

**Spoof Detection for Automatic Speaker Verification
Third Meeting
Report**

I. Organizer:

Name:	Prof Masashi Unoki
Position:	Professor
Institution:	Japan Advanced Institute of Science and Technology

II. Program:

Date: 2-4 October 2024
Venue: Japan Advanced Institute of Science and Technology
1-chōme-1 Asahidai, Nomi, Ishikawa 923-1211, Japan Jawa

Program Agenda:

2 October 2024 – ASEAN-IVO Technical Meeting	
09:30 - 10:00	Registration
10:00 - 10:10	Welcome remarks by Prof. Masashi Unoki
10:10 - 10:20	Welcome remarks and introduction by ASEAN-IVO Project leader (Dr. Kasorn Galajit)
10:20 - 10:30	Self-Introduction (everyone)
10:30 - 10:40	Photo session
10:40 - 11:00	NECTEC's research progress (Dr. Kasorn and Dr. Pakinee) Title: ThaiSpoof improvement: Spoof detection in Thai Language
11:00 - 11:30	Break
11:30 - 11:50	SIIT's research progress (Dr. Sasiporn) Title: Automatic Speaker Verification using Deep Learning
11:50 - 12:20	NICT's research progress (Dr. Xugang Lu) Title: Cross-domain spoof speech detection
12:20 - 13:30	Lunch
13:30 - 14:00	UCSY's research progress (Dr. Win Pa Pa) UCSYSpoof: A Myanmar Language Dataset for Voice Spoofing
14:00 - 14:30	Presentation from UCSY's student (Hay Mar Soe Naing) Analysis of Pathological Features for Spoof Detection

14:30 - 15:00	UBD's research progress (Dr. Yassin) Title: Employing Transfer Learning Enabled Convolutional Neural Networks for Deep Fake Voice Detection
15:00 - 15:30	Presentation from UBD's student (Navod Nieranjan Thilakarathne) Title: Exploring a Cutting-Edge Convolutional Neural Network for Speech Emotion Recognition
15:30 - 16:00	Break
16:00 - 16:30	ITB research progress (Dr. Dessi Puji Lestari) Title: An Indonesian Language Dataset for Voice Spoofing
16:30 - 17:00	Presentation from ITB's student (Sarah Azka) Title: Initial Experiments on Detecting Voice Spoofing in Indonesian Using CNN-Based Models
17:00 - 17:30	Discussion + Q/A
3 October 2024 – ASEAN-IVO Technical Meeting	
10:00 – 10:25	JAIST's research progress (Prof. Masashi Unoki) Title: Deepfake Speech Detection: Approaches from Acoustic Features Related to Auditory Perception to DNNs
10:25 – 10:50	JAIST's research sharing (Dr. Li Kai) Title: Contributions of Jitter and Shimmer in the Voice for Fake Audio Detection
10:50 - 11:00	Break
11:00 - 11:30	JAIST's students research sharing (Zaman) Title: A Survey of Audio Classification Using Deep Learning
11:30 - 12:00	JAIST's students research sharing (Anuwat) Title: Potential of Speech-pathological Features for Deepfake Speech Detection
12:00 - 13:30	Lunch
13:30 - 14:00	JAIST's research progress (Dr. Candy Olivia Mawalim) Recent progress in ASVspoof challenges
14:00 - 14:30	Presentation from JAIST's student (Aulia Adila) Uncovering Spoof Voices in Non-Native Asian Speech
14:30 - 15:00	Coffee Break
15:00 - 16:00	Group discussion on current research issues in deepfake detection
16:00 - 16:15	Meeting conclusion for press tracking and discussion part

16:15 - 18:00	Each institute prepares for a showcase that they lead on the last day
18.00 -20.30	Dinner host by JAIST and NECTEC
4 October 2024 – JAIST-ASEAN Deepfake Voice Detection Hub Symposium	
10:00 - 10.40	Lab visiting, dividing members into 2 groups. First group visit Unoki lab and lab facilities at 9 th floor, second group visit Okada lab at the first floor.
10.40 - 11:20	Switch between group: First group visit Okada lab and second group visit Unoki lab
11.20 -11.30	Break
11:30 - 12:00	Opening remark and introduction to JAIST-ASEAN Deepfake Voice Detection Hub (Dr. Candy Olivia Mawalim)
12.00 - 12:30	Board members meeting: Future collaboration research proposal discussion
12:30 - 12:30	Lunch
13:30 - 15:00	Keynote session (Dr. Jessada) Spoof Detection for Automatic Speaker Verification (Dr. Waree) Research Sharing From SIIT, Thailand
15:00 - 15:30	Break
15:30 - 17:00	Presentation and demonstration open to JAIST members (All members) Introduction – overview of project (Dr. Candy)(15 min) Thai representative (Dr.Kasorn)(15 mins) Indonesia representative (Sarsh)(15 mins) Myanmar representative (Dr WinPaPa)(15 mins) Brunei representative (Dr Yassin)(15 mins) Japan representative (Dr Unoki) (15 mins)
17:10 - 17:20	Closing remark (Dr Unoki)

III. Participants:

No.	Name	Organization	Remark
1	Masashi Unoki	JAIST	
2	Candy Olivia Mawalim	JAIST	
3	Kai Li	JAIST	

4	Khalid ZAMAN	JAIST	
5	Nopparut LI	JAIST	
6	Aulia Adila	JAIST	
7	Shogo Okada	JAIST	
8	Anuwat Chaiwongyen	JAIST	
9	Hung	JAIST	Help in technical issues
10	Pakinee Aimmanee	SIIT	
11	Waree Kongprawechnon	SIIT	Self-support
12	Sasiporn Usanavain	SIIT	
13	Kasorn Galajit	NECTEC	
14	Jessada Karnjana	NECTEC	Self-support
15	Suradej Duangpummet	NECTEC	
16	Win Pa Pa	UCSY	
17	Hay Mar Soe Naing	UCSY	
18	Dk Hayati Pg Hj Mohd Yassin	UBD	
19	Navod NERANJAN Thilakarathne	UBD	
20	Xugang Lu	NICT	Self-support
21	Dessi Puji Lestari	ITB	
22	Sarah Azka Arief	ITB	
23	Wang Yan -Secretary of NICT	NICT	Self-support
24	Surasak Boonkla	NECTEC	
25	Khaing Zar Mon	SIIT	Online
26	Pantarat Vichathai	SIIT	Online
27	Puchit Bunpleng	SIIT	Online
28	Patharapol Laolakkana	SIIT	Online
29	Kosin Kalarat	SIIT	Online
30	Myat Aye Aye Aung	UCSY	Online
31	Aye Mya Hlaing	UCSY	Online

We have 23 members are onsite, and 8 participants are online.

Since the last day is open session, we also have the audience from JAIST both professor and student to join the symposium, keynote speaker and showcases session.

IV. Summary of the activities corresponding to the objectives.

A. Objective of this event

The Third Meeting of the ASEAN Language Project aims to achieve two key objectives: progress tracking and demonstration.

The first activity is progress tracking. We have three sub- objectives as follows,

Research sharing: We will share the research completed by each institution. Therefore, each institute can learn from other experiences, as well as give advice (Oct 2 whole day).

Collaborative Problem-Solving: Members will discuss any obstacles encountered during their work and brainstorm solutions together (Oct 3 afternoon).

Following Up on Outputs: The meeting will also review deliverables submitted by each institute (Oct 3 afternoon).

The second activity is demonstration. The goal is to demonstrate the dataset preparation for ASEAN language. We have three sub- objectives as follows,

Unoki Lab Visit: Participants will visit the Unoki Laboratory (Acoustics Information Science) at JAIST. This lab offers a unique opportunity to learn about standard data preparation techniques due to its specialized facilities, including a soundproof room for clean speech recordings and high-quality recording equipment. These resources are not readily available at NECTEC or other member institutes. Therefore, members will study dataset preparation using soundproof room facilities. (Oct 3 morning).

Showcase Dataset Preparation Process: Dr. Candy will give a demonstration on preparing datasets for ASEAN languages. The clean dataset preparation will be demonstrated in the soundproof room (Oct 3 morning). The spoof dataset is provided using different techniques. Only NECTEC and JAIST have experience with spoof dataset preparation. NECTEC and JAIST will teach others about dataset preparation. This part will be held as a workshop because participants need to implement and experience the preparation (Oct 4 whole day).

Deepfake Voice Detection Research: A workshop will cover research on deepfake voice detection specific to Asian languages. Before the meeting each institute is assigned to study deepfake techniques, and those techniques will be shared among members (Oct 4 whole day).

B. Activities corresponding to the objective.
The Summary of the objective and its status

Objective	Status
1. Progress tracking	Done
Research Sharing	Done
Collaborative Problem-solving	Done
Following up on Output	Done
2. Demonstration	Done
Unoki and Okada Lab visiting	Done
Showcase dataset preparation process	Done
Deepfake detection research	Done

Activities corresponding to the objective.

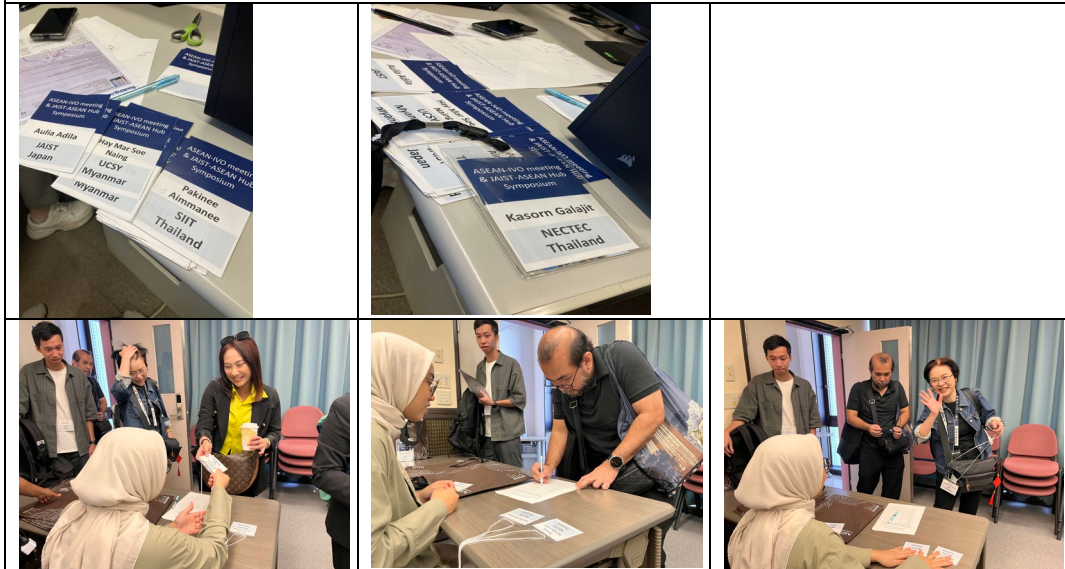
1. Objective 1 – Progress tracking.

The meeting was held in hybrid mode on the first day and second day morning.

The second day afternoon was closed session and the last day we open for audience.

The first day meeting start with registration

The meeting registration



Professor Masashi UNOKI and following self-introduction.




Warm welcome and Self-introduction and Photo session



The project status report by project leader Dr Kasorn GALAJIT to update about the project. From the project launched, now project have 33 members in total from 5 countries and 8 institute. We introduce new members and report one member from CU has already retired from his career. Prof Dessi from Indonesia ask whether we do still can add more members since there are someone that work for her on this project. We say yes to her and inform to member that someone can be only member under institute, and individual person cannot be member of project. That is if there is any member of project graduate from their institute, we should withdraw the membership. The project leader summarized the scope of the project to members and on-going activities. She presented the project detail, updated the status and expected output f to the audience.

Sub-objective 1.1 – Research Sharing.

After that there are a presentation from each institute for research sharing

<u>Research sharing</u>	
Dr kasorn and Dr Pakinee from NECTEC	Dr kasorn and Dr Pakinee from NECTEC
	
Dr Sasiporn From SIIT	Dr Xugang Lu from NICT
	
Dr Win PaPa from UCSY	Hay Mar Soe Naing from UCSY



Sub-objective 1.2 – Collaborative Problem-solving

Members will discuss any obstacles encountered during their work and brainstorm solutions together.

Here is the following issue from what we discussed.

1. Can we collect data by our own?

Answer. Now our data is gathered from internet mostly which we face a problem of target group (native, non-native speaker, environment of data collection).

Answer : we can do this, but we must hire the third parties to do that since we cannot pay to participant who donate voice. So, if any country has plan to collect data themselves- please clarify the budget and plan and submit to

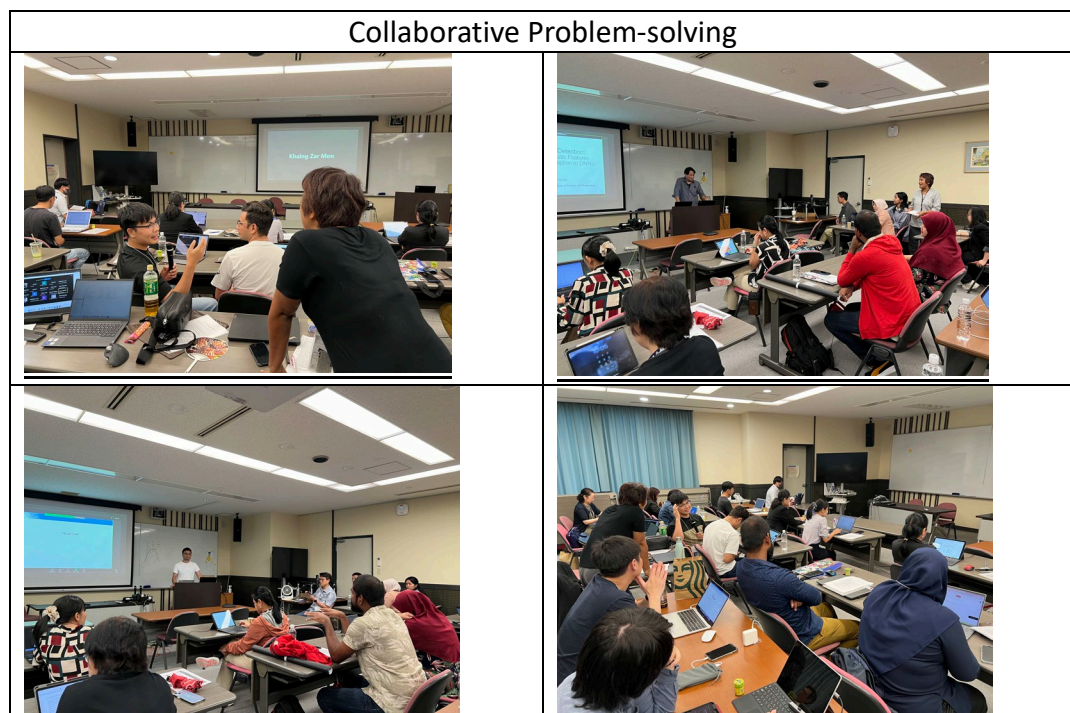
project leader

2. Since data collection need times as well as to create dataset since we have budget left can this project continue for one year.

Answer : Project members discussed and agree on the conclusion that we should ask for project extension for one year.

3. Now we have dataset, but it is still individually, how to share the spoofed technique and combine data to be ASEAN language dataset

Answer : we will set the online meeting more often to share the spoof techniques and dataset combination.



Sub-objective 1.3 – Following up on output.

The spoof detection project target 10 Conference papers and 2 journals within 2 years. From last year on ASEAN IVO Forum, we report that we got 9 Conference papers and 1 journal. List as follows

Conference

No:	Paper title:	Author names	Affiliation	Conference name:	The date of the conference	The venue of the conference
1	ThaiSpoof: A Database for Spoof Detection in Thai Language	<ul style="list-style-type: none"> Kasorn Galajit, Thunpist Kosolruiwatt, Candy Olivia Mawallim, Pakinee Aimanee, Waree Kongprawechnon, Win Pa Pa, Anuwat Chaivongyen, Teeradai Racharak, Havati Yassin, Jessada Karnjana, Surasak Boonkha, Masashi Unoki, 	<ul style="list-style-type: none"> NECTEC, National Science and Technology Development Agency, Sirindhorn International Institute of Technology (SIIT), Japan Advanced Institute of Science and Technology, University of Computer Studies,Yangon, Universiti Brunei Darussalam, Brunei Darussalam 	The 18th International Joint Symposium on Artificial Intelligence and Natural Language Processing (ISAI-NLP 2023) (https://isai-nlp-aiot2023.aiat.or.th/)	27-29/11/2023	Bangkok, Thailand
2	Spoof Detection using Voice Contribution on LFCC features and ResNet-34	<ul style="list-style-type: none"> Khaine Zar Mon, Kasorn Galajit, Candy Olivia Mawallim, Jessada Karnjana, Tsuyoshi Isshiki, Pakinee Aimanee, 	<ul style="list-style-type: none"> Sirindhorn International Institute of Technology, NECTEC, National Science and Technology Development Agency, Japan Advanced Institute of Science and Technology, Tokyo Institute of Technology 	The 18th International Joint Symposium on Artificial Intelligence and Natural Language Processing (ISAI-NLP 2023) (https://isai-nlp-aiot2023.aiat.or.th/)	27-29/11/2023	Bangkok, Thailand
3	Using Novel Hybrid Convolutional Neural Network for Dysarthria Diagnosis	<ul style="list-style-type: none"> Navod Neranian, Thilakarathne, Kasorn Galajit, Jessada Karnjana, Win Pa Pa, Candy Olivia Mawali, Havati Yassin, 	<ul style="list-style-type: none"> Universiti Brunei Darussalam, Brunei Darussalam NECTEC, National Science and Technology Development Agency, Japan Advanced Institute of Science and Technology, University of Computer Studies,Yangon 	The 10th IEEE CSDE 2023, the Asia-Pacific Conference on Computer Science and Data Engineering 2023,(IEEE CSDE 2023) https://ieee-csde.org/csde2023/	4-6/12/2023	Yanuca Island , Fiji
No:	Paper title:	Author names	Affiliation	Conference name:	The date of the conference	The venue of the conference
4	Speech Watermarking for Tampering Detection Using Singular Spectrum Analysis with a Psychoacoustic Model	<ul style="list-style-type: none"> Phondanai Khanti, Pannathorn Sathirasattayanon, Pathranit Kaewcharuay, Nanthayod Termkoh, Ekachai Phaisangittisagul, Kasorn Galajit, Jessada Karnjana 	<ul style="list-style-type: none"> Sirindhorn International Institute of Technology, NECTEC, National Science and Technology Development Agency, Kasetsart University 	The 26th Conference of the Oriental COCOSDA https://www.ococosda2023.com/	4-6/12/2023	Delhi, India
5	A Large Vocabulary End-to-End Myanmar Automatic Speech Recognition	<ul style="list-style-type: none"> Hay Mar Soe Naing Win Pa Pa 	<ul style="list-style-type: none"> University of Computer Studies,Yangon, 	M3Oriental Workshop of ACM Multimedia Asia 2023 The ACM Multimedia Asia 2023 https://sites.google.com/view/m3oriental	8/12/2023	Tainan city, Taiwan
6	Deepfake-speech detection with pathological features and multilayer perceptron neural network	<ul style="list-style-type: none"> Anuwat Chaivongyen, Suradej Duangpummert, Jessada Karnjana, Waree Kongprawechnon, Masashi Unoki 	<ul style="list-style-type: none"> Japan Advanced Institute of Science and Technology, NECTEC, National Science and Technology Development Agency, Sirindhorn International Institute of Technology, 	The 15th annual conference organized by Asia-Pacific Signal and Information Processing Association (APSIPA2023) https://www.apsipa2023.org/	31/11/2023-3/12/2023	Teipei , Taiwan
No:	Paper title:	Author names	Affiliation	Conference name:	The date of the conference	The venue of the conference
7	M-Diarization: A Myanmar Speaker Diarization using Multi-scale dynamic weights	<ul style="list-style-type: none"> Myat Aye Aye Aung, Win Pa Pa, Hay Mar Soe Naing. 	<ul style="list-style-type: none"> University of Computer Studies,Yangon,, 	The 26th Conference of the Oriental COCOSDA https://www.oco-cosda2023.com/	4-6/12/2023	Delhi, India

Journal

Published Journal Papers:

No:	Paper title:	Author names	Affiliation	Journal name:	The publisher of the Journal	The volume number and Pages
1	Contributions of Jitter and Shimmer in the Voice for Fake Audio Detection	KAI LI, XUGANG LU, MASATO AKAGI, MASASHI UNOKI.	<ul style="list-style-type: none"> Japan Advanced Institute of Science and Technology, Advanced Speech Technology Laboratory, National Institute of Information and Communications Technology, 	IEEE Access	IEEE Access	VOLUME 11, 2023 Received 21 June 2023, accepted 29 July 2023, date of publication 3 August 2023, Digital Object Identifier 10.1109/ACCESS.2023.3301616

Therefore, we still need 1 Conference paper and 1 journal. Here is the update list from each institute.

Journal papers

JAIST

- Li, Kai, Xugang Lu, Masato Akagi, and Masashi Unoki. "Contributions of Jitter and Shimmer in the Voice for Fake Audio Detection." *IEEE Access* (2023).
- Zaman, Khalid, Melike Sah, Cem Direkoglu, and Masashi Unoki. "A survey of audio classification using deep learning." *IEEE Access* (2023).
- Ota, Yasuji, and Masashi Unoki. "Anomalous sound detection for industrial machines using acoustical features related to timbral metrics." *IEEE Access* (2023).

SIIT

- Chaiwongyen, A., Duangpummet, S., Karnjana, J., Kongprawechnon, W. and Unoki, M., 2024. Potential of Speech-pathological Features for Deepfake Speech Detection. *IEEE Access*.

UCSY

- Myat Aye Aye Aung, Hay Mar Soe Naing, Win Pa Pa, "[End-to-End Neural Diarization for Unknown Number of Speakers with MultiScale Decoder](#)", International Journal of Intelligent Engineering and Systems, Vol.17, No.5, 2024

Conference papers

JAIST

- Chaiwongyen, Anuwat, Norranat Songsriboonsit, Suradej Duangpummet, Jessada Karnjana, Waree Kongprawechnon, and Masashi Unoki. "Contribution of timbre and shimmer features to deepfake speech detection." In *2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 97-103. IEEE, 2022.
- Li, Kai, Dung Kim Tran, Xugang Lu, Masato Akagi, and Masashi Unoki. "Data-driven Non-uniform Filterbanks Based on F-ratio for Machine Anomalous Sound Detection." In *2023 31st European Signal Processing Conference (EUSIPCO)*, pp. 201-205. IEEE, 2023.
- Cheng, Haowei, Candy Olivia Mawalim, Kai Li, Lijun Wang, and Masashi Unoki. "Analysis of Spectro-Temporal Modulation Representation for Deep-Fake Speech Detection." In *2023 Asia Pacific Signal and Information Processing*

Association Annual Summit and Conference (APSIPA ASC), pp. 1822-1829. IEEE, 2023.

UCSY

- Aye Mya Hlaing, Win Pa Pa, “Generative Adversarial Network based Neural Vocoder for Myanmar End-to-End Speech Synthesis”, ICNLSP2024, Oct 2024, Italy
- Hay Mar Soe Naing, Win Pa Pa, Aye Mya Hlaing, Myat Aye Aye Aung, Kasorn Galajit, Candy Olivia Mawalim, “UCSYSpooF: A Myanmar Language Dataset for Voice Spoofing Detection”, O-COCOSDA 2024, Oct 2024, Taiwan
- Myat Aye Aye Aung, Hay Mar Soe Naing, Aye Mya Hlaing, Win Pa Pa, Kasorn Galajit, Candy Olivia Mawalim, “Analysis of Pathological Features for Spoof Detection”

NECTEC

- Pantarat Vichathai, Puchit Bunpleng, Patharapol Laolakkana, Sasiporn Usanavasin, Phondanai Khanti, Kasorn Galajit, Jessada Karnjana, “Speech Watermarking for Tampering Detection using Singular Spectrum Analysis with Quantization Index Modulation and Psychoacoustic Model”, O-COCOSDA 2024, Oct 2024, Taiwan

UBD

- Thilakarathne, Navod Neranjan, Kasorn Galajit, Candy Olivia Mawalim, and Hayati Yassin. "Exploring a Cutting-Edge Convolutional Neural Network for Speech Emotion Recognition." In *2024 5th International Conference on Industrial Engineering and Artificial Intelligence (IEAI)*, pp. 110-116. IEEE, 2024.

2. Objective 2 – Demonstration. The goal is to demonstrate the dataset preparation for ASEAN language. We have three sub- objectives

Sub-objective 2.1 – Unoki and Okada lab visiting.

Participants visited the Unoki Laboratory (Acoustics Information Science) at JAIST. This lab offers a unique opportunity to learn about standard data preparation techniques due to its specialized facilities, including a soundproof room for clean speech recordings and high-quality recording equipment. These resources are not readily available at NECTEC or other member institutes. Therefore, members will study dataset preparation using soundproof room facilities. Unoki lab has 2 types of soundproof room one with for subjective evaluation which cost around 3 million yen and the one with non-reverberation which cost around 4 million yen. Therefor Miss Aulia Adila uses this facility to do on the dataset collected by JAIST side. Member can learn more about how to collect data with the proper method, As well as the guidance for data collection. From This visiting, Brunei and Indonesia team learn and manage to try for data

collection.

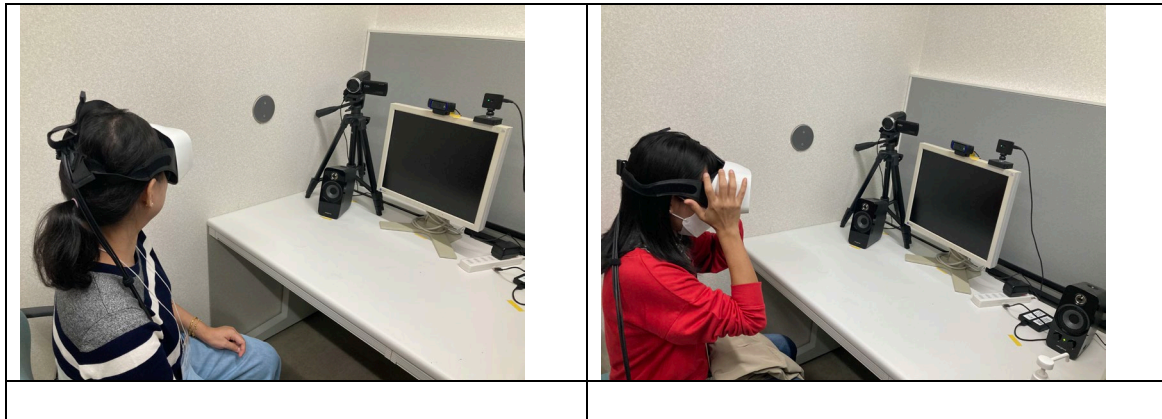
Unoki Lab Visiting



Participants visited the Okada Laboratory at JAIST. This lab offers a unique opportunity to learn about Multimodal Interaction and Social Signal Modeling. Therefore, we can study further for using speech processing together with another multimedia such as audio, vision, language for future study. Here Okada sensei demonstrate the visual multimodal meeting.

Okada Lab Visiting

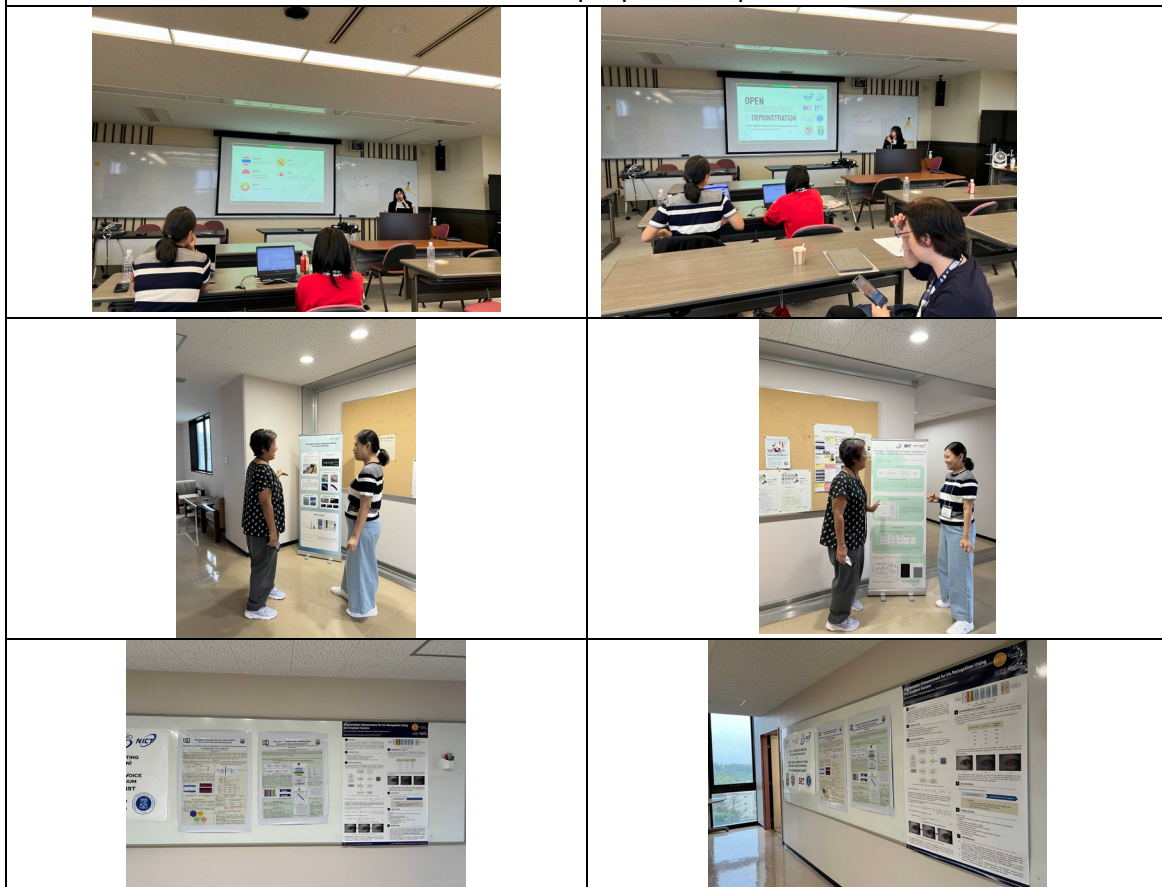




Sub-objective 2.2 – Showcase dataset preparation process.

Initially Dr Candy and Miss Audila shows the dataset provided using the soundproof room and team members demonstrate dataset preparation. This part hold as a workshop because participants need to implement and experience the preparation.

Showcase dataset preparation process





Sub-objective 2.3 – Deepfake Voice Detection Research.

This Session is open session, we announce to both JAIST students and professor. We have Dr Waree and Dr Jessada as keynote speaker to keep a talk about the Deepfake voice detection research. The representative from each countries showed their research to audience. Here, we have a chance to discuss with JAIST professor and student for future research. We set up some plan to submit proposal for funding in the future.

Deepfake Voice Detection Research



