Vision and Sensor Based Human Activity Recognition using Machine Learning Approaches

Vision Computing Platform Cluster : Dr. Jungpil SHIN

The University of AIZU, JAPAN

E-mail: *jpshin@u-aizu.ac.jp*

Introduction

Ⅰ. 소개

- 1999년3월: 일본큐슈대학Graduate School of Information Science and Electrical Engineering에서 박사학위
- 1999년4월~현재: 일본아이즈 대학 교수로 재직, (Pattern Processing Lab)
- “2000 Outstanding Intellectuals of the 21th Century”, “2008 Man of The Year representing JAPAN” 수상 외 20회의 수상경력
- 350편이상의IEEE, ACM, SCI 저널과 국제학회 논문게제및 발표
- IEEE, ACM 등의 100회이상의 국제저널, 국제학회 책임, 운영위원 및 심사위원
- IEEE Senior Member
- 아이즈대학 첨단 연구소(CAIST) 비젼 클러스터 리더 역임중

Introduction

Awards (selected)

1. Awarded a **Best Paper** Award from 3rd Eurasian Conference on Educational Innovation 2020 (IEEE ECEI 2020), M200134, Feb 5-7, 2020, Hanoi, Vietnam.
2. Awarded a **Best Paper** Award from 2nd IEEE International Conference on Knowledge Innovation and Invention (IEEE ICKII 2019), July 13-16, 2019, Seoul Korea.
3. Awarded a **Best Paper** Award from The 3rd ACM Conference on Applications in Information Technology (ICAIT-2018), Nov. 1-3, 2018, Aizuwakamatsu, Japan.
4. Awarded a **Best Paper** Award from The 8th International Conference on Awareness Science and Technology (iCAST 2017), pp. 93 - 97, Nov. 8-10, 2017, The Splendor Hotel, Taichung, Taiwan.
5. Awarded a **Professional Invitation** from Marquis "2006-2007 (6th), 2009-2010 (7th) and 2011-2012 (8th) Edition of **Who's Who in Medicine and Healthcare**®" 2006-2008. Listed in Marquis "2006-2007 (6th), 2009-2010 (7th) and 2011-2012 (8th) Who's Who in Medicine and Healthcare".
6. Awarded a **Professional Invitation** from Marquis "**Who's Who in the World**" Mar. 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2018, 2019. Listed in Marquis "Who's Who in the World (2006, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2018, 2019)".
7. Awarded an **Aizu-IT Encouragement Prize** of from Aizu-Wakamatsu City by Aizu-Wakamatsu City president http://www.u-aizu.ac.jp/events/aizuit2009.html Jan 20, 2010
8. Awarded a **Best Paper** Award from *The 2nd International Conference on the Frontiers of Information Technology, Application and Tools* (FITAT2009) Sep.2009.
9. Awarded a **Best Presentation** Award from *The 19th Intelligent System Symposium* (FAN09) Sep.2009.
10. Awarded a Professional Invitation from American Biographical Institute "**2008 Man of The Year representing JAPAN**" Nov. 2008.
11. Awarded "**International Educator of the Year 2006**" from International Biographical Centre, Cambridge, Great Britain, in the mid 2006.
12. Awarded "**2000 Outstanding Intellectuals of the 21th Century**" from International Biographical Centre, Cambridge, Great Britain, in the mid 2006.
13. Awarded a **Best Paper** from 2004 Conference on Chinese Language Computing.
14. Awarded a **Encouragement Prize** from **IEEE Society** Nov. 2004. T. Tamura and Jungpil Shin, "Moving Object Extraction Using Frame Difference,'' Proc. Joint Conf. of Electrical and Electronics Eng.
15. Awarded a **Promotion Prize** from Information Processing Society in Kyushu, Japan, in May 1996.

Editor and Reviewer of International Journal, Reviewer of Research Fund (selected lists.)

1. Guest Editor of The Journal Sensors MDPI (ISSN 1424-8220, IF 3.275), Special Issue entitled "Vision and Sensor-Based Sensing in Human Action Recognition", deadline: Aug. 2021. https://www.mdpi.com/journal/sensors/special_issues/Human_Action
2. Vice Editors-in-Chief of KIPS Transactions on Computer and Communication Systems (KTCCS, ISSN: 2287-5891 / eISSN: 2734-049X)
3. Editors of KIPS Transactions on Software and Data Engineering (KTSDE, ISSN: 2287-5905 / eISSN: 2734-0503)
4. Associate Editors-in-Chief of International Journal of Digital Content Technology and its Application (JDCTA, ISSN: 1975-9339).
5. Associate Editors-in-Chief of International Journal of Convergence Information Technology (JCIT, ISSN: 1975-9320)
6. Associate Editors-in-Chief of Journal of Next Generation Information Technology (JNIT, ISSN: 2233-9388)
7. Editor of International Journal of Intelligent Information Processing (ISSN: 2093-5765).
8. Editor of International Journal of Robots, Education and Art (IJREA, ISSN: 2233-4572)
9. Editor of International Journal of Intelligent Information Processing (IJIIP, ISSN: 2093-1964)
10. Editor of International Journal of Emerging Multidisciplinary Research (ISSN: 2546-1583)
11. Editor of Daehan Society of Industrial Management, Korea
12. Editor of Engineering Science & Technology (Universal Wiser), Singapore
13. Reviewer of JSPS Kakenhi Fund
14. Reviewer of NRF (National Research Fund of Korea)
15. Reviewer of Multimedia Tools and Application (Springer)
16. Reviewer of IEEE Access
17. Reviewer of IEEE Transactions on Cybernetics
18. Reviewer of IEEE Transactions on Systems, Man, and Cybernetics

Chairman on International Conference and Workshop (selected lists.)

1. The 2005 Conference of the Japanese Society for Planetary Science (JSPS)
   Finance Chair (held in University of Aizu, Aizu-Wakamatsu City, Japan Sep. 19 - 23, 2005)
2. The IEEE 6th International Conference on Computer and Information Technology (IEEE CIT2006)
   Publicity Chair and Member of the Program Committee (held in Korea University, Seoul, Korea, Sep. 20-22, 2006)
3. The IEEE International Conf. on Computing Technology and Infor. Management (IEEE ICCM, NCM and ICNIT) Editor (held in Seoul, Korea, Apr. 24-26, 2012.)
4. The 2012 International Conference on Digital Contents and Applications (DCA 2012)
   Editorial Committee (held in Kangwondo, Korea, Dec. 16-19, 2012)
5. The IEEE 5th International Conference on Awareness Science and Technology (IEEE iCAST 2013)
   Publicity Chair, (held in Aizu-Wakamatsu, Japan, Nov. 2 - 4, 2013)
6. 18th International Symposium on Advanced Intelligent Systems (ISIS2017)
   Program Committee Chair (held in Daegu, Korea, Oct. 11-14, 2017)
7. ACM 2017 Research in Adaptive and Convergent Systems (ACM RACS 2017),
   International Advisory Chair (held in Krakow, Poland, Sep. 20-23, 2017)
8. The Convergence Society of Small and Medium Business (SMBs)
   An executive director in 2010-current.

Introduction

Ⅱ. 전문분야 및 국제공동연구실적

1. 전문분야, 현 연구/관심 분야:정보통신(소프트웨어)

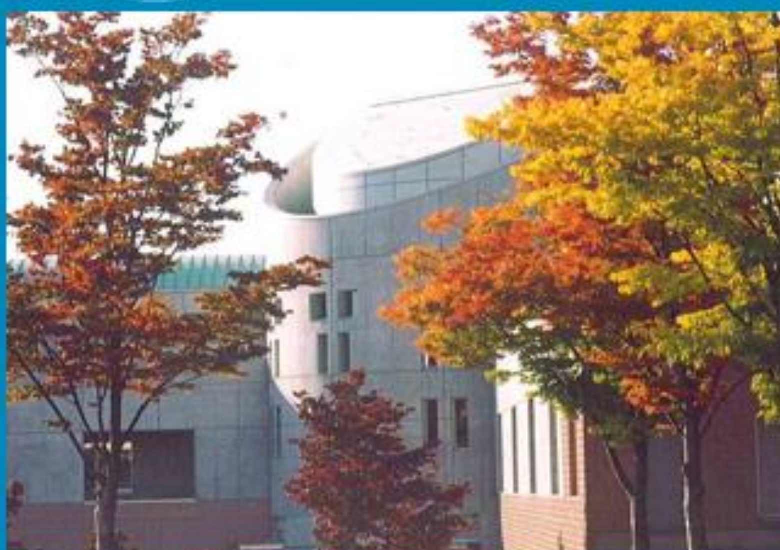
패턴인식, 기계학습, 인공지능, 신경장애인식, 발달장애(ADHD), 의료신호처리, 문자인식, 서명인증, 행동패턴인식, 제스쳐인식, 대화형인터페이스, 인간컴퓨터 상호작용.

2. 국제공동연구:

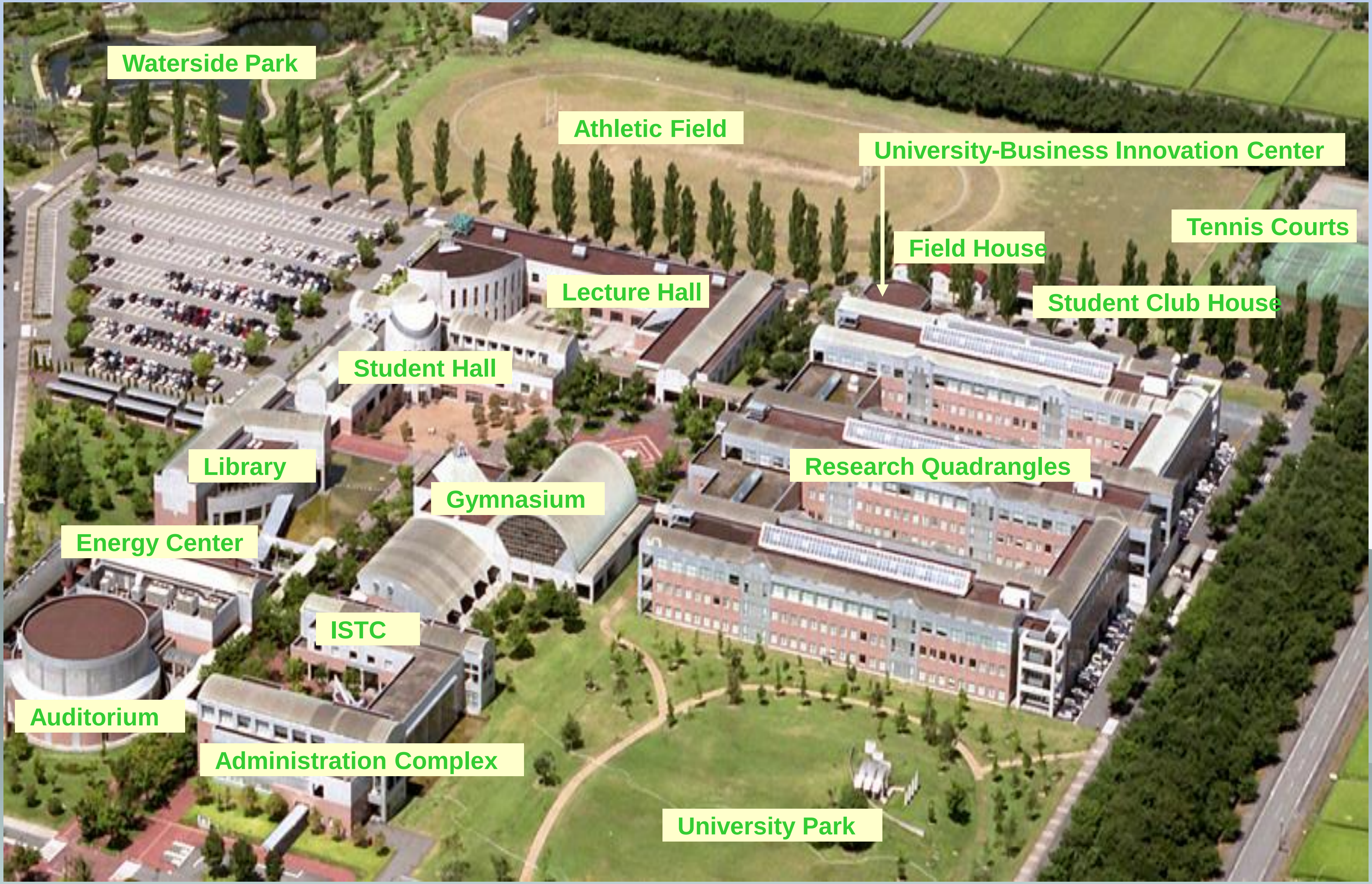
1. June, 2010- May, 2012: I have achieved the cooperative research among (1) University of Seoul in Korea, (2) University of Aizu, and (3) Korean Company (SQUARENET Inc.) – International University-Industry Cooperation. (Title: Laser based Intelligent Input Device for Interactive Screen). – about 110,000,000 Won / 2 Years (totally about 400M Won)
2. July 2006 - June. 2008: Japan-Korea Joint Research Program from JSPS.
3. 2021-Current : Cooperative Research with WACOM Co. (Title: **Developing Engine of Character Recognition System for IoT Paper**.)

The University of AIZU, JAPAN

The University of Aizu




Facilities (Ground: Approx. 200,000㎡, Buildings: Approx.47,000㎡)



Pattern Processing Lab

Theme: Human Computer Interaction, Non-touch Interface, Parkinson's Disease Diagnosis, ADHD Diagnosis, Human Gesture Recognition, Handwriting Analysis, Recognition and Synthesis, User Authentication, and so on.

Dr. Jungpil Shin
Pattern Processing Lab.
The University of AIZU, JAPAN
E-mail: *jpshin@u-aizu.ac.jp*

福島民報社　2021年10月03日(日)

会津大 研究室を歩く 41

慎重弼研究室（パターン処理学）

手の動きを自動認識

手書きからパーキンソン病の症状の有無を自動検出する実験

慎重弼（シン・ジュンピル）研究室は、人間の手の動きをコンピューターが自動認識することで、医療や本人認証、文字入力など多岐わたる分野に応用できる仕組みづくりに取り組んでいる。

医療分野のシステムは、手書きなどからパーキンソン病の症状の有無を高確率で推測する。症状特有の筆跡や筆圧、文字の震えなどを分析し、人工知能（AI）が十秒間ほどで判断する。また、文字の乱れや枠内のはみ出しなどから子どもの注意欠陥多動性障害（ADHD）を自動検出するシステムもある。いずれも医療現場に対し、客観的な情報を示せる。

研究室は手や関節の動きから動的な生体認証の研究も進めている。指紋認証など静的な認証と組み合わせることでより安全性の高い本人認証が可能になる。

このほか、手のジェスチャーを三次元で識別して、文字入力できるシステムを開発した。キーボードを用いず、非接触の操作を実現できる。慎教授は「人の動きを認識するシステムを開発することで、コロナ禍の新たな生活様式などに対応できる」と話した。

慎　重弼教授（パターン処理学講座）54

身近な動作である手書きや手の動きから、パーキンソン病や注意欠陥多動性障害（ADHD）、自閉スペクトラム障害（ASD）などを自動検出する研究をしています。ジェスチャーによる非接触型インターフェースの実現、IoTペーパーでの手書き文字の書き順と自動成績採点の研究も進めています。

春日　勇太さん（コンピュータ理工学研究科博士前期課程）23

授業で研究室の説明を聞き、そこで慎先生と仲良くさせていただいたのがきっかけです。研究では病気の診断をテーマに扱っており、身近でかつインパクトのある課題に取り組めている点が魅力です。今後は病気の診断など、現在は定性的に行われていることをデータを用いて定量的に診断できるようにしたいです。

キャンパス通信

▼会津IT秋フォーラム開催　29日、会津から県内外に最先端のICTに関する情報を発信、地域産業活性化に寄与することを目的とした「会津IT秋フォーラム2021」が開催されます。詳しくはIT秋フォーラムWebページをご覧ください。

Easy input without touching

We developed a system of non-touch character input. It can be done by simply holding up your hand. It is also helping to prevent Corona infection.

Published in the newspaper, Fukushima Minyu, at Jan/5/2022

2022年
1月5日

第42178号（日刊）　昭和21年3月15日第三種郵便物認可
福島民友
4−29　代表電話 024・523・1191　www.minyu−net.com

元気を出そう

手をかざすだけ…感染予防にも一助

触らず簡単文字入力

会津大・慎教授 システム開発

会津大は、空間に映し出された文字に手をかざすことで、文字を入力できる非接触型のシステムを開発した。特別な練習をしなくても、スマートフォンの文字入力と同じ感覚で操作できるのが特長。市役所や図書館など多くの人が利用する場所に導入することで、新型コロナウイルスなどの感染症対策に役立てることができるとしている。

開発したのは、パターン処理学を専門とする慎重弼教授(54)。2019年に論文を発表した。県外の病院で共用パソコンのキーボードを介して新型コロナウイルスのクラスター（感染者集団）が発生した例も報道されていることから、慎教授は「医療現場や食品会社など厳しい衛生対策が求められる職場での活用も期待できる」と話している。

上下左右に加え、奥行きを認識できる複眼カメラを使用。カメラが接続された画面の前に立つと、肩口にスマホの文字入力の画面のように「あかさたなー」が並んだ文字盤が表示される。手をパーの形で動かし、グーにして行を決定。「あいうえお」などから文字を選ぶ画面に移り、入力したい文字の方向に手を動かして入力を完了する。平仮名だけでなく、アルファベットや数字の入力も可能。手を前後に動かすことで文字盤の大きさを調節できるのも特長の一つだ。

慎教授は、カメラの映像から手の付け根と指先を正確に認識できるアルゴリズム（計算手法）を確立し、システムを完成させた。実験では98・6％の精度で文字を入力できたという。

当初は指1本を立てると「あ行」、2本だと「か行」といったように、指の本数で「あかさたなー」を決めるシステムの開発を進めていたが、手の動きだけで直感的に操作できるシステムに改良したという。慎教授は「元々はSF映画の世界に憧れて非接触型のシステムの開発を始めた。新型コロナ禍で非接触型の需要が高まっているので実用化を目指したい」と意気込む。

（若松支社・高崎慎也）

❶カメラが接続された画面の前に立つと、肩口に文字盤が表示される

❷手をパーの形で動かし、グーにして「あかさたなー」から行を決定

❸「かきくけこ」などから文字を選ぶ画面に移り、入力したい文字の方向に手を動かして入力を完了する

慎重弼教授

Aizu University Activity: **Invited Lecture in Aizu City**

福島民報　3月5日　2010年

コンピューターの可能性講演
ペン活用技術紹介
会津産学懇話会が定例会

会津産学懇話会（田苗博会長）の三月定例会は三日、会津若松市の会津若松ワシントンホテルで開かれ、会津大コンピュータ理工学部の慎重弼上級准教授が「ペンとコンピュータ」をテーマに講演した。

慎准教授は普段使っているペンの感覚を基に、コンピューターでの署名検証や文字フォント生成、漢字学習、書道学習システムなどを研究しており、最新の成果について分かりやすく紹介した。

くせのある自分の文字をいかしながらきれいな文字にしていく技術や、書道の書体だけでなく墨の濃淡やにじみ加減まで調整できる技術も発表した。出席者は慎准教授の興味深い話に聞き入り、コンピューターの新たな可能性を感じていた。

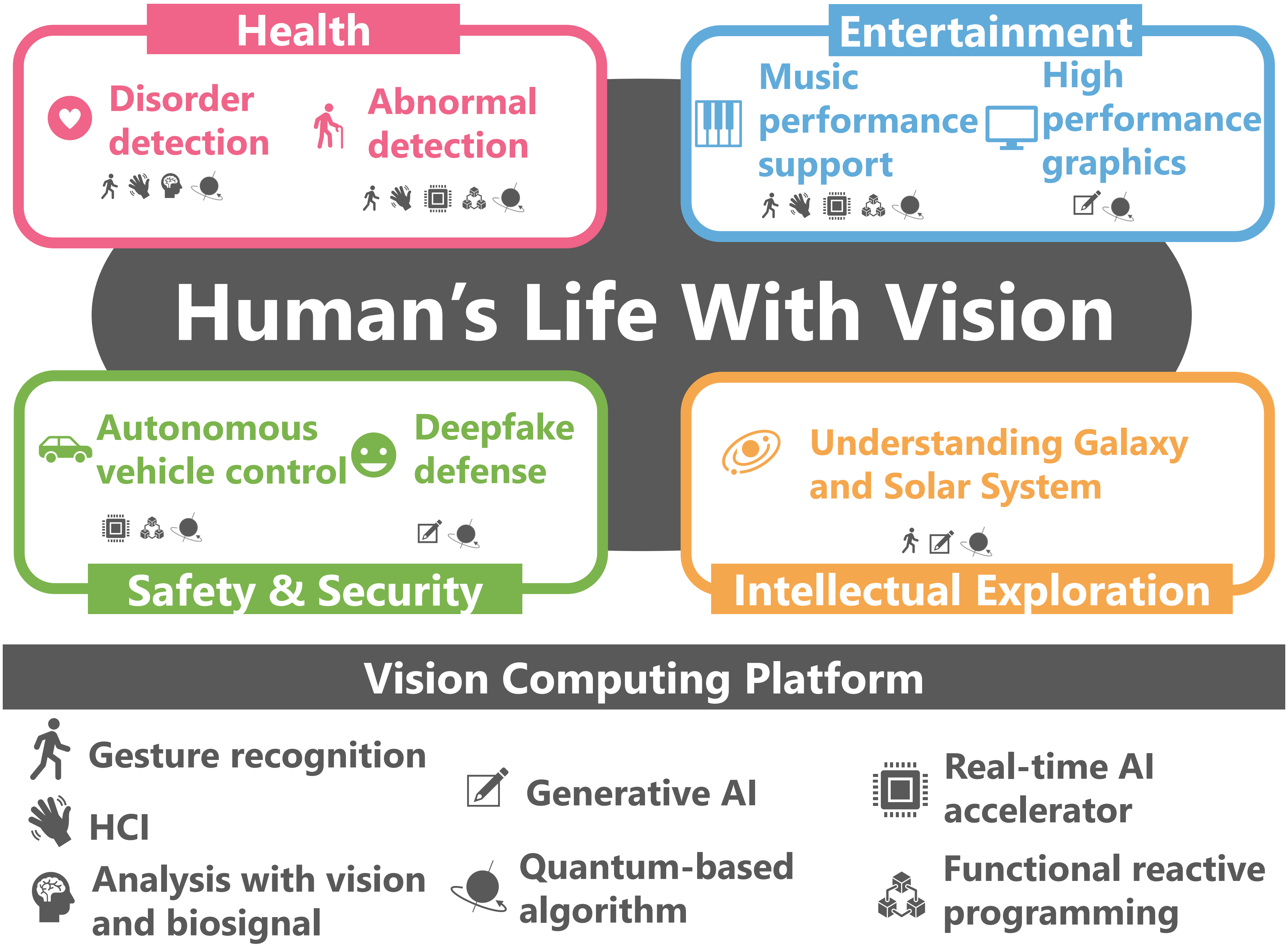
「ペンとコンピュータ」について講演する慎准教授

"An Introduction of Pen-Computer in Aizu City" in the Newspaper, Fukushima-Minpo (福島民報, daily newspaper of Fukushima Prefecture, Japan), March. 5, 2010.

Vision Computing Platform Cluster

Research Center for Advanced Information Science and Technology (CAIST)
VISION Cluster

Leader: **Dr. Jungpil SHIN**
The University of AIZU, JAPAN
E-mail: *jpshin@u-aizu.ac.jp*

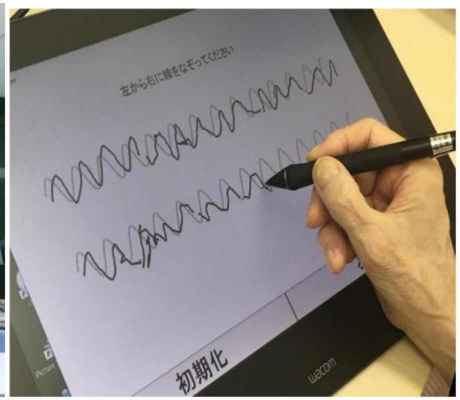

Health
Disorder detection
Abnormal detection
Entertainment
Music performance support
High performance graphics
Human's Life With Vision
Autonomous vehicle control
Deepfake defense
Safety & Security
Understanding Galaxy and Solar System
Intellectual Exploration
Vision Computing Platform
Gesture recognition
HCI
Analysis with vision and biosignal
Generative AI
Quantum-based algorithm
Real-time AI accelerator
Functional reactive programming

Technical Themes To Contribute to Human's Life

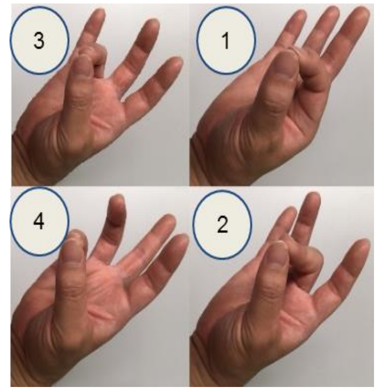
- **Gesture Recognition**
  - Music performance support
  - Abnormal activity detection
  - Pose detection
- **Disorder Detection**
  - Parkinson disease detection
  - Developmental disorder detection
- **Autonomous Vehicle Control**
  - Reinforce learning
  - Functional reactive programming/SLAM
- **Visualization**
  - Image generation
  - Galaxy simulation
  - Deepfake detection
- **Computational Infrastructure and Hardware**
  - Deep learning computational environment
  - Quantum-based algorithm
  - Real-time AI accelerator

Implementation of Vision Algorithms

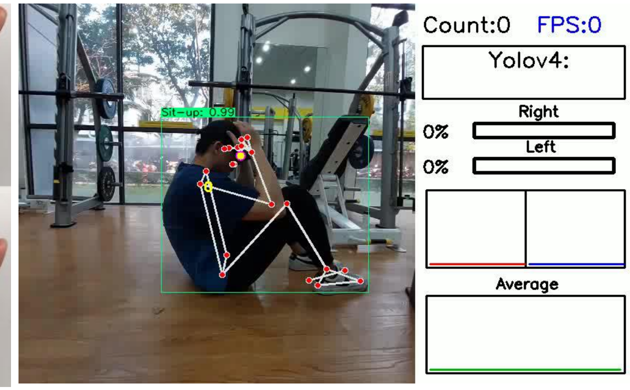
Gesture Recognition



Disease Diagnosis



User Identification



Activity Recognition

Roles & Related Research Theme

- Coordinator of Cluster
- Gesture Recognition, Disorder Detection, Machine Learning and Deep Learning Algorithm

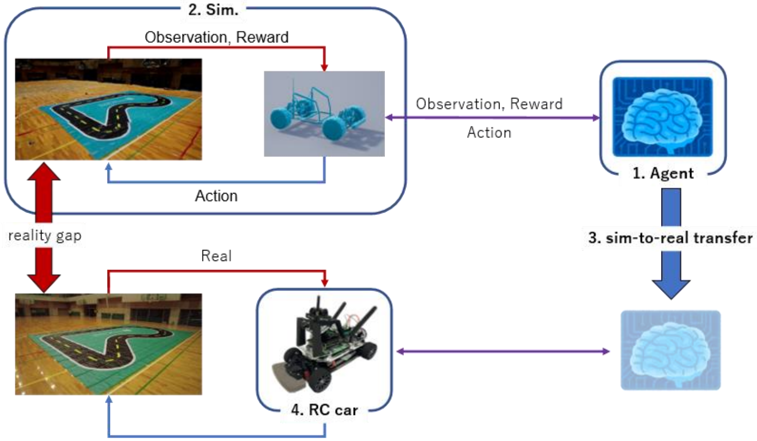
Relationships and Current Results

- Research related to Human Activity Recognition based on members' experiences
- 28 Major Journal Papers, 10 Conference Papers (2022)
- 3 Special Issues Editor of Major Journal, 1 Lecture Note Volumn Editor, 2 Patents (2022)

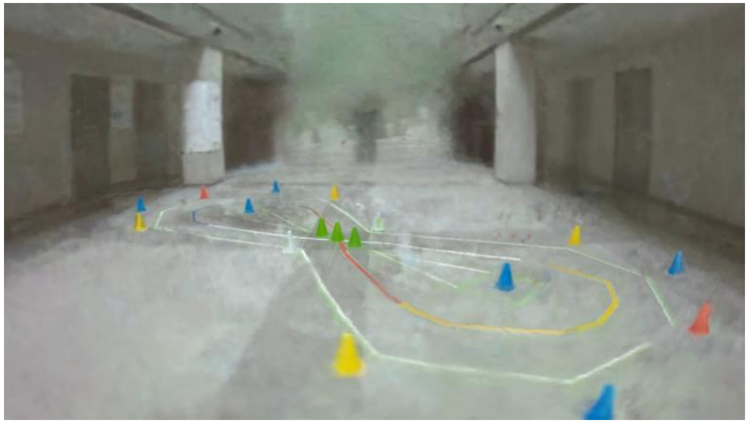
Research Plan

- Implementation and test of applications
- Conduct the theme, Vision and Sensor-Based System for Human Activity Recognition using Machine Learning (4 Sub themes).

AI-based Autonomous Driving



Sim-to-real transfer for autonomous driving



NeRF-based driving simulator

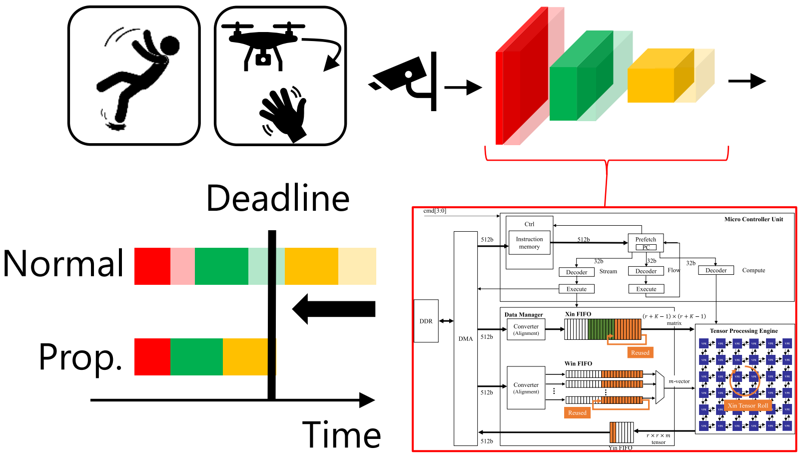
Roles & Related Research Theme
- Acceleration of vision-based networks with DNN compilers
- Simulators, real-time inference, and reinforcement learning for autonomous driving (AD)

Relationships and Current Results
- Investigation of latest DNN technologies related to hardware, vision, and AD.
- Decision on research themes related to AD based on members' experiences
- Development of AI-based autonomous driving cars
- Compiler technologies for DNNs

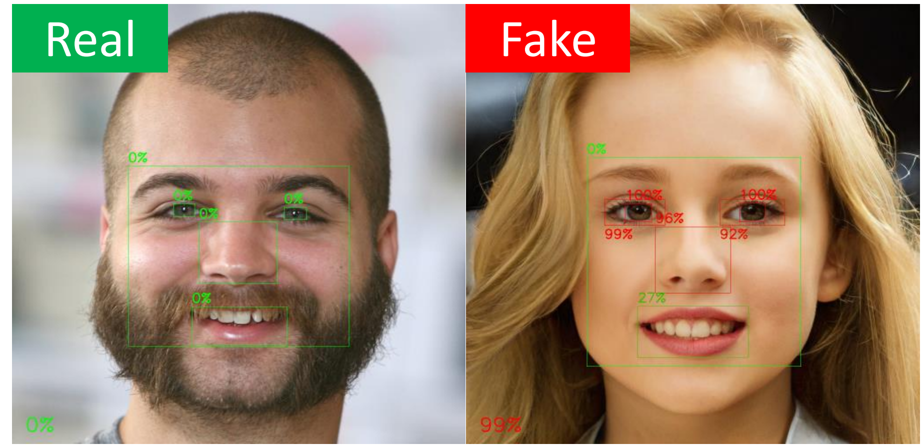
Research Plan
- Development of Reinforcement learning for ADs on low-friction surface road
- Driving simulator to realize real-environment using NeRF-related technologies
- Acceleration of DNNs with compiler technologies

Acceleration of Vision Applications

Deadline-driven Inference on RT Accelerator



Deep Fake Detection and Proactive Defense



Roles & Related Research Theme

- Acceleration of human activity recognition based on real-time DNN accelerators
- Image recognitions for safe and secure society (e.g., deep fake detection)

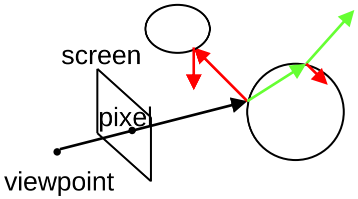
Relationships and Current Results

- Discussion about the acceleration of gesture recognition is underway
- Highly energy-efficient CNN accelerator has been developed
- Face part-based deep fake detection method has been developed

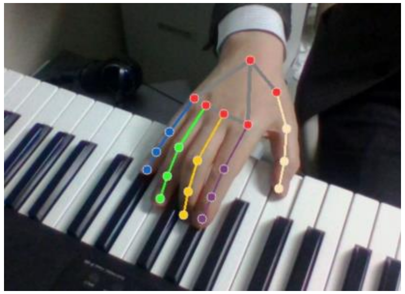
Research Plan

- Realizing deadline-driven inference that can realize hard real-time image recognition by shortening the inference time according to the deadline
- To protect social media images so that they cannot be used for deep fake generation

Image Generation / Application of Hand Pose Estimation



Acceleration of Image Generation



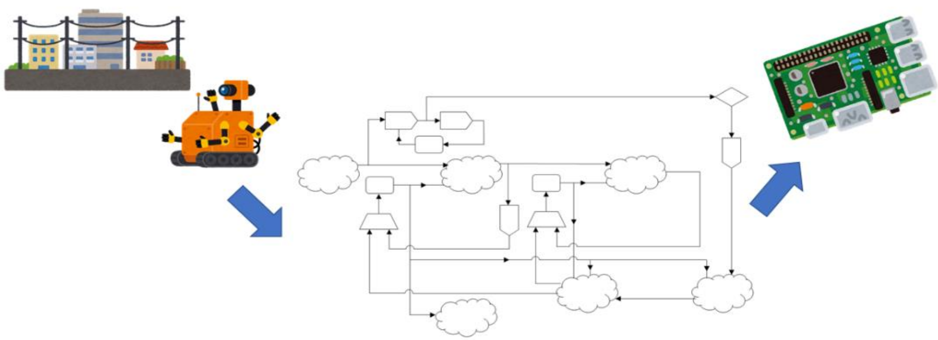
Application of Hand Pose Estimation to Computer Music

Roles & Related Research Theme
- Acceleration of ray tracing by optimizing BVHs with reinforcement learning
- Fingering estimation for keyboard instruments based on hand pose analysis

Relationships and Current Results
- Sharing knowledge on deep networks and reinforcement learning with other members
- Developing basic programs for BVH construction and rendering
- Investigating possible applications of hand pose estimation in computer music

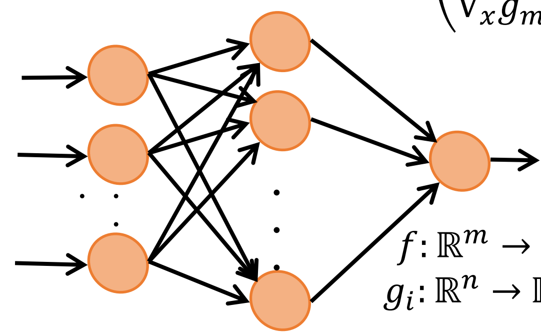
Research Plan
- Designing and evaluating methods for fast ray tracing with machine learning techniques
- Seeking for faster hand pose estimation and improving accuracy for fingering estimation

Symbolic and verification technique for vision computing



Implementation of SLAM with FRP

$$\nabla_x f(g(x)) = \nabla_{g(x)} f(g(x)) \begin{pmatrix} \nabla_x g_1(x) \\ \vdots \\ \nabla_x g_m(x) \end{pmatrix}$$



$f: \mathbb{R}^m \to \mathbb{R}$

$g_i: \mathbb{R}^n \to \mathbb{R} \ (1 \leq i \leq m)$

Verification of properties of DNN

Roles & Related Research Theme

- Implementation of SLAM with FRP to ease the porting of bottlenecks to FPGA
- Formalization and verification of DNN as a model of vision transformation with Coq

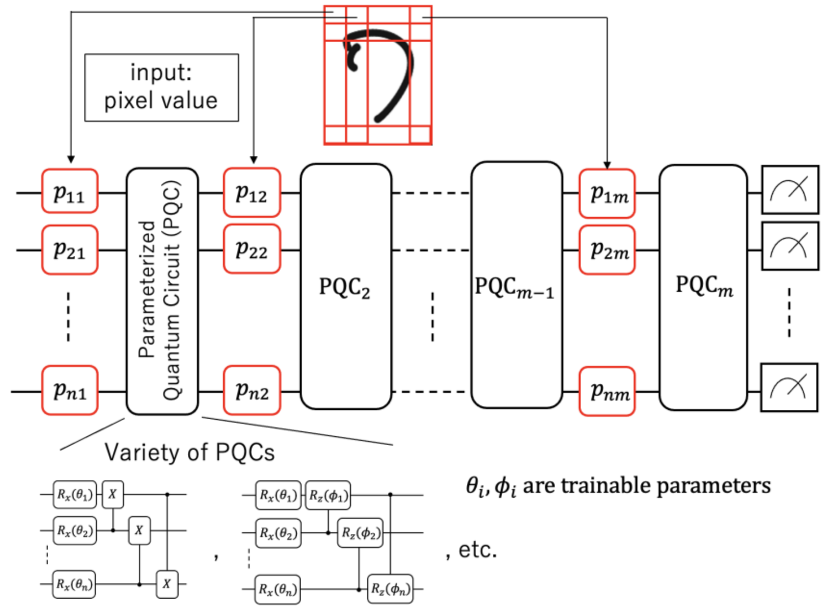
Relationships and Current Results

- Sharing knowledge on DNNs and reinforcement learning with other members
- Designing a synchronous model of SLAM with FRP on C++
- Formalizing gradients and derivatives of vectors for back propagation with Coq

Research Plan

- Seeking the design of SLAM program with FRP to ease the porting of bottlenecks to FPGA
- Developing verification methods of properties making DNNs for vision computing properly

Quantum Machine Learning



Quantum Machine Learning

Roles & Related Research Theme

- Investigate direct and indirect connections to machine learning from the aspect of quantum algorithms and explore themes that can be linked.

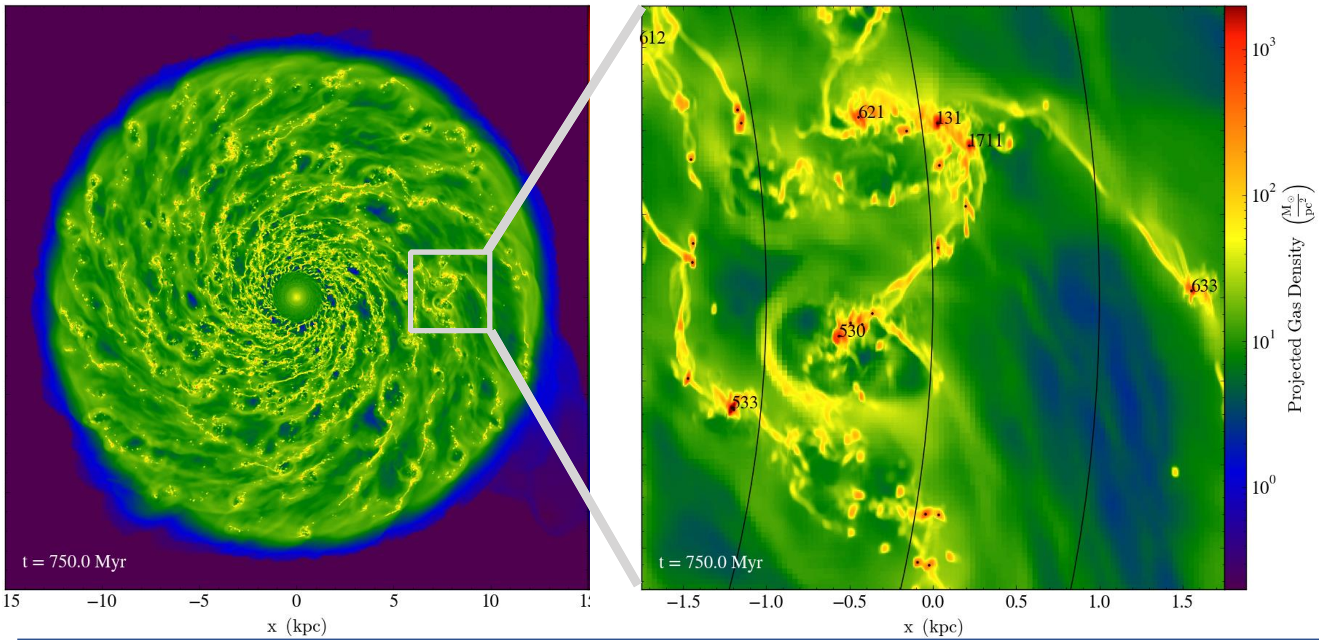
Relationships and Current Results

- Quantum machine learning is not yet at a practical stage, but is still exploring its relationship to machine learning and visionary computing

Research Plan

- Research and investigation of the technology underlying machine learning and visionary computing in terms of quantum algorithms.

Visualization



Detection and tracing the interstellar gas clouds in pictures and move of the numerical hydrodynamical simulation of a galaxy

Roles & Related Research Theme

- Performing large-scale galaxy simulations using Supercomputers and analyzing the simulation pictures and movies using state-of-art image recognition techniques.

Relationships and Current Results

- Introducing how to apply and to use external computational resources

Research Plan

- Development and acceleration of detection and tracking methods for specific objects in images and videos

Important Features Selection and Classification of Adult and Child from Handwriting Using Machine Learning Methods

Vision Computing Platform Cluster : Dr. Jungpil SHIN

The University of AIZU, JAPAN

E-mail: *jpshin@u-aizu.ac.jp*

[1] **Jungpil Shin**, Md. Maniruzzaman, Yuta Uchida, Md. Al Mehedi Hasan, Akiko Megumi, Akiko Suzuki, Akira Yasumura, "Important Features Selection and Classification of Adult and Child from Handwriting using Machine Learning Methods", ***Applied Sciences*** (MDPI), Vol. 12, Issue 10, https://doi.org/10.3390/app12105256, May, 2022. Impact Factor: 2.736.

Outline of the Study

- Motivation of the Study
- Objective of the Study
- Materials and Methods
  - Dataset Preparation
  - Proposed Methodology
  - Acquired Features
  - Feature Extraction
  - Feature Selection
  - Feature Scaling
  - Hyper-parameter tuning
- Experimental Result
- Conclusion
- Extension of the Current Study
- References

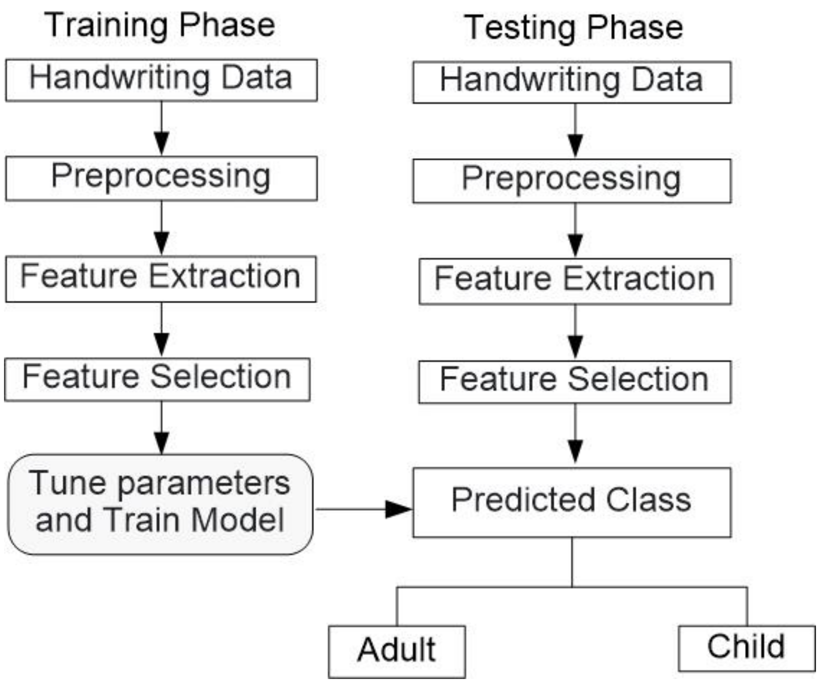
Objective of the Study

- Extract different types features from handwritten text and pattern datasets.
- Select the most efficient features for classifying adults and children.
- Propose an ML-based approach for automatically classifying people as adults or children based on their handwritten text and patterns.

Proposed Methodology

Steps of Proposed method:

- Divide the handwriting dataset into two sets: training phase (80%) and testing sets (20%).
- Preprocess the handwriting data.
- Extract various features from input data.
- Select the potential features using sequential forward floating selection (SFFS).
- Select the optimal parameters using a grid search.
- Classify adult and child using SVM and RF.



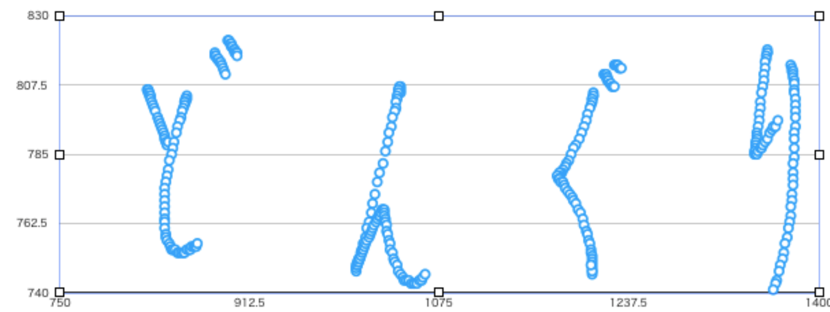
Dataset Preparation

Text Dataset

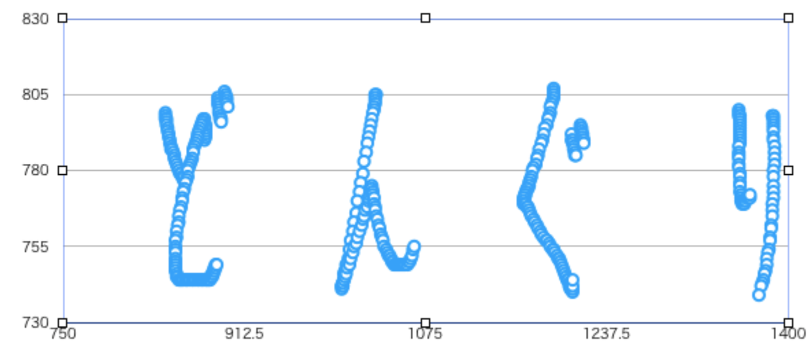
- The resolution is 2560 × 1440
- Total number of subjects: 57
  - Adult(age 19~27) : 26
  - Child(age 12) : 31
  - Each subject wrote the same of 30 tasks (word) of hiragana.

Summary of the dataset

Group name	Age (years)	No. of subjects	No. of written words	Total Samples
Child	12	31	30	930 (31×30)
Adult	19 -27	26	30	780 (26×30)



Adult handwriting



Child handwriting

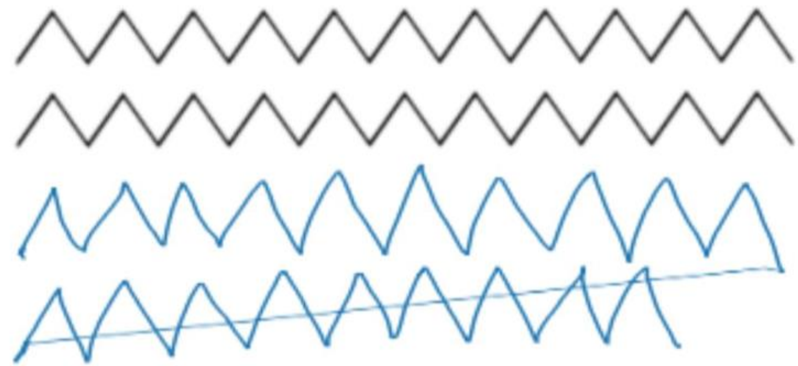
Dataset Preparation (Cont.,)

Handwritten Pattern Dataset

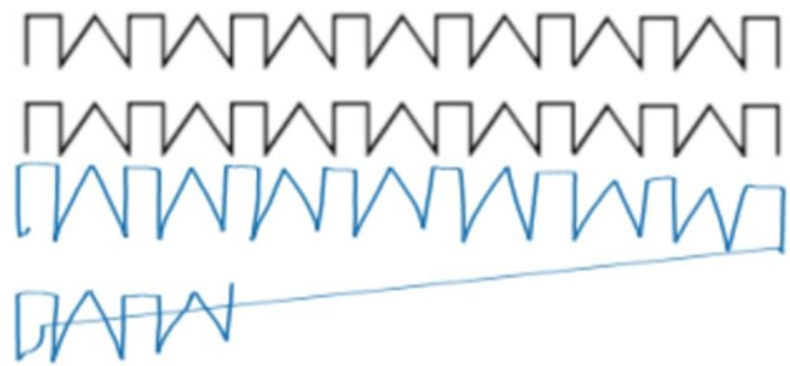
- The resolution is 2560 × 1440
- Total number of subjects: 81
  - Adult(age 19-43) : 42
  - Child(age 8-13) : 39
  - Each subject ask to draw 4 lines as zigzag and periodic lines (trace and predict).
  - Repeat the tasks into 3 times

Summary of the dataset

Group	Age (years)	No. of subjects	No. of task	No. of repeat	Total samples
Child	8~13	39	4	3	468 (39×4×3)
Adult	19~32	42	4	3	505 (42×4×3)



Zigzag line sample and painted

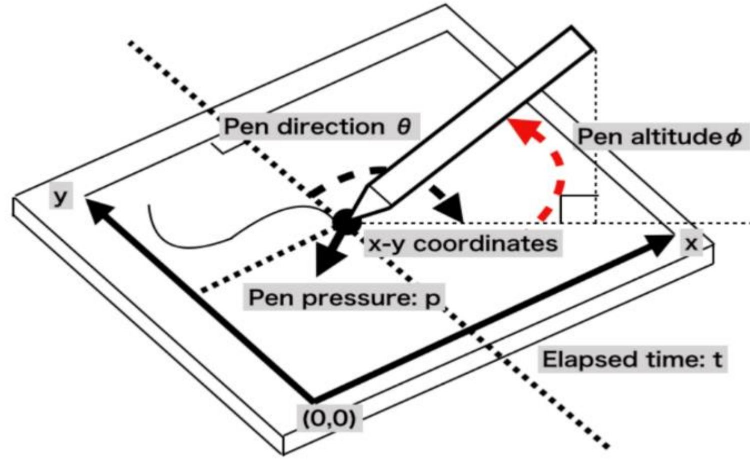


PL line sample and painted

Dataset Preparation (Cont.,)

Device

- The handwritten dataset was obtained using a pen tablet system.
- The tablet was connected to a laptop PC running Windows 10.



Collect 6 raw features

- Elapsed time from the start of the test
- Writing pressure
- x coordinate, y coordinate
- Vertical component height
- Horizontal component direction

List of Extracted Feature Names

SN	Feature	Description	SN	Feature	Description
1	Width	Max (X)-Min (X)	16	PCSDPos	SD of increase in pressure between two-time points
2	Height	Max (Y)-Min (Y)	17	PCMax	The maximum increase in handwriting pressure between two-time points
3	Length	The total length of the drawing	18	PCAvgNeg	Mean decrease in pressure between two-time points
4	Velocity	(Length)/(total) drawing time	19	PCSDNeg	SD of decrease in pressure between two-time points
5	PIVH	The maximum speed recorded at any time point	20	PCMin	Maximum reduction in handwriting pressure between two-time points
6	PIVL	Minimum speed recorded at any time point (PMS>0)	21	Error	Number of outliers and triangle square errors based on angles
7	PIAH	Maximum acceleration recorded at any time point	22	PeakPresMean	Mean pressure at minima
8	PIAL	Minimum acceleration recorded at any time point (PMA>0)	23	ErrorStopTime	Mean stuck time at the starting minima point just before the error
9	GripAngleMeanW	Mean of grip angle values for the entire drawing task (Horizontal)	24	AngleMean	Mean of angles at maxima and minima
10	GripAngleMeanL	Mean of grip angle values for the entire drawing task (Vertical)	25	AngleVar	The variance of angles at maxima and minima
11	GripAngleSDW	SD of grip angle values for the entire drawing task (Horizontal)	26	ReglineSlope	The slope of the regression line
12	GripAngleSDL	SD of grip angle values for the entire drawing task (Vertical)	27	ReglineIntercept	The intercept of the regression line
13	Pressure Mean	Mean of recorded pressure values for the entire task	28	LoopCount	Time spent writing divided by the number of peaks
14	PressureSD	SD of recorded pressure values for the entire task	29	AngleSpeed	Mean of velocities at the edge of the peaks and valleys
15	PCAvgPos	Mean increase in pressure between two-time points	30	ErrorRate	(Error)/ (All Peaks) Error rate

Feature Scaling & Tuned Hyperparameters

Feature Scaling:

- We standardization the extracted features using the following formula:

$$z = \frac{\mathrm{X} - \boldsymbol{\mu}}{\boldsymbol{\sigma}}$$

- Where X is the original feature vector; **μ** is the mean of that feature vector, and **σ** is its SD. The value of z lies between 0 to 1.

List of hyperparameters utilized this study

CT	Hyper-parameters
SVM	C: [[0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000], kernel: [‘linear', ‘rbf’, ‘poly’, ‘sigmoid’], and gamma: [0.00001, 0.0001, 0.001, 0.01, 0.1, 1]
RF	max_depth: [2,3,5,10, None], n_estimators: [50,100,200,300,400], min_samples_split: [2, 3, 10], min_samples_leaf: [1, 3, 10], bootstrap: [True, False], criterion: [‘gini’, ‘entropy’]

- To improve the accuracy of classifiers, it is necessary to determine the optimal value of hyperparameters.
- Tuned the hyperparameters using grid search method.
- We choose the hyperparameters that will provide the highest accuracy

Experimental Result

- For handwritten text dataset, the overall prevalence of adult and child was **45.6%** and **54.4%** and their averages age were 27.3±10.5 and 12.5 ±0.3 years.
- About 42.7% and 59.3% of adult and child were female.
- For handwritten pattern dataset, 64.6% and 35.4% of adult and child were female.

Baseline characteristics of adult and child

Variables	**Overall**	**Adult**	**Child**	**p-value[1]**
Handwritten text dataset				
Total, n (%)	57	26 (45.6)	31 (54.4)	
Gender, Female (%)	27 (47.4)	11 (42.7)	16 (59.3)	0.483
Age, Mean±SD	19.2 ±10.2	27.3 ±10.5	12.5 ±0.3	**<0.001**
Handwritten pattern dataset				
Total, n (%)	81	42 (51.9)	39 (48.1)	
Gender, Female (%)	48 (59.3)	31 (64.6)	17 (35.4)	**0.0025**
Age, Mean±SD	17.8±7.9	23.9±4.9	11.8±1.6	**<0.001**

- Age and gender (except gender for text data) were significantly associated with adult and child for both handwritten text and pattern dataset.

Experimental Result (Cont.,)

Expt.-1: Evaluation for Text Dataset

Table 5: Performance scores of SVM and RF

CT	# of Features	ACC	Rec	Prec	f1-score	AUC
SVM	15	87.7	92.4	85.9	89.1	0.919
RF	18	**93.5**	**95.7**	**92.2**	**93.9**	**0.983**

- RF-based classifier achieved an accuracy of 93.5% and AUC of 0.980 which are comparatively better than SVM.

Table 6: Performance scores of SVM and RF for 11 common features

CT	ACC	Recall	Prec	f1-score	AUC
SVM	87.4	90.8	86.6	88.7	0.947
RF	**91.5**	**93.0**	**91.5**	**92.3**	**0.967**

- RF also produced higher classification accuracy of 91.5%, and AUC of 0.967 compared to SVM.

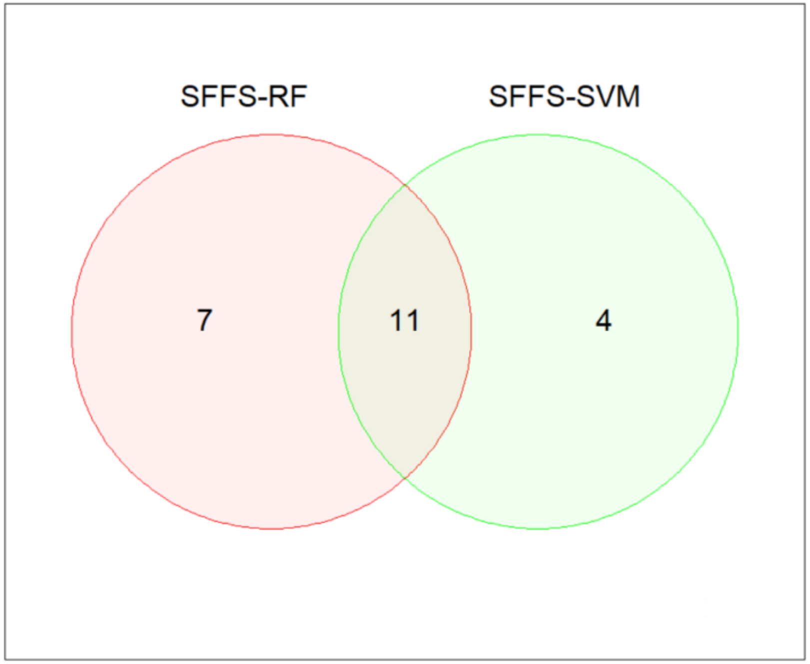


Fig. 4: Identification of common features from SFFS-RF and SFFS-SVM

- Finally, we may conclude that RF provide better performance scores than SVM for the prediction of the adult and child handwritten text dataset.

Experimental Result (Cont.,)

Expt.-2: Evaluation for Pattern Dataset

Table 7: Accuracy of SVM & RF

Line Types	# of features	RF	# of features	SVM
Zigzag trace	19	83.7	26	71.4
Zigzag predict	13	85.7	3	75.5
Pl trace	24	73.5	7	79.6
Pl predict	**9**	**89.8**	**12**	**87.7**
All	25	85.6	28	82.1

- RF produced 89.8% accuracy with only 9 features for the prediction of PL line. Whereas, SVM produced 87.7% accuracy for the selected 12 features.

Table 8: Accuracy of SVM & RF for common features

Line Types	# of features	RF	SVM
Zigzag trace	18	79.5	71.4
Zigzag predict	2	73.4	69.3
Pl trace	7	83.6	79.5
Pl predict	**9**	**89.8**	**79.5**
All	23	84.1	85.1

- RF also produced higher accuracy of 89.8% with 9 common features, whereas, SVM provided 79.5% accuracy. Finally, we can say that RF performed better incase of our experiment than SVM.

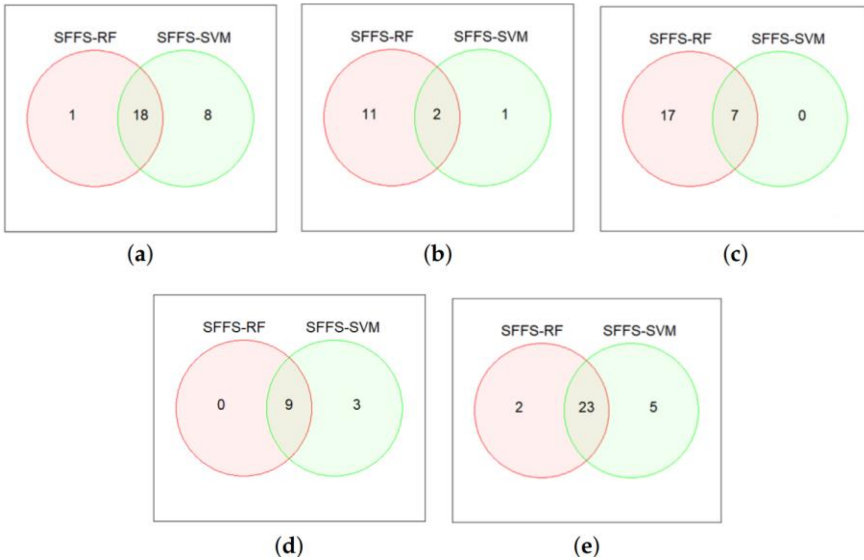


Fig. 4: Identification of common features from SFFS-RF & SFFS-SVM: Feature from (a) trace of zigzag line, (b) prediction of zigzag line, (c) trace of PL line, (d) prediction of PL line, and (e) all handwritten patterns

Table 9: Performance scores of SVM & RF for common features

Line types	CF	RF				SVM			
		Rec	Prec	f1-score	AUC	Rec	Prec	f1-score	AUC
Zigzag trace	18	80.6	86.2	83.3	0.870	67.7	84.0	75.0	0.820
Zigzag predict	2	83.8	76.4	80.0	0.784	77.4	75.0	76.1	0.732
Pl trace	7	83.8	89.6	86.6	0.926	74.1	92.0	82.1	0.872
Pl predict	**9**	**87.0**	**93.1**	**90.0**	**0.903**	**83.8**	**83.8**	**83.8**	**0.811**
All	23	85.1	84.3	84.7	0.923	83.1	87.5	85.2	0.919

Experimental Result (Cont.,)

Table 11: Accuracy comparison with the existing similar studies

Authors	Year	Data types	ACC (%)
Guimaraes et al. [4]	2017	Sentences	Prec: 95
Rizwan et al. [17]	2021	FG-NET (Face)	94.6
Özkan and Turan [18]	2018	FG-NET	78.5
Goshvarpour [19]	2019	ECG	94.6
Ilyas [20]	2020	Auditory perception	92.0
Reade [21]	2015	FG-Net	82.0
Tin [22]	2012	FG-Net	92.5
Our proposed	2022	Handwritten text	93.5%
		Handwritten pattern	89.8%

- As shown here, Our proposed SFFS with RF (SFFS-RF) model produced higher accuracy compared to SFFS with SVM (SFFS-SVM) to classify adult and child based on their handwritten text and handwritten pattern

Conclusion and Extension of this work

Conclusion:

- Classification of adults and children by handwriting text can be done with 93.5 % accuracy and 89.8% accuracy for handwriting pattern.
- By increasing the number of subjects, it is expected that the accuracy will increase and decrease the number of features.
- This study will provide evidence of the possibility of classifying adults and child based on their handwritten texts and patterns.

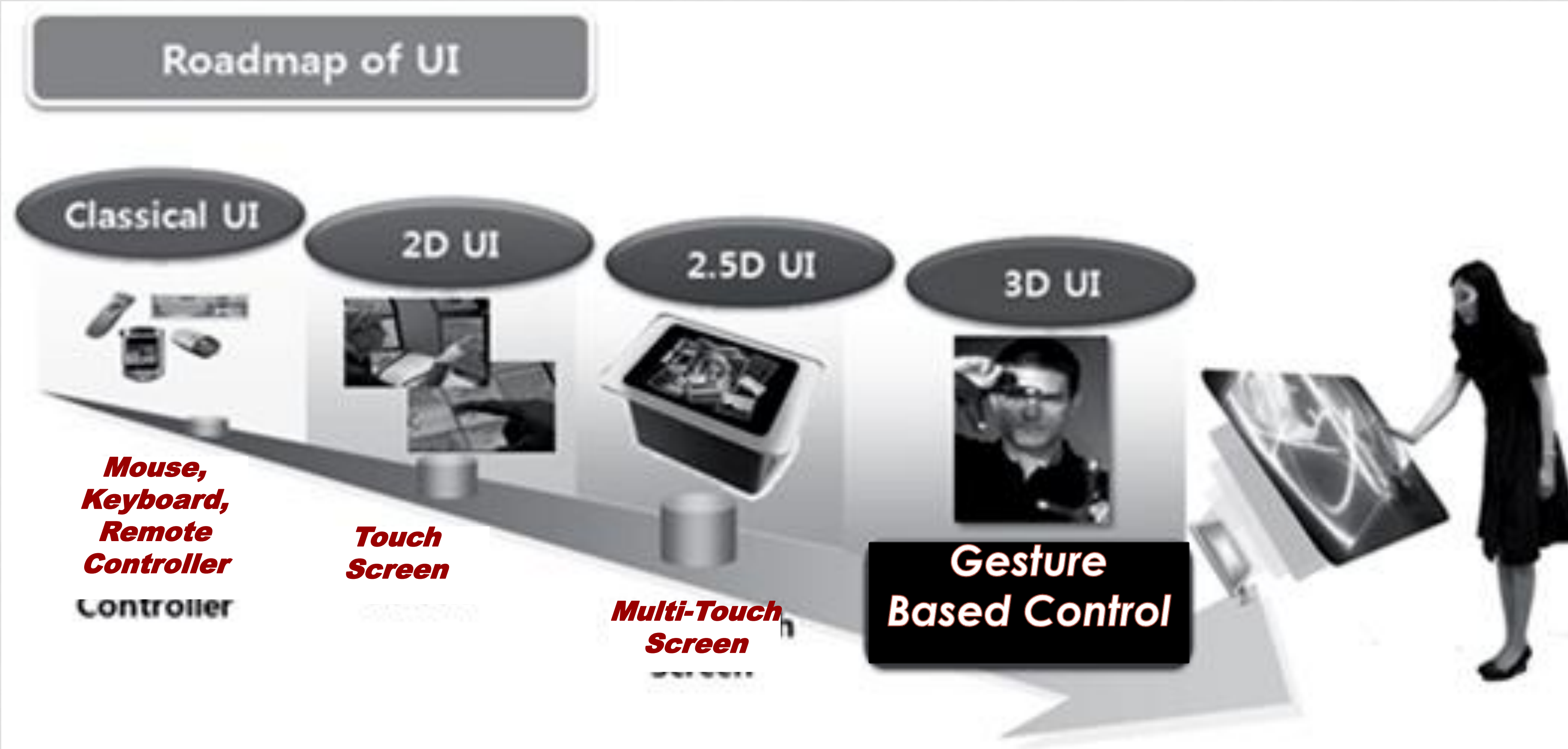
Extension of this study:

- Extract more features from the input data
- Add more subjects to get the better performance of classifiers.
- Apply different ML and deep learning-based approaches to classify adults and child's and compared their performances with our current study

MINORITY REPORT

(BY STEVEN SPIELBERG)

The Trend of User Interface



Hand Gesture and Numeral Recognition with Kinect Sensor

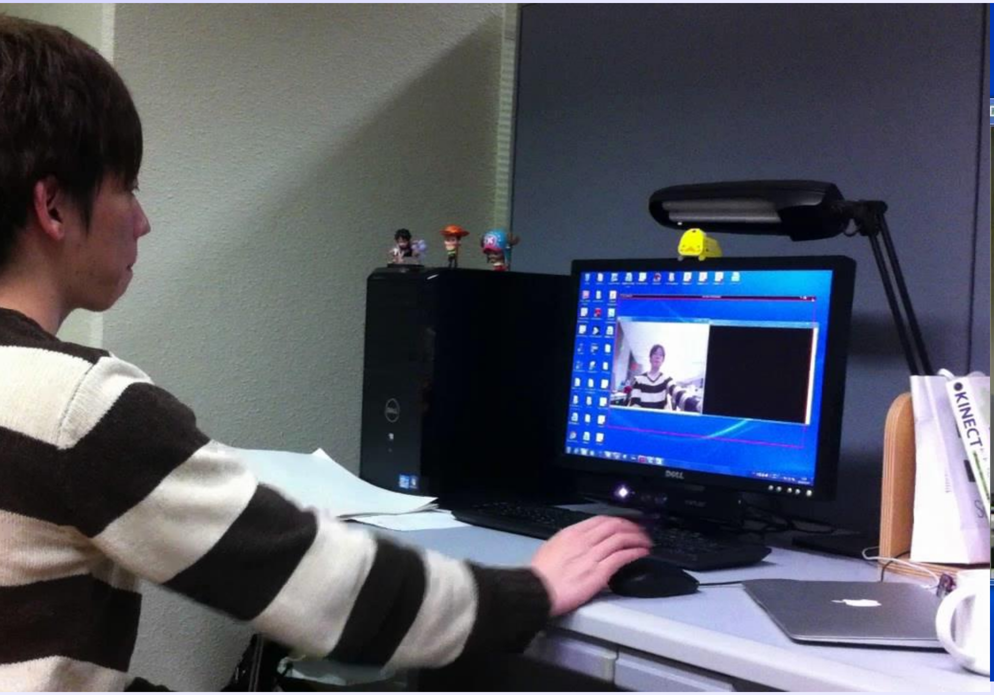
Purpose

- Recognition of written numeral and Calculation in the 3D space by the handwriting.

Motivation

- Hand gesture recognition is an important research issue in the field of Human-Computer-Interaction, because of its extensive applications in virtual reality, sign language recognition, and computer games.

Hand Gesture and Character Recognition Based on Kinect Sensor [1]

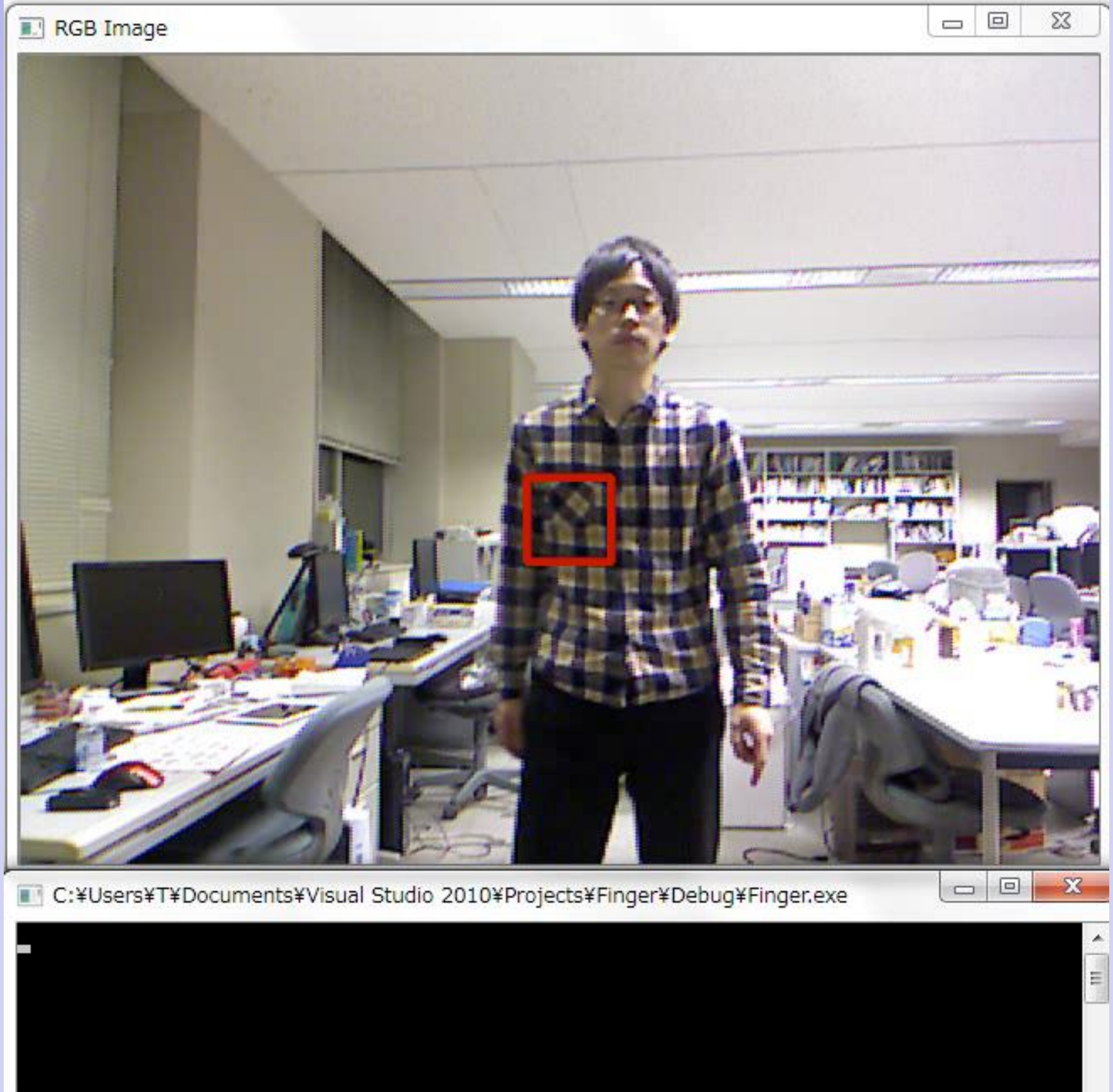


[1] T. Murata, **Jungpil Shin**, "Hand Gesture and Character Recognition Based on Kinect Sensor," *International Journal of Distributed Sensor Networks* [SCI], Vol. 2014, Article ID 278460, July 2014. http://dx.doi.org/10.1155/2014/278460

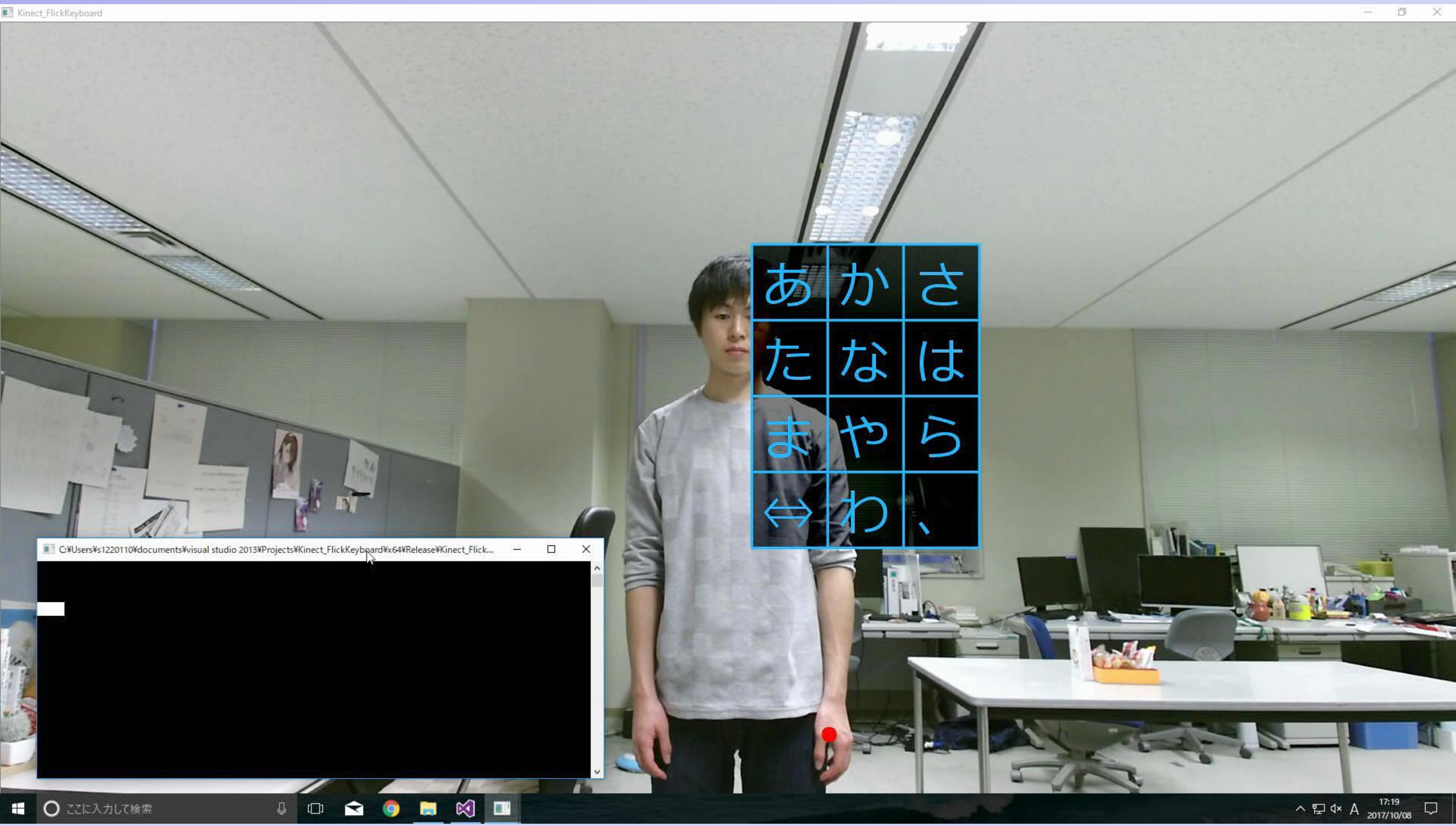
Detected left & right area

Input

Output



Gestural Flick Input-based Non-touch Interface for Character Input:



Hand Gesture Based User Authentication System using 3D motion capture

Research Objectives and Proposed Method

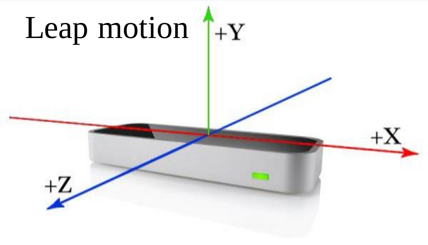
- **Optimal feature selection** based user authentication system (about 600 features extracted)
- Hand gestures are used instead of text-based or pattern-based even biometric

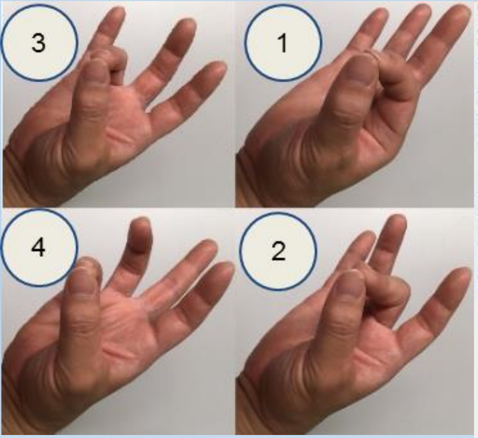
The problem of the previous research

- The leap motion has different parameters and it is **difficult to access** all of these to authenticate the users.
- Biometrics and key pattern authentication are not secure because attackers can easily crack patterns.

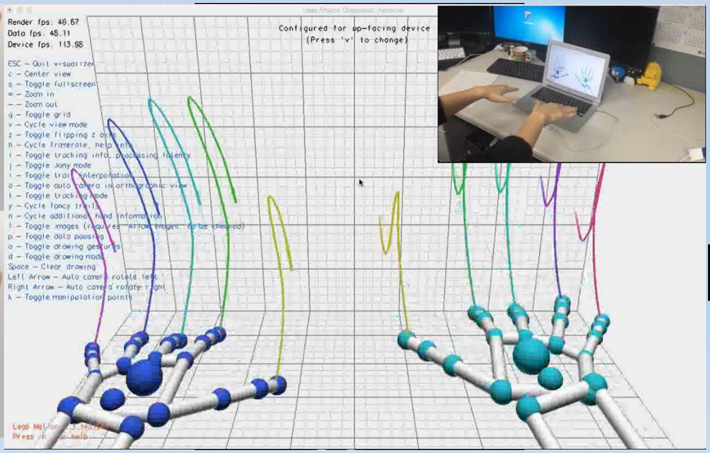
Novelty

- **Use dynamic hand gesture** data to recognize the users
- Machine-learning approaches will be used to **authenticate** the users
- Build a **convenient, low cost motion capture** system that can be used for recognizing the person while performing **daily life activities**.

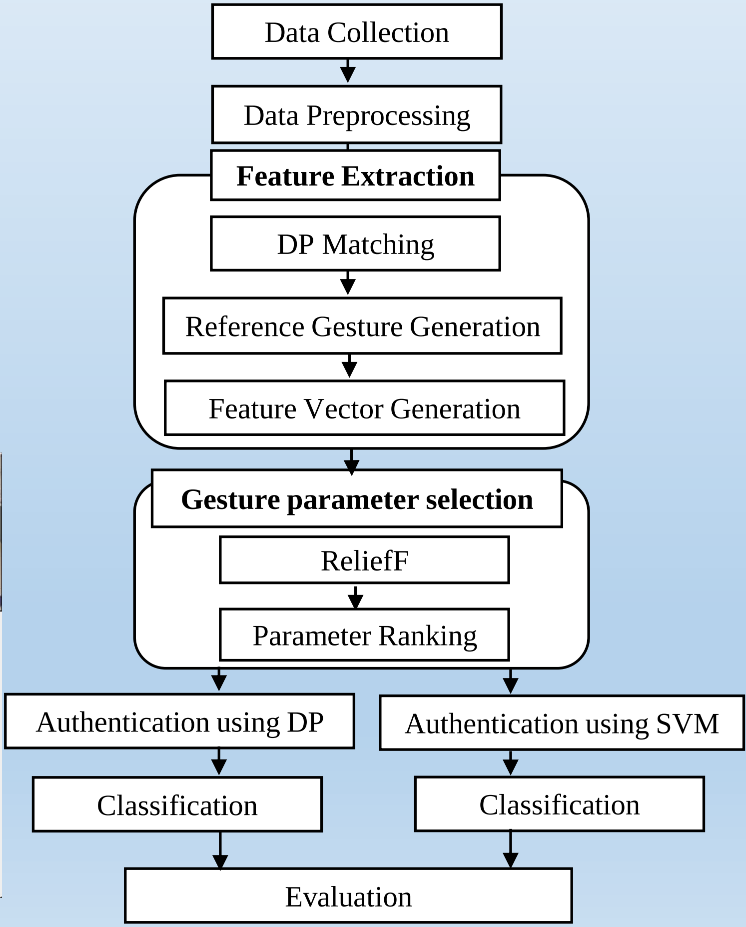




(hand gesture)

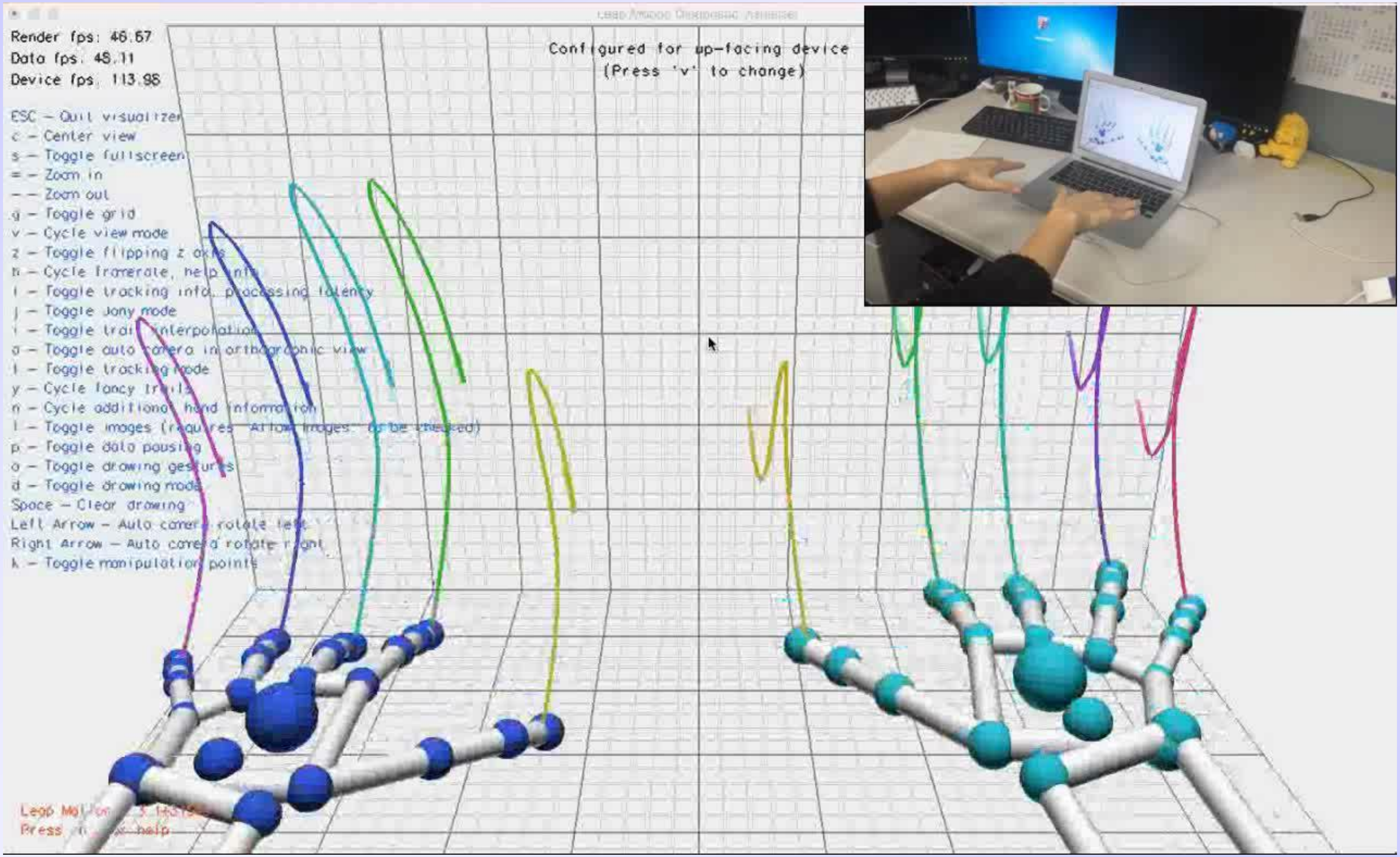


(hand gesture simulation)



Proposed model

Leap Motion controller



Deep Learning-based Hand Pose Estimation for Human-Computer Interaction: Related Research Achievement

- **List of Achievement based on Hand Gestures (References)**

1. Shin, Jungpil, and Cheol Min Kim. "***Non-touch character input system based on hand tapping gestures using Kinect sensor***." *IEEE Access,* Vol. 5 (2017): 10496-10505.
2. Rahim, Md Abdur, Jungpil Shin, and Md Rashedul Islam. "***Gestural flick input-based non-touch interface for character input***." *The Visual Computer* (2019): 1-14.
3. Rahim, Md Abdur, Jungpil Shin, and Md Rashedul Islam. "***Hand gesture recognition-based non-touch character writing system on a virtual keyboard***." *Multimedia Tools and Applications* (2020): 1-24.
4. Murata, Tomoya, and Jungpil Shin. "***Hand gesture and character recognition based on kinect sensor***." *International Journal of Distributed Sensor Networks* 10, no. 7 (2014): 278460.
5. Rahim, Md Abdur, and Jungpil Shin. "***Hand Movement Activity-Based Character Input System on a Virtual Keyboard***." *Electronics,* Vol. 9, no. 5 (2020): 774.
6. Rahim, Md Abdur, Md Rashedul Islam, and Jungpil Shin. "***Non-Touch Sign Word Recognition Based on Dynamic Hand Gesture Using Hybrid Segmentation and CNN Feature Fusion***." *Applied Sciences*, Vol. 9, no. 18 (2019): 3790.
7. Rahim, Md Abdur, Jungpil Shin, and Md Rashedul Islam. "***Human-machine interaction based on hand gesture recognition using skeleton information of Kinect sensor***." *the 3rd international conference on applications in information technology*, pp. 75-79. 2018.
8. I. M. Rashedul, U. K. Mitu, R. A. Bhuiyan, and J. Shin. "***Hand gesture feature extraction using deep convolutional neural network for recognizing American sign language***." In *2018 4th International Conference on Frontiers of Signal Processing (ICFSP)*, pp. 115-119. IEEE, 2018.,

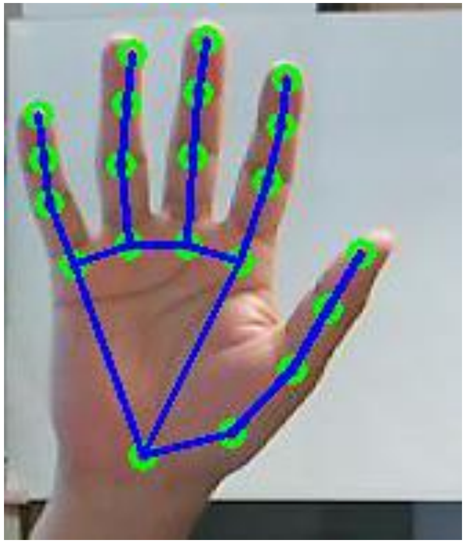
Related Research Achievement

Reference	Method	Advantages	Accuracy
Ref. [1]	Hand Tapping Gestures	Non-touch character input	94.3% (Finger alphabet)
Ref. [2]	Recognize the position and state of the hands	Gestural Flick input introduced for non-touch character input	98.61% (characters) 97.50% (motions gestures)
Ref. [3]	Image processing, deep learning, and support vector machine.	Virtual Keyboard based non-touch character input	Avg. Recognition accuracy 97.93%
Ref. [4]	DP matching	Numerical & alphabetical character recognition	95.0 % and 98.9%
Ref. [5]	Analyze the accelerometer, gyroscope, and EMG signal	Movement Activity-Based Character Input	96.57%
Ref. [6]	Hybrid segmentation and CNN	Sign Word Recognition	97.28%
Ref. [7]	Skeleton Information	Numeric number	96%
Ref. [8]	Deep Neural Network	American Sign Language	94.57%

- **Depth cameras/sensors** like Kinect, Leap motion, RGB-D are used

3 Technologies for HCI using Human Hand

- **Hand pose estimation**
  - Determine 3D locations of hand joints
  - Require ***depth*** images
  - May also need RGB images



Keypoints

- **Hand tracking**
  - Associate the same hand over consecutive video frames
  - Works on RGB image

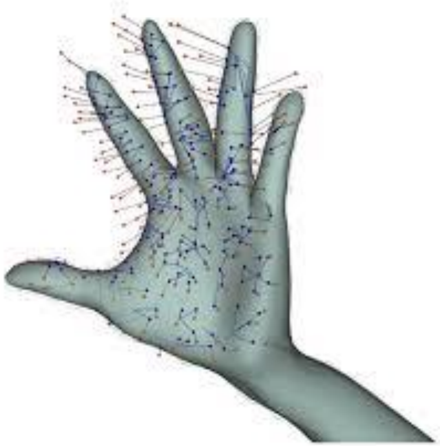
- **Modelling a Hand**

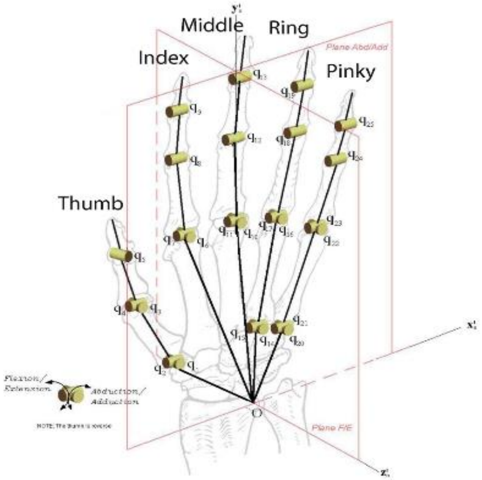
Shape model

Describe the shape of a hand

Kinematic model

Describe DoF of hand joints





Shape and Kinematic model

Steps for Hand Pose Estimation

- **Detection**
  - Find the local region with the hand in the image and crop it
- **Mapping**
  - Map the hand image to the locations of its joints.
- **Refinement**
  - Apply proposed method to refine the estimation

Method

- directly learning a map between the image and pose

RGB image → **Machine Learning** → **3D joint locations (or model parameters)**

Research Target

RGB image

Machine Learning

3D joint locations (or model parameters)

Recognize Hand Gesture Activity

Character input

Non-touch Input Command

Sign Language Recognition

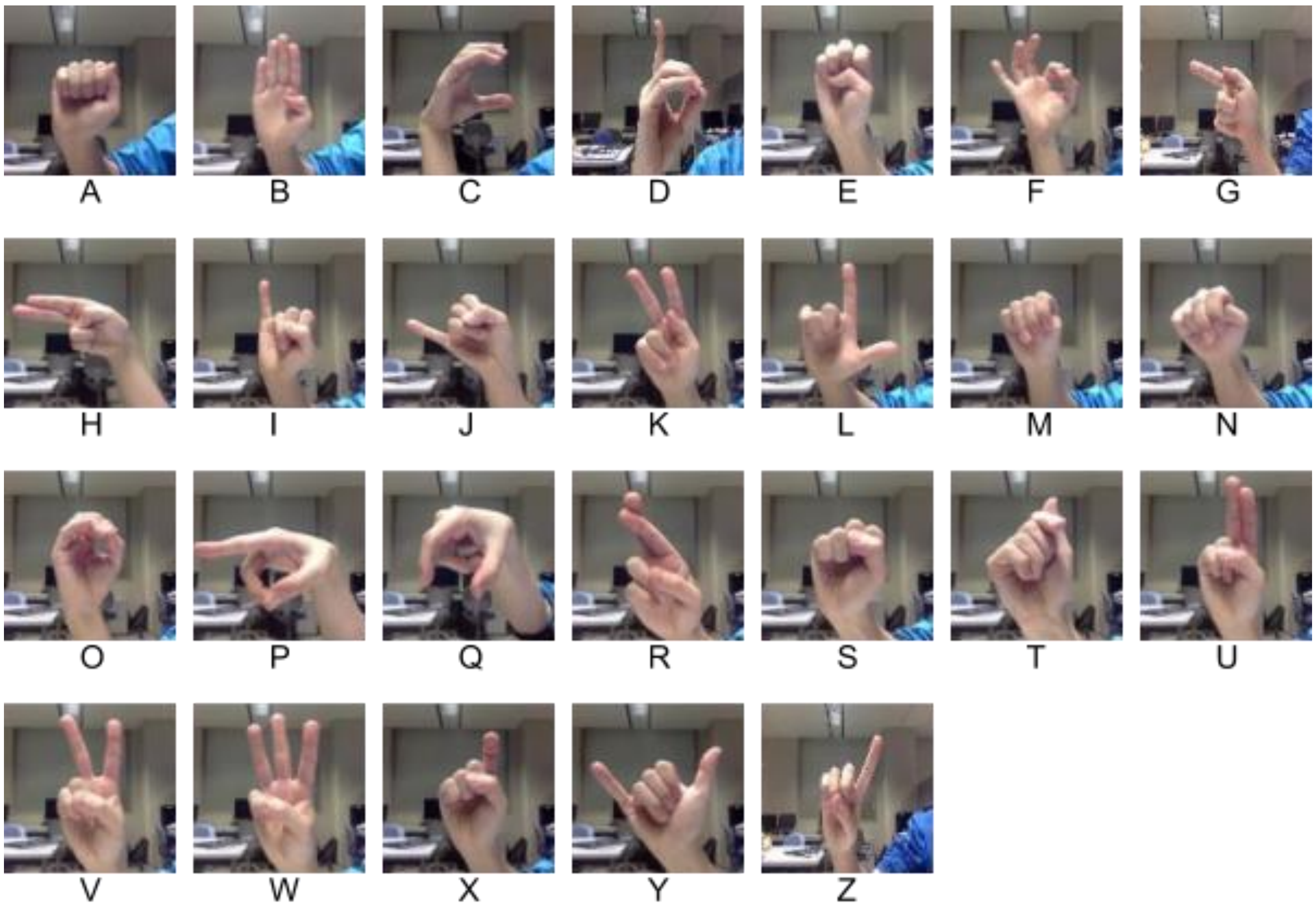
Robotic instruction

Clinical and health recognition

Interacting with an object/motion detection

American sign language alphabet

- American sign language is a sign language used in English-speaking countries such as the United States and Canada, and can express the letters of the alphabet from A to Z with one hand.



Previous study

Table 1. Previous works performances at a glance including the approach and classifier used.

Reference	Approach	Classifier	Recognition Rate
[1]	Smart Bands [1]	KNN with DTW	89.20%
[2]	Smart Bands [1]	CNN	83.20%
[3]	Kinect [1]	DP Matching	98.90%
[4]	Video Camera [2]	DP Matching	75.00%
[5]	Kinect [1]	VGG19	94.80%
[6]	Kinect [1]	Random Forest	90.00%
[7]	Direct Images [2]	InceptionV3	90.00%
[8]	Images from Massey Dataset [2]	RBM	99.31%
[9]	Images from Finger Spelling A Dataset [2]	RBM	97.56%
[10]	Images from NYU Dataset [2]	RBM	90.01%
[11]	Images from ASL Fingerspelling Dataset of Surrey University [2]	RBM	98.13%
[12]	Leap Motion Camera [1]	ANN	94.44%
[13]	Leap Motion in Virtual Reality Environment [1]	HMC	86.10%
[14]	Leap Motion Controller [1]	CNN	80.10%
[15]	Leap Motion Controller [1]	SVM	72.79%
[16]	Leap Motion Controller [1]	DNN	88.79%
[17]	Skin Color Modeling [2]	CNN	93.67%
[18]	Direct Images [2]	DNN with Squeezenet	83.28%
[19]	Skeletal Data and Distance Descriptor [1]	TreeBag & NN	90.70%
[20]	Geometrical Features [1]	ANN	96.78%
[21]	Neuromorphic Sensor [1]	ANN	79.58%
[22]	Multiview Augmentation & Inference Fusion [1]	CNN	93.00%

1 Followed **sensor-based** approaches; 2 Followed **vision-based** approaches.

Purpose

- Using a WEBCAM to recognize American sign language alphabet
- Applying joint estimation data to character recognition

American sign language alphabet

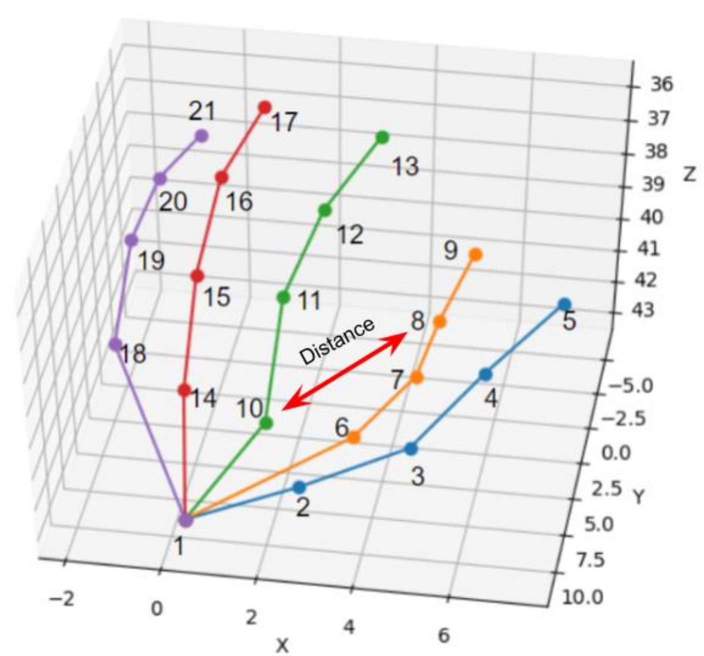
Web Camera

Distance based feature

- Using the distance between two joints as a feature
- We used the Euclidean distance between two points
- Using z score normalization

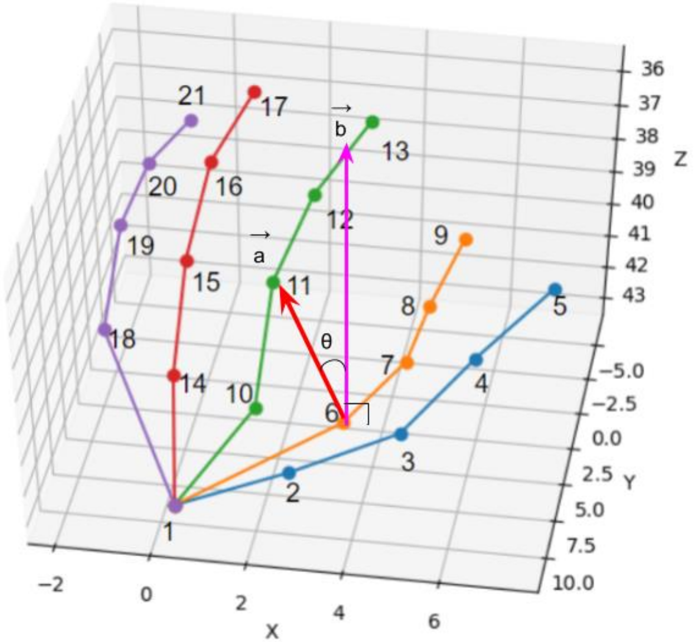
$$zscorenormalization = \frac{data - data_{mean}}{data_{std}}$$

- A total of 190 features are obtained.



Angle based feature

- Using angle-based features for recognition.
- We used the angle between the vector obtained from the two points and the xyz axis.
- A total of 630 features are obtained.



Results Analysis

Table. Obtained experimental results for considered datasets by using SVM and light GBM.

Classifier	Dataset	All Distance Features (190)	All Angle Features (630)	Both Distance and Angle Features (820)
SVM	ASL Alphabet	81.20%	87.06%	**87.60%**
	Massey	98.56%	99.23%	**99.39%**
	Finger Spelling A	96.97%	97.63%	**98.45%**
Light GBM	ASL Alphabet	79.11%	86.01%	**86.12%**
	Massey	96.51%	97.25%	**97.80%**
	Finger Spelling A	94.50%	96.06%	**96.71%**

Table . Average hand pose estimation time (Average HPET), average feature extraction time (Average FET), prediction time per sample (PTPS), recognized frames per second (RFPS), and required memory to load final trained model (Req. Memory) for all three datasets using SVM. All times are measured in seconds.

Dataset	Samples	Avg. HPET	Avg. FET	PTPS	RFPS	Req. Memory
Massey	1815	0.011	0.003	0.014	71	4.04 MB
Finger Spelling A	65,774	0.01	0.002	0.015	66	47.97 MB
ASL Alphabet	780,000	0.011	0.002	0.016	62	115.14 MB

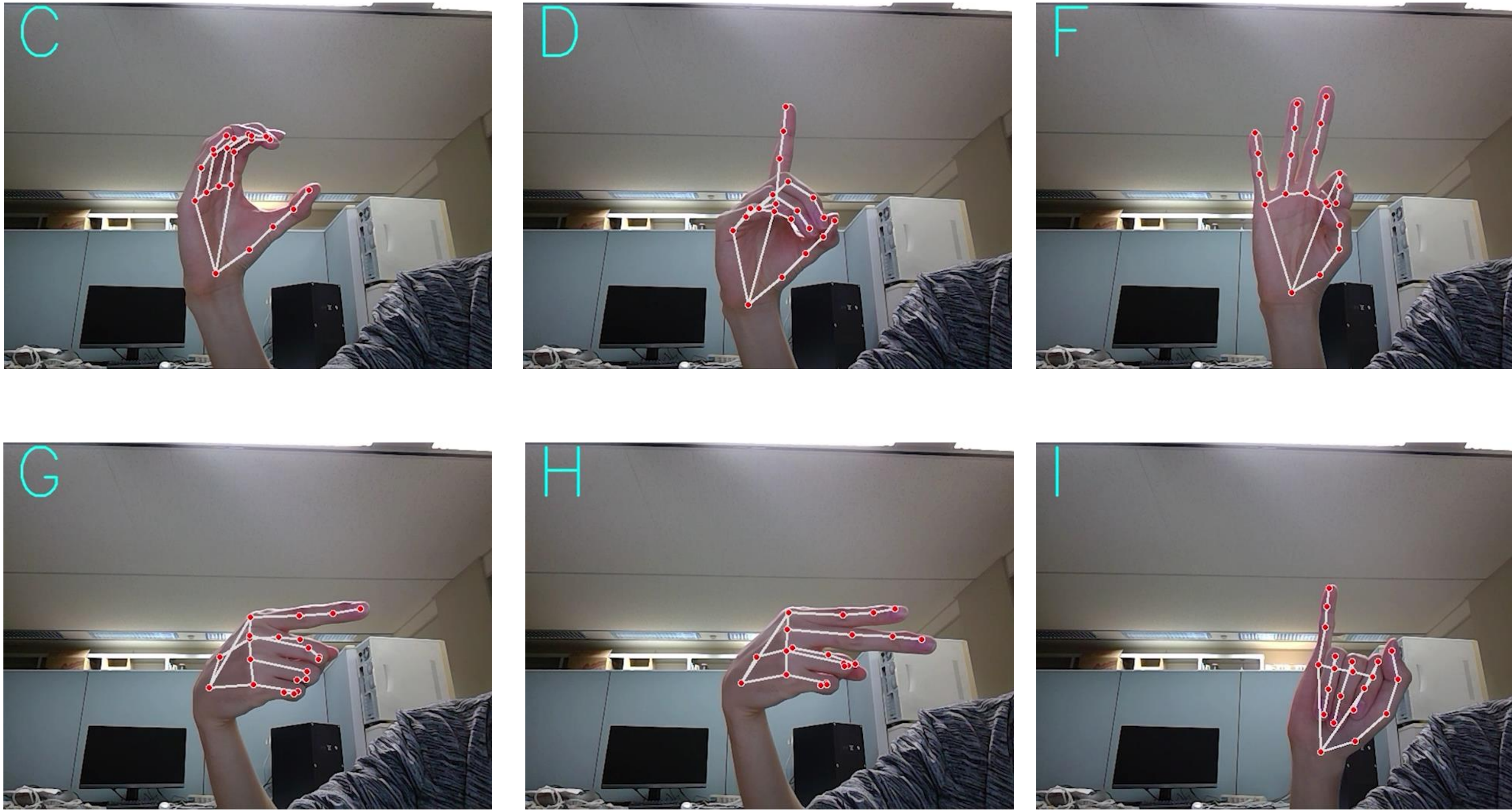
Comparison with Previous Studies

Table. Comparison with other existing works.

Dataset	Approach	Accuracy
Massey Dataset	CNN [53]	72.00%
	RBM [26]	99.31%
	Proposed	**99.39%**
Finger Spelling A Dataset	Random Forest [33]	90.00%
	InceptionV3 [34]	90.00%
	DNN with Squeezenet [40]	83.28%
	ANN [43]	79.58%
	CNN [44]	93.00%
	RBM [26]	98.13%
	Proposed	**98.45%**

[1] **Jungpil Shin**, Akitaka Matsuoka, Md Al Mehedi Hasan, Azmain Yakin Srizon, "American Sign Language Alphabet Recognition by Extracting Feature from Hand Pose Estimation", *Sensors* (MDPI) [SCI], Vol. 21, Issue 17, https://doi.org/10.3390/s21175856, Aug, 2021. Impact Factor: 3.576

Extraction Results



Demo Video

30 sec

Conclusion

- This research has presented a non-touch interface technique
- Introduce novel techniques that allow tracking hand gestures
- Describes experimental results, new frameworks, evaluations and interactive applications
- Gesture flick input system is proposed for character input.
- Hand gesture images are captured and machine learning techniques are implemented.

The achieved accuracy has become more successful and it leads to better results than state-of-the-art systems.

Fitness Movement Types and Completeness Detection Using a Transfer-Learning-Based Deep Neural Network

Vision Computing Platform Cluster : Dr. Jungpil SHIN

The University of AIZU, JAPAN

E-mail: *jpshin@u-aizu.ac.jp*

[1] Kuan-Yu Chen, **Jungpil Shin**, Md. Al Mehedi Hasan, Jiun-Jian Liaw, **Okuyama Yuichi**, **Yoichi Tomioka**, "Fitness Movement Types and Completeness Detection Using Transfer Learning Based Deep Neural Network", ***Sensors*** (MDPI) [SCI], Vol. 22, Issue 15, https://doi.org/10.3390/s22155700, July, 2022. Impact Factor: 3.847

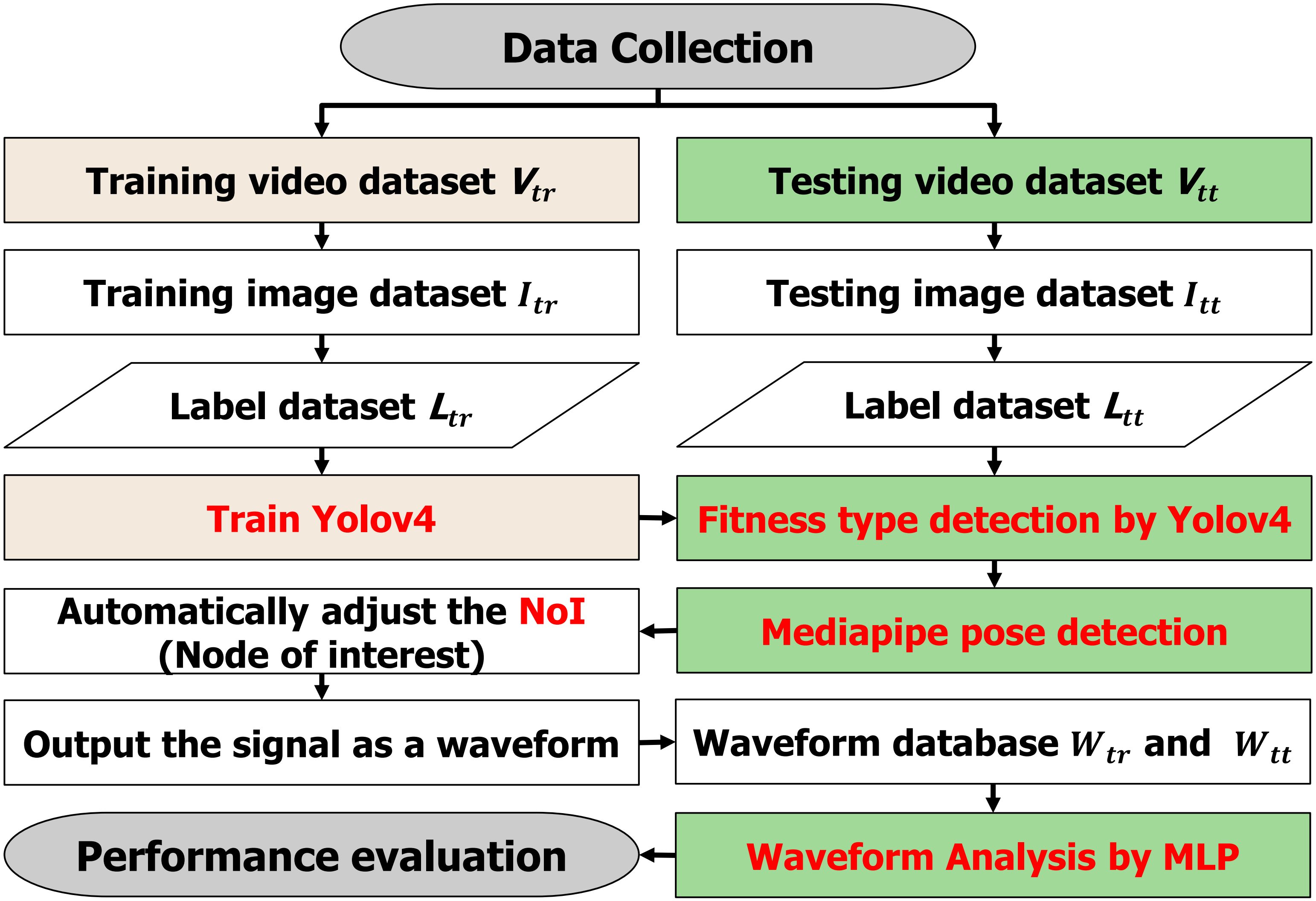
Outline

- Introduction
  - Motivation
- Proposed method
  - Data Collection
  - Yolov4、Mediapipe、MLP
  - Evaluation Method
- Experiment
  - Experiment setup
  - Results and Comparison
- Conclusion

Introduction – Fitness recognition

- Fitness movements can be detected using image recognition.
- Fitness recognition already has good applications.
- Like games and AI fitness coaches, it can assist people in fitness.
- Therefore, the use of images to detect fitness movements has become a very important research.

Proposed method – Method flow



Proposed method – Database

1.Squat

3.Push-up

5. Standing

2.Pull-up

4.Sit-up

6. Biceps-curl

Proposed method – Database

7. Bulgarian-split-squat

9. Lateral-raise

11. Dumbbell-rowing

8. Bench-press

10. Overhead-press

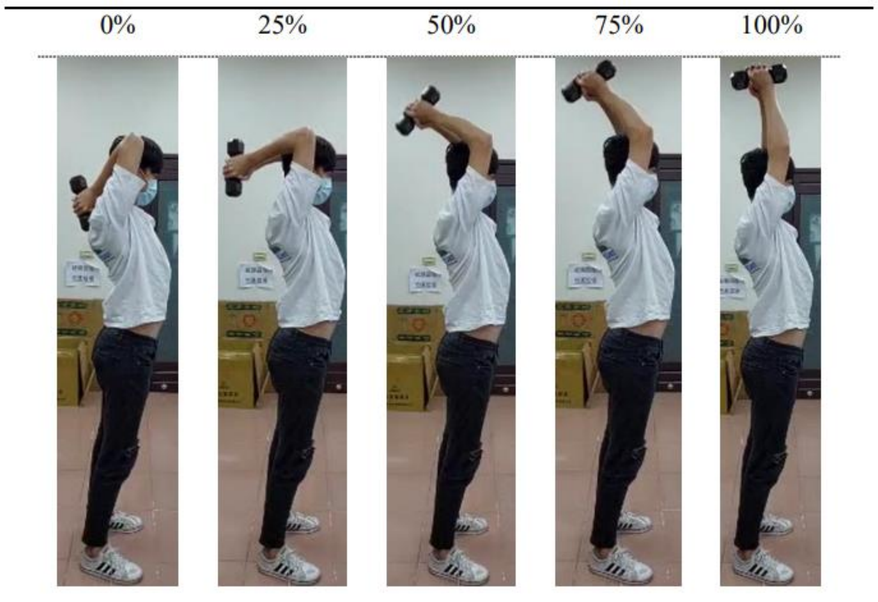
12. Triceps-extension

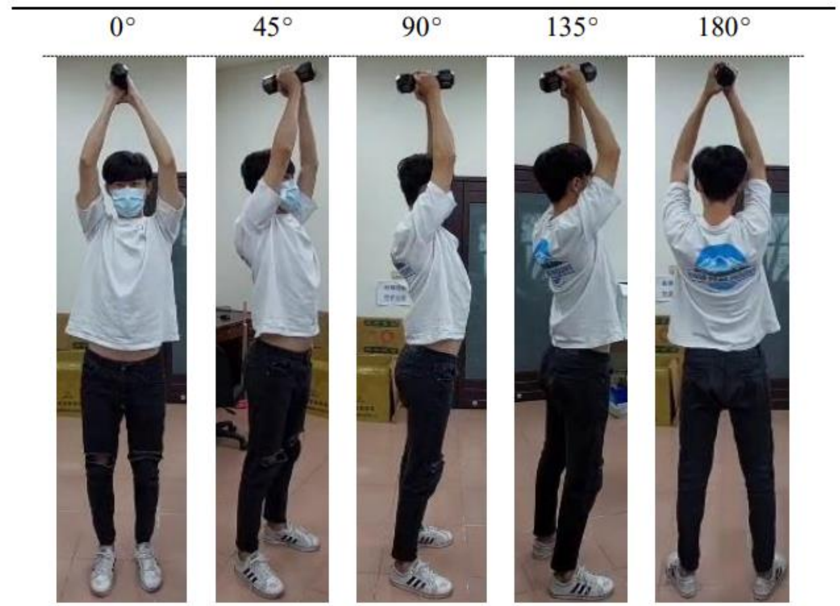
Proposed method – Training database

- We collected images containing 12 fitness movements.
- The images of the fitness movement are taken from 10 users.
- The database contains 12 types fitness movements by 10 users.

Proposed method – Training database

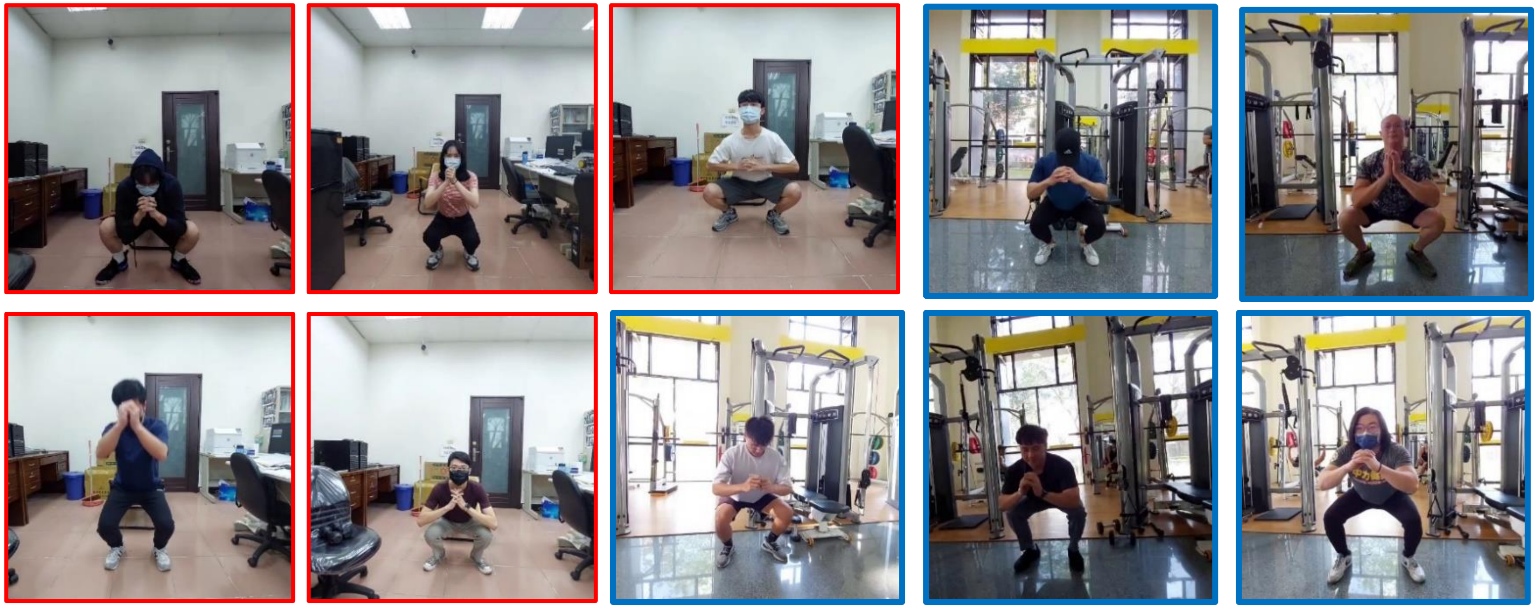
- The images of each step of each fitness movement will be added to the database.
- Each user will be photographed with more than 3 angles of fitness movements.





Proposed method – Test database

- 10 test-videos from 10 users
- Test-video database from 10 users:
  - 5 videos from the gym
  - 5 videos from the home room
- Each user does 12 type fitness movements.



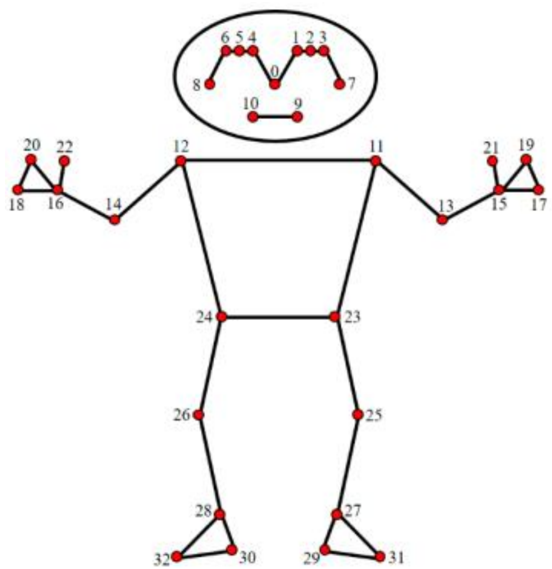
Training and testing for Yolov4

- Training-images contains Online images and images of 10 users.
- Testing-images contains images of 10 users.
  - Training-images: 12301
  - Testing-images: 3823

Fitness	L_{tr} (online)	L_{tr}	L_{tt}
Squat	198	644	121
Pull-up	239	1442	427
Push-up	317	913	264
Sit-up	373	977	328
Standing	132	1065	454
Biceps-curl	273	405	154
Bulgarian-split-squat	311	577	365
Bench-press	304	924	471
Lateral-raise	162	299	152
Overhead-press	202	724	365
Dumbbell-rowing	305	598	347
Triceps-extension	148	769	384
Total	2964	9337	3823

Combining Yolov4 with Mediapipe

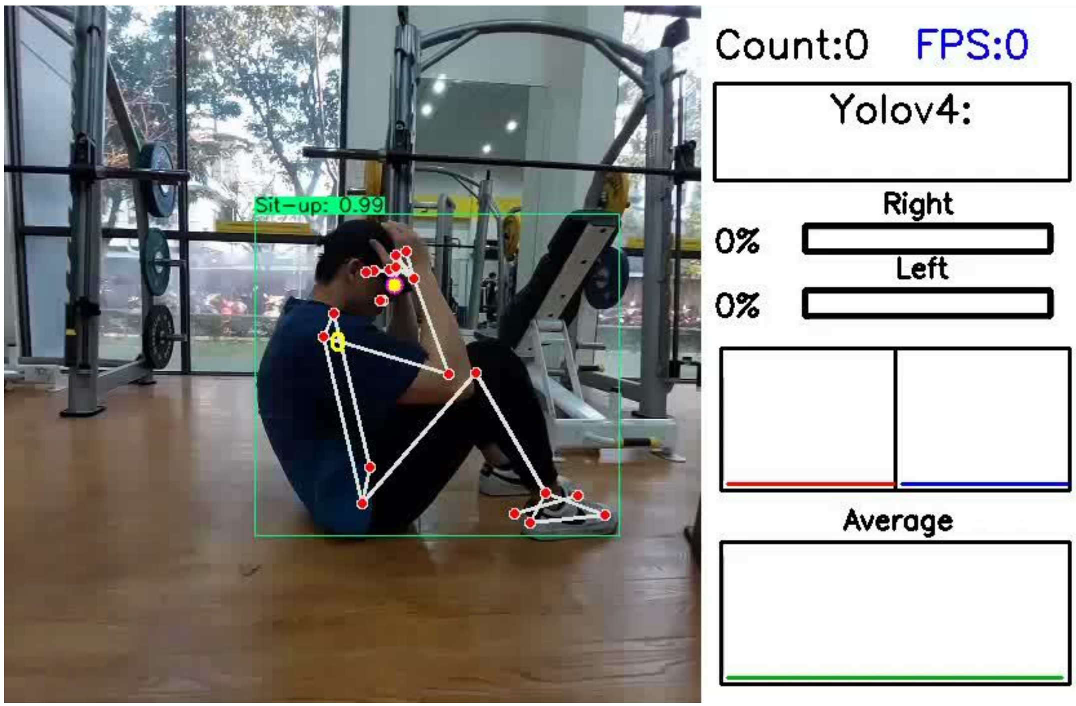
- After completing the Yolov4 detection, use Mediapipe to detect the nodes of the human body.
- Then calculate the node of interest (NoI) of each fitness movement, and store the one-dimensional signal of NoI to create a database *W*



0. Nose
1. Left_eye_inner
2. Left_eye
3. Left_eye_outer
4. Right_eye_inner
5. Right_eye
6. Right_eye_outer
7. Left_ear
8. Right_ear
9. Mouth_left
10. Mouth_right
11. Left_shoulder
12. Right_shoulder
13. Left_elbow
14. Right_elbow
15. Left_wrist
16. Right_wrist
17. Left_pinky
18. Right_pinky
19. Left_index
20. Right_index
21. Left_thumb
22. Right_thumb
23. Left_hip
24. Right_hip
25. Left_knee
26. Right_knee
27. Left_ankle
28. Right_ankle
29. Left_heel
30. Right_heel
31. Left_foot_index
32. Right_foot_index

Yolov4 and Mediapipe result

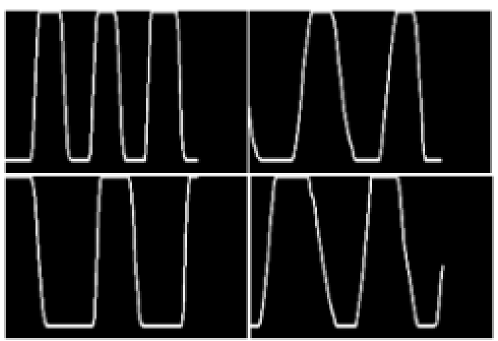
Yolov4 and mediapipe result video



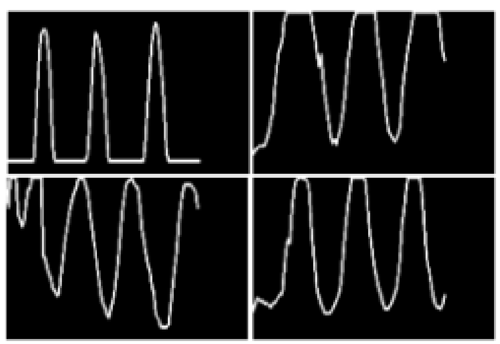
Training and testing for MLP

- These signals are divided into Complete, No-Complete, No-movement, and divided into training and testing.

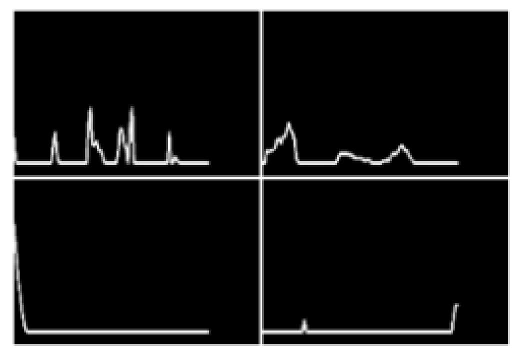
	W_{tr}	W_{tt}
Complete	223	203
No-Complete	372	261
No-movement	62	123
Total	657	587



Complete



No-Complete



No-movements

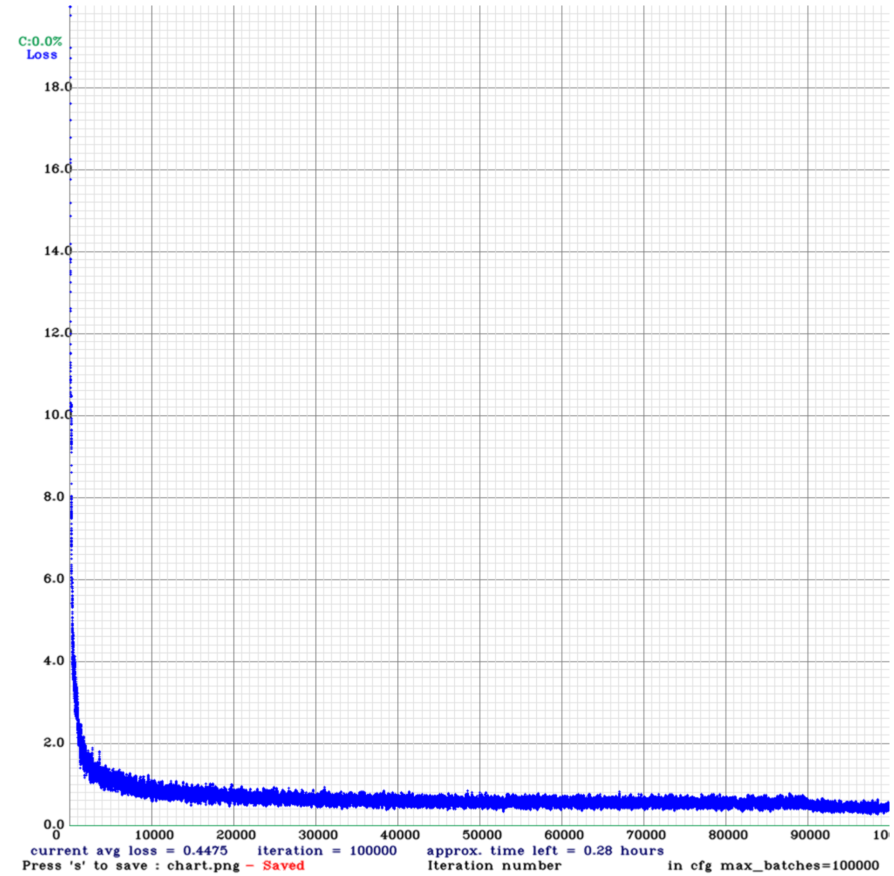
Evaluation index

- Training time
- Accuracy
- mAP
- Precision
- Recall
- F1-score

- The computer used is:
  - CPU: i7-1185G7
  - GPU: GTX-1650Ti

Experimental setup for Yolov4

- Yolov4
  - batchsize=64
  - Imagesize=416*416
  - Class=12
  - Iteration=10000~100000
  - Filter=51 (filter = (class+5)*3)
  - classes = 12
- Result
  - Training time is 253 hours.
  - The final loss is 0.4475
  - Iteration is 100000.
  - Total image: 12301
  - Users: 10



Experimental setup for MLP

- Class
  - Complete
  - No-Complete
  - No-movement

Parameters	Value
Class	3
Batch size	64
Data size	100
Max batches	100
Learning rate	0.001
Decay	0.0001

Experimental results

- Using the results from the testing set.
- When iteration is 50000, get the best results.

IoU = 0.5 、Total=3832									
Iteration	TP	FP	FN	IoU	mAP	Precision	Recall	F1-score	Training_Time(Ours)
10000	3552	431	280	75.90%	99.13%	89.18%	92.69%	90.90%	24
20000	3762	84	70	85.80%	99.74%	97.82%	98.17%	97.99%	48
30000	3653	187	179	85.33%	99.55%	95.13%	95.33%	95.23%	72
40000	3746	143	86	87.15%	99.89%	96.32%	97.76%	97.03%	96
50000	3777	81	55	87.90%	99.71%	97.90%	98.56%	98.23%	120
60000	3653	226	179	82.23%	99.72%	94.17%	95.33%	94.75%	147
70000	3387	350	445	81.67%	97.38%	90.63%	88.39%	89.50%	176
80000	3672	249	160	85.24%	99.89%	93.65%	95.82%	94.72%	204
90000	3560	64	272	88.93%	99.63%	98.23%	92.90%	95.49%	228
100000	3757	70	75	91.64%	99.94%	98.17%	98.04%	98.11%	253

Results and Comparison

- The result of timely classification of fitness movements using Yolov4, Accuracy is 98.56%.
- Compared to other methods, we get the highest performance.

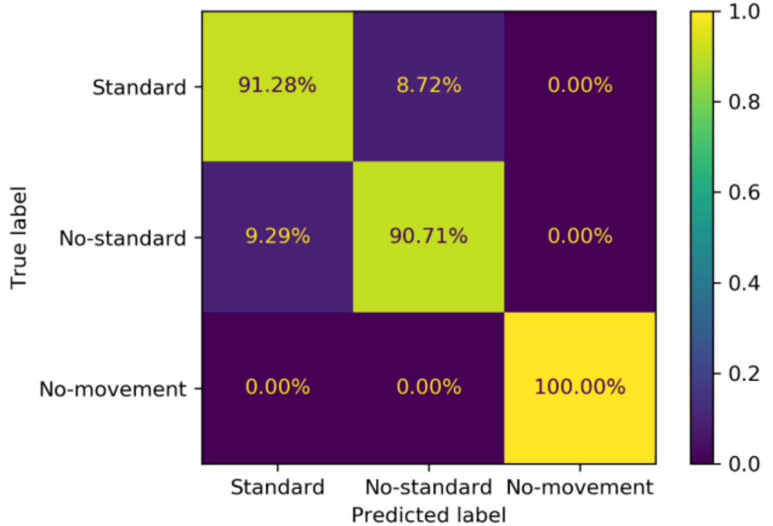
Evaluation index	IoU Threshold=0.5	IoU Threshold=0.75
mAP	99.71%	99.08%
Accuracy	98.56%	96.97%
Precision	97.90%	98.67%
Recall	98.56%	96.97%
F1-score	98.23%	97.82%

Evaluation index	Accuracy
Ours	98.56%
Yongpan Zou et al[31]	96.07%
C. Crema et al[32]	94.36%
Ali Bidaran et al[14]	92.9%

Results and Comparison

- The classification result of fitness completeness using MLP is 92.84%.
- Get higher performance compared to other methods.

Evaluation index	MLP
Accuracy	92.84%
Precision	92.85%
Recall	92.84%
F1-score	92.83%



Evaluation index	Accuracy
Ours	92.84%
Yongpan Zou et al[29]	90.7%
Jiangkun Zhou et al[11]	59.7%

Conclusion

- This paper proposes a deep transfer learning method, establishes a complete fitness database for training Yolov4, and combined with Mediapipe.
- With sufficient training data, Detection can achieve quite high accuracy and speed, Accuracy is 98.56%, FPS is 17.5.
- Using MLP to classify fitness completeness, Accuracy is 92.84%.
- Our proposed method achieves highest performance on classification and completeness detection of fitness.

Thank you for your attention

Vision Computing Platform Cluster : Dr. Jungpil SHIN
The University of AIZU, JAPAN
E-mail: *jpshin@u-aizu.ac.jp*

会津大学
THE UNIVERSITY OF AIZU