

# Project Title:

## Background :

The human voice is one of the biometric indicators representing a person's identity. It has been successfully used in automatic speaker verification (ASV) systems in many applications. The ASV system is currently vulnerable to spoofing attacks in which someone disguises as another and illegitimately accesses a secure system. Hence, countermeasures against spoofing attacks are necessary to verify whether the claimed voice is a genuine or fake representation before verifying or identifying the speaker's identity. Therefore, this research focuses on spoofing detection in automatic speaker verification.

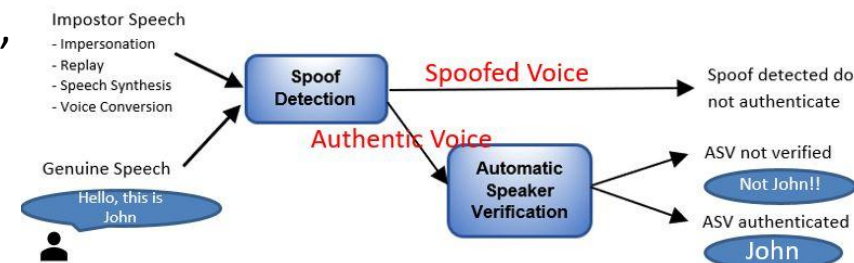
## Targets:

The objectives of this project are listed as follows:

- (1) To explore and investigate significant of speech features for spoof detection,
- (2) To optimize percentage of voice and non-voice segments in features used in spoofing detection,
- (3) To investigate pathological feature for spoof detection,
- (4) To minimize detection error,
- (5) To improve an accuracy of ASV, and
- (6) To study multi-lingual spoof detection.

## Speaker:

Kasorn GALAJIT(Project Leader and Speaker)  
National Electronics and Computer Technology Center, Thailand



# Project Title:

35 members , 5 countries, 8 institutes  
 (Thai – NECTEC, SIIT, Myanmar- USCY,NSPU, Brunei –UBD, Indonesia – IBD, Japan- NICT, JAIST)

## Project Members :

Name / Position / Institution	Name / Position / Institution	Name / Position / Institution	Name / Position / Institution	Name / Position / Institution
Masashi Unoki (PhD) / Professor / JAIST, Japan	Waree Kongprawechnon (PhD) / Assoc. Prof. / SIIT, Thailand	Nanthayod Termkoh / Undergrad student / SIIT, Thailand	Win Lai Lai Phyu / Asst. Lecturer / UCSY, Myanmar	Kasorn Galajit (PhD) / Project Leader / NECTEC, Thailand
Candy Olivia Mawalin (PhD) / Asst. Prof. / JAIST, Japan	Pakinee Aimmanee / Assoc. Prof. / SIIT, Thailand	Kosin Kalarat / PhD student / SIIT, Thailand	Myat Aye Aye Aung / Asst. Lecturer / UCSY, Myanmar	Suradej Duangpummet (PhD) / Researcher / NECTEC, Thailand
Kai Li / PhD / JAIST, Japan	Sasiporn Usanavasin(PhD) / Assoc. Prof. / SIIT, Thailand	Pantarat Vichathai/ Undergrad student / SIIT, Thailand	Win Pa Pa (PhD) / Professor / NSPU, Myanmar	Jessada Karnjana (PhD) / Researcher / NECTEC, Thailand
Anuwat Chaiwongyen / PhD / JAIST, Japan & SIIT, Thailand	Khaing Zar Mon / Master student / SIIT, Thailand	Puchit Bunpleng/ Undergrad student / SIIT, Thailand	Aye Mya Hlaing (PhD) / Lecturer / UCSY, Myanmar	Surasak Boonkla (PhD) / Researcher / NECTEC, Thailand
Khalid ZAMAN/ PhD Candidate / JAIST, Japan	Patthranit Kaewcharuay / Undergrad student / SIIT, Thailand	Patharapol Laolakkana/ Undergrad student / SIIT, Thailand	Hay Mar Soe Naing/ Lecturer/ UCSY, Myanmar	Phondanai Khanti / Assistant Researcher / NECTEC, Thailand
Aulia Adila / Master Student / JAIST, Japan	Ananda Garin Mills / Undergrad student/ SIIT, Thailand	Dk Hayati Pg Hj Mohd Yassin (PhD) / Lecturer / UBD, Brunei	Dessi Puji Lestari (PhD) / Professor / ITB, Indonesia	Xugang Lu (PhD) / Senior Researcher / NICT, Japan
Nopparut Li / Master Student / JAIST, Japan	Pannathorn Sathirasattayanon / Undergrad student / SIIT, Thailand	Navod NERANJAN Thilakarathne / PhD Candidate / UBD, Brunei	Sarah Azka Arief/ Master student/ ITB, Indonesia	Kwanchiva Thangthai (PhD) / Project Leader / NECTEC, Thailand

## Project Duration :

3 years, 1 April 2023 – 31 March 2026

## Project Budget:

80,000 USD (40,000 USD per year)

- **NECTEC**, JAIST, SIIT, UBD make the improvement of “ThaiSpoof” Dataset for spoof detection for Thai language
- Incorporating new spoofed voice data generated using the multilingual speech (MMS) model. Researchers can use this dataset to study spoof detection
- **SIIT, JAIST**, NECTEC Study “Potential of Speech-pathological Features for Deepfake Speech Detection”  
Investigate the significant of pathological feature for deepfake detection
- **UBD**, NECTEC, JAIST, study Exploring a Cutting-Edge Convolutional Neural Network for Speech Emotion Recognition.  
This study explorer the emotion from speech signal.

- **UCSY**, NECTEC, JAIST create UCSYSpoof: A Myanmar Language Dataset for Voice Spoofing Detection  
The study gives knowledge of voice spoofing in Myanmar language  
Also Pathological studies for Myanmar language
- **JAIST**, NECTEC, SIIT study Deepfake-speech detection with pathological features and multilayer perceptron neural network  
This study can show the significant of pathological for spoof detection
- **ITB** JAIST study Anomalous Machine Sound Detection Based on Time Domain Gammatone Spectrogram Feature and IDNN Model
- This study explorer the feature of the sound in time domain to find the anomalous machine sound detection.

# Project Activities: Budget plan in detail

**Budget:** 40,000 USD/year

Meeting/workshop (2 times a year)

IVO Forum

Publication, presentation at international conferences

Kick-off meeting **2,500USD** (Thailand May 2023) ,

Second meeting **6,000USD** (Indonesia Dec 2023)

Third meeting **12,600USD** (Japan Oct 2024)

Forth meeting (Tentative on Indonesia Feb 2026)

**Publication (First year)**

ISAI 2 papers 800USD

SCDE 1 Paper 2,700USD

Cocosda 2 papers 2,1260USD

ACM 1 paper 310USD

APSIPA2023 1 paper 0 USD

IEEE Access 1 Journal 0 USD

Total **5970 USD**

**Publication (2nd year)**

ISAI Thailand 2 papers 400USD

ICNLSP 1 Paper 350USD

Cocosda 3 papers 2,020USD

APSIPA2025 2,000USD

M3 conference 5,000USD

IEEE Access 1 Journal 2,000 USD

Total **11,770 USD**

**Approximately used budget**

- First year Total **15,400 USD**
  - Meeting 8,500
  - Publication 5,970
  - Forum 1,000
- 2nd year Total **25,370 USD**
  - Meeting 12,600
  - Publication 11,770
  - Forum 1,000
- 3rd year Total **0 USD**

**Publication (2nd year)**

ISAI Thailand 1 paper

ICAIT Myanmar 1 paper

AWIST Japan 1 paper



## JAIST-ASEAN Deepfake Detection Hub Symposium and ASEAN IVO Meeting



Additional funding from JAIST

[https://candyolivia.github.io/jaist\\_asean\\_symposium\\_2024/](https://candyolivia.github.io/jaist_asean_symposium_2024/)

- Talk on Deepfake Speech Detection: Approaches from Acoustic Features (Prof. Masashi Unoki)
- Discussions on the role of jitter and shimmer in fake audio detection (Dr. Li Kai)
- Student presentations on audio classification and speech-pathological features for deepfake detection (Zaman, Anuwat)
- Recent progress in ASVspoof challenges (Dr. Candy Mawalim, Aulia Adila)
- Group discussion on current research issues in deepfake detection (All participants)



Prof. Unoki's talk



JAIST student presentation



Group discussion



Welcome dinner

# R&D results:

R&D results: **Done**

- “ThaiSpoof ”: Dataset for spoof detection for Thai language  
This dataset is published on AI for Thai for sharing
- “UCSYSpoof”: Dataset for spoof detection for Myanmar language
- Dataset for spoof detection for Indonesian language

R&D results: **Ongoing**

- ASEAN language dataset : Dataset is already provided, now on model demonstration
- Webpage for spoof detection.

Choose File

PA\_D\_0000001.flac

Clear

0:00 / 0:04

Predict

Result of PA\_D\_0000001.flac

Model	Reference	Trained With	Reported Accuracy	Prediction	Confidence
PA HT20%	10.1109/ISAI-NLP60301.2023.10354625	ResNet34 + LFCC Feature: LFCC (57,600) Classifier: ResNet34 dataset: ASVSpoof 2019 SpoofType: Replay	EER: 0.41 ACC: 90.84 Precision: 90.86 Recall: 90.82 F1: 90.84 DCF: 9.16	Real	98.75%
LA HT20%	10.1109/ISAI-NLP60301.2023.10354625	ResNet34 + LFCC Feature: LFCC (57,746) Classifier: ResNet34 dataset: ASVSpoof 2019 SpoofType: Logical Access	EER: 0.41 ACC: 99.59 Precision: 99.59 Recall: 99.59 F1: 99.59 DCF: 0.41	Real	99.77%
LFCC-MMS+Genuine	10.1109/ISAI-NLP60301.2023.10354956	LCNN + LFCC-MMS + Genuine test Feature: LFCC Classifier: LCNN dataset: MMS + Genuine Spoof_Type: Mixed	EER: 0.34 ACC: 99.46 Precision: 98.64 Recall: 99.79 F1: 99.48 DCF: 0.26	Fake	69.10%
MFCC-MMS+Genuine	10.1109/ISAI-NLP60301.2023.10354956	LCNN + MFCC-MMS + Genuine Feature: MFCC Classifier: LCNN dataset: MMS + Genuine Spoof_Type: Mixed	EER: 0.14 ACC: 96.20 Precision: 99.61 Recall: 92.47 F1: 95.91 DCF: 0.11	Real	99.99%

# Scientific Contribution:

Presentations at International Conferences:

Expected output in 2 years

10 Conference papers

Current output

18 Conference papers

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"> <li>Kasorn Galajit,</li> <li>Thunpisit Kosolsriwiwat,</li> <li>Candy Olivia Mawalim,</li> <li>Pakinee Aimmanee,</li> <li>Waree Kongprawechnon,</li> <li>Win Pa Pa,</li> <li>Anuwat Chaiwongyen,</li> <li>Teeradaj Racharak,</li> <li>Hayati Yassin,</li> <li>Jessada Karnjana,</li> <li>Surasak Boonkla,</li> <li>Masashi Unoki,</li> </ul>	<ul style="list-style-type: none"> <li>NECTEC, National Science and Technology Development Agency,</li> <li>Sirindhorn International Institute of Technology,</li> <li>Japan Advanced Institute of Science and Technology,</li> <li>University of Computer Studies,Yangon,</li> <li>Universiti Brunei Darussalam},</li> <li>Brunei Darussalam</li> </ul>	The 18th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP 2023) ( <a href="https://isai-nlp-aiot2023.aiat.or.th/">https://isai-nlp-aiot2023.aiat.or.th/</a> )	27-29/11/2023	Bangkok, Thailand
2	Spoof Detection using Voice Contribution on LFCC features and ResNet-34	<ul style="list-style-type: none"> <li>Khaing Zar Mon,</li> <li>Kasorn Galajit,</li> <li>Candy Olivia Mawalim,</li> <li>Jessada Karnjana,</li> <li>Tsuyoshi Isshiki,</li> <li>Pakinee Aimmanee,</li> </ul>	<ul style="list-style-type: none"> <li>Sirindhorn International Institute of Technology,</li> <li>NECTEC, National Science and Technology Development Agency,</li> <li>Japan Advanced Institute of Science and Technology,</li> <li>Tokyo Institute of Technology Tokyo,</li> </ul>	The 18th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP 2023) ( <a href="https://isai-nlp-aiot2023.aiat.or.th/">https://isai-nlp-aiot2023.aiat.or.th/</a> )	27-29/11/2023	Bangkok, Thailand
3	Using Novel Hybrid Convolutional Neural Network for Dysarthria Diagnosis	<ul style="list-style-type: none"> <li>Navod NERANJAN Thilakarathne,</li> <li>Kasorn Galajit,</li> <li>Jessada Karnjana,</li> <li>Win Pa Pa,</li> <li>Candy Olivia Mawali</li> <li>Hayati Yassin,</li> </ul>	<ul style="list-style-type: none"> <li>Universiti Brunei Darussalam},</li> <li>Brunei Darussalam</li> <li>NECTEC, National Science and Technology Development Agency,</li> <li>Japan Advanced Institute of Science and Technology,</li> <li>University of Computer Studies,Yangon,</li> </ul>	The 10th IEEE CSDE 2023, the Asia-Pacific Conference on Computer Science and Data Engineering 2023,(IEEE CSDE 2023) <a href="https://ieee-csde.org/csde2023/">https://ieee-csde.org/csde2023/</a>	4-6/12/2023	Yanuca Island, Fiji



# Scientific Contribution:

Presentations at International Conferences:

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"> <li>Phondanai Khanti,</li> <li>Pannathorn Sathirasattayanon,</li> <li>Patthranit Kaewcharuay,</li> <li>Nanthayod Termkoh,</li> <li>Ekachai Phaisangittisagul,</li> <li>Kasorn Galajit,</li> <li>Jessada Karnjana</li> </ul>	<ul style="list-style-type: none"> <li>Sirindhorn International Institute of Technology,</li> <li>NECTEC, National Science and Technology Development Agency,</li> <li>Kasetsart University</li> </ul>	The 26th Conference of the Oriental COCOSDA <a href="https://www.ococosda2023.com/">https://www.ococosda2023.com/</a>	4-6/12/2023	Delhi, India
5	A Large Vocabulary End-to-End Myanmar Automatic Speech Recognition	<ul style="list-style-type: none"> <li>Hay Mar Soe Naing</li> <li>Win Pa Pa</li> </ul>	<ul style="list-style-type: none"> <li>University of Computer Studies,Yangon,</li> </ul>	M3Oriental Workshop of ACM Multimedia Asia 2023 The ACM Multimedia Asia 2023 <a href="https://sites.google.com/view/m3oriental">https://sites.google.com/view/m3oriental</a>	8/12/2023	Tainan city, Taiwan
6	Deepfake-speech detection with pathological features and multilayer perceptron neural network	<ul style="list-style-type: none"> <li>Anuwat Chaiwongyen,</li> <li>Suradej Duangpummet,</li> <li>Jessada Karnjana,</li> <li>Waree Kongprawechon,</li> <li>Masashi Unoki</li> </ul>	<ul style="list-style-type: none"> <li>Japan Advanced Institute of Science and Technology,</li> <li>NECTEC, National Science and Technology Development Agency,</li> <li>Sirindhorn International Institute of Technology,</li> </ul>	The 15th annual conference organized by Asia-Pacific Signal and Information Processing Association (APSIPA2023) <a href="https://www.apsipa2023.org/">https://www.apsipa2023.org/</a>	31/11/2023-3/12/2023	Teipei, Taiwan

# Scientific Contribution:

Presentations at International Conferences:

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	Myat Aye Aye Aung, Win Pa Pa, Hay Mar Soe Naing.	<ul style="list-style-type: none"> <li>University of Computer Studies,Yangon,,</li> </ul>	The 26th Conference of the Oriental COCOSDA <a href="https://www.oco-cosda2023.com/">https://www.oco-cosda2023.com/</a>	4-6/12/2023	Delhi, India
8	Exploring a Cutting-Edge Convolutional Neural Network for Speech Emotion Recognition.	Thilakarathne, Navod Neranjan, Kasorn Galajit, Candy Olivia Mawalim, and Hayati Yassin.	<ul style="list-style-type: none"> <li>Universiti Brunei Darussalam, Brunei Darussalam,</li> <li>NECTEC, National Science and Technology Development Agency,</li> <li>Japan Advanced Institute of Science and Technology</li> </ul>	The 2024 5th International Conference on Industrial Engineering and Artificial Intelligence (IEAI)	24-26/04/2024	Bangkok, Thailand
9	UCSYSpoof: A Myanmar Language Dataset for Voice Spoofing Detection	Win Pa Pa, Hay Mar Soe Naing, Aye Mya Hlaing, Kasorn Galajit, Candy Olivia Mawalim	<ul style="list-style-type: none"> <li>University of Computer Studies,Yangon,</li> <li>NECTEC, National Science and Technology Development Agency,</li> <li>Japan Advanced Institute of Science and Technology</li> </ul>	The 27th International Conference of the Oriental COCOSDA 2024	17-19/10/2024	HainChu, Taiwan

# Scientific Contribution:

Presentations at International Conferences:

No:	Paper title:	Author names	Affiliation	Conference name:	The date of the conference	The venue of the conference
10	Analysis of Pathological Features for Spoof Detection	Myat Aye Aye Aung, Win Pa Pa, Hay Mar Soe Naing, Aye Mya Hlaing, Kasorn Galajit, Candy Olivia Mawalim	<ul style="list-style-type: none"> <li>University of Computer Studies, Yangon,</li> <li>NECTEC, National Science and Technology Development Agency,</li> <li>Japan Advanced Institute of Science and Technology</li> </ul>	The 27th International Conference of the Oriental COCOSDA 2024	17-19/10/2024	HainChu, Taiwan
11	Speech Watermarking for Tampering Detection using Singular Spectrum Analysis with Quantization Index Modulation and Psychoacoustic Model	Pantarat Vichathai, Puchit Bunpleng, Patharapol Laolakkana, Dr. Sasiporn Usanavasin, Phondanai Khanti, Dr. Kasorn Galajit, Dr. Jessada Karnjana	<ul style="list-style-type: none"> <li>Sirindhorn International Institute of Technology,</li> <li>NECTEC, National Science and Technology Development Agency,</li> </ul>	The 27th International Conference of the Oriental COCOSDA 2024	17-19/10/2024	HainChu, Taiwan
12	Generative Adversarial Network based Neural Vocoder for Myanmar End-to-End Speech Synthesis	Aye Mya Hlaing and Win Pa Pa	<ul style="list-style-type: none"> <li>University of Computer Studies, Yangon, Myanmar</li> </ul>	7th International Conference on Natural Language and Speech Processing <a href="https://www.icnlsp.org/2024/welcome/">https://www.icnlsp.org/2024/welcome/</a>	19-20/10/2024	Trento, Italy

# Scientific Contribution:

Presentations at International Conferences:

No:	Paper title:	Author names	Affiliation	Conference name:	The date of the conference	The venue of the conference
13	Parallel and Limited Data Voice Conversions on Myanmar Language Speech for Spoofed Detection	Hay Mar Soe Naing, Win Pa Pa, Sheng Li	<ul style="list-style-type: none"> <li>University of Computer Studies,Yangon,</li> <li>National Institute of Information and Communications Technology,</li> </ul>	The 2nd Workshop of Multimodal, Multilingual and Multitask Modeling Technologies for Oriental Languages (M3Oriental) <a href="https://sites.google.com/view/m3oriental">https://sites.google.com/view/m3oriental</a>	3/12/2024	Auckland, New Zealand
14	Automatic Speaker Verification on Myanmar Spoofing Voice Data using GMM-UBM and TDNN	Win Pa Pa, Win Lai Lai Phyu, Hay Mar Soe Naing	<ul style="list-style-type: none"> <li>University of Computer Studies,Yangon,</li> </ul>	The 2nd Workshop of Multimodal, Multilingual and Multitask Modeling Technologies for Oriental Languages (M3Oriental) <a href="https://sites.google.com/view/m3oriental">https://sites.google.com/view/m3oriental</a>	3/12/2024	Auckland, New Zealand
15	Anomalous Machine Sound Detection Based on Time Domain Gammatone Spectrogram Feature and IDNN Model	Dessi Puji Lestari,Primanda Adyatma Hafiz, Candy Olivia Mawalim, Sakriani Sakti, and Masashi Unoki	<ul style="list-style-type: none"> <li>Bandung Institute of Technology, Bandung, Indonesia</li> <li>Japan Advanced Institute of Science and Technology</li> </ul>	The 16th APSIPA Annual Summit and Conference (ASC)2024 <a href="http://www.apsipa2024.org/index.html">http://www.apsipa2024.org/index.html</a>	3-6/12/2024	Macau, China

# Scientific Contribution:

Presentations at International Conferences:

No:	Paper title:	Author names	Affiliation	Conference name:	The date of the conference	The venue of the conference
16	Spoof Detection in Automatic Speaker Verification Using ResNet-34 and Early-Stage Cepstral Coefficient Fusion	Kosin Kalarat, Sasiporn Usanavasin, Thanaruk Theeramunkong, Kasorn Galajit, Jessada Karnjana	<ul style="list-style-type: none"> <li>Sirindhorn International Institute of Technology,</li> <li>NECTEC, National Science and Technology Development Agency,</li> </ul>	The 13th Asean workshop on information science and technology 2025 <a href="https://sites.google.com/view/awist2025/home">https://sites.google.com/view/awist2025/home</a>	Status: accepted 6-8/11/2025	Ishikawa, Japan
17	Spoof Detection for Automatic Speaker Verification	Win Lai Lai Phyu, Win Pa Pa, Kasorn Galajit	<ul style="list-style-type: none"> <li>University of Computer Studies,Yangon,</li> <li>NECTEC, National Science and Technology Development Agency,</li> </ul>	6th International Conference on Advanced Information Technologies (ICAIT2025) <a href="https://www.uit.edu.mm/icait-2025/">https://www.uit.edu.mm/icait-2025/</a>	Status: accepted 3/11/2025	Yangon, Myanmar
18	Deepfake Speech Detection Based on Pathological Entropy Features and Deep Learning	Anuwat Chaiwongyen, Piriya Utamachant, Khalid Zaman, Waree Kongprawechon	<ul style="list-style-type: none"> <li>Japan Advanced Institute of Science and Technology</li> <li>Sirindhorn International Institute of Technology,</li> </ul>	The 20th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP 2025) <a href="https://isai-nlp2025.aiat.or.th/">https://isai-nlp2025.aiat.or.th/</a>	Status: accepted 12-14/11/2025	Phuket, Thailand

# Scientific Contribution:

Published Journal Papers:

Expected output in 2 years

10 Conference papers  
2 Journals

Current output

18 Conference papers  
3 Journals

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"><li>Japan Advanced Institute of Science and Technology,</li><li>Advanced Speech Technology Laboratory, National Institute of Information and Communications Technology,</li></ul>	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
2	Title: End-to-End Neural Diarization for Unknown Number of Speakers with Multi-Scale Decoder	Aung, Myat Aye Aye, Win Pa Pa, Hay Mar Soe Naing	<ul style="list-style-type: none"><li>University of Computer Studies, Yangon</li></ul>	International Journal of Intelligent Engineering & Systems 17, no. 5 (2024).	The Intelligent Networks and Systems Society	Vol 17, Issue 5, p870 ISSN: 2185-310X DOI: 10.22266/ijies2024.1031.66



# Scientific Contribution:

Published Journal Papers:

## Expected output in 2 years

10 Conference papers  
2 Journals

## Current output (approximately 1.5 year)

15 Conference papers  
3 Journals

No:	Paper title:	Author names	Affiliation	Journal name:	The publisher of the Journal	The volume number and Pages
3	Potential of Speech-pathological Features for Deepfake Speech Detection	ANUWAT CHAIWONGYEN, SURADEJ DUANGPUMMET, JESSADA KARNJANA, WAREE KONGPRAWECHNON and MASASHI UNOKI	<ul style="list-style-type: none"> <li>Japan Advanced Institute of Science and Technology,</li> <li>Sirindhorn International Institute of Technology,</li> <li>NECTEC, National Science and Technology Development Agency,</li> </ul>	IEEE Access	IEEE Access	Received 20 June 2024, accepted 15 August 2024, date of publication 22 August 2024, date of current version 10 September 2024. Digital Object Identifier 10.1109/ACCESS.2024.3447582

## Additional Journal papers

- 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).

## ThaiSpoof: Spoof Detection in Thai Language

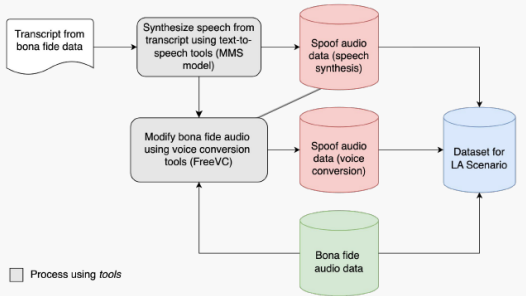
Label	Dataset Type	Degree	Utterances
Genuine	Genuine Dataset	-	4,583
Spoof	Text to Speech - TTS	-	4,583
		10 ch/oct	4,583
		40 ch/oct	4,583
		160 ch/oct	4,583
	Pitch Shifting	320 ch/oct	4,583
		+ 4%	4,583
		+ 10%	4,583
		+ 20%	4,583
		-4%	4,583
		-10%	4,583
		- 20%	4,583
	Massively Multilingual Speech - MMS	-	4,583

Label	Subset Type	No. of Speaker	No. of Utterance
Genuine	Genuine	3 (4K utts/each)	12,000
Spoofed	Vocoder-based [3]	3	23,932
	FreeVC-based	3	24,000
	Text-to-speech [3]	1	8,000
	GMM VC	2	8,000
	GMM DIFFVC	2	8,000

### Building the Dataset

For Logical Access Scenario

- Generated using Massively Multilingual Speech (MMS) model for TTS-based spoofs and the FreeVC system for voice conversion
- Data covers three major Indonesian accents: Indonesian, Javanese, and Bataknese



Bona fide and spoof audio in LA dataset

Dataset	Type	Accent	Number of Utterances
Common Voice	Bona fide	multi	4,540
	Spoof (VC)	multi	90,800
	Spoof (TTS)	bbc	4,540
	Spoof (TTS)	ind	4,540
	Spoof (TTS)	jav	4,540
Prosa.ai	Bona fide	multi	2,000
	Spoof (VC)	multi	120,000
	Spoof (TTS)	bbc	2,000
	Spoof (TTS)	ind	2,000
	Spoof (TTS)	jav	2,000
	Spoof (TTS)	jav	2,000
Total			236,960

### ENIT DATASET: ENGLISH NATIVE AND INDONESIAN-THAI NON-NATIVE SPEECH

Type	Sets	# unique spk.		# utterances		# spoof attack
		Male	Female	Bonafide	Spoof	
Native	Train.	56	56	5,590	33,540	8
	Dev.	8	8	800	4,800	8
	Eval.	16	16	1,600	9,600	8
Non-native	Train.	10 / 12	4 / 5	5,696	33,164	3
	Dev.	1 / 1	1 / 2	542	4,275	3
	Eval.	3 / 3	2 / 3	1,752	10,501	3

### Data Splitting

- Train-dev-eval
- 70-10-20
  - No overlap speaker

## Conclusion:

- “ThaiSpoof ”: Dataset for spoof detection for Thai language  
This dataset is published on AI for Thai for sharing
- “UCSYSpoof”: Dataset for spoof detection for Myanmar language
- Dataset for spoof detection for Indonesian language
- The knowledge of the contributions of Jitter and Shimmer in the Voice for Fake Audio Detection
- The knowledge of the contributions of voice and non-voice for detect spoofing in MFCC and LFCC
- The knowledge of characteristic of abnormal voice.
- The knowledge of speech recognition in Myanmar language
- The Knowledge of the significant of pathological for spoof detection
- The Knowledge of characteristic of tampering speech which can be used for spoof detection

## Future works:

R&D results:

Ongoing

- ASEAN language dataset : Dataset is already provided, now on model demonstration
- Webpage for spoof detection.

Choose File PA\_D\_0000001.flac Clear

0:00 / 0:04

Predict

Result of PA\_D\_0000001.flac

Model	Reference	Trained With	Reported Accuracy	Prediction	Confidence
PA HT20%	10.1109/ISAI-NLP60301.2023.10354625	ResNet34 + LFCC Feature: LFCC (57,600) Classifier: ResNet34 dataset: ASVSpooF 2019 Spoof type: Replay	EER: 0.41 ACC: 90.84 Precision: 90.86 Recall: 90.82 F1: 90.84 DCF: 9.16	<span>Real</span>	98.75%
LA HT20%	10.1109/ISAI-NLP60301.2023.10354625	ResNet34 + LFCC Feature: LFCC (57,746) Classifier: ResNet34 dataset: ASVSpooF 2019 Spoof type: Logical Access	EER: 0.41 ACC: 99.59 Precision: 99.59 Recall: 99.59 F1: 99.59 DCF: 0.41	<span>Real</span>	99.77%
LFCC-MMS+Genuine	10.1109/ISAI-NLP60301.2023.10354956	LCNN + LFCC-MMS + Genuine test Feature: LFCC Classifier: LCNN dataset: MMS + Genuine Spoof type: Mixed	EER: 0.34 ACC: 99.46 Precision: 98.64 Recall: 99.79 F1: 99.48 DCF: 0.26	<span>Fake</span>	69.10%
MFCC-MMS+Genuine	10.1109/ISAI-NLP60301.2023.10354956	LCNN + MFCC-MMS + Genuine Feature: MFCC Classifier: LCNN dataset: MMS + Genuine Spoof type: Mixed	EER: 0.14 ACC: 96.20 Precision: 99.61 Recall: 92.47 F1: 95.91 DCF: 0.11	<span>Real</span>	99.99%