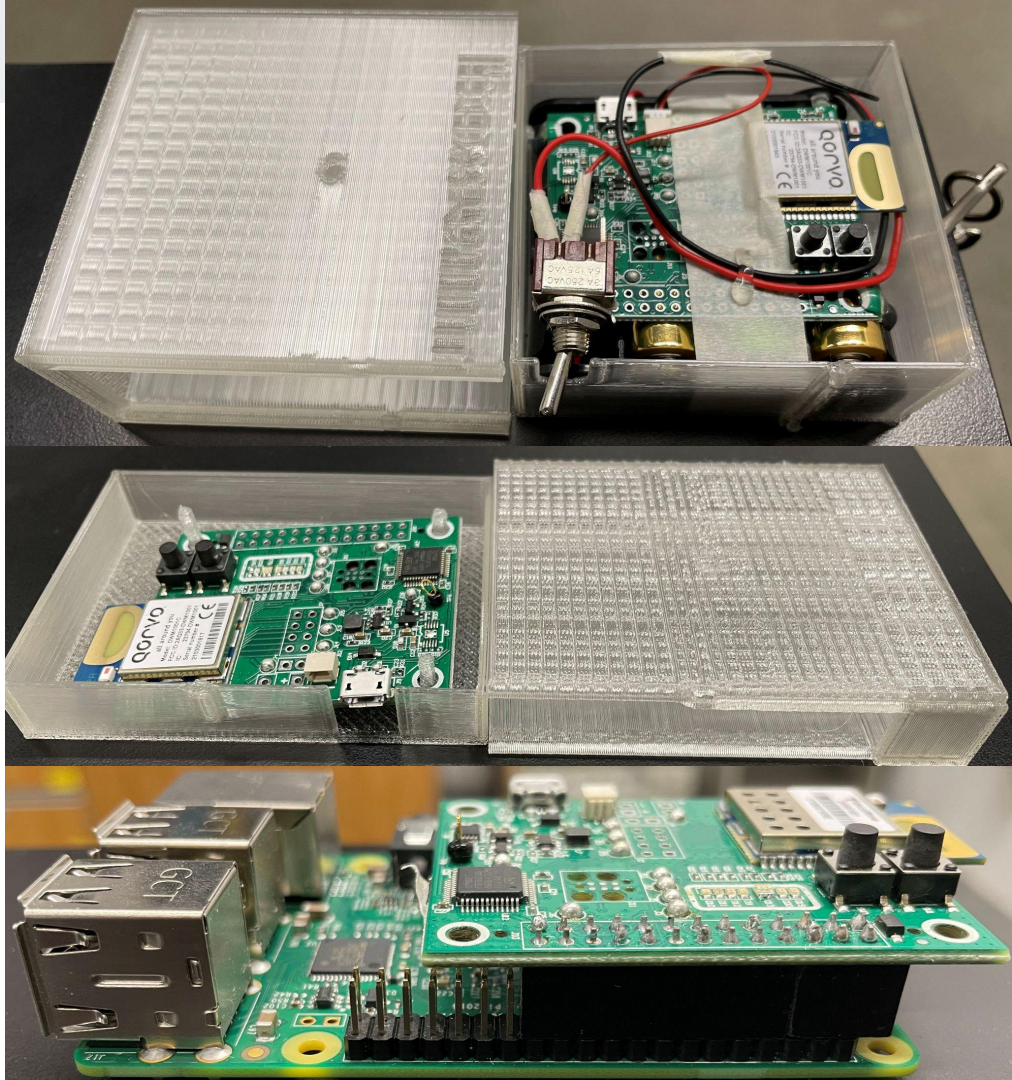
System Stack



Hardware

DWM1001 Development Board

- Emulates both
  - Transmitter/Tag
  - Receiver/Anchor
- One used as gateway to Raspberry Pi using MQTT protocol (used for IoT)



Software: Gateway & Raspberry Pi

- Gateway Built with combination of DWM1001 Board and Raspberry Pi
- Gateway delivers received position information to the MQTT server
- Raspberry Pi using miniconda for Python package management
- Raspberry Pi reads stream of data from MQTT server and stores into influxDB

Software: Database

- InfluxDB: time-series database, capable of high flux of insertion and query of time-sequential data
- PostgreSQL: traditional relational DBMS used to store tag and user data
- Both running in AWS

Software: Web Application

- Backend: Built with Java, Spring Boot, to support following APIs
  - Assigning unoccupied tags to incoming users
  - Detaching tags for leaving users
  - Export Position Data in a form of Excel file
- Frontend: Built with React for easy access of functionalities

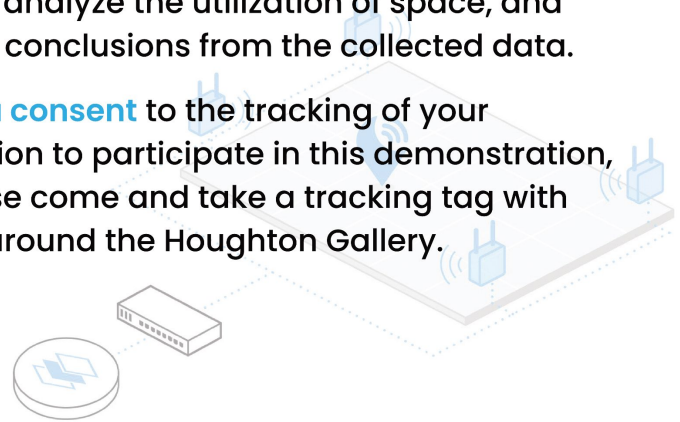
Demo Run at Houghton Gallery

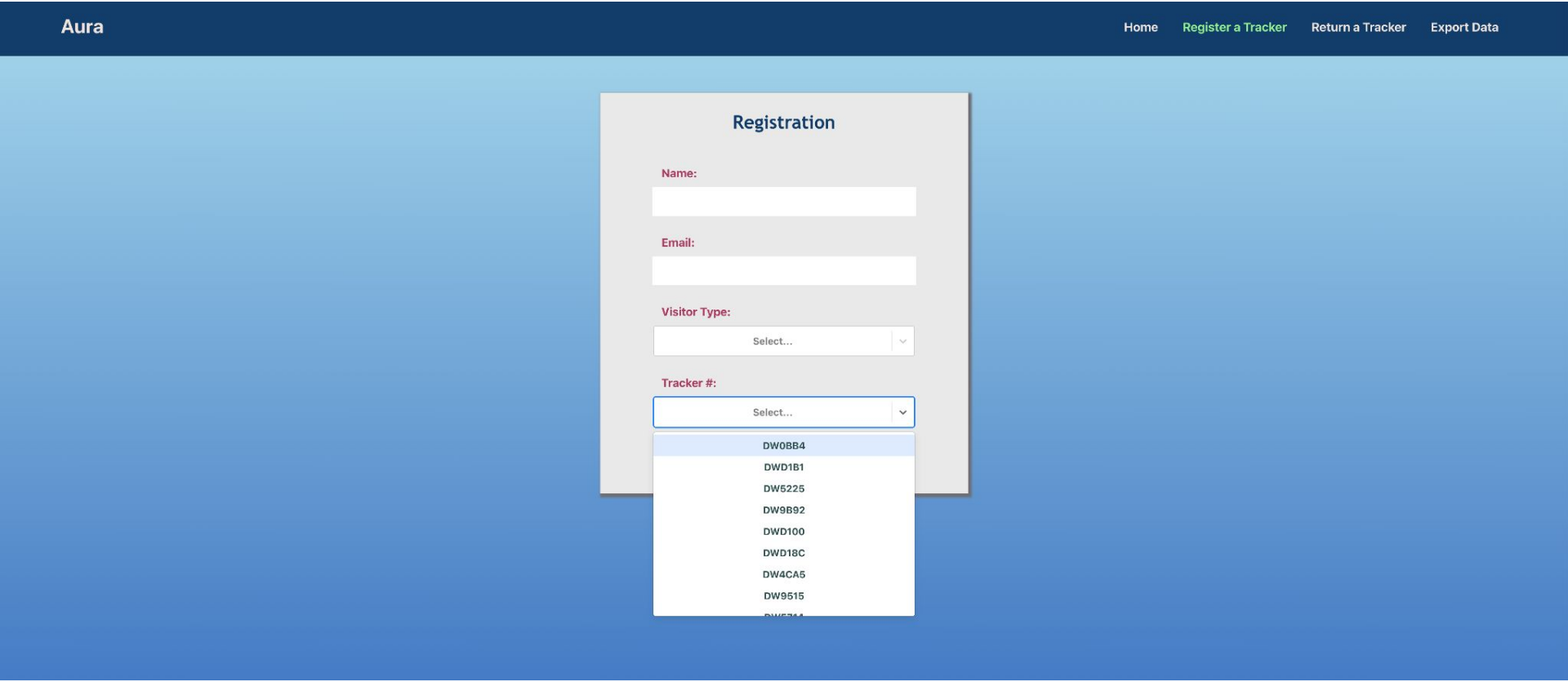
AURA

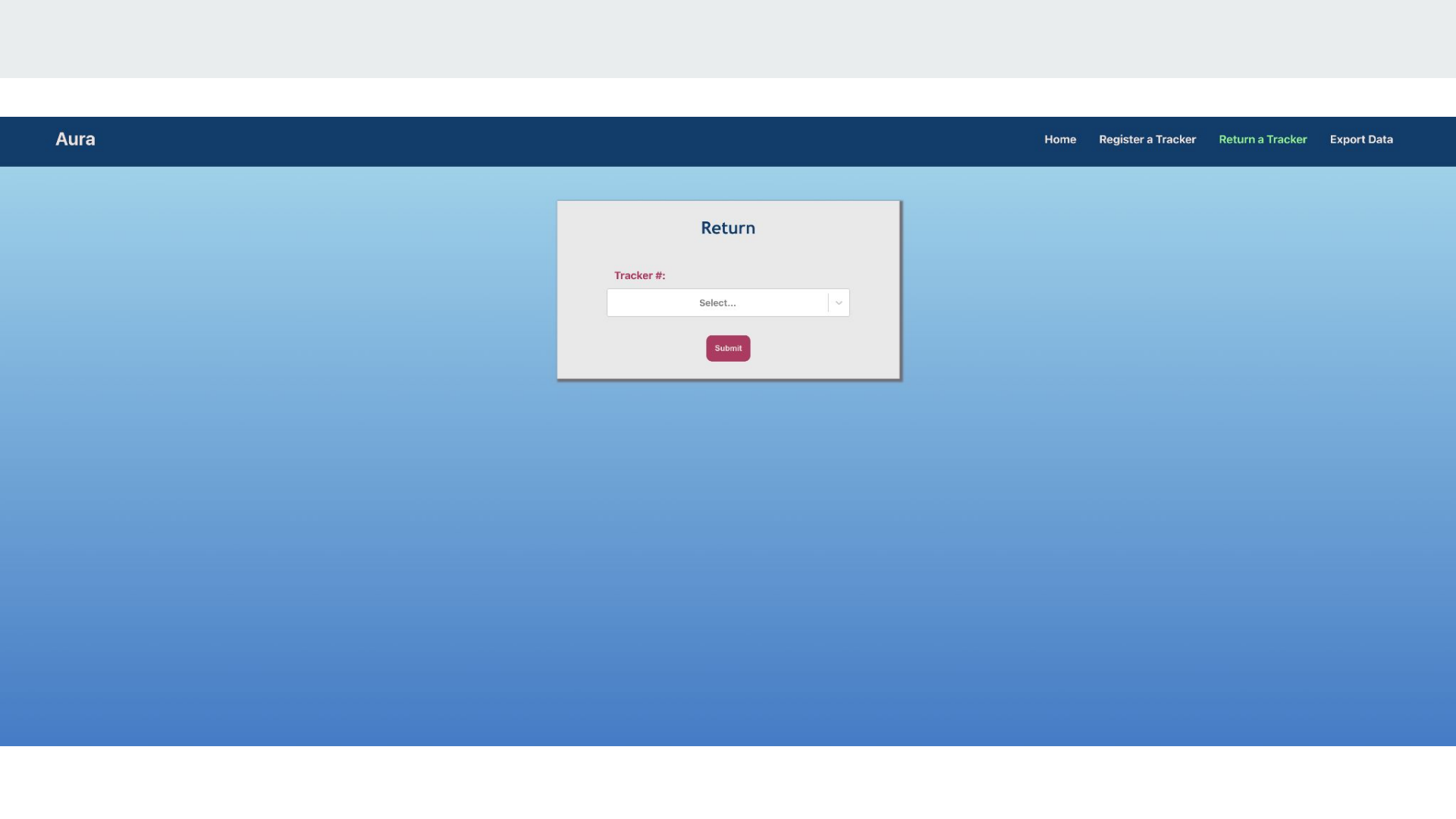
Applications of Ultra-Wideband RTLS in Architecture

AURA explores the generative nature of human pathing behavior, and visualizes two key aspects of traffic: **flow and density**.
We explore these two aspects in relation to time, analyze the utilization of space, and draw conclusions from the collected data.

If you consent to the tracking of your location to participate in this demonstration, please come and take a tracking tag with you around the Houghton Gallery.

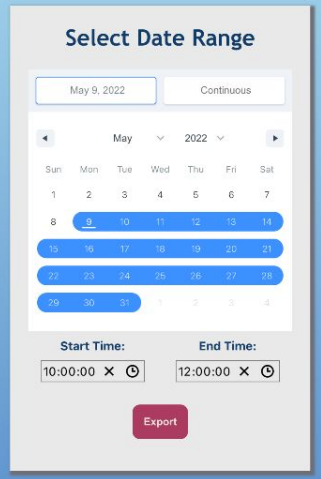


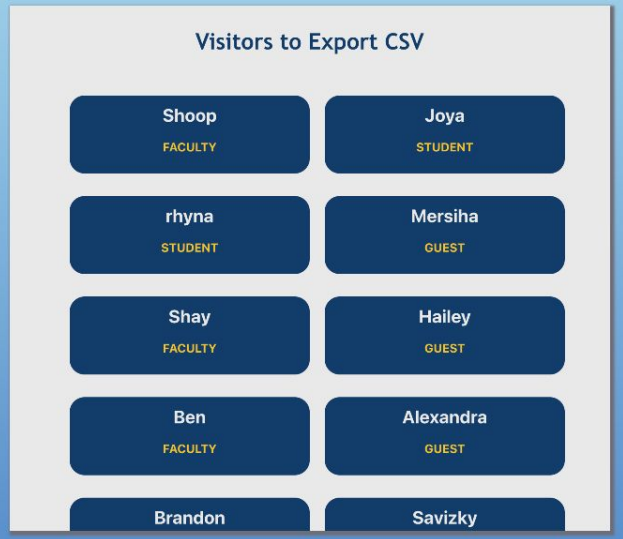

Aura
Home
Register a Tracker
Return a Tracker
Export Data
Registration
Name:
Email:
Visitor Type:
Select...
Tracker #:
Select...
DW0BB4
DWD1B1
DW5225
DW9B92
DWD100
DWD18C
DW4CA5
DW9515

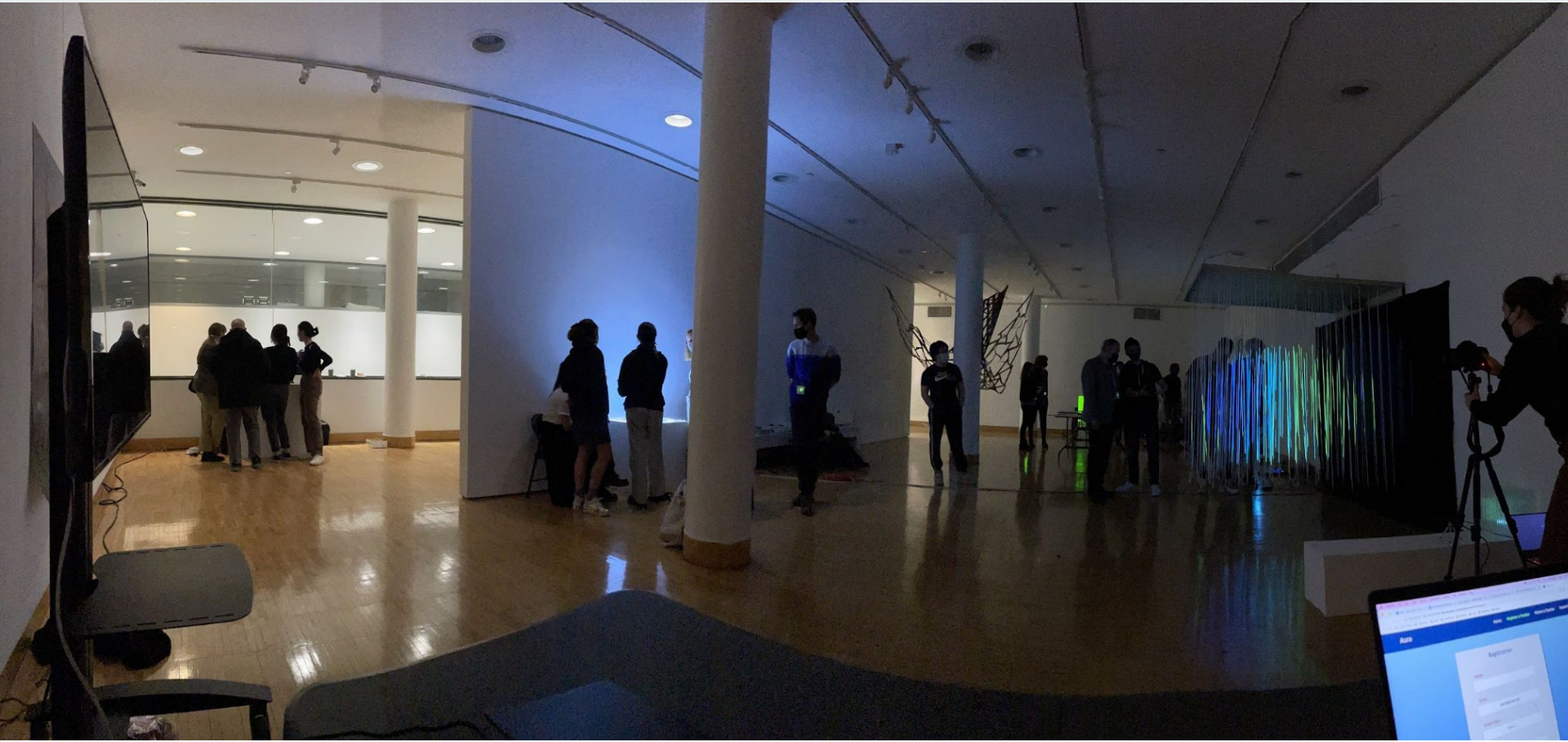

Aura
Home
Register a Tracker
Return a Tracker
Export Data
Return
Tracker #:
Select...
Submit

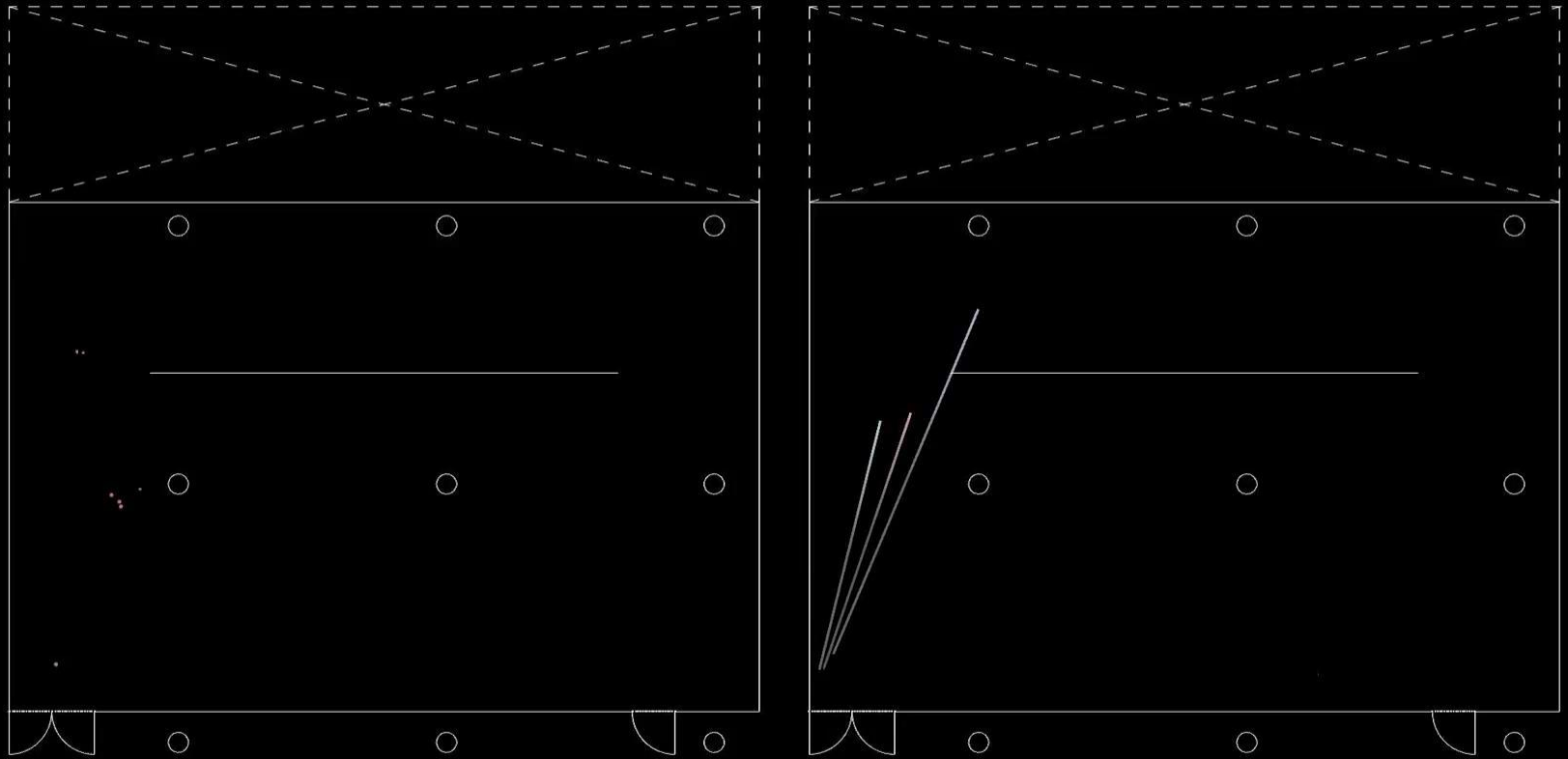
Aura

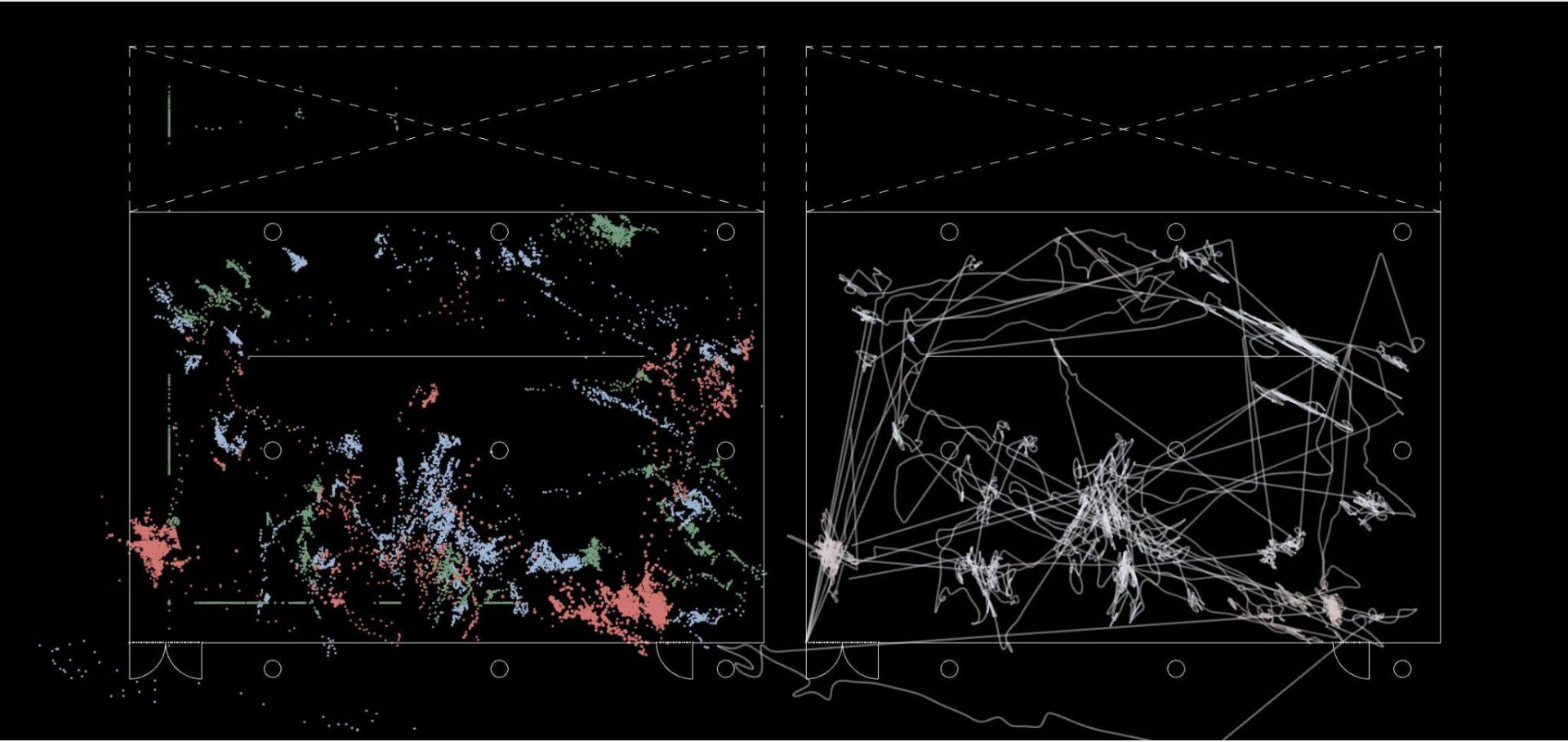
Home | Register a Tracker | Return a Tracker | Export Data











Social Science

- Benevolent Uses
  - Improving floor plans of frequently congested areas
  - Better public infrastructure
    - public fountains, bathrooms, info centers etc.
- Malevolent Uses
  - Collected information could be sold as data
  - Invasion of privacy in exchange for access to service
- Better awareness from both users and providers are required

Future Work

- Software
  - Scaling the system to support massive numbers of users, and process the increased data
  - More personal information to perform more rigorous data analysis
    - Ethnic patterns
    - Age patterns
    - Gender patterns
- Interdisciplinary
  - Create better visualization pipeline with more robust tools to remove computational bottleneck
  - Come up with a systemic approach to redesign floor plans to be more effective through means such as machine learning

Acknowledgements

- Professor Keene
- School of Architecture
  - Kyungmin Park
  - Tianyang Sun

CARBON CURATOR

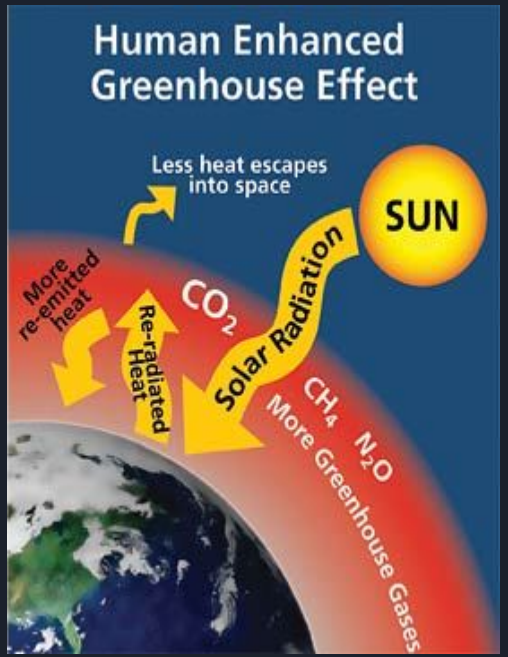
By: Danny Hong and Arthur Skok

Problem Statement

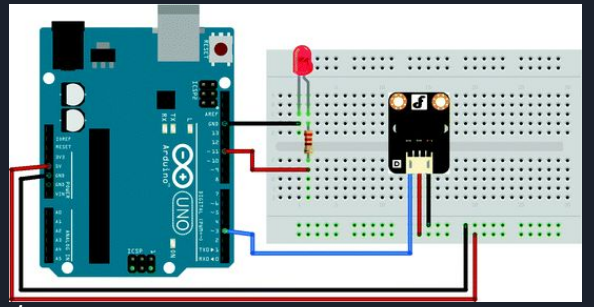
Scientific advancements have separated many people from the environment they live in.

Many people perform tasks that have harmful effects on the environment without realizing that there are more eco friendly alternative behaviors to do.

We propose a portable device that can help users become more aware of how their actions have



What is it composed of?



The project is composed of a Hardware and Software Component

- **Hardware Component**
  - Consists of an Arduino which is hooked up to the Temperature/Humidity, and GPS sensors.
  - Also has a battery pack connected to get power.
  - Has a Bluetooth component in order to have internet connection between the Arduino and the code.
- **Software Component**
  - Consists of the user interface experience between the user and the device.
  - Implements storage for user inputted data combined with sensory data

Current Iteration: Hardware Specs

ARDUINO SPECS	TEMPERATURE SENSOR SPECS	GPS SENSOR SPECS
- 256 KB of flash memory 8 KB used by bootloader - 16 MHz clock speed - 54 digital input/output pins - 5 V operating voltage - 37 grams of weight, 53 mm by 100 mm size (for portability)	- 3.3 to 6 V DC power supply - 0 - 100% relative humidity operating range - 0.1% relative humidity resolution (precision) - +/- 2% relative humidity accuracy - -40 to 80^{o} C operating range - 0.1^{o} C temperature precision - +/- 0.5^{o} C temperature accuracy	- 3.6 V power supply - 50 db antenna gain - 9600 baud rate - 2.5 meter precision

Current Iteration: Software

Communication Between Devices

- Data is encoded into packets on the arduino side of the code
  - Broadcasted via the usage of characteristics on the HM-10 BLE device
  - Packets decoded on application side

Data Utilization

- After decoding, data is stored locally and treated differently based on packet type:
  - Temperature data is parsed into temperature and humidity readings
  - GPS data is parsed into longitude and latitude
- Distance accumulates and resets, real outside temperature fetched from openweather API key and compared to readings
- Depending on what values are calculated, states are set throughout the application for rendering and alerts

Current Iteration: Software (Continued)

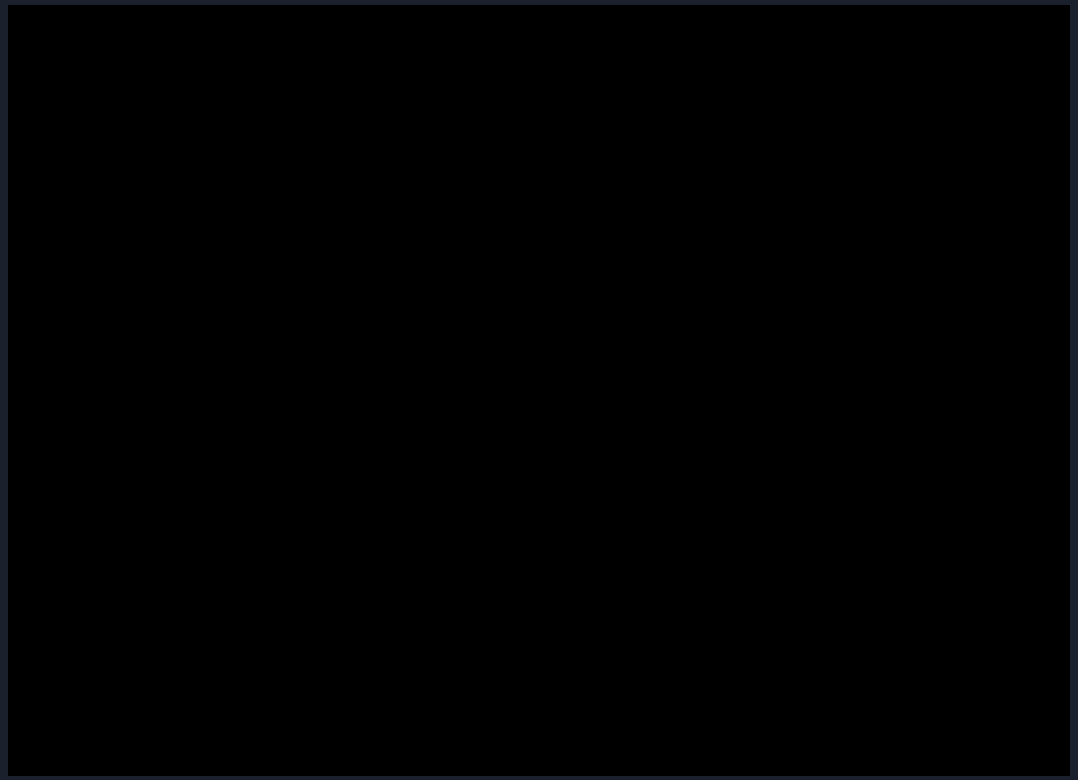
Communicating to the User

- If specific thresholds are met, a notification via alert is sent to the user to tell them that a question is waiting for them at the prompt screen
- Furthermore, the navigation button rerenders as an exclamation point until the user answers the program in affirmative or negative to signify pending questions
    - Depending on the user response, the application seeks further clarification for what the user is doing

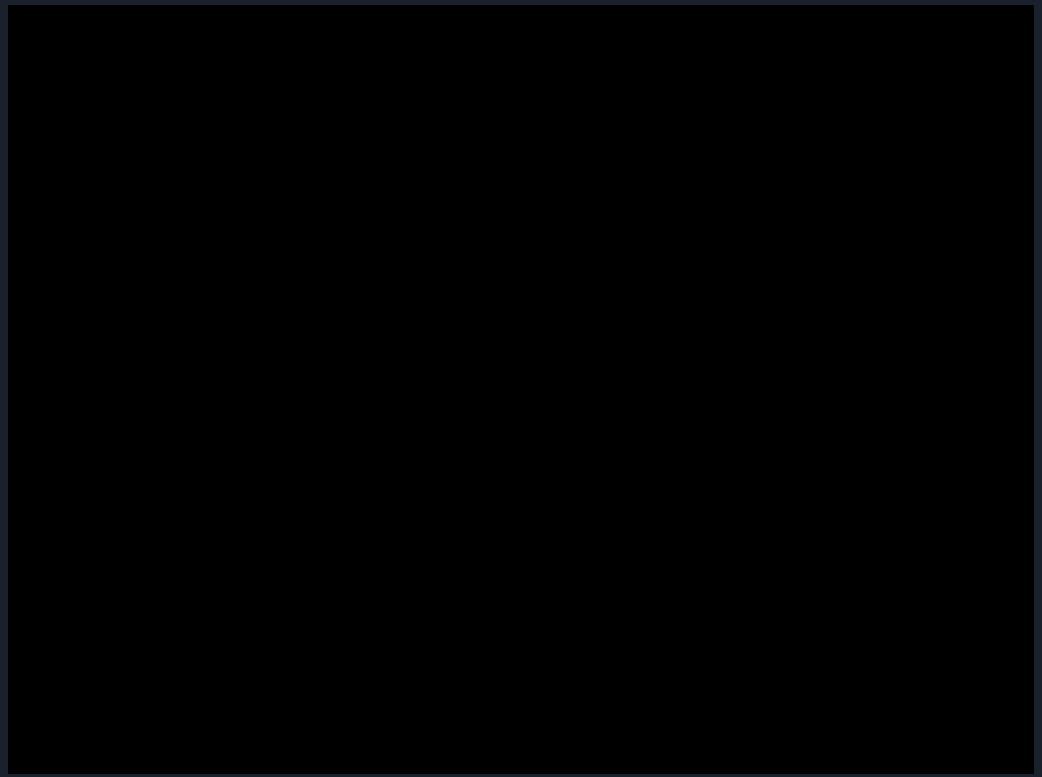
Storing Responses

- Certain answers are then stored locally
    - However this data needs to stay in memory even as the application goes out of scope, turned off, or anything shy of an uninstall
- We decided to implement SQLite for our application for persistent data storage in this case
- The data being stored in the SQLite database is rendered to the user in the prompt screen so they may see their previous results

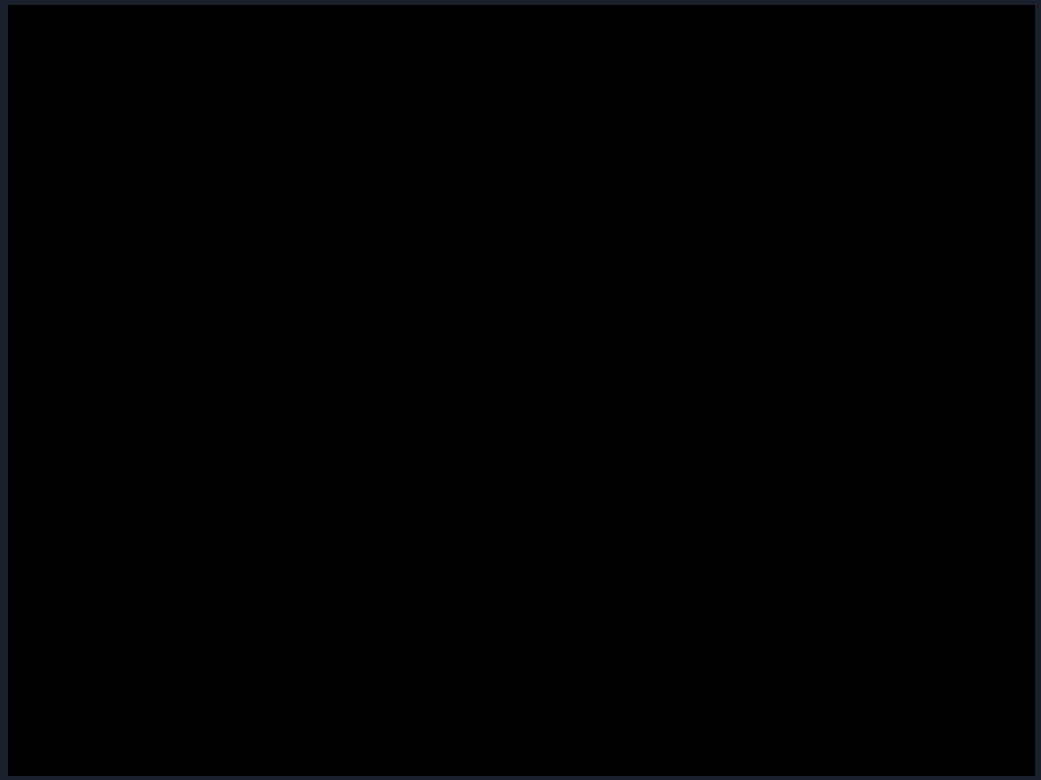
User Logging Data based on Travel



User Logging Data based on Temperature (Alerts also showcased regardless of screen)



Pending Prompt Changes Render Status if User Ignores Alert



...And it's portable!

...sort of

What is next?

- **Additional user prompts:**
  - What are you doing?
    - Purchases, working
- **Recommendations:**
  - Alternative transit options:
    - Fetch nearby transit features from google maps API using stored GPS data at user end travel entries
  - Alternative purchasing options:
    - Goods can be hard coded to give an environmental grade
  - Reducing excessive heating or cooling
- **Hardware:**
  - Making our design more compact and portable
    - Further refining our physical carriage to take up less space, perhaps by mounting components facing away from each other

Acknowledgements

Thank you to Professor Keene and Professor Simson for guidance in this project throughout the year!

A Hybrid Approach for Image Vectorization for Semi-Geometric Images

by
Jonathan Lam, Derek Lee, Victor Zhang

Prof. Sam Keene

May 13, 2022

Problem statement

- Converting a raster (pixel-based) image to vector (shape-based) image
- Develop hybrid method that combines benefits of previous methods

Edge tracing

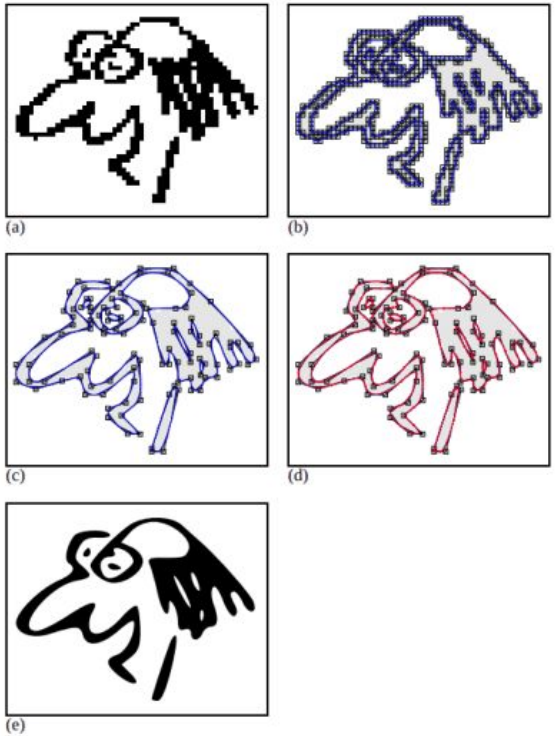


Figure: Illustration of the Potrace [1] vectorization process

Blue-noise sampling

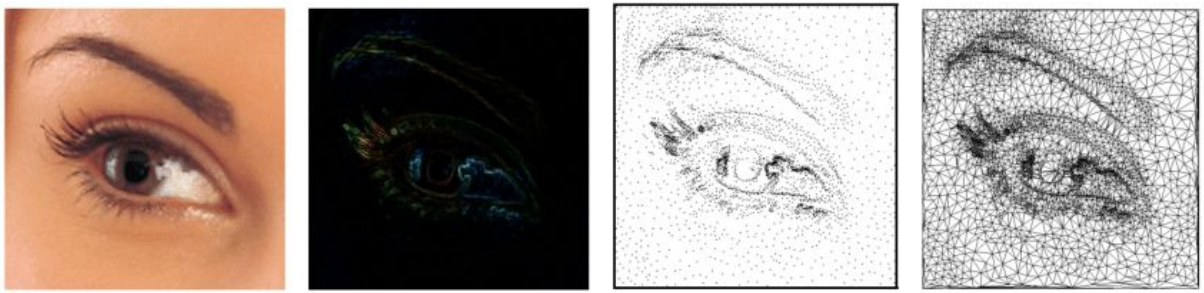


Figure: Illustration of the BNS vectorization process [2]

Hybrid approach

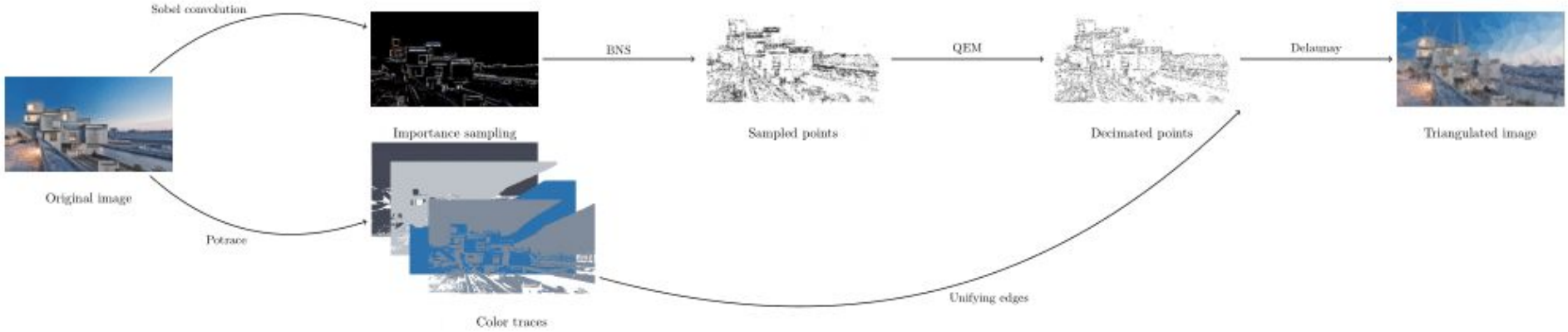


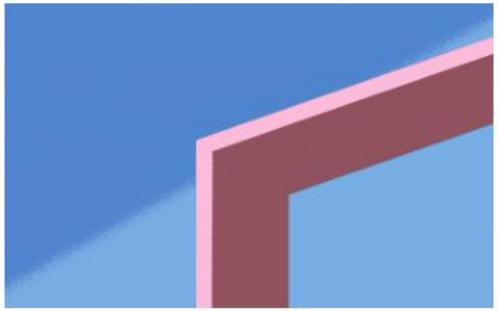
Figure: Architecture diagram

Results

(a) Original image

(b) BNS image

(c) Hybrid image



(d) Potrace image

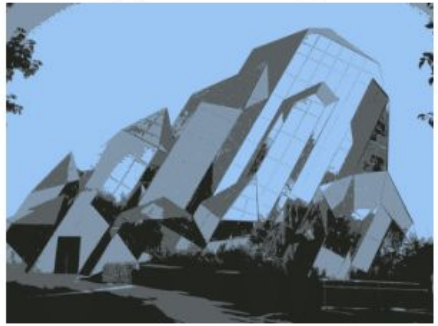
Figure: Set of images for experiment 3

Results

(a) Original image

(b) BNS image

(c) Hybrid image



(d) Potrace image

Figure: Set of images for experiment 4

Results

(a) Original image

(b) BNS image

(c) Hybrid image

(d) Potrace image

Figure: Set of images for experiment 5

Results

(a) Original image

(b) BNS image

(c) Hybrid image

(d) Potrace image

Figure: Set of images for experiment 8

Future work

- Alternative methods to strengthen edges
- Curve simplification
- Machine learning preprocessing
- Mathematical model of pipeline
- Improved evaluation metrics

Conclusions

- Implemented framework for vectorizing images
- Based on blue-noise sampling and Potrace
- Larger file size in exchange for better performance on accuracy (MSE)

References

[1] Peter Selinger. "Potrace: a polygon-based tracing algorithm". In: *Potrace (online), http://potrace.sourceforge.net/potrace.pdf (2009-07-01)* 2 (2003).

[2] Jiaojiao Zhao, Jie Feng, and Bingfeng Zhou. "Image vectorization using blue-noise sampling". In: *Imaging and Printing in a Web 2.0 World IV*. Vol. 8664. International Society for Optics and Photonics. 2013, 86640H.

NFT Marketplace

Brian, Amy, & Andrew

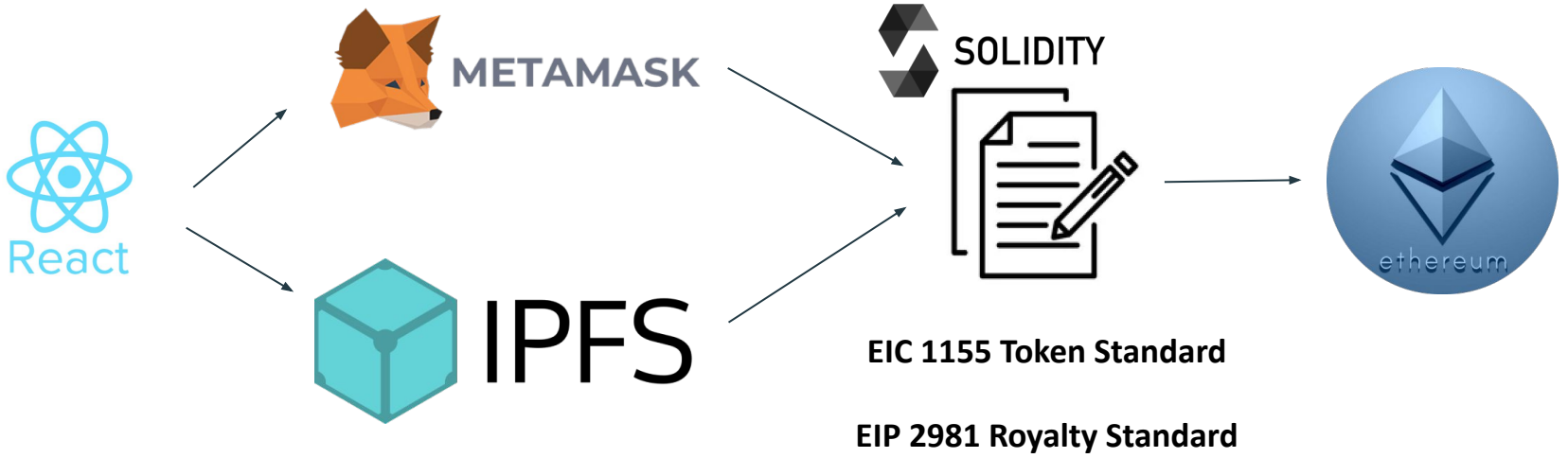
Background

- Art graduates struggle to earn a living following graduation due to lack of reputation & startup funds.
- Students are unable to present art in fairs, because strict entry deadlines conflict with school work.
- Art market is expanding into the digital space with the emergence of NFTs & Smart Contracts.

Solution / Objective

- Ethereum-based NFT marketplace that introduces students to their respective markets using the Cooper Union name brand.
- Restrict minting to Cooper Union students alone to give students a platform for their work to be seen and purchased.
- Proof-of-concept to display possibilities of applying an NFT marketplace to an academic institution and initiate conversations within Cooper Union.

Technologies Used

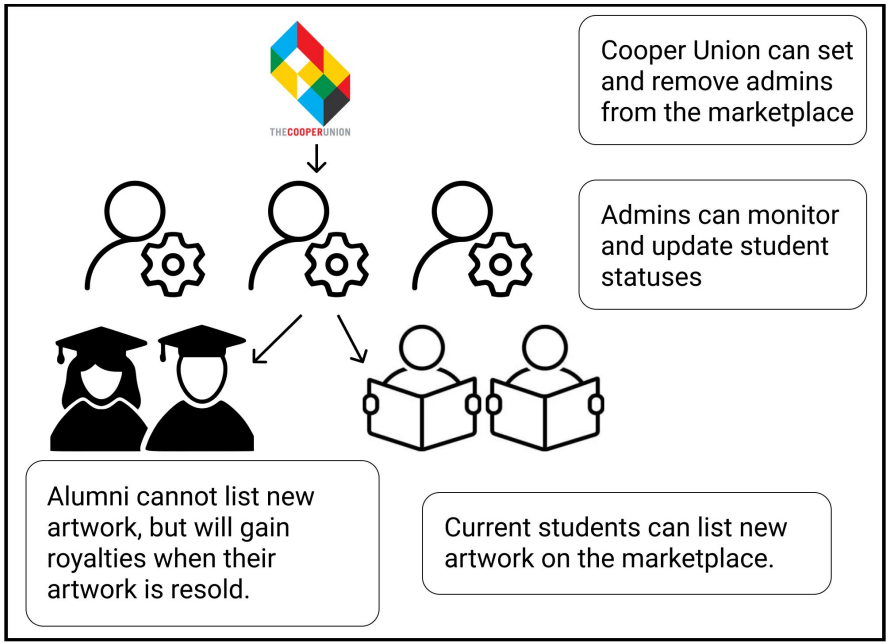


EIC 1155 Token Standard

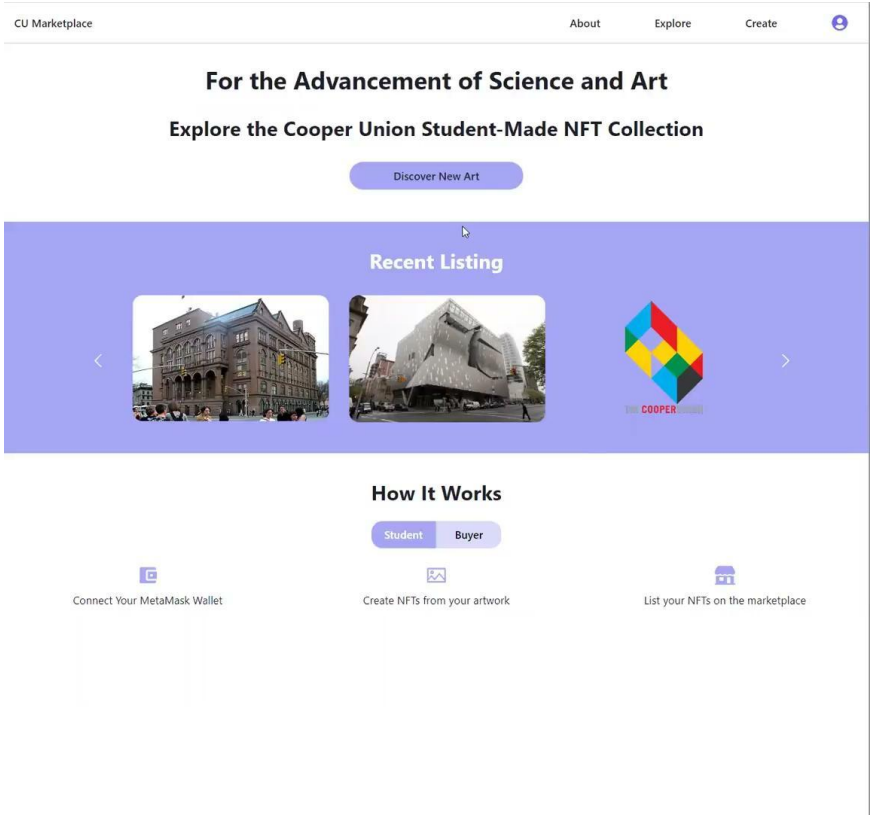
EIP 2981 Royalty Standard

Marketplace Hierarchy

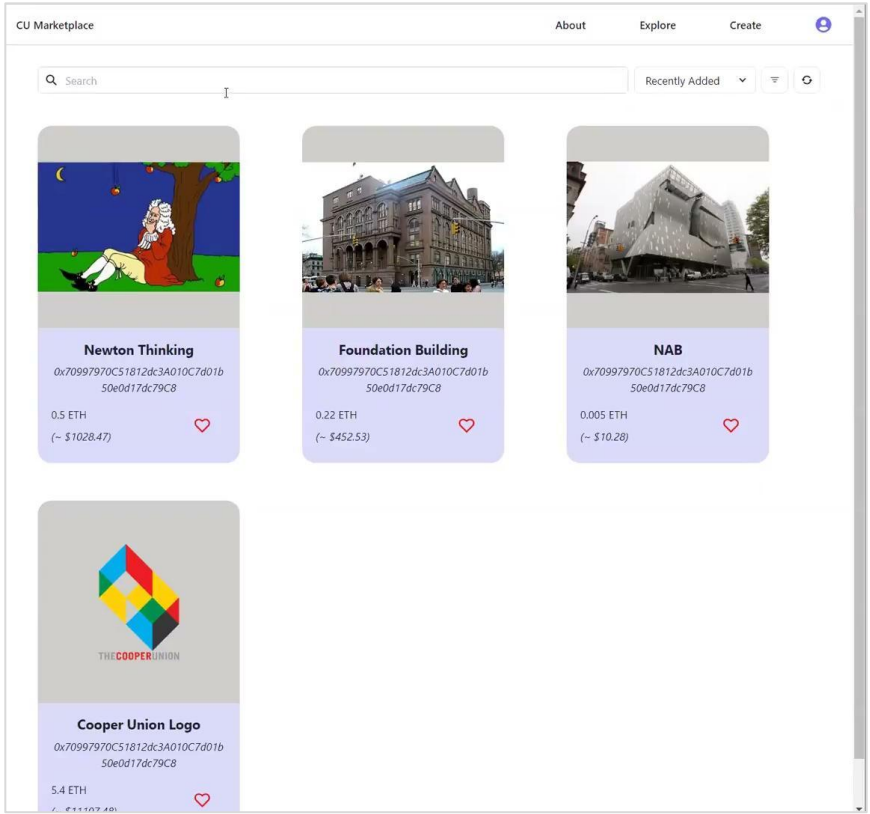
- Marketplace roles are built into the smart contracts and tied to Ethereum wallet addresses.



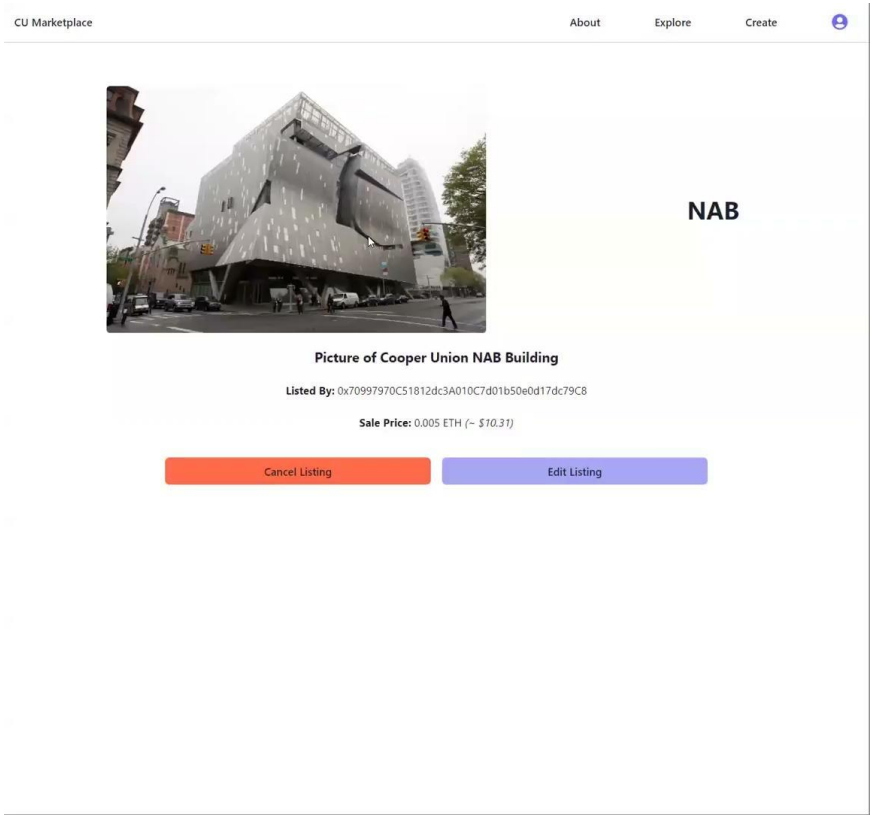
Demo (Home Page)



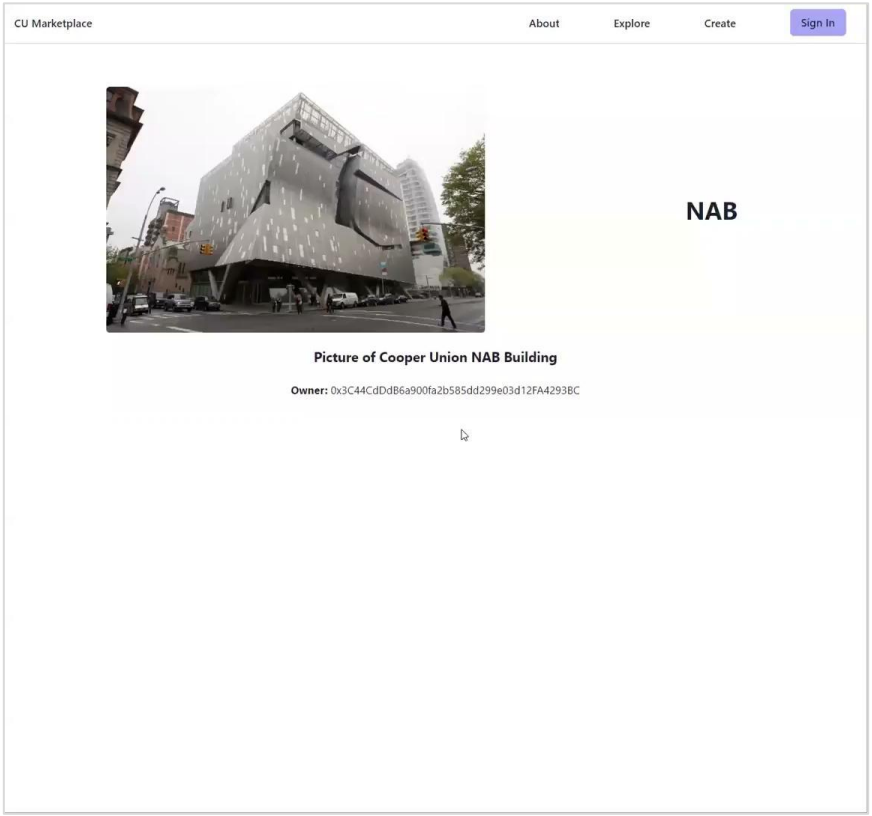
Demo (Explore Page)



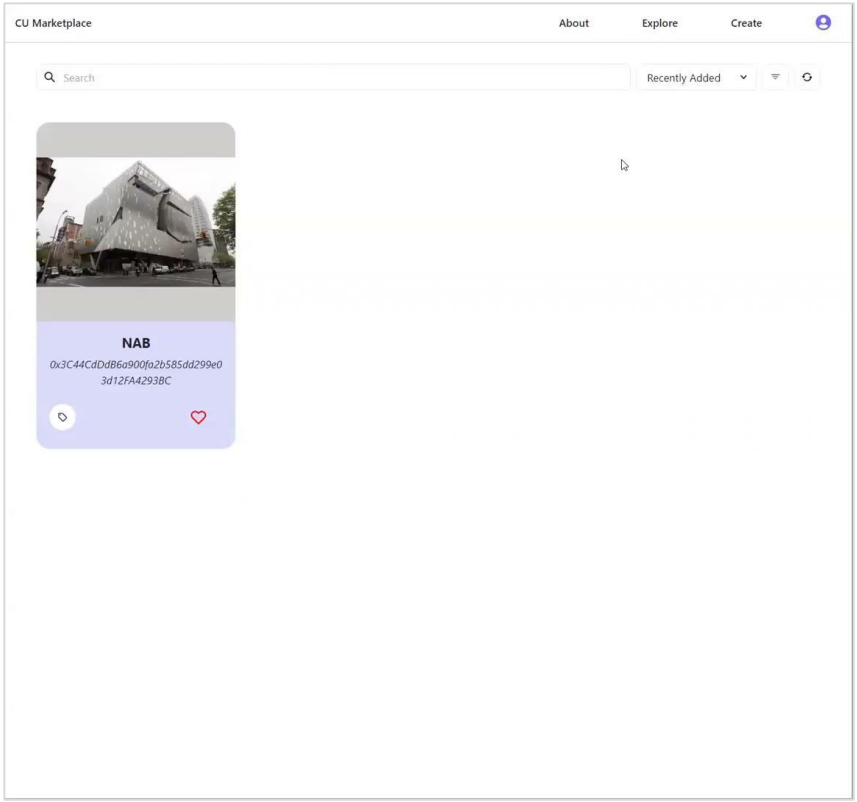
Demo (NFT Page)



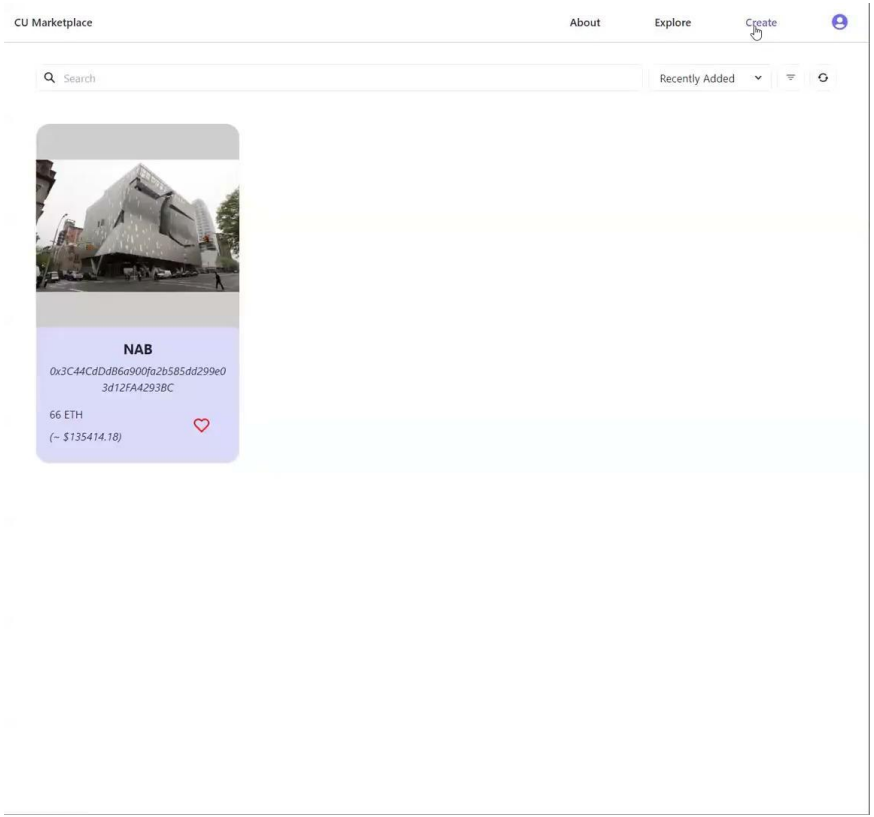
Demo (Sign in with MetaMask)



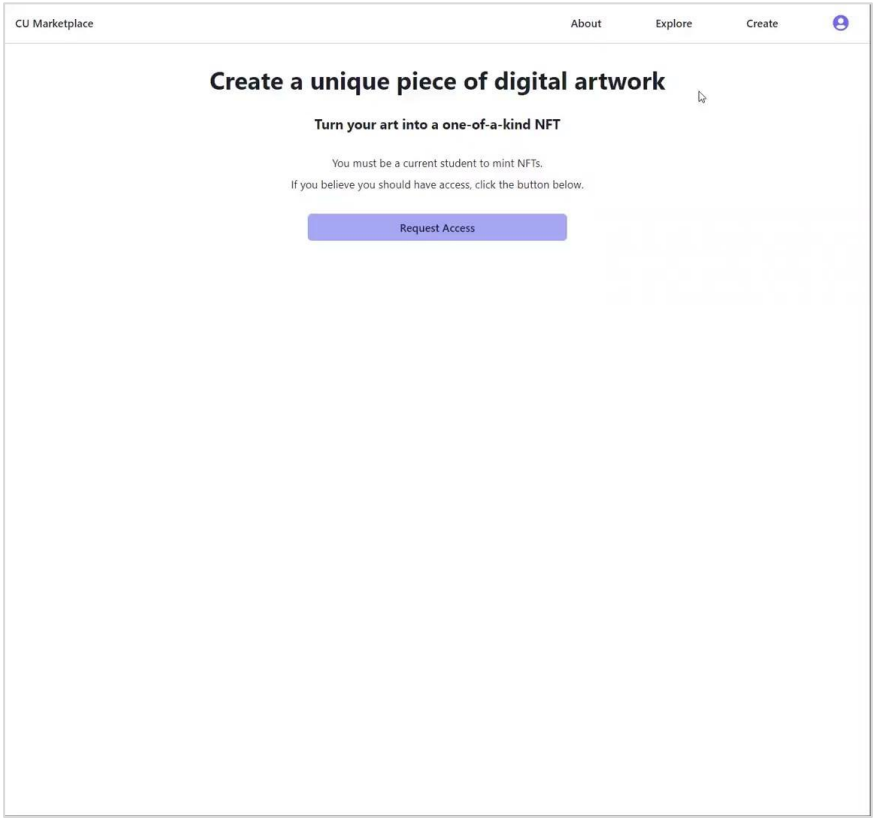
Demo (My NFTs Page)



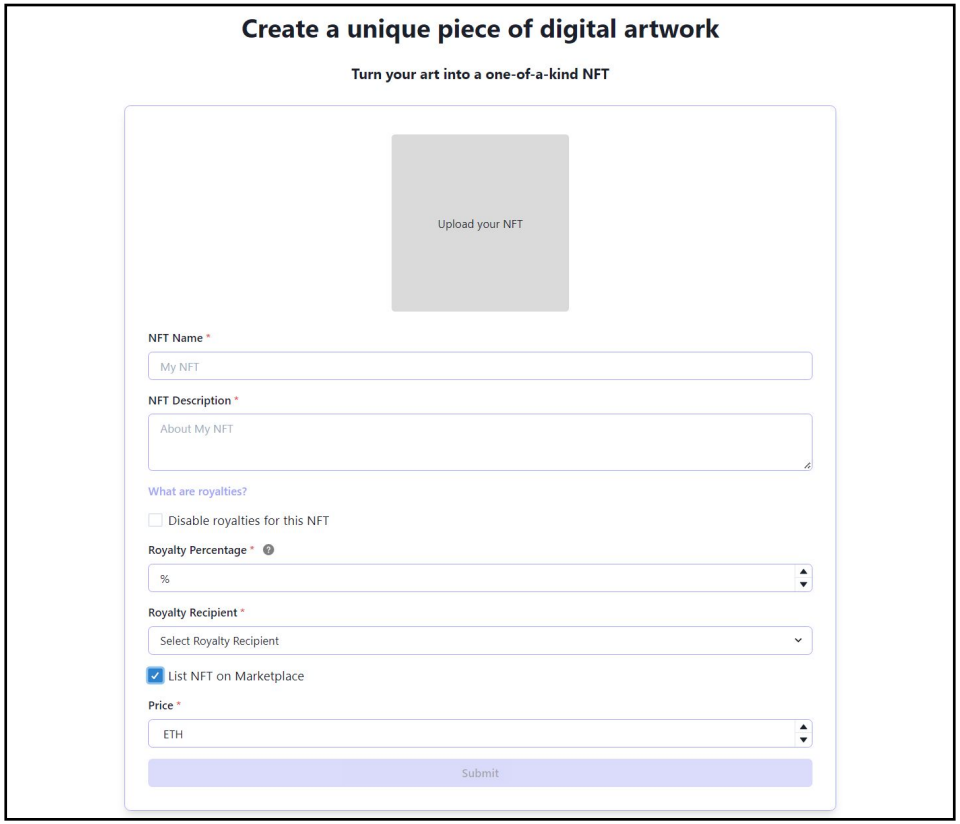
Demo (NFT Creation - Not a Student)



Demo (Update Student's Status)



Demo (NFT Creation)

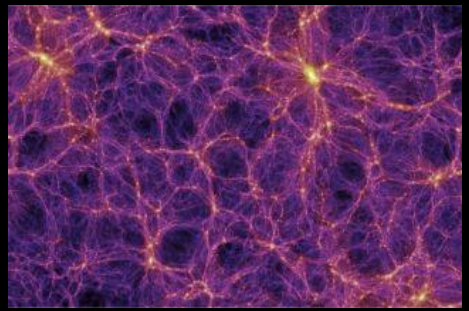


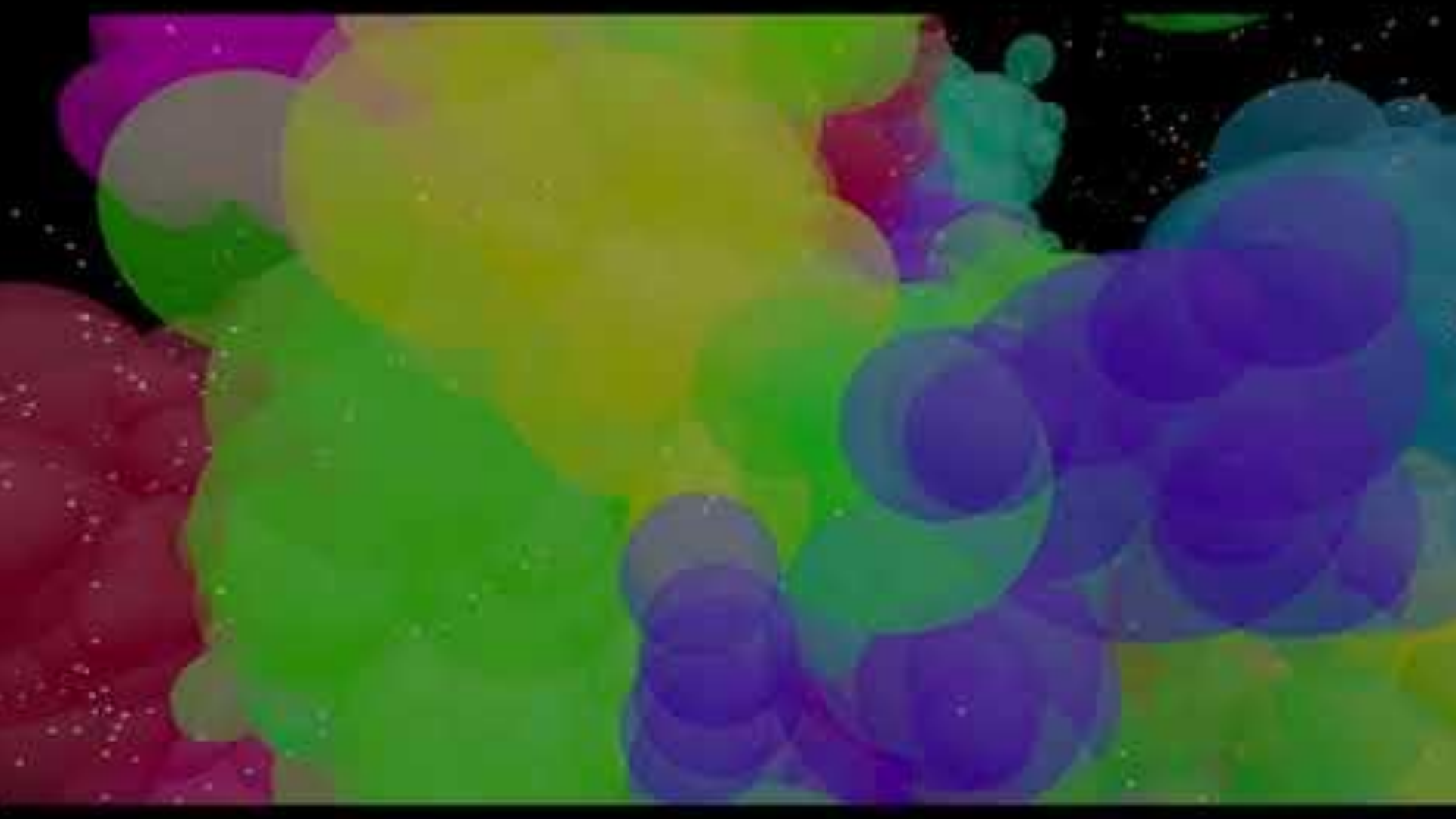
Future Work

- Implement system where Cooper allots an allowance of ETH to each student for gas fees (to promote equity in student listing opportunities).
- Allow alumni to donate NFTs for Cooper to sell on marketplace.
- Add alternate sale methods (i.e auction).
- Research eco-friendly blockchains for final product.

Machine Learning And Data Visualization for Cosmic Voids

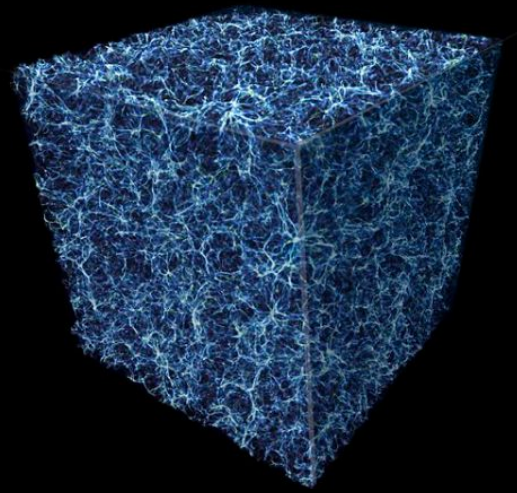
Bonny Wang
Allister Liu
Prof. Pisani
Prof. Keene

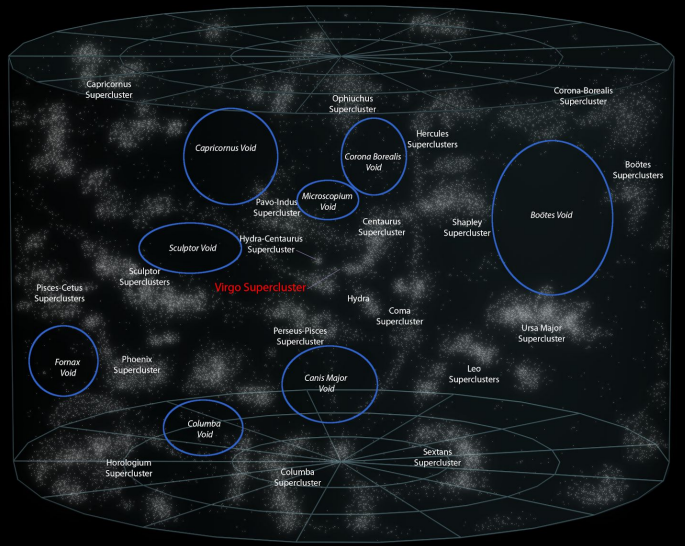




What are Voids?

- The large under-dense regions in the Universe dominated by dark energy
- A novel probe for our understanding of the Universe
- Yet an understudied field

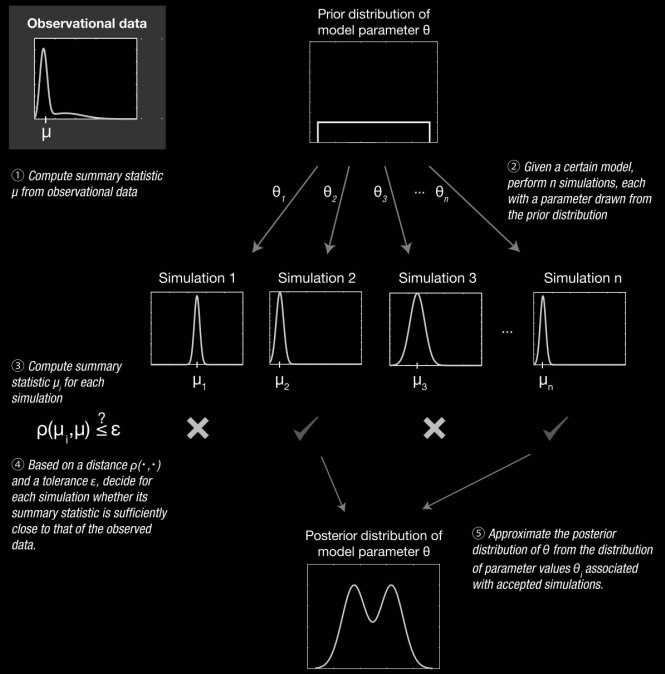




Problem Statement

- The intractability of the likelihood function in astrophysical simulations
- Computation

Traditional Method for Scientific Inference



- The likelihood function is intractable in complex simulations
- Approximate Bayesian computation is time-consuming and computationally expensive
- The exact value of the posterior probability is unknown and may not be normalized

↓

Density Estimation-based

Likelihood-Free Inference

Moment Networks

- A simple hierarchy of fast neural regression models
- Addresses the "**curse of dimensionality**"
- Directly skipping to the estimations of mean, standard deviation of the predicted parameter

$$L = || (\theta - F(x))^2 - G(x) ||^2$$

mean of posterior distribution

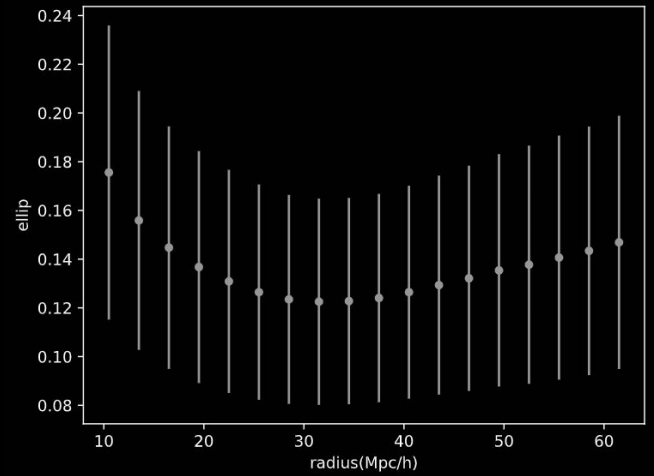
standard deviation of posterior distribution

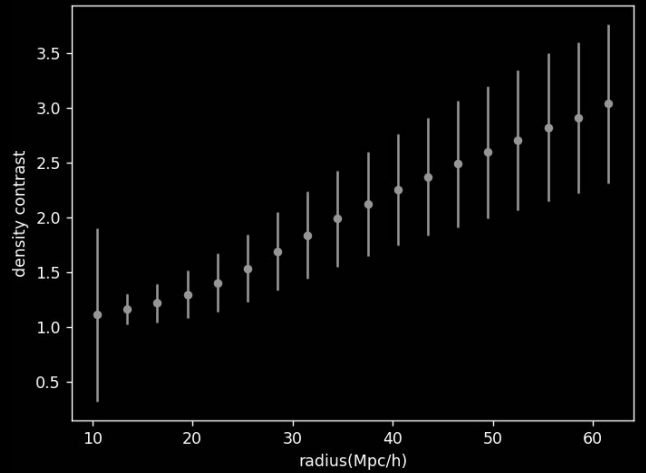
Dataset

GIGANTES

- An extensive and realistic void catalog
- Contains over 1 billion cosmic voids with their different cosmological parameters – over 20 TB
- Created by running the void finder algorithm VIDE on simulations

Relationship between what we want to explore and the previous study





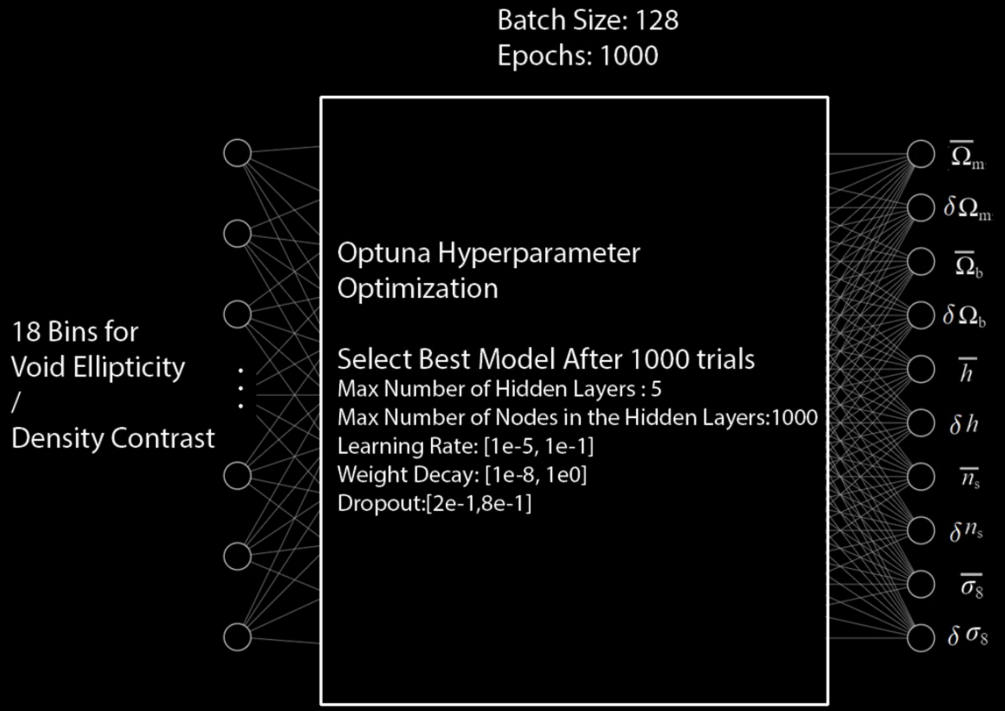
Void Ellipticity:

Describes the elliptical shape of a void

Void Density Contrast:

The ratio of the minimum density of the particle on the ridge of the void to the minimum density of the void

Our Work with GIGANTES using Moment Networks



– Matter density

– Baryon density

– Dimensionless Hubble constant

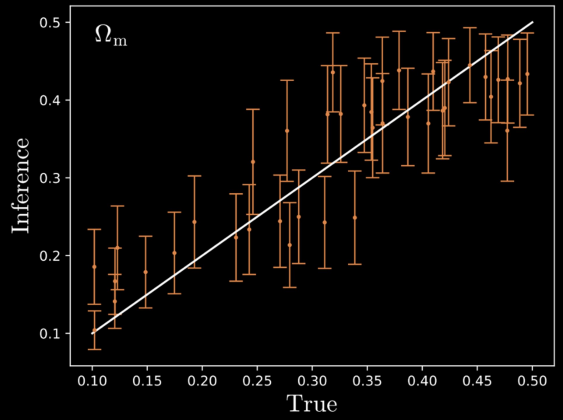
– Scalar spectral index

– Present root-mean-square matter fluctuation

$$L = ||(\theta_{true} - \overline{\theta}_{true})^2 - \delta\theta||^2$$

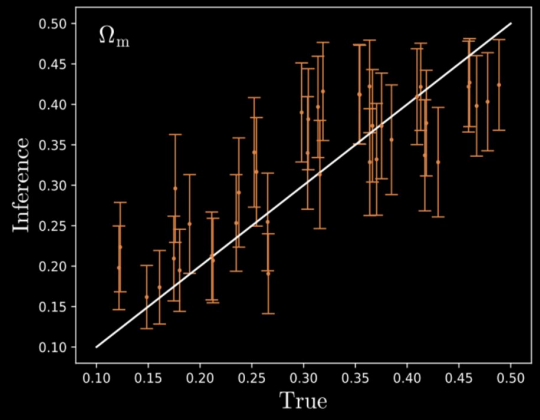
Results

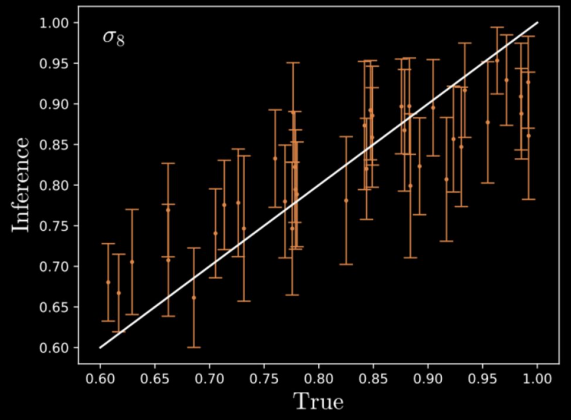
Using Void Ellipticity to predict



Parameter	RMSE (root-mean-square deviation)	R-Squared	χ^2_ν
Ω_m	0.0594	0.729	1.010

Using Void Density Contrast to predict

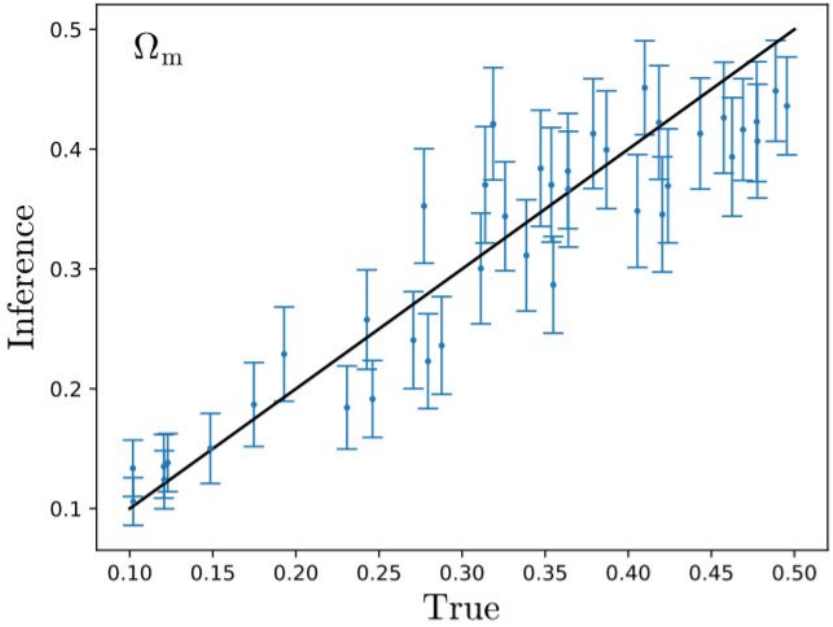


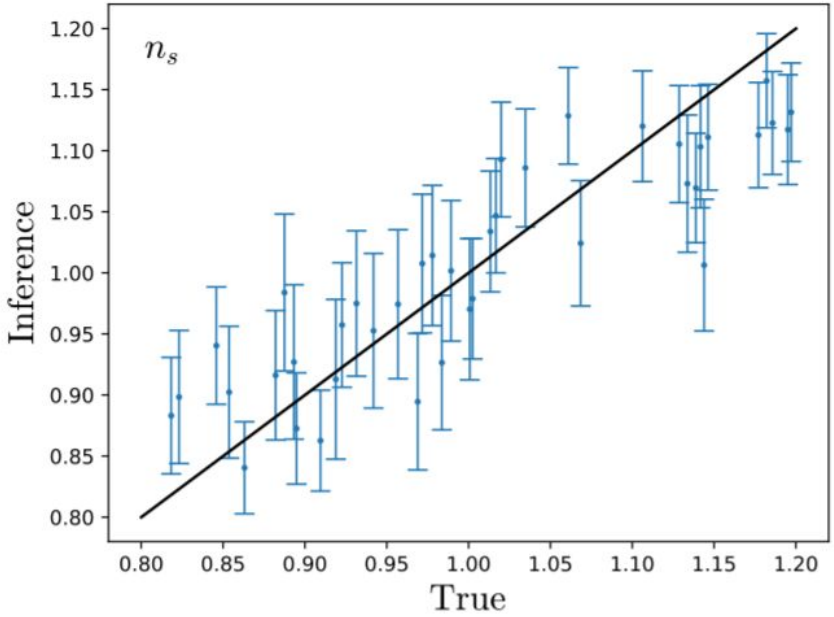


Parameter	RMSE (root-mean-square deviation)	R-Squared	χ^2_ν
Ω_m	0.0586	0.736	1.09
σ_8	0.0822	0.428	0.842

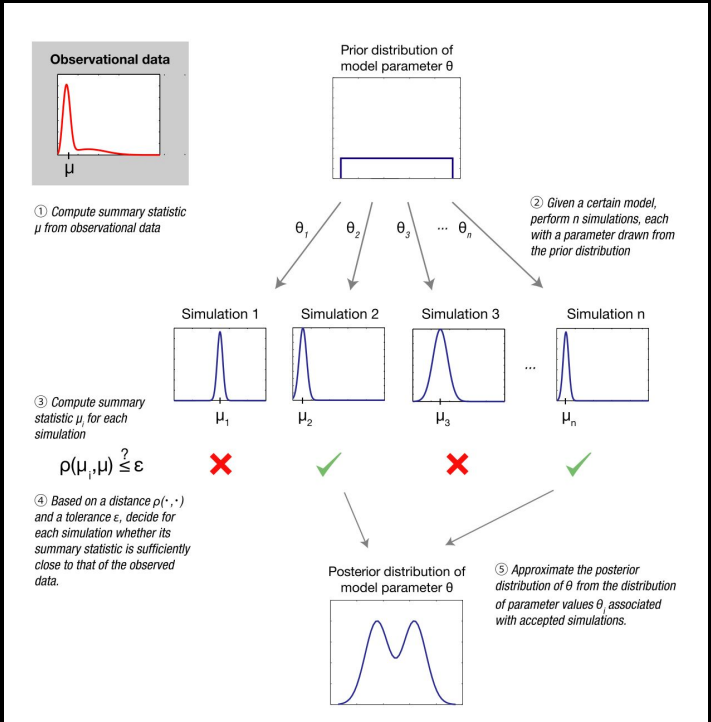
Ω_m: matter density parameter

σ_8: The present root-mean-square matter fluctuation averaged over a sphere of radius $8h^{-1}$Mpc


$\Omega_{\rm m}$
Inference
True
0.1
0.2
0.3
0.4
0.5
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50


n_s
Inference
True
0.80
0.85
0.90
0.95
1.00
1.05
1.10
1.15
1.20

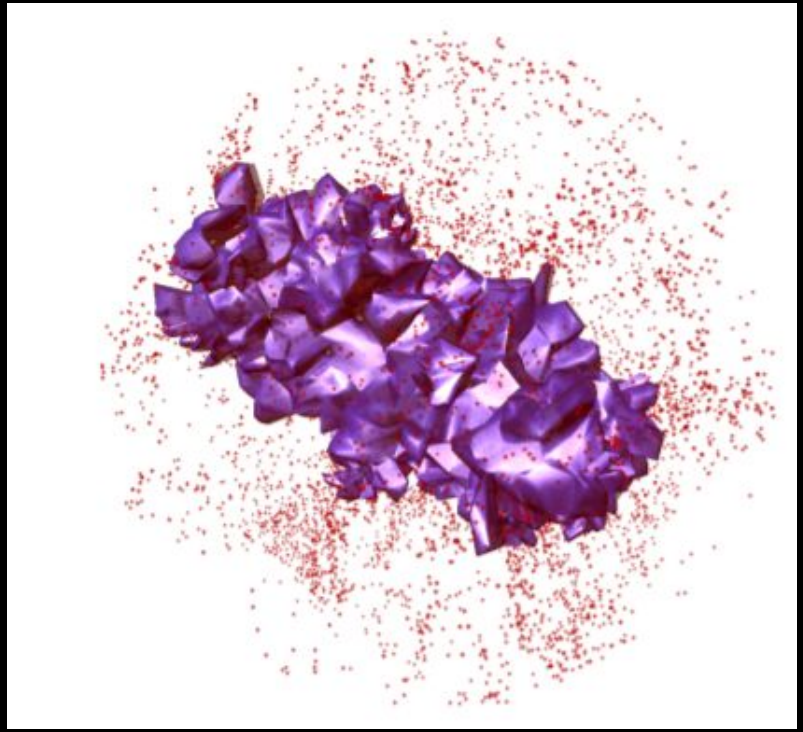
Traditional Method for Likelihood-free inference



- Approximate Bayesian computation
- Extremely time-consuming and computationally expensive (sometimes impractical) due to the high dimensionality of astrophysical data

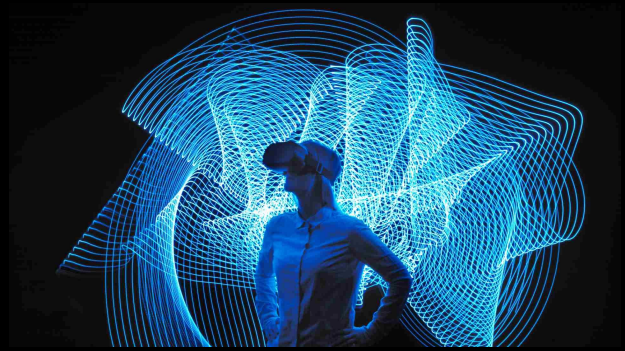
Visualize the conceptual voids –
as a stepping stone to learn more
about our mysterious universe.

How do we visualize voids? – As of NOW



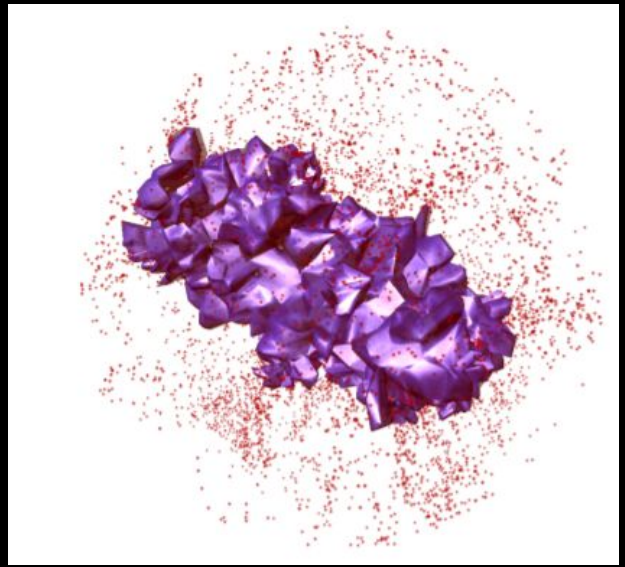
Acquaint with the Invisible

Introduce the general public to the world of invisible through VR technology – fly through and interact with cosmic voids.



Void Visualization

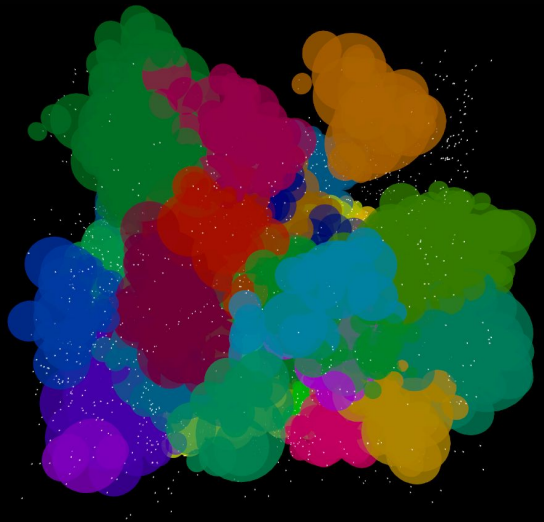
- 3D VR scenes of voids
  - Better representation of the void's spatial complexity than a plain 2D graph
  - Much more immersive and interactive experience
  - Unity game engine – provides native support for Virtual Reality
  - HTC Cosmos Elite VR headset

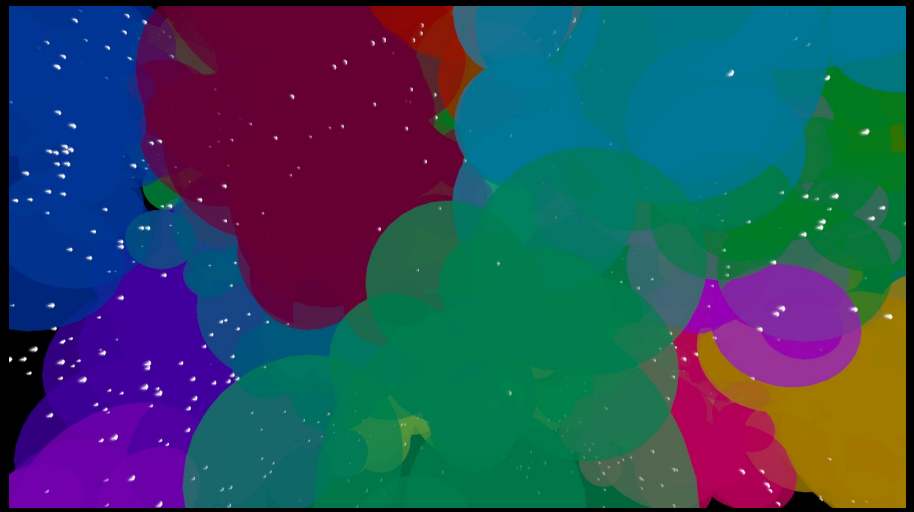


GIGANTES Void Catalog

Use VIDE to Generate Void Shape (Coordinates of Voronoi Cells)

Create VR Scene in Unity with Galaxy data(Coordinates) and Void data(Coordinates and Volume)





Void Catalog From the True Parameters

Use VIDE to Generate Void Shape (Coordinates of Voronoi Cells)

Create VR Scene in Unity with Galaxy data(Coordinates) and Void data(Coordinates and Volume)

Void Catalog From the Predicted Parameters

	Models From The Expected Parameter	Models From The Predicted Parameter
Using Void Ellipticity to Predict Ω_M	 $\Omega_M = 0.2673$	 $\Omega_M = 0.2897$
Using Void Density Contrast to Predict Ω_M and σ_8	 $\Omega_M = 0.2673$ $\sigma_8 = 0.731$	 $\Omega_M = 0.2467$ $\sigma_8 = 0.727$


predicted

MATLAP

Manga
Automatic
Translation
Language
Assistance
Program

Thodori Kapouranis Steven Lee

Motivation

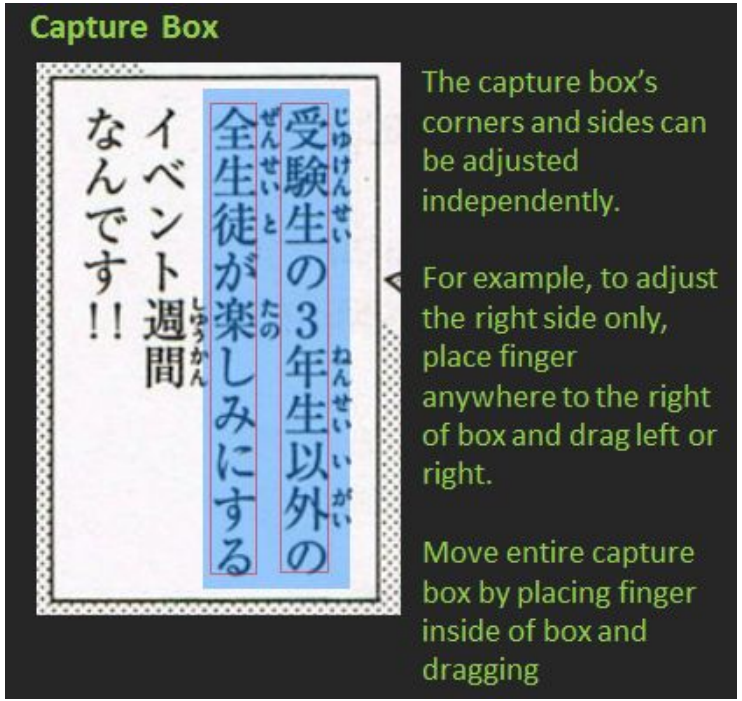
- Rise in interest in Japanese language
- Study material is saturated at the beginner level
- Large difficulty jump to start studying from native material

Objectives

- Desktop Application for mining vocabulary from comic books
- Automatic text detection
- Easy to use UI
- Less time note-taking, more time reading

Background - Typical OCR readers

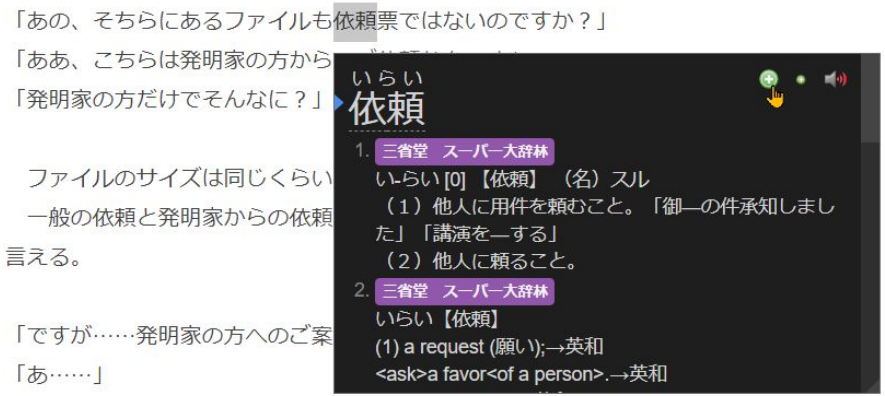
- Manual, On demand OCR detection

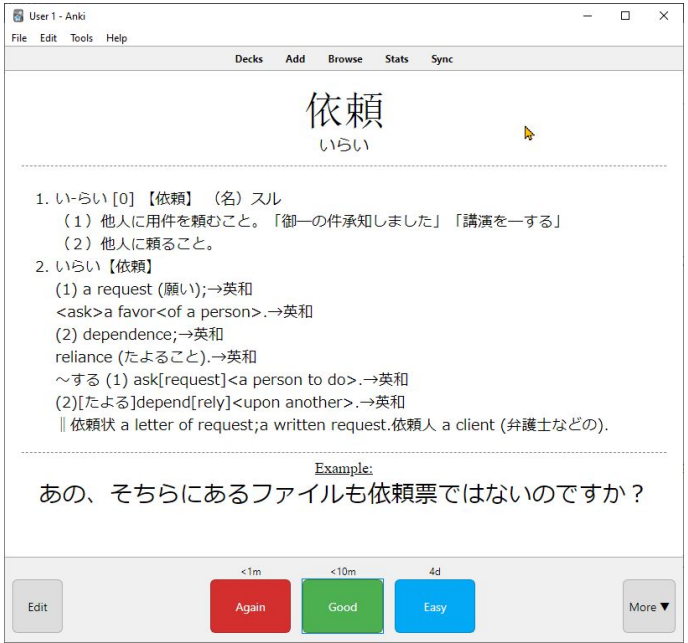


[OCR Manga Reader for Android]

Background - Study tool integration

- Applications that make creating study notes easier



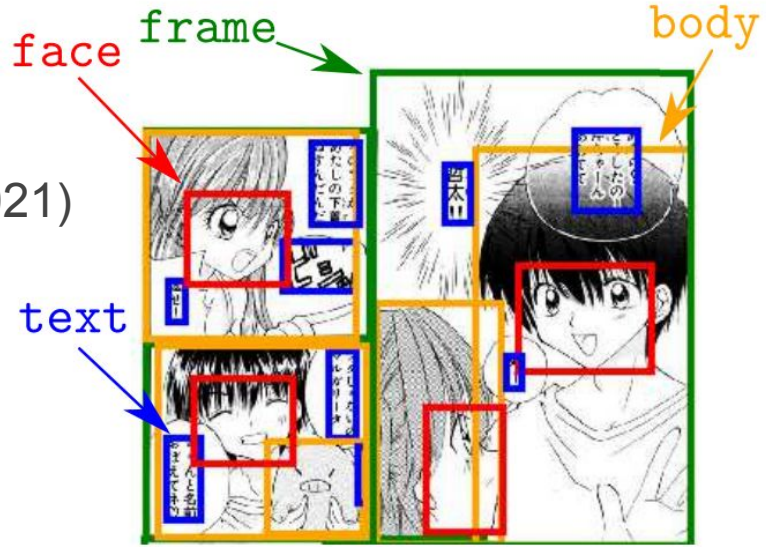


[Google Chrome Extension ‘Yomichan’]

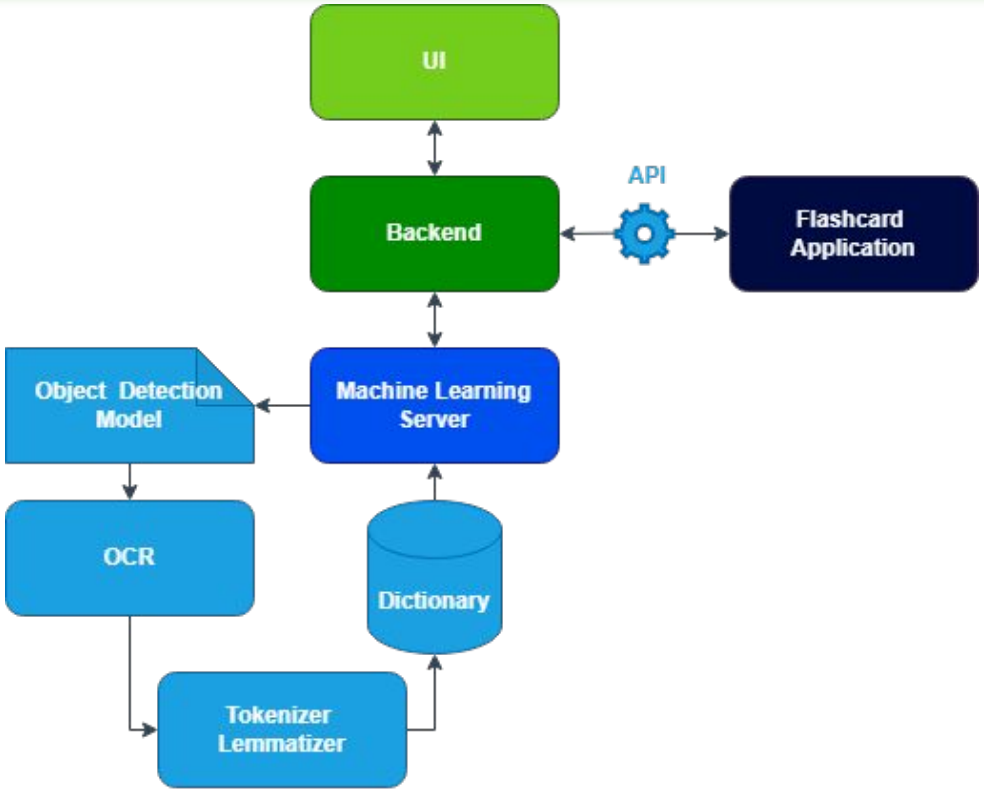
Background - Comic Object Detection

Building a Manga Dataset "Manga109" with Annotations for Multimedia Applications (2020)

“Towards Fully Automated Manga Translation” (2021)

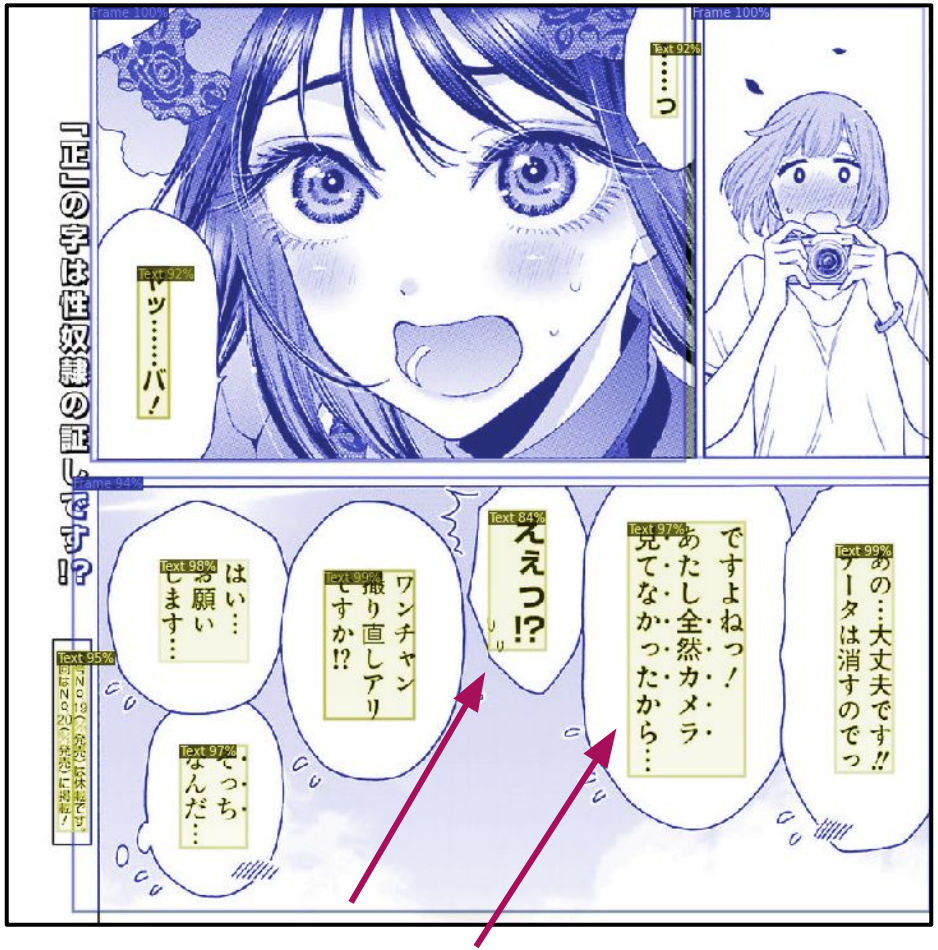


System Overview

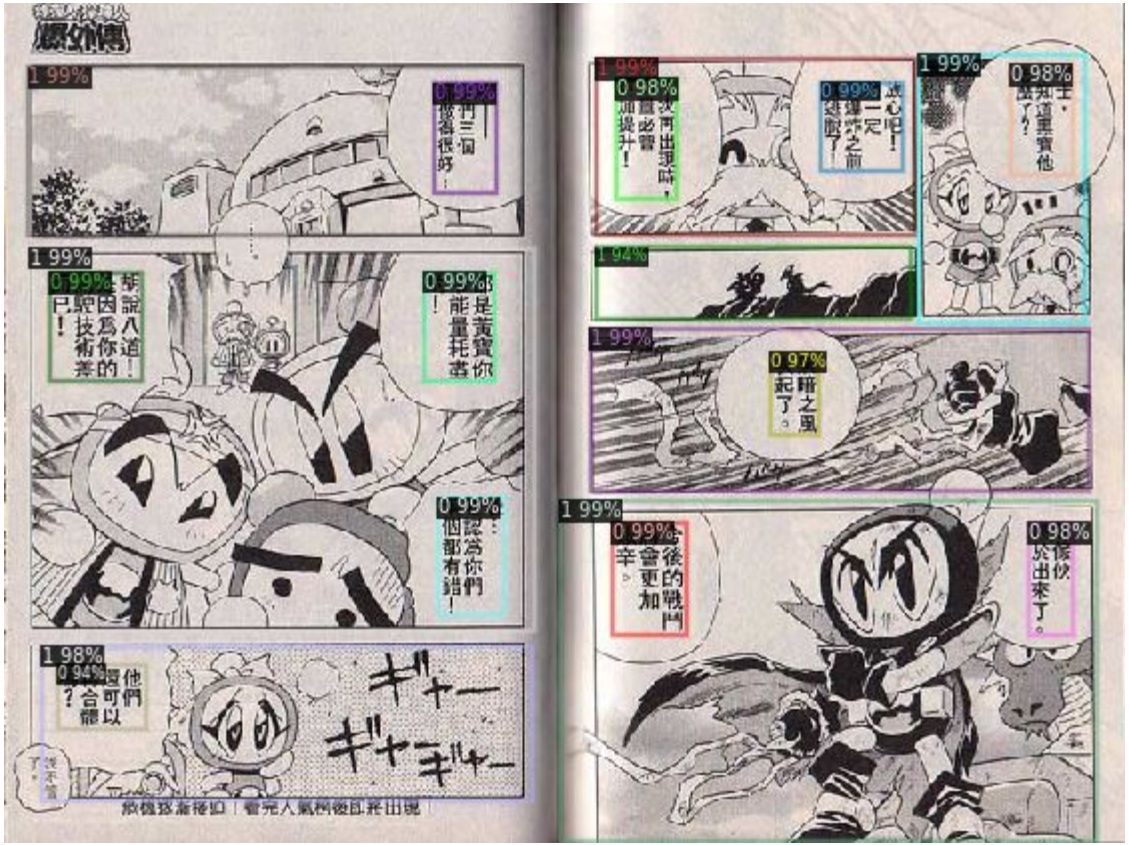


https://www.youtube.com/watch?v=xcKrOIKNEjo&feature=emb_title

Model Results - High quality scan



Model Results - Chinese language & bad scan



Real-time Personalized Sound Signal Separation & Augmentation

...

Jungang Fang
Jinhan Zhang

Background & Problem Statement

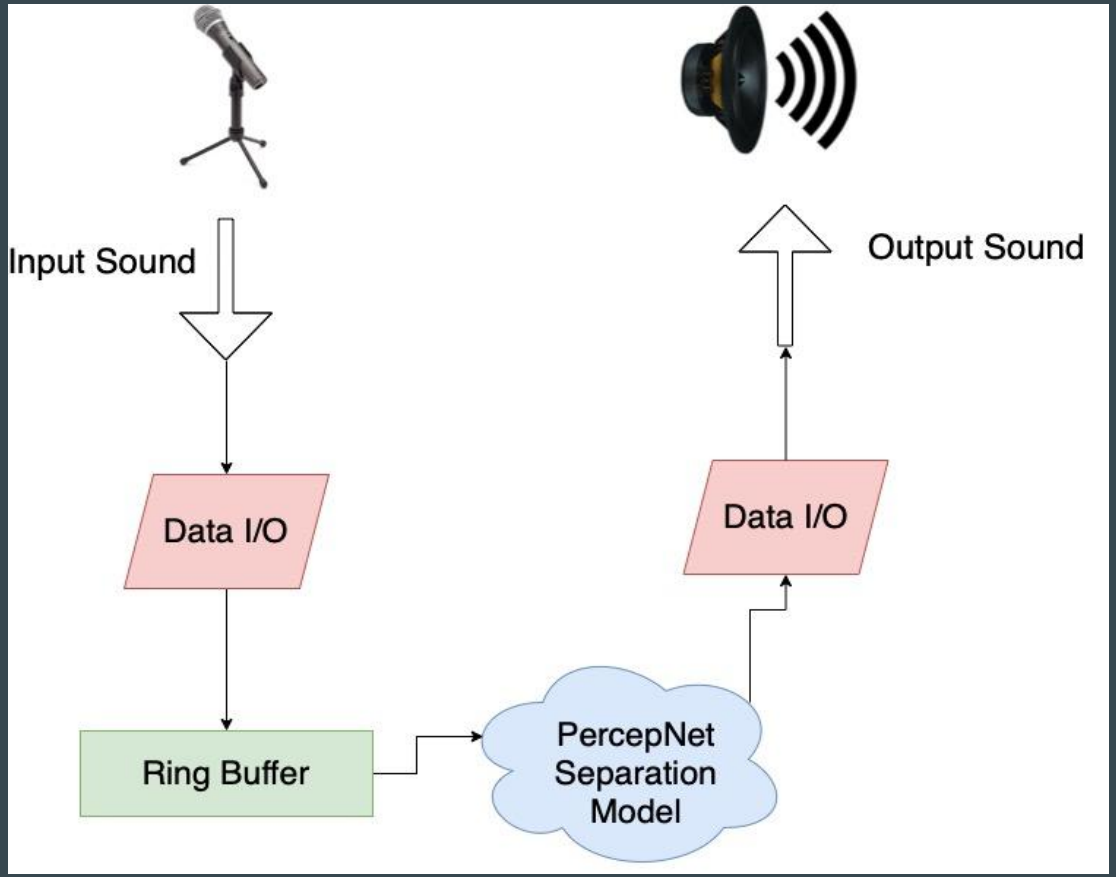
Goal:

To improve single source speech quality by denoising and speech recognition in real-time.

- bandwidth: wideband
- sample rate: 48000 Hz
- frame length: 10 ms
- trying to achieve total algorithmic latency(framesize + stride time): ≤ 60ms

C++ Implementation of PercepNet Pipeline

- Taking real-ime input from 3M worktunes headset microphone,
- process the input sound data frame by frame with the model, and generate output sound signal



Our Progress – PercepNet Implementation

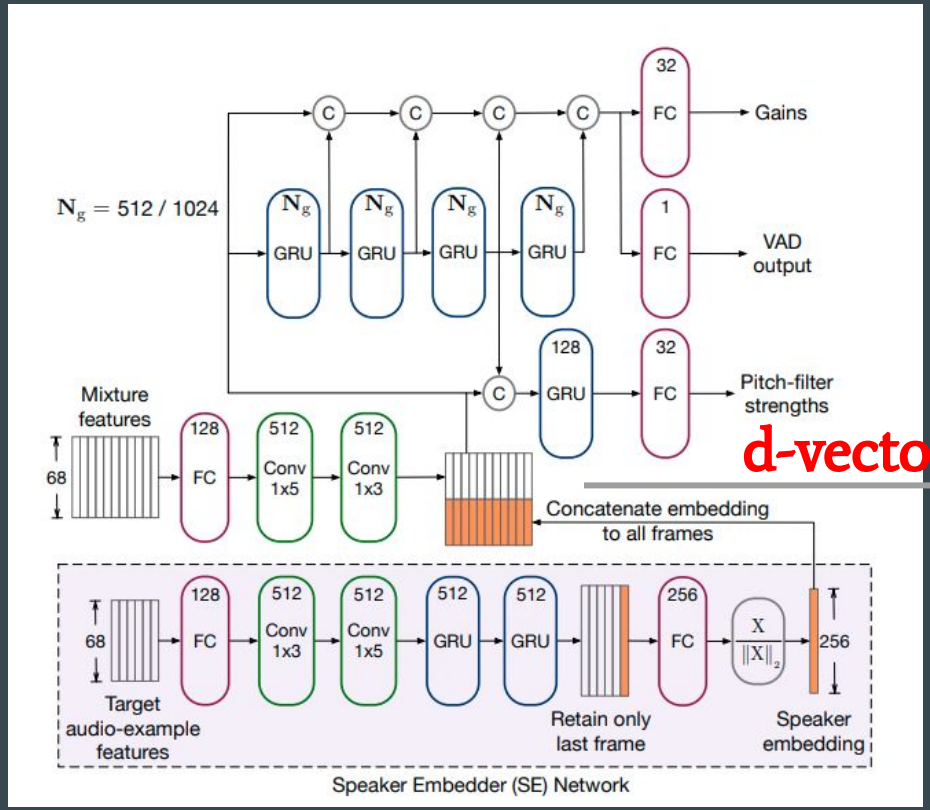
- Fully implemented PercepNet based on a partial implementation lacking the key feature ‘Envelope Postfiltering’ of the PercepNet.
- Implemented Envelope Postfiltering by modifying the gain produced by the DNN:

$$\hat{g}_b^{(w)} = \hat{g}_b \sin\left(\frac{\pi}{2}\hat{g}_b\right)$$

- Compensate for the global gain:

$$G = \sqrt{\frac{(1+\beta)\,\frac{E_0}{E_1}}{1+\beta\left(\frac{E_0}{E_1}\right)^2}}$$

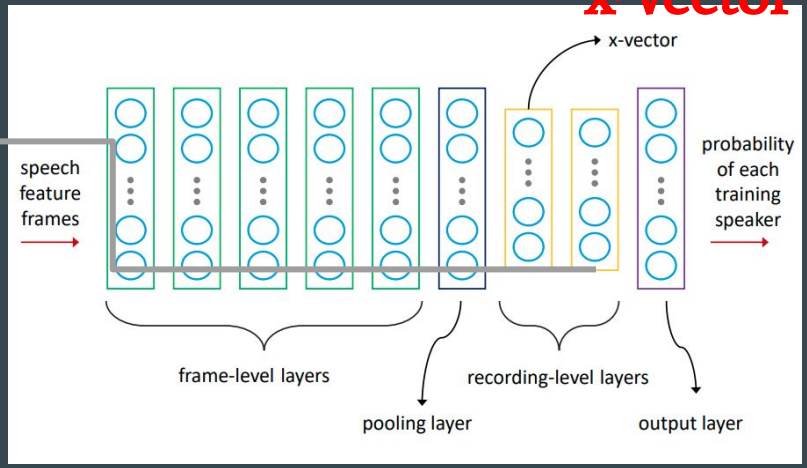
Our Progress – Personalized Layer Deep Speaker Embedding



d-vector

Deep Speaker Embedding is the extracted feature of a target speaker generated by the feature extracting model trained by the Deep Neural Network.

x-vector



Our Progress – Model Training with X-vector

Personalized PercepNet: d-vector

- Cosine scoring
- Required a large number of in-domain training speakers

Our Work: x-vector

- Improved from i-vector, probabilistic linear discriminant analysis (PLDA) scoring.
- Improve performance on smaller, publicly available datasets
- instead of training the system to separate same-speaker and different speaker pairs, the DNN learns to classify training speakers

Datasets

- Training PercepNet DNN

- Training Personalized Embedding DNN
  - Kaldi Speech Recognition Toolkit

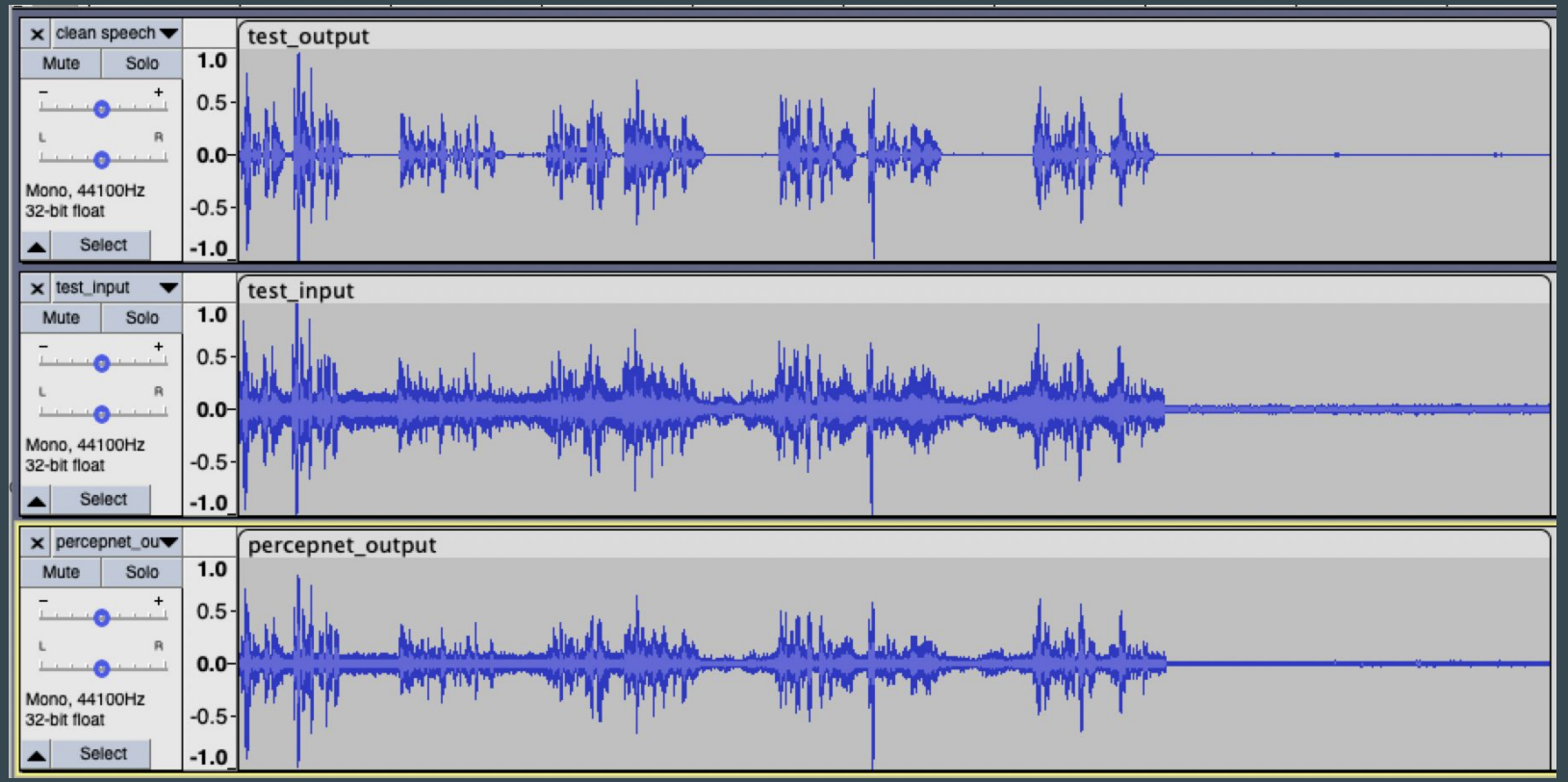
- Testing PercepNet and x-vector Personalized PercepNet with PESQ

Our Progress – Model Testing with PESQ Objective Testing

- PESQ stands for Perceptual Evaluation of Speech Quality
- Objective test for scoring audio quality
- The test compares an audio output to the original voice file
- higher scores indicating better quality

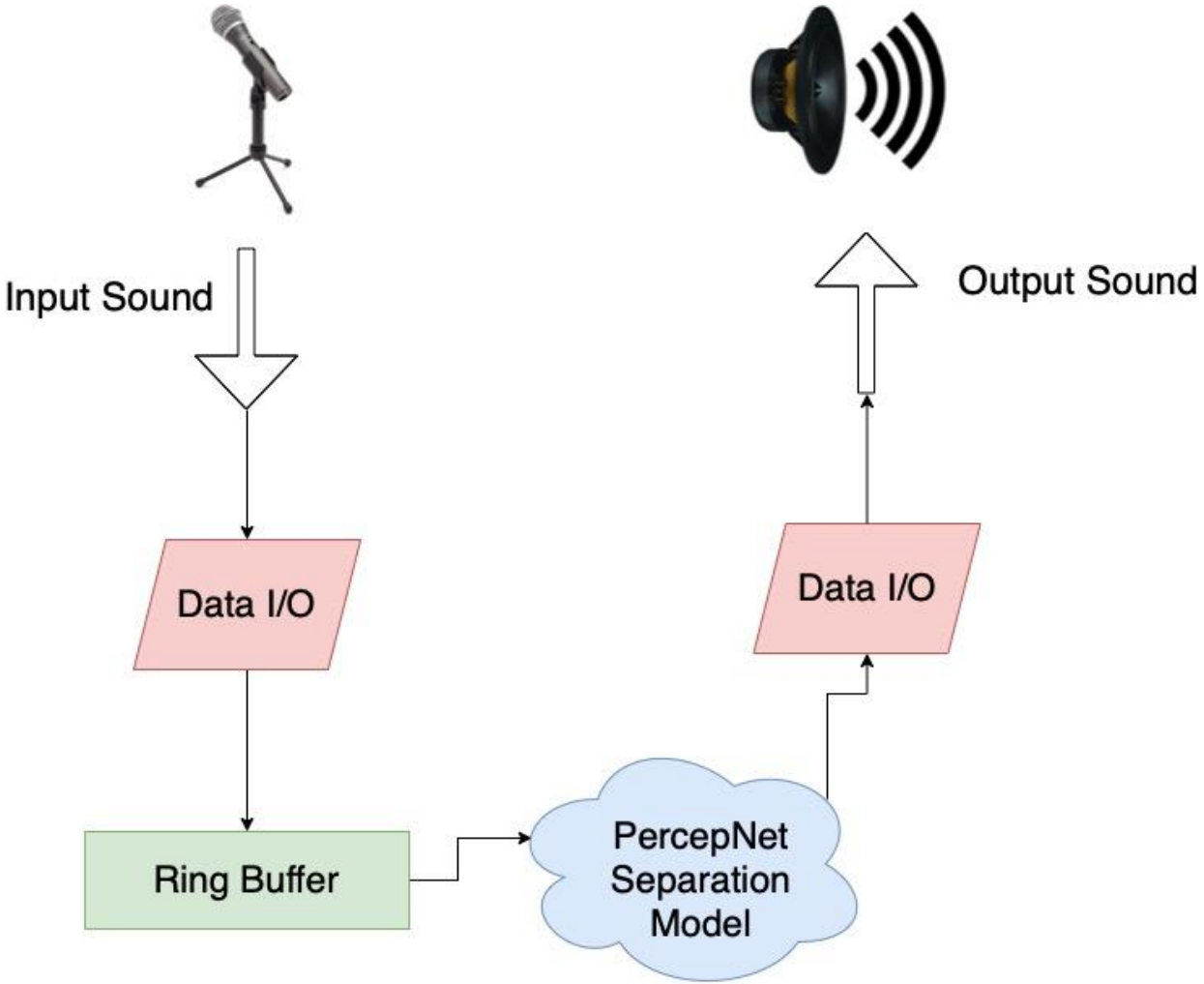
	PercepNet	x-vector Personalized PercepNet
PESQ-WB	1.55	2.36

Waveforms


clean speech
Mute
Solo
Mono, 44100Hz
32-bit float
Select
1.0
0.5
0.0
-0.5
-1.0
test_output
test_input
test_input
Mute
Solo
Mono, 44100Hz
32-bit float
Select
1.0
0.5
0.0
-0.5
-1.0
percepnet_output
Mute
Solo
Mono, 44100Hz
32-bit float
Select
1.0
0.5
0.0
-0.5
-1.0

Future Work

- Generate C++ implementation for the personalized layer of the model and add it to the current pipeline
- Test the model with personalized layer in real-time
- Reduce pipeline’s latency


Input Sound
Output Sound
Data I/O
Data I/O
Ring Buffer
PercepNet
Separation
Model

Existing Solution

- Traditional solution
  - Use multiple mics
    - The mic closer to the mouth will capture the sound energy of human voice
    - The mic further away from the mouth will capture the noise signal
    - Softwares subtract signals from each other, and generate almost clean signals
  - Difficulty
    - Multiple hardwares, hard to fit everything in a mobile device
    - Some situations cannot be handled
- DSP solution
  - Traditional DSP solutions try to find the pattern of noise signals and filter the signal frame by frame
  - Difficulty
    - Low variety, can only be used in a certain conditions, where noise signals have patterns

Separation Model

- PercepNet
  - 10ms frames
  - Ratio mask to separate signals
  - Envelope postfiltering
- Incorporate Personalized Layer to PercepNet
  - Produce a representative embedded vector
- Customize window size
  - A window size to be able to provide enough information
  - Do not cost too much latency
- Reduce the computational latency
  - Graph surgery
    - Rewire the network's components