Human Activity Analysis and Recognition using Machine Learning Techniques

機械学習を用いた人間行動解析および認識

Instructors: Prof. Jungpil SHIN

INTRODUCTION

The recent advancement in technology and sensors has enabled the detection of anomalies and the recognition of daily human activities for surveillance.

Human Activity Recognition (HAR) has been utilized in diverse application domains including healthcare, surveillance, sports and event analysis, elderly care, and **Human-Computer Interaction (HCI).**

In our seminar, we mainly challenged **the experiment of HAR** and **the development of HCI apps** using Deep/Machine Learning for Supporting people who have disabilities and security.

Finger Tapping Person Identification

Person identification (PI) is one of the most vital tasks for the purpose of security. In this experiment, I tried to identify a person from the skeleton and EMG data From **Finger Tapping (Hand Gesture Movement).**

Skeleton-Based Identification

We extracted the hand-skeleton coordinates using Mediapipe technology and the 1039 feature from coordinates to train the Machine Learning Model.

We got the **more than 99%** classification accuracy.

Experiments	Models	Accuracy	Precision	Recall	F1-score
Task.1 User-defined password	RF	**99.38**	**99.43**	**99.38**	**99.37**
	SVM	**99.38**	**99.43**	**99.38**	**99.37**
	MLP	**99.38**	**99.43**	**99.38**	**99.37**
Task.2 Same password	RF	98.74	98.86	98.75	98.75
	SVM	**99.37**	**99.43**	**99.38**	**99.37**
	MLP	98.74	98.86	98.75	98.71

EMG Sensor-Based Identification

Also, We analyzed the muscle activity that was recorded by EMG and extracted 640 features to train the Machine Learning Model.

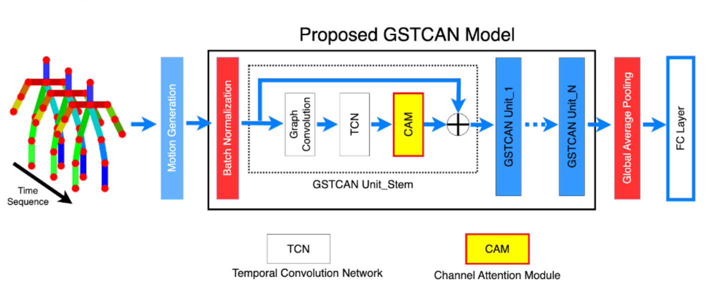
We got also **more than 98%** classification accuracy.

Experiments	Default SVM	Tuned SVM	Feature Selection + Tuned SVM
Task.1 User-defined password	92%	96.7%	**100.0%**
Task.2 Same password	90%	98.1%	**98.8%**

Fall Detection

Our goal is to develop an indoor fall detection system that will be subject and environment-independent.

We proposed a graph-based spatial-temporal convolutional and attention neural network (GSTCAN) with an Attention model.



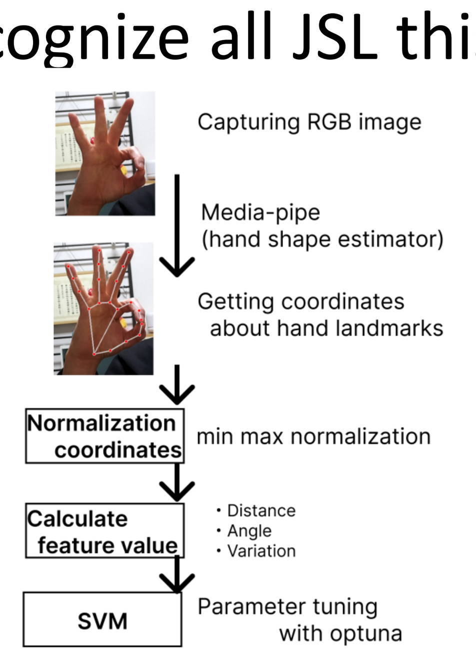
Algorithm	Dataset	Accuracy [%]	Precision [%]	Sensitivity [%]	Specificity [%]	F-Score [%]
Hontago [65]	ImViA	96.86	97.01	96.71	96.81	96.77
Wang [80]	ImViA	96.91	97.65	96.51	97.37	97.08
Chamle [81]	ImViA	79.31	79.41	83.47	73.07	81.39
Proposed GSTCAN Model	ImViA	99.93	99.57	99.67	n/a	99.92

Japanese Sign Language Recognition(JSL)

We focused on static JSL and achieved 97% accuracy last year. I added **dynamic** JSL and tried to recognize all JSL this year.

Algorithm

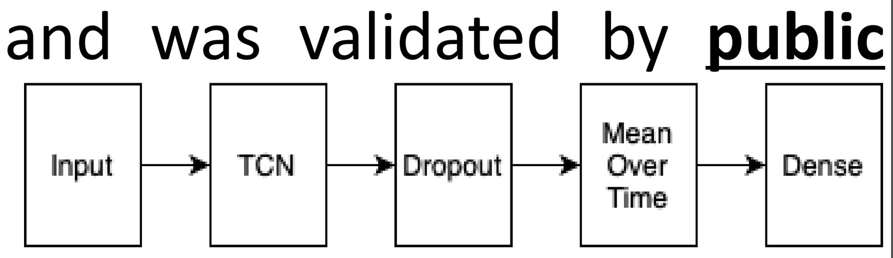
- Get hand coordinates via mediapipe
- Calculate features including movement
- Training the Model by feature
- App uses the model to detect sign



We got the **99.01 % accuracy**.

EMG Gesture Recognition

We tried EMG Gesture Recognition. We proposed the model that consisted of **Temporal Convolution Network** and **MeanOverTime layer**, and was validated by **public huge database** (ninapro).

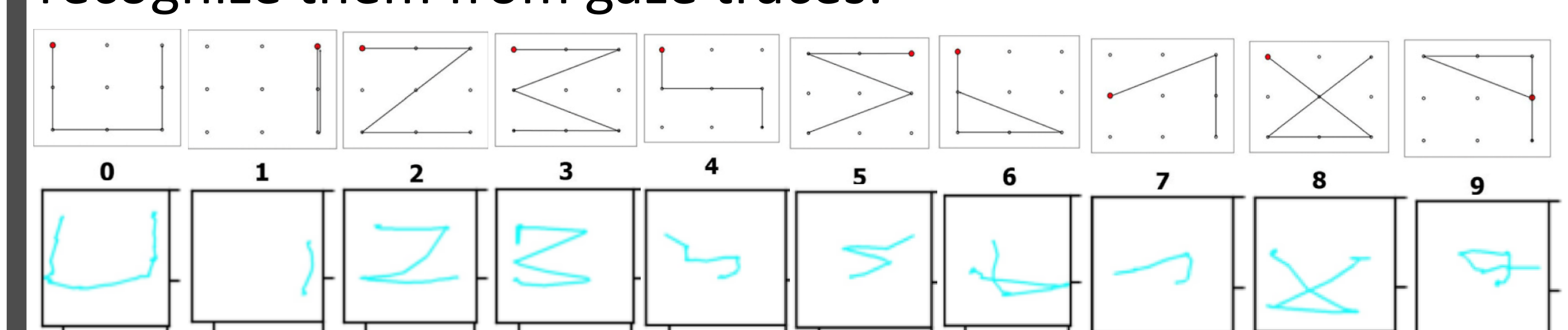


The proposed Model achieved **91.48% accuracy** and surpassed SoTA result.

Author	Year	Classes	Accuracy (%)
Weidong et al	2016	52	76.1
wei et al.	2019	52	88.2
Sidharth et al.	2021	53	91.1
Proposed	2023	53	**91.5**

Eye Writing

Eye writing is an HCI tool that translates eye movements into letters using automatic recognition by computers. In this experiment, we focused on the digits and tried to recognize them from gaze traces.



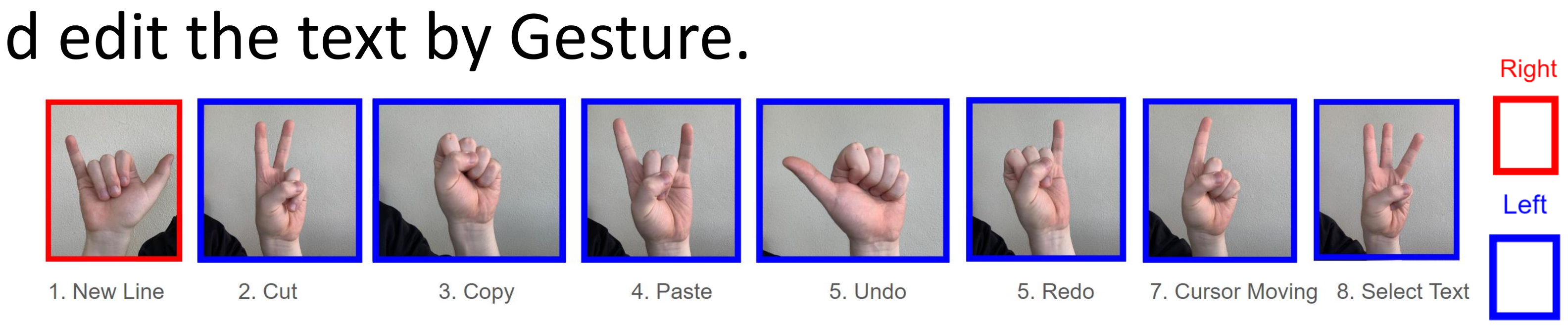
(User-independ case)	Epoch	Train acc	Val acc	Test acc
CNN(ResNet18)	24	99.67%	34.21%	48.75%
LSTM	43	94.21%	50.53%	65.75%
Ensemble				64.75%

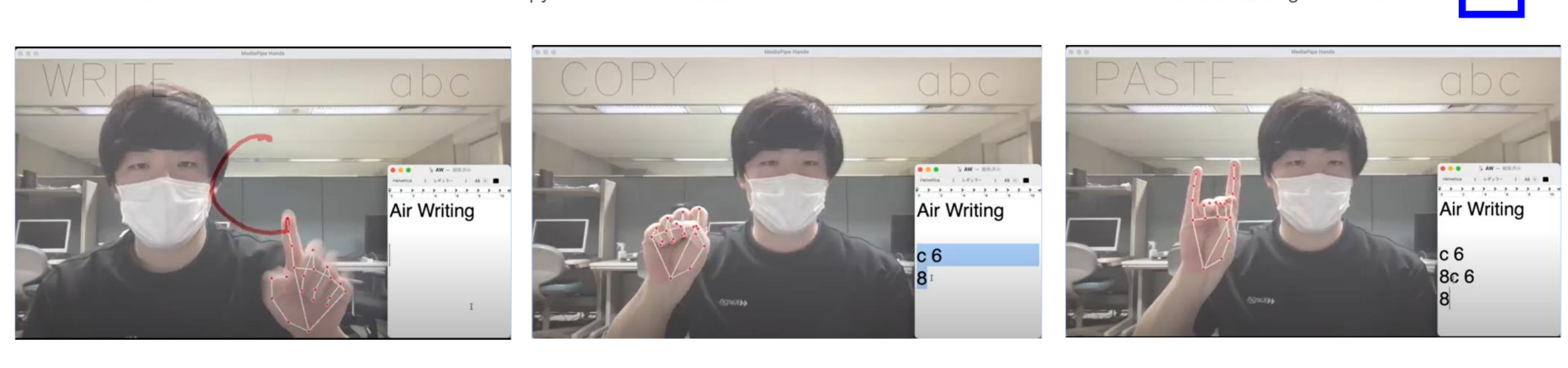
(User-depend case)	Epoch	Train acc	Val acc	Test acc
CNN(ResNet18)	33	99.67%	94.21%	91.00%
LSTM	42	96.53%	95.53%	95.50%
Ensemble				96.75%

Air-Writing and Gesture-based Interface

Last two years, We have made some **non-touch interface** apps such as the Sign-Language-based interface, Flick input-based interface, and Air-Writing interface.

We **merged those systems** and **added new text editor functions** such as copy, paste, undo and so on. This interface can input the letters and digits by Air-Writing, and edit the text by Gesture.





AY2023 CREATIVE FACTORY SEMINAR

Let's launch the environment of Integrated Visualization and Analysis for Lunar and Martian Rovers using JupyterLab, Python !

JupyterLab(Python)で月火星ローバーデータの統合表示・解析環境 を作ってみよう！

Instructors:
DEMURA H., OHTAKE M.,
OGAWA Y., YAMADA R., HONDA C.

Introduction

Rover Explorations of the Moon and Mars have been planed and done. To help plan a rover traverse and understand of exploration data, environment of Integrated visualization and analysis is needed. In this seminar, we focus on a rover Mars 2020 Perseverance in the vicinity of the delta in "Jezero Crater" in Fig. 1 as an example of various data types.

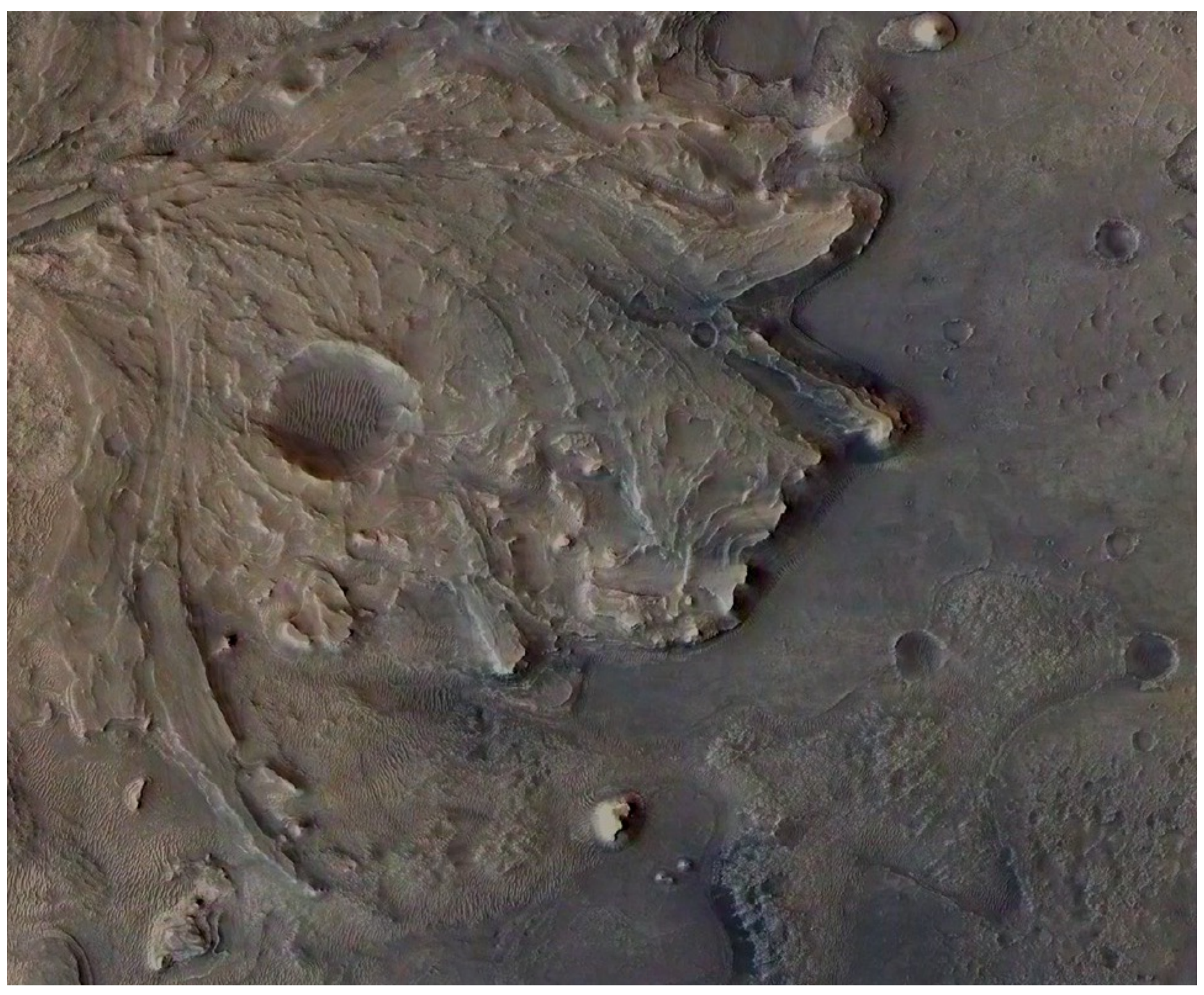


Fig. 1 Delta in the vicinity of Jezero Crater

Output

- Implement an environment in JupyterLab for visualization and analysis along the traverse of data from Mars 2020 Perseverance
- Demonstrate 3D display of the data

- This output is a synchronized display in 2D and 3D (Fig. 2)

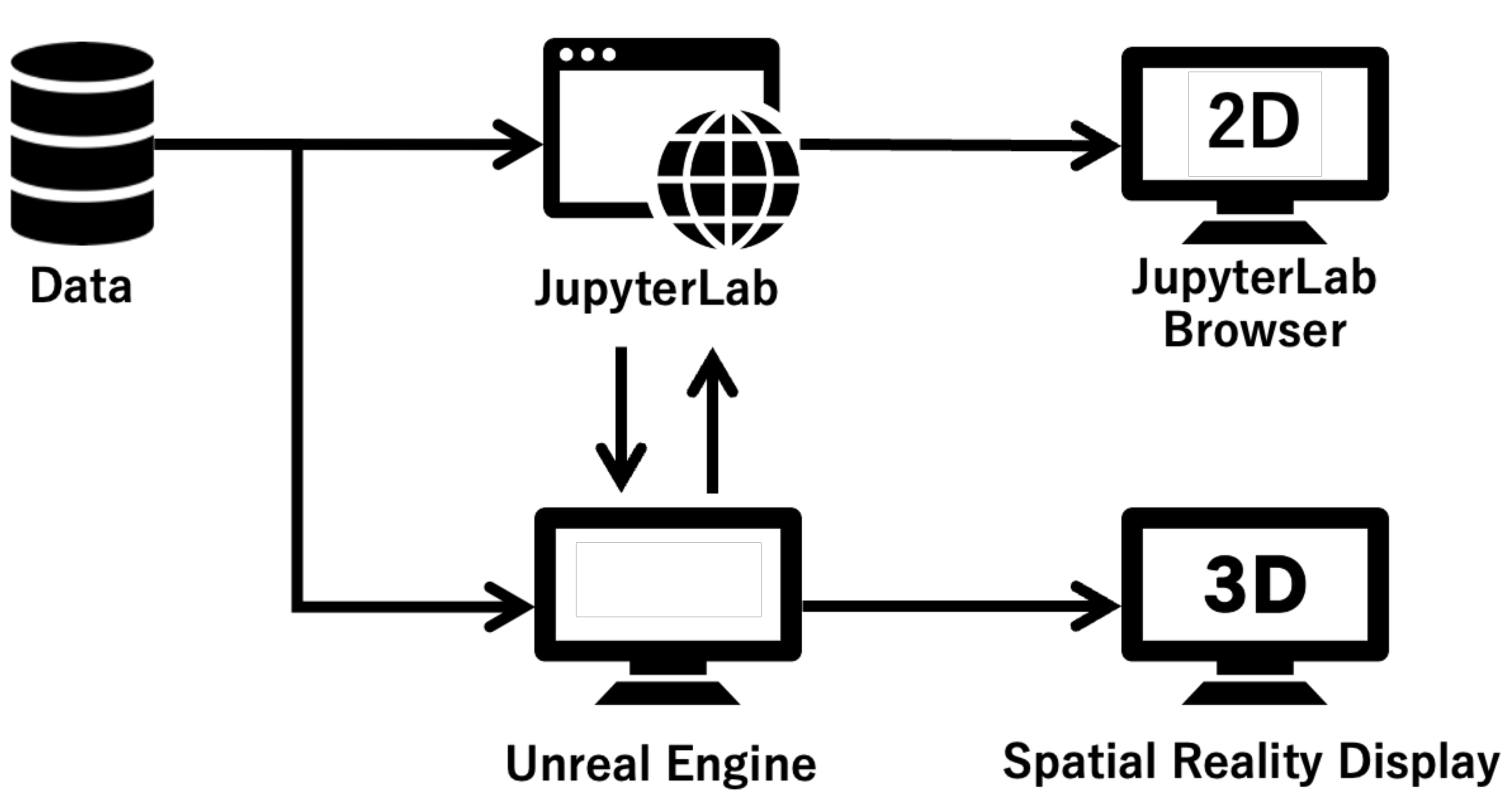


Fig. 2 Overview of this system

Results

Fig. 3 shows (a) bird's-eye view of delta in the vicinity of Jezero Crater and (b) a filed of view in Wildcat Ridge. Red line shows a traverse of Mars 2020 Perseverance Rover. The broken line in yellow indicates that the gap is too large to climb by rover. Real rover chooses another path.

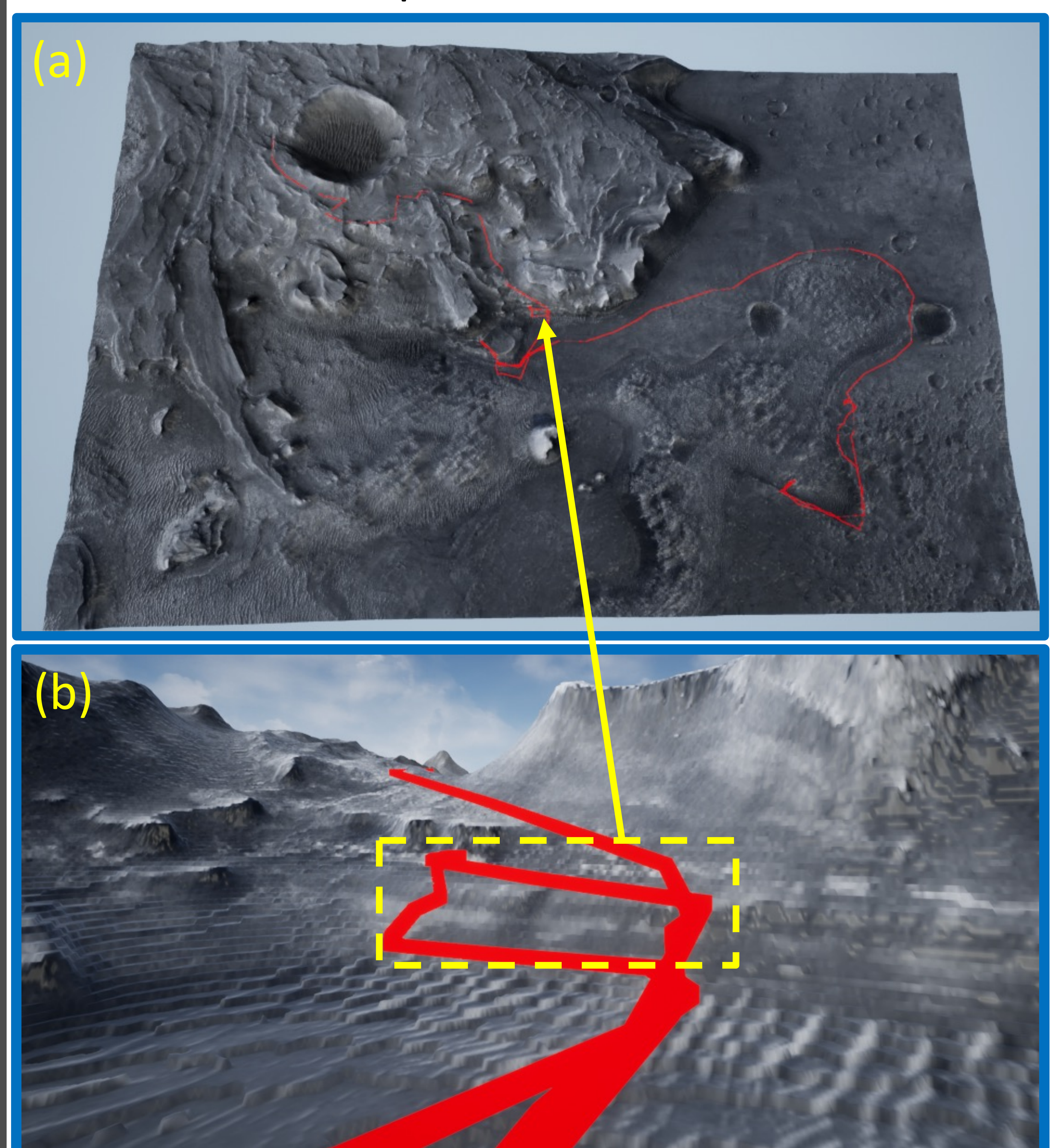


Fig. 3 Bird's view of rover traverse and rover view in "Wildcat Ridge"

Demonstration

Multimodality Medical Data Analysis for Predicting Neurological Recovery from Coma After Cardiac Arrest

DING Jiayi, TIAN Xingjian, ZHANG Xiaoyi, YANG Haoyi, XUE Bowen, LIANG Ziwei, HONG Yixiao

Instructors: ZHU,X.

THE UNIVERSITY OF AIZU

Background and Goal

More than 6 million cardiac arrests happen every year worldwide, with survival rates ranging from 1% to 10%. Severe brain injury is the most common cause of death for patients surviving initial resuscitation, and most survivors admitted to an intensive care unit (ICU) are comatose. The possibility that a patient will eventually recover consciousness needs to be assessed in the first few days after a cardiac arrest. Patients with a high probability will receive continued treatment, while others will normally be taken off life support and die. Therefore, it is important to improve the accuracy of the prognosis.

Clinical neurophysiologists have come to recognize numerous patterns of brain activity that help to predict prognosis following cardiac arrest .

Brain monitoring with electroencephalography aims to remove subjectivity in neurologic prognostication following cardiac arrest. Automated analysis of continuous EEG data has the potential to improve prognostic accuracy and to increase access to brain monitoring where experts are not readily available. International Cardiac Arrest RE search consortium (I-CARE) assembled a large representative set of EEG data and neurologic outcomes from comatose patients who underwent EEG monitoring following cardiac arrest.

The goal of this venture factory is to use the data of Physionet 2023 Challenge to use longitudinal EEG recordings for predicting good and poor patient outcomes after cardiac arrest.

Method

In this seminar, we applied single-modality and multi-modality to achieve the prediction task. The original data consists of two parts: EEG and Clinical data. We use EEG data to train CNN model and Clinical data to train BP network. Then we fuse the two models with weighted values and obtain a total of three results.

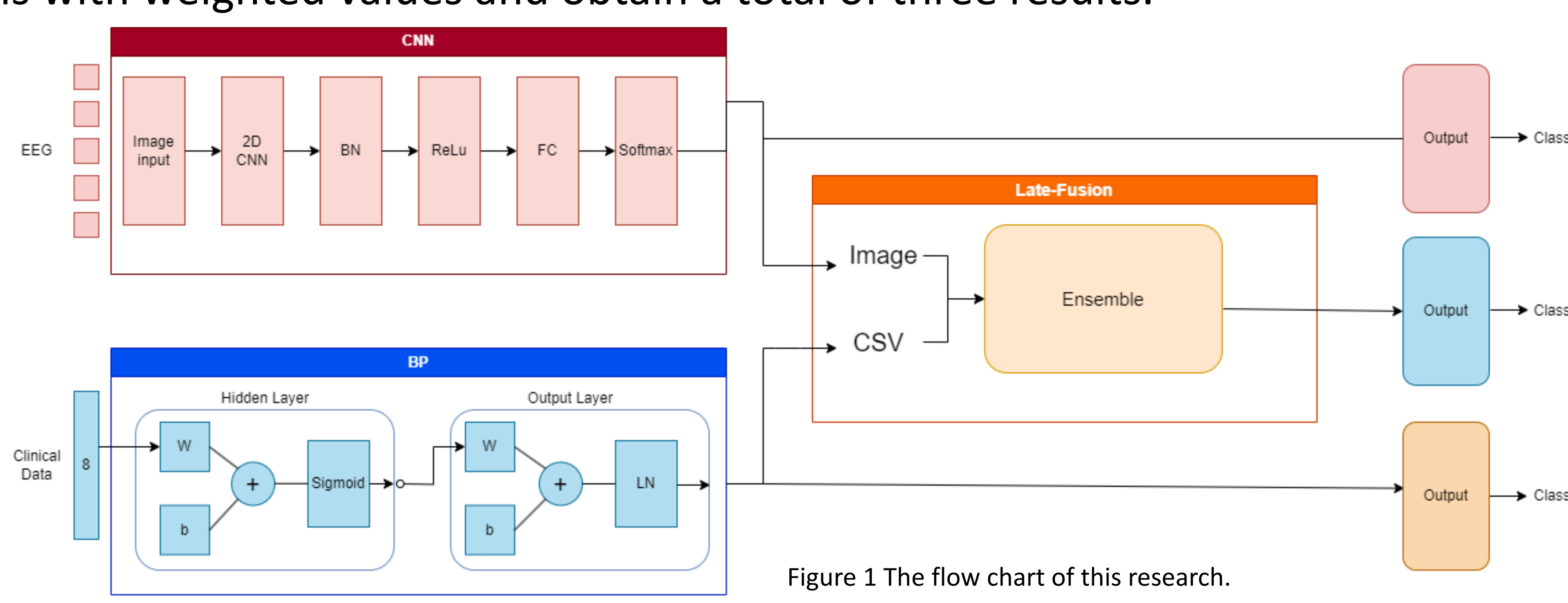


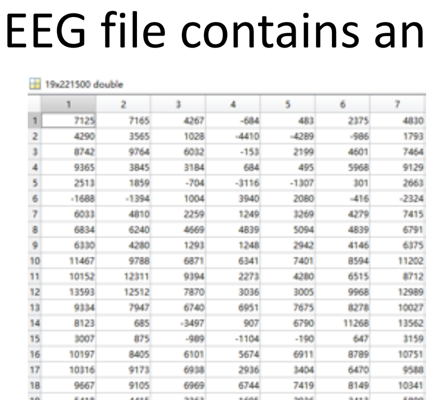
Figure 1 The flow chart of this research.

Data Processing

We use the EEG data and Textual data from I-CARE.

1 EEG Data

Each EEG file contains an array of signal, we covert it to plot.



⇨

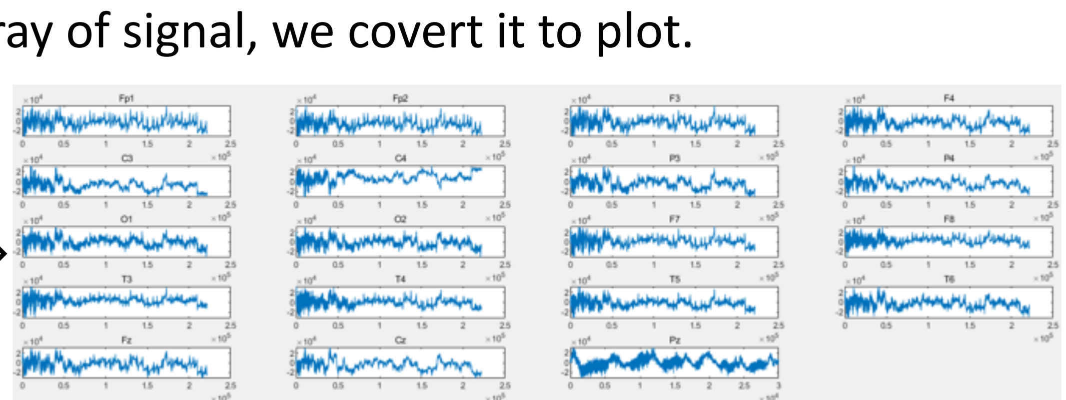


Figure 2 Raw data of EEG.

Figure 3 Preprocessed EEG data.

2 Clinical Data

Each patient has a clinical data including admission records, the time between cardiac arrest and ROSC and so on. We combine all patients' data to csv file.

```
Patient: 0284
Hospital: A
Age: 53
Sex: Male
ROSC: nan
OHCA: True
Shockable Rhythm: True
TTM: 33
Outcome: Good
CPC: 1
```

⇨

```
Patient,Hospital,Age,Sex,ROSC,OHCA,Shockable Rhythm,TTM,Outcome,CPC
0284, A, 53, Male, nan, True, True, 33, Good, 1
0286, F, 85, Female, 7, False, False, nan, Good, 1
0296, A, 48, Male, nan, True, True, 36, Good, 1
0303, D, 51, Male, 24, True, True, 33, Good, 1
0306, E, 73, Male, 20, True, False, 33, Good, 1
0312, D, 51, Male, nan, nan, True, nan, Good, 1
0313, B, 56, Male, 3, False, False, 33, Good, 1
0319, A, 62, Male, nan, False, True, 33, Good, 2
0320, F, 55, Male, 20, True, True, 33, Good, 1
0326, A, 46, Male, nan, True, True, 33, Good, 1
0328, B, 56, Male, 40, False, True, 33, Good, 1
```

Figure 4 Original Clinical data.

Figure 5 Preprocessed Clinical data.

Result

As indicated in the methods part, we use three different models. The following are the results we obtained:

Single modal (CNN)
Training with EEG data
- Accuracy:42.7%
- Precision: 0.47
- Recall:0.6
- F1-score:0.5

Single modal (BP)
Training with Clinical data
- Accuracy:60%
- Precision:0.6
- Recall:0.8
- F1-score:0.67

Multimodal
Training with EEG & Clinical
- Accuracy:60%
- Precision:0.625
- Recall:0.625
- F1-score:0.58

Figure 5 The results of this research.

Conclusion

In this seminar, we performed three kinds of models to finish the prediction task. But we didn't get a good result. We think there are several reasons.

1. The distribution of brain wave data is uneven, and the number of data in some results is too small to effectively obtain image features.
2. Directly using EEG waveforms is not conducive to CNN feature extraction. In the future, we should try to obtain frequency domain characteristics.
3. Late fusion cannot obtain the potential connection between medical image and clinical data, and cannot effectively fuse two kinds of data.

Object Recognition with Tactile Data Glove
データグローブを利用した物体認識

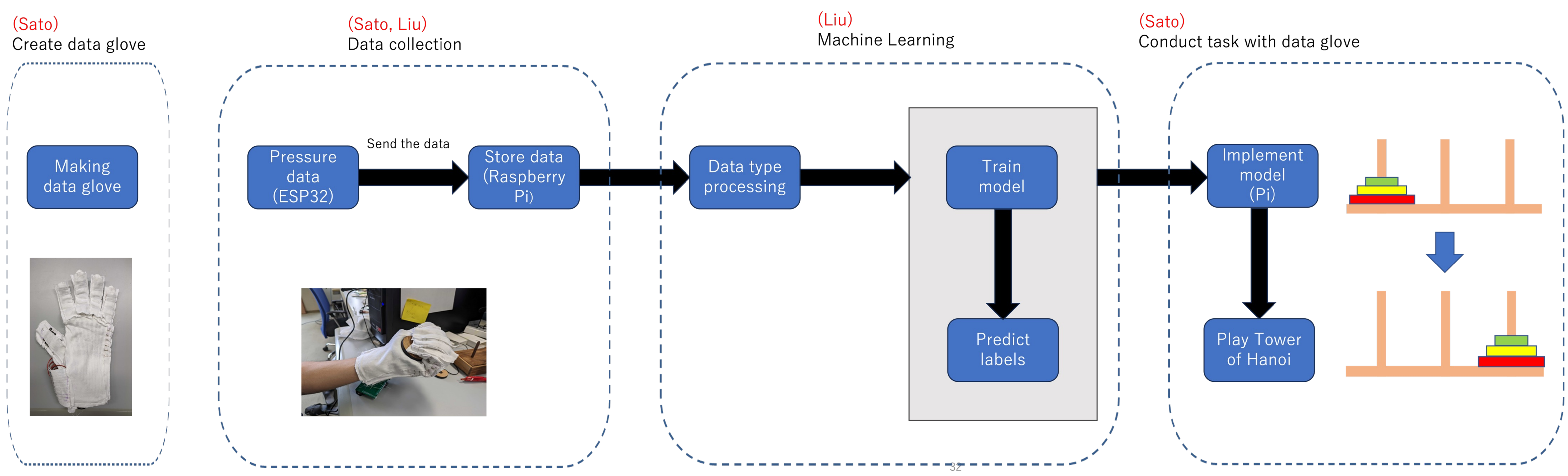
Instructors: Lei JING, Xiang LI, Suzuki Daisuke

SEMINAR INTRODUCTION

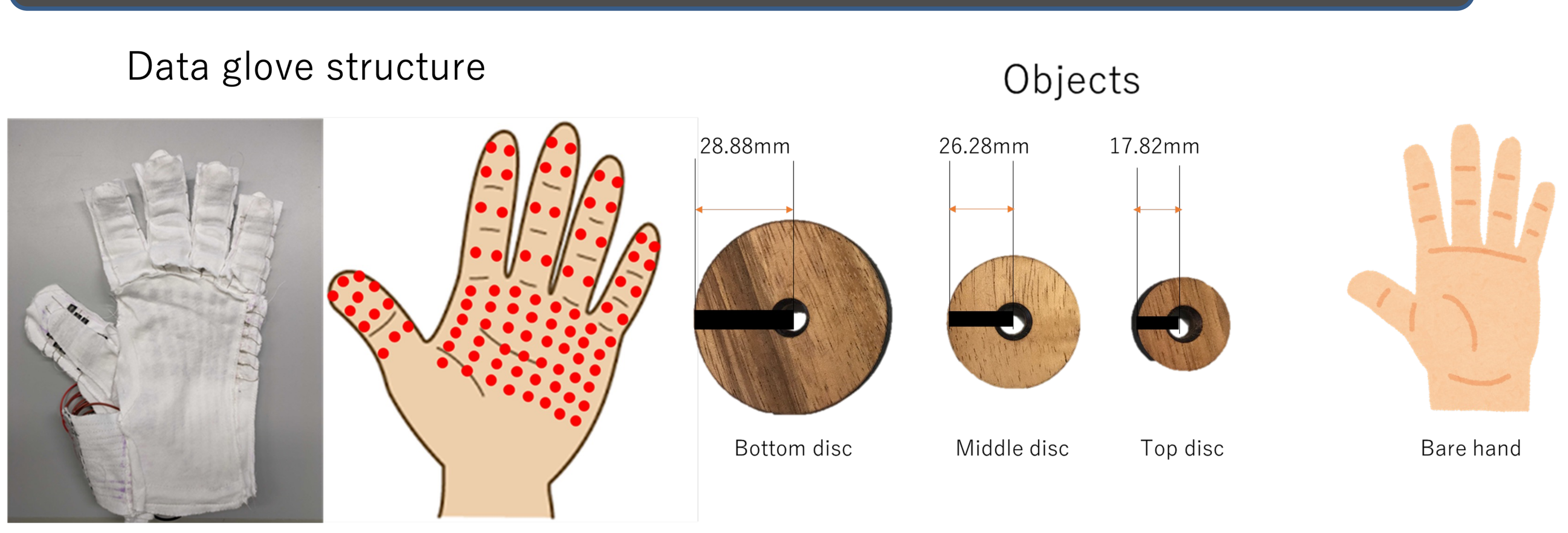
Data gloves means the gloves attached sensors around it such as pressure sensors to get the hand data directly. Using this technology, we can solve camera problems like a special environments and blind spot. In this seminar, we made a 92 tactile sensors data glove. A tactile sensor consisted of conductive thread and pressure-sensitive conductive sheet (velostat). Velostat has characteristics that as the pressure on the velostat increases, the resistance decreases. We collected the pressure data when grabbing a objects for two people with ESP32 Hanoi game is used to show the usability of object recognition function of the data glove. Then the data is processed and the data of 4 different cases, BARE, BOTTOM, MID and TOP, are trained using the MLP model and the 1DCNN model. By comparing the accuracy of the results after training, it can be clearly seen that 1DCNN has a higher rate of correctness.

After getting the label of the output result, then input the label into the code set for the rules of the game, after the corresponding calculation, it will give the choice of which disk should be moved in the next step, until the finalization of the task.

Project Outline and Task Assignment



Data Set



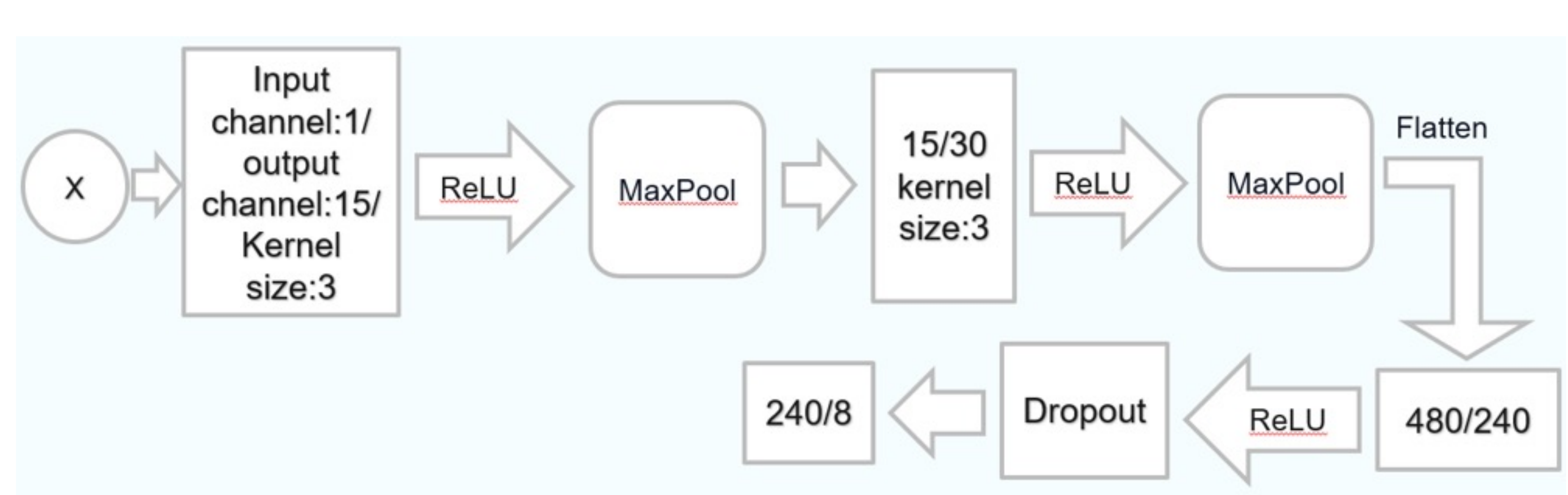
- All finger have 16 sensor each
- Palm of sensors have 48 sensors
- Obtain 12-bit ADC values
- Save to csv file

- Collected four different positions for each disks
- Training and testing data were collected separately.

Accuracy Results

Training with 1DCNN model (training data:1229 csv files, Test data :307csv files)	Training with MLP model (training data:1229 csv files, Test data :307csv files)
• Accuracy in training data: 99.92% • Test data : 99.52%	• Accuracy in training data: 98.7% • Test data : 96.26%

1DCNN Model



Conclusion

- We estimated tactile data glove system.
- We proposed object recognition model for tactile data glove.
- The accuracy of our model is over 90%.
- The accuracy of 1DCNN model is higher rate than MLP model.

AY2023 CREATIVE FACTORY SEMINAR

Methods, Tools, and Devices to Design and Produce 3D Objects: Haptic Modeling and Rapid Prototyping ("3D Printing")

立体形状デザインと造形のための手法と技術

Michael Cohen; 公園 マイケル & Yoshioka Rentaro; 吉岡 廉太郎

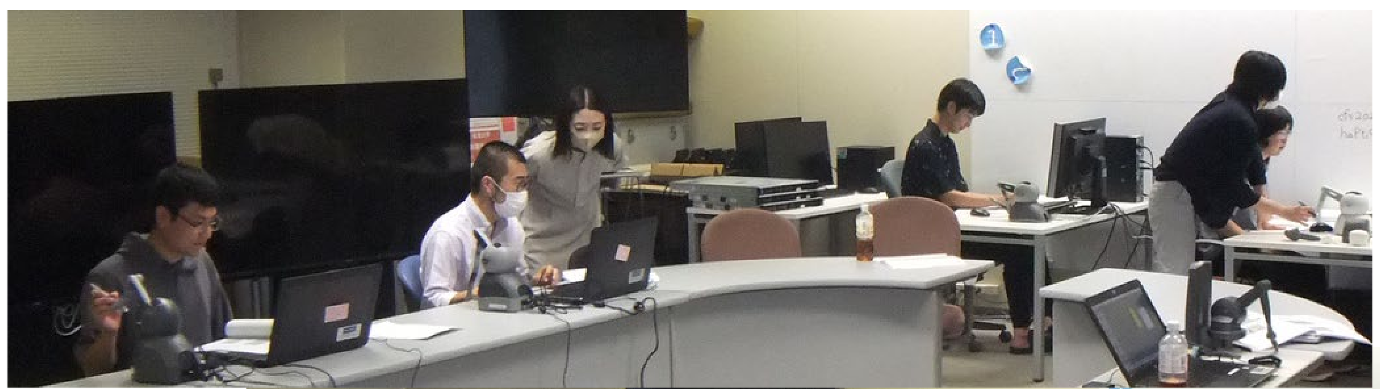
In this seminar, students use 3-dimensional haptic sculpting devices and CAD software for virtual objects, "Geomagic Touch" and "Freeform," to make virtual models by "carving" virtual clay displayed on a computer monitor and felt through a force-feedback stylus. Students then render objects based on their models via 3D printing. The seminar encourages student creativity, giving participants experience with unique interfaces that encourage organic artistic expression.

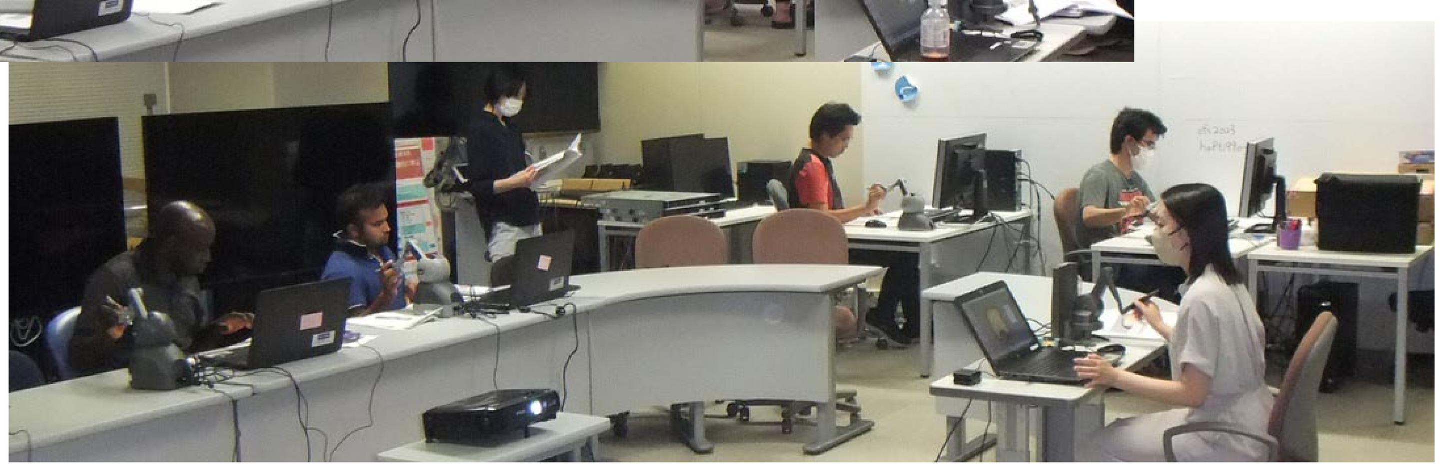
このセミナーでは、3次元触覚デバイスとCADソフト"Geomagic Touch"と"Freeform"を使って、コンピュータの画面上で実際に粘土を削る感覚でモデルを作成し、そのモデルを3Dプリンターに出力し、形にします。このセミナーでは、学生達の創作意欲を刺激する、特別なインターフェースを操作することで、最先端の技術に触れ、アイデアを形にする事が出来ます。

① 3D Modeling Introduction

3D Systems Geomagic® Touch™

2-day tutorial

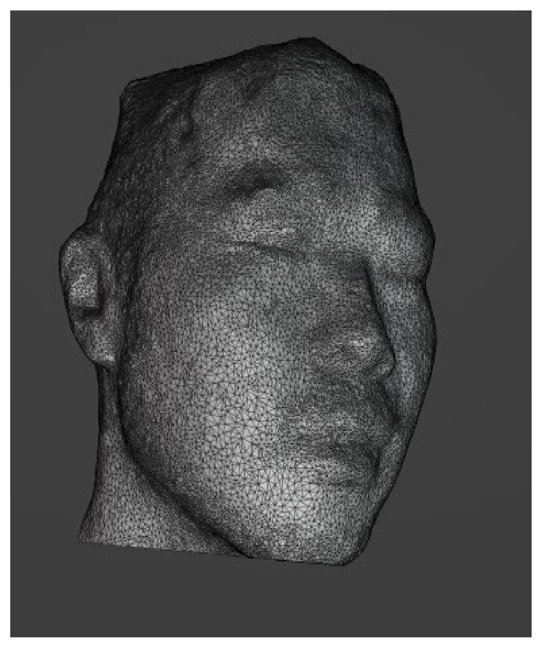




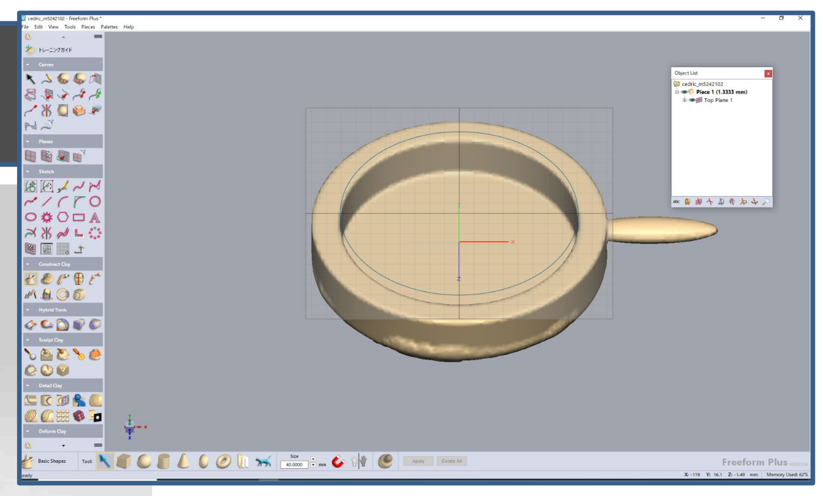
② Individual Modeling Sessions

All students make something.

Model Renderings



③ 3D Printing



3D Prints

AY2023 CREATIVE FACTORY SEMINAR

Performance Improvement of Applications Using FPGA Boards

Instructors: SAITO Hiroshi, KOHIRA Yukihide, TOMIOKA Yoichi

Objective

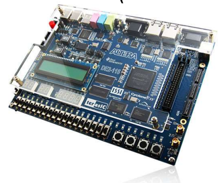
To accelerate an application using a field programmable gate array (FPGA) board

Through this CFS, students learn
1. How to model an application
2. How to use a design tool
3. How a synthesized circuit or a program code works on an FPGA board
4. Evaluation of the developed circuit or code

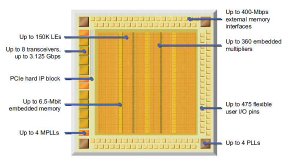
FPGA

A reconfigurable device
- Designers can change circuit structure freely
- Low design cost
- Well used in embedded systems

Terasic DE2-115 (source:terasic)



Cyclone IV FPGA (source:Intel)



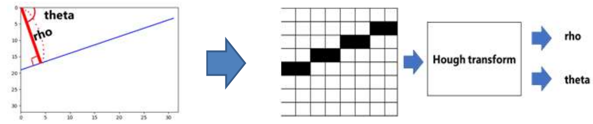
Acceleration of Hough Transform

Designer: Suzuki Yasuyuki (m5271038)

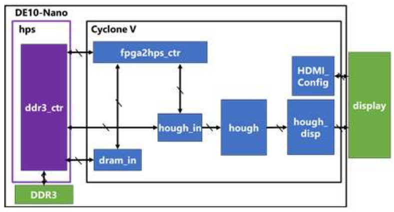
Motivation:
- Hough transform detects straight lines in given bainary images
- In autonomous driving, Hough transform is used to detect traffic lanes
  - Important to reduce execution time

Purpose: Acceleration of Hough transform on FPGA

Functional specification:



Architecture:
- Board: Terasic DE10-Nano
- FPGA: Intel Cyclone V



Results:

Software execution:706 μs
- C program
- Compiler option –O
- Intel(R) Xeon(R) Gold 6248R CPU @ 3.00GHz

Accelerator: 109 μs
- From gate-level simulation using ModelSim
- Clock cycle time is 15 ns
- Synthesized by using Quartus Prime 20.1 Lite edition

6.48 times faster

Acceleration of Haze Removal

Designer: HOSHI Masaya (m5261115)

Motivation:
- Haze Removal removes white haze from foggy images
- Dark channel used in Haze Removal is computationally expensive and takes a long time to process

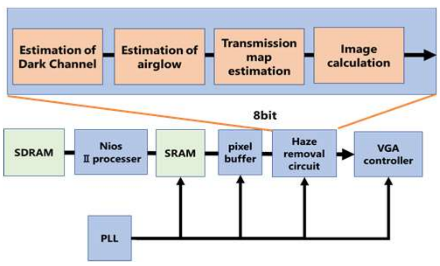
Purpose: Acceleration of Haze Removal on FPGA

Functional specification:

$$I(x,y) = J(x,y)t(x,y) + A(1 - t(x,y))$$

$$t(x,y) \approx 1 - \frac{DC(x,y)}{A} \qquad DC(x,y) = \min_{(u,v)\in\Omega(x,y)} \min_{c\in R,G,B} I^c(x,y)$$

Architecture:
- Board: Terasic DE2-115
- FPGA: Intel Cyclone IV E



Results:

Software execution:3.68 ms
- C program
- Compiler option –O3
- Intel(R) Core(TM) i7-8650U CPU @ 1.90GHz

Accelerator: 1.56 ms
- From gate-level simulation using ModelSim
- Clock cycle time is 10 ns
- Synthesized by using Quartus Prime 20.1 Lite edition

2.36 times faster

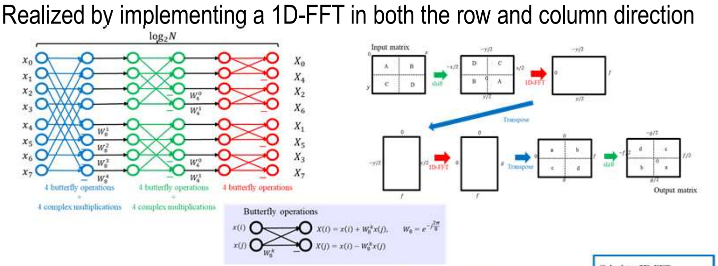
Acceleration of 2D FFT

Designer: KURAMOCHI Masaki (m5271017)

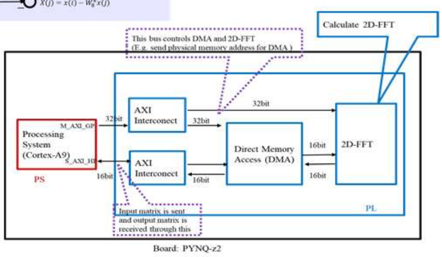
Motivation:
- High-speed 2D Fast Fourier Transform (FFT) is required for image analysis and signal processing that require real-time performance

Purpose: Acceleration of 2D FFT on FPGA

Functional specification:
Realized by implementing a 1D-FFT in both the row and column direction



Architecture:
- Board: PYNQ-z2
- FPGA: AMD Zynq



Results:

Software execution:450.0 μs
- C program
- Compiler option –O3
- ARM Cortex-A9 650 MHz Dual core

Accelerator: 22.4 μs
- Report from Vitis HLS 2022.2
- Clock cycle time is 10 ns

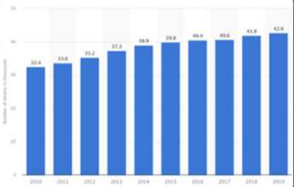
20.09 times faster

AY2023 CREATIVE FACTORY SEMINAR

Developing Spark In-Memory Big Data Analytical Framework to Find Spatiotemporal Trends in Japan's Air Pollution Database

Instructors: RAGE Uday Kiran

1. Motivation



- Estimated annual deaths due to air pollution is 7 million
- Average annual deaths in Japan due to air pollution is 40,000
- SORAMAME [1] sensor network was setup in Japan to monitor air pollution in Japan
- Useful patterns/information that can facilitate the environmentalists to improve human life lie hidden in the SORAMAME data.
- Our objective is to find the locations where people was recently exposed harmful revels pollutants.

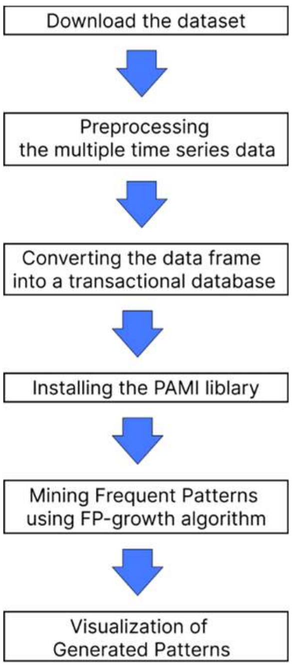
2. Challenge

- Data is very noisy
- High dimensional data containing 1800+ sensors
- Very large dataset covering 5+ Years

➢ Thus, we need to devise knowledge discovery techniques to deal with noisy, high-dimensional, and big data.

3. Approach

1. Download the data from [2], extract the pm2.5 measurement values, and combine them into one file.
2. The data contains negative values, abnormal values that are clearly measurement errors, and NaNs. These values prevent accurate pattern mapping, so remove them.Also, the data has a timestamp, but since it was originally sorted chronologically, I want to remove this property.
3. Installing the PAMI [3]. PAttern MIning (PAMI) is a Python library containing several algorithms to discover user interest-based patterns in a wide-spectrum of datasets across multiple computing platforms.
4. Convert the data to a transactional database for pattern mapping.
5. Perform pattern mapping using the Parallel FP-growth algorithm to obtain frequent patterns [4].
6. Based on the frequently occurring pattern data obtained earlier, we will visualize the patterns on the map by coloring them.



4. Experimental Setup

- **Execution Environment**:
  - free version of Google Colaboratory [5]
  - Intel Xeon CPU with 2 vCPUs
  - 13GB of RAM
- **Dataset:**
  - Time period: 5 years
  - Pollutant: PM2.5
  - Number of sensors: 1831 sensors

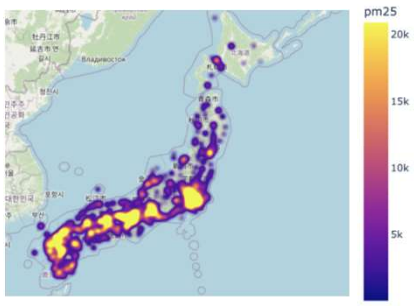
Please use the following QR-code to execute or improve our work

Figure 2: Sample air pollution dataset

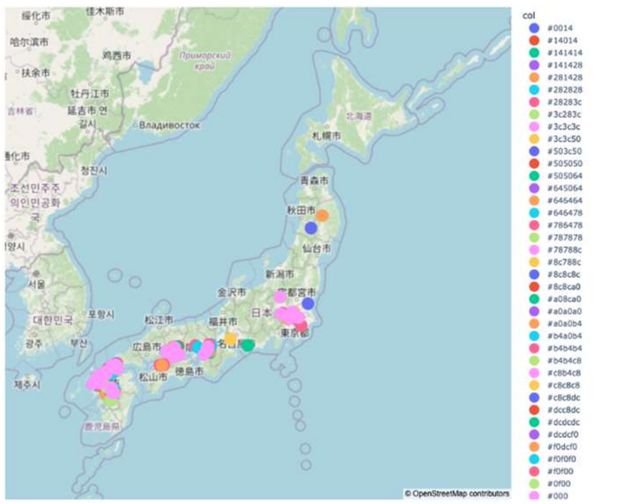
	TimeStamp	Point(139.0794379 36.3727776)	Point(139.1051411 36.3963822)	Point(139.0960211 36.4047323)	Point(139.0428727 36.3816035)	Point(138.9955116 36.33801589999999)	Point(139.342672 36.4105658)
0	2018-01-01 01:00:00	NaN	NaN	5.0	13.0	18.0	20.0
1	2018-01-01 02:00:00	NaN	NaN	11.0	12.0	22.0	15.0
2	2018-01-01 03:00:00	NaN	NaN	7.0	12.0	19.0	16.0
3	2018-01-01 04:00:00	NaN	NaN	5.0	11.0	16.0	11.0
4	2018-01-01 05:00:00	NaN	NaN	6.0	11.0	10.0	8.0
...	...	...	...	...	...	...	...
46000	2023-04-25 18:00:00	NaN	NaN	NaN	22.0	3.0	15.0
46001	2023-04-25 19:00:00	NaN	NaN	NaN	21.0	2.0	19.0
46002	2023-04-25 20:00:00	NaN	NaN	NaN	20.0	10.0	19.0
46003	2023-04-25 21:00:00	NaN	NaN	NaN	19.0	2.0	15.0
46004	2023-04-25 22:00:00	NaN	NaN	NaN	19.0	1.0	17.0

5. Result

- **Heat Map of PM2.5 values:**
  - Observation: High pollution was mostly observed in the places located south-east of Japan



- **Generated Pattern Map:**
  - The same pattern can be seen mostly in Tokyo, Osaka, Okayama, and northern Kyushu.



References

[1] 環境省大気汚染物質広域監視システム(そらまめくん)
https://soramame.env.go.jp/
[2] pm2.5 dataset
https://www.dropbox.com/s/wa8d1sujzlx56hh/ETL_DATA_new.csv
[3] PAMI GitHub
https://github.com/UdayLab/PAMI
[4] Parallel FP-Growth algorithm
https://github.com/UdayLab/PAMI/blob/main/notebooks/parallelFPGrowth.ipynb
[5] Google Colaboratory
https://colab.research.google.com/notebook