Preface

The Association for Vietnamese Language and Speech Processing (VLSP) was officially founded in 2020, as a chapter of Vietnam Vietnam Association for Information Processing. VLSP consortium is an initiative which came from Institute of Information Technology - Vietnam Academy of Science and Technology, to establish a community working on speech and text processing for the Vietnamese language. The first national project KC01.03/06-10 from 2007-2009 which received strong support from the Ministry of Science and Technology, gathered eight active research groups from universities and institutes in Vietnam and overseas. A main goal of the first project is to set up a long term strategy on Vietnamese language processing in order to provide and to involve the community to enrich shared language resources and tools for R&D purposes. Since 2012, the VLSP Consortium has organized a series of workshops, led by Assoc. Prof. Luong Chi Mai from IOIT-VAST for a branch of speech processing, and Dr. Nguyen Thi Minh Huyen from HUS-VNU Hanoi - for text processing. From the years 2012 to 2019, six events have taken place with different forms of activities such as technical reports, activity reports, discussion panels, and shared tasks on VLSP, in conjunction with large international conferences organized in Vietnam.

The Association for VLSP organized the seventh edition of the VLSP workshop series in 2020 as its first own workshop, and continues to organize these events annually. The ninth international workshop on Vietnamese Language and Speech Processing will be organized in collocation with the 25th Conference of the Oriental COCOSDA, which will be held in Hanoi in November 2022.

VLSP 2012 was the first and kick-off international workshop on Vietnamese Language and Speech Processing, organized in conjunction with IEEE-RIVF 2012. The first VLSP evaluation campaign happened in the second VLSP workshop, associated with the IEEE- RIVF 2013 Conference. This first campaign dealt with two different tasks. One concerns the very essential tools for Vietnamese language processing, word segmentation and part-of- speech (POS) tagging. The other concerns one of the most important Natural Language Processing (NLP) applications: the machine translation task. The third VLSP workshop (2015) was organized for the first time as a satellite event of PAKDD 2015, with a call for participation of only scientific and technical papers in the NLP domain.

VLSP 2016, the fourth international workshop, was once again organized in conjunction with the IEEE-RIVF conference. From this event onwards, organizing shared tasks on Vietnamese processing became the main activities of the workshop series in order to promote the development of essential tools and resources for VLSP. The organization of various evaluation campaigns with sponsors from academia and industry permit the building and offer to the VLSP community gold datasets for Vietnamese text and speech processing.

In 2016, two popular tasks were considered: named-entity recognition (NER) and sentiment analysis (SA). The NER evaluation dealt with three types of entity: person, organization and location. The SA task consists of classifying reviews of technology products as negative, positive or neutral.

The fifth international workshop on Vietnamese Language and Speech Processing – VLSP 2018 was carefully prepared as one important event for CICLing 2018 – the 19th International Conference on Computational Linguistics and Intelligent Text Processing. Four shared tasks were organized for text processing as well as for speech processing. For text processing, the campaign dealt with two tasks NER and SA as in 2016. The NER task used texts extracted from online newspapers and social media texts. The SA task focused this time on aspect-based sentiment analysis of reviews and comments on hotels and restaurants. For speech processing, two tasks have been organized for the first time:

- Automatic Speech Recognition (ASR): Speech recognition with Northern, Central, and Southern dialects.

- Text-to-Speech (TTS): Speech synthesis for one or more regions (Northern, Central, and Southern).

Technical papers of the best systems for each task in VLSP 2018 have been published in a special issue of the Journal of Computer Science and Cybernetics (Volume 34, number 4).

The sixth VLSP workshop was organized in conjunction with the 16th Conference of the Pacific Association for Computational Linguistics (PACLING 2019). The VLSP 2019 evaluation campaign dealt with four shared-tasks, including two speech processing tasks (ASR and TTS) as in 2018 and two new tasks for text processing:

- Vietnamese universal dependency parsing (UDP): the task of determining syntactic dependencies between words in a sentence;
- Hate speech detection on social networks (HSD): the task of classifying hate contents on social network sites.

The TTS 2019 challenge was designed for understanding and comparing research techniques in building Vietnamese corpus-based speech synthesizers on the same data, while the ASR challenge evaluated ASR systems under various conditions of input signal.

Due to the COVID-19 pandemic, the seventh VLSP workshop in 2020 took place as an independent event in Hanoi. Six challenges have been organized:

- Automatic Speech Recognition for Vietnamese (ASR): the task includes two evaluation sub-tasks. For the first sub-task, all participants are required to use only the provided data to develop ASR models including acoustic and language models. For the second one, participants can use all available data sources to develop their ASR models without any limitation.
- Vietnamese Text-To-Speech on Common Datasets (TTS): the task of building a TTS system with a training voice from the same speech database released by the organizers.
- Vietnamese Universal Dependency Parsing (UDP): the same task as in 2019, with larger datasets.
- Vietnamese Relation Extraction (RE): the task of identifying and determining the semantic relations between pairs of named entity mentions within a single sentence.
- English-Vietnamese Machine Translation (MT): text translation from English to Vietnamese in the news domain.
- Reliable Intelligence Identification on Vietnamese Social Network Sites (ReINTEL): the task of identifying a piece of information shared on social network sites (SNSs) as reliable or unreliable.

The participants of the evaluation campaign were asked to present their system in a dedicated paper. This year 2020 equally marks the first time that the VLSP proceedings is hosted by the ACL Anthology (https://aclanthology.org/venues/vlsp/).

In 2021, the eighth VLSP workshop was once again organized as an independent event, hosted by the University of Information Technology (VNU-HCM). Because of the COVID-19 pandemic situation, the workshop is finally organized totally online. But VLSP 2021 was still the largest in terms of number of shared tasks and number of participating teams. As for previous editions, the workshop attracted not only academic research institutions but also R&D teams from the industry sector.

The VLSP 2021 international workshop have successfully organized seven shared tasks, including three challenges on speech processing and four challenges on text processings:

- Automatic Speech Recognition for Vietnamese (ASR): ASR 2021 featured two evaluation tasks. The ASR-T1 focused on a full pipeline development of the ASR model from scratch.

All participants are required to use only datasets provided by the organizers to develop models (including acoustic and language models). The ASR-T2 task focused on spontaneous speech in different real scenarios e.g., meeting conversation, lecture speech. For this subtask, participants are free to use all available data sources to develop their models, without any limitation.

- Vietnamese Text-To-Speech (TTS): the task of building Vietnamese spontaneous speech synthesizers on the same data.
- Vietnamese Speaker Verification (SV): featured two evaluation tasks. The SV-T1 task focused on the development of SV models with limited data. The SV-T2 task focused on testing the robustness of SV systems. For this second subtask, participants can use datasets from any available source for model development.
- Named Entity Recognition (NER): the task of recognizing named entities in Vietnamese documents with an extended set of entities compared to VLSP 2018.
- Vietnamese Machine Reading Comprehension (ViMRC): the task of extraction-based machine reading comprehension on Vietnamese Wikipedia-based texts.
- Vietnamese and English-Vietnamese Textual Entailment (vnNLI): the task of recognizing textual entailment relations between 2 sentences.
- Image Captioning (vieCap4H): the task of automatic generation of Vietnamese descriptions of a given image in the healthcare domains.

150 systems from Vietnam but also from other countries such as Singapore, Japan, Korea, Australia and the United States participated in these challenges. In this special issue of the VNU Journal of Science: Computer Science and Communication Engineering, we introduce 22 technical papers of the best systems and a summary paper for each of seven tasks in VLSP 2021. Each technical paper presents the main technology used to develop the system, analyzes the data used for training and testing, and discusses obtained results. All the papers in this special issue are processed under the standard review process of the Journal with the emphasis of content to technology questions.

Luong Chi Mai and Nguyen Thi Minh Huyen
Guest Editors
VNU Journal of Science: Computer Science and Communication Engineering