

Hai Son Le

Computer and Informatics Engineer
Ph.D. Computer Science

+84 983 133 385
lehaizon13@gmail.com

Professional Experiences

- 2018–now **Researcher**, *FPT Technology Research Institute*, Vietnam.
- 2015–2018 **Lead Software Engineer**, *VNG Corporation*, Vietnam.
- 2013–2015 **Researcher**, *Institute of Information Technology (IOIT), Vietnam Academy of Science and Technology (VAST)*, Vietnam.
- 2015 **Content Based Image Retrieval (CBIR) Project**, *IOIT, VAST*, Vietnam, Building a CBIR system in the collaboration project with Linguistic Modeling Department, Institute of Information and Communication Technologies, Bulgarian Academy of Sciences.
Good performance
- 2014 **Gesture Recognition with Kinect Project**, *IOIT, VAST*, Vietnam, Building a system that uses Kinect to recognize gestures for controlling TV with Infrared signals.
Good performance
- 2013 **Automatic Speech Recognition(ASR) and Statistical Machine Translation (SMT) Projects**, *IOIT, VAST*, Vietnam, Building an ASR system (English) and SMT systems (French-English) for International Workshop on Spoken Language Translation (IWSLT) 2013.
Good performance
- 2010–2013 **SMT Projects**, *Laboratoire d'Informatique et de Mécanique pour les Sciences de l'Ingénieur (LIMSI), Université Paris-Sud*, Orsay, France, Building SMT systems (English-French, French-German...) for IWSLT, Workshop on Statistical Machine Translation (WMT) and Quaero - a European Research and Development Program).
In top 3 rankings among teams of Carnegie Mellon, Karlsruhe, Aachen, Edinburgh, Stanford...
- 2011 **Kaggle Competition Project**, *LIMSI, Université Paris-Sud*, Orsay, France, Building a system based on semi-supervised feature learning with constructed kernels from combinations of non-linear weak rankers for Kaggle competition on Semi-Supervised Feature Learning Challenge 2011.
Second and third rankings
- 2009 **Internship**, *LIMSI, Université Paris-Sud*, Orsay, France.
Subject : Modélisation du langage dans un espace continu
Supervisors : Alexandre Allauzen and François Yvon
- 2006–2008 **Internship**, *Department of Pattern Recognition and Knowledge Engineering, IOIT, VAST*, Vietnam.
Subject: Vietnamese Digits Automatic Speech Recognition
Supervisor : Chi-Mai Luong

Education

- 2009–2013 **Ph.D. Computer Science**, *Laboratoire d'Informatique et de Mécanique pour les Sciences de l'Ingénieur (LIMSI), Université Paris-Sud*, Orsay, France.
Thesis : Continuous Space Models with Neural Network in Natural Language Processing
Supervisors: Alexandre Allauzen and François Yvon
Rapporteurs: Laurent Besacier and Holger Schwenk
Examineurs: Yoshua Bengio and Hermann Ney and Michèle Sebag

- 2008–2009 **Master of Research (Master 2 Recherche) MVA (Mathématiques / Vision / Apprentissage)**, *Ecole Normale Supérieure de Cachan, Cachan, France, Mention très bien (high honors)*.
- 2003–2008 **Engineer Diploma**, *Faculty of Information Technology, Hanoi University of Technology, Hanoi, Vietnam, High honors*.
- 2000–2003 **High School Diploma**, *Hanoi – Amsterdam High School, specialty Mathematics, Hanoi, Vietnam, High honors*.
- 2002 **Prize in the 2002 National Mathematical Olympiad**, Hanoi, Vietnam.

Skills and Qualifications

- Programming languages C, C++, JavaScript, Python, Perl, Matlab
- Languages Vietnamese: Mother tongue
French: good communication skills (oral, written)
English: good communication skills (oral, written)

Research Fields

- Natural Language Modeling
- Language Statistical Machine Translation
- Processing Automatic Speech Recognition
- Computer Vision Object Recognition
- Image Segmentation
- Deformable Models
- Machine Learning Graphical Models
- Neural Network with Deep Learning
- Reinforcement Learning
- Kernel Methods

Publications

In 2015:

- Van-Hieu Vu, **Hai-Son Le**, Olga Kanishcheva, Galia Angelova. Fine tuning SIMPLE based Content Based Image Retrieval system. In *Proceedings of the Sixth International Symposium on Information and Communication Technology*, 2015.

In 2014:

- **Hai-Son Le**, Ngoc-Quan Pham, Duc-Dung Nguyen. Neural Networks with Hidden Markov Models in Skeleton-Based Gesture Recognition. In *Proceedings of the Sixth International Conference on Knowledge and Systems Engineering*, 2014.
- Ngoc-Quan Pham, **Hai-Son Le**, Duc-Dung Nguyen, Truong-Giang Ngo. A Study of Feature Combination in Gesture Recognition with Kinect. In *Proceedings of the Sixth International Conference on Knowledge and Systems Engineering*, 2014.

In 2013:

- Alexandre Allauzen, Nicolas Pécheux, Quoc Khanh Do, Marco Dinarelli, Thomas Lavergne, Aurélien Max, **Hai-Son Le**, François Yvon. LIMSIO@WMT13. In *Proceedings of the Eighth Workshop on Statistical Machine Translation*, 2013.
- Ngoc-Quan Pham, **Hai-Son Le**, Tat-Thang Vu, Chi-Mai Luong. The Speech Recognition and Machine Translation System of IOIT for IWSLT 2013. In *Proceedings of the Tenth International Workshop on Spoken Language Translation (IWSLT)*, Heidelberg, Germany, 2013.
- **Hai-Son Le**, Ilya Oparin, Alexandre Allauzen, Jean-Luc Gauvain, and François Yvon. Structured Output

Layer Neural Network Language Models for Speech Recognition. In *IEEE Transactions on Audio, Speech and Language Processing* VOL. 21, NO. 1, January 2013.

In 2012:

- **Hai-Son Le**, Alexandre Allauzen, and François Yvon. Continuous space translation models with neural networks. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology (NAACL)*, Montréal, Canada, 2012.
- **Hai-Son Le**, Alexandre Allauzen, and François Yvon. Measuring the influence of long range dependencies with neural network language models. In *Proceedings of the NAACL-HLT 2012 Workshop: Will We Ever Really Replace the N-gram Model? On the Future of Language Modeling for HLT*, Montréal, Canada, 2012.
- **Hai-Son Le**, Thomas Lavergne, Alexandre Allauzen, Marianna Apidianaki, Li Gong, Aurélien Max, Artem Sokolov, Guillaume Wisniewski, and François Yvon. LIMSI @ WMT12. In *Proceedings of the Seventh Workshop on Statistical Machine Translation*, Montréal, Canada, 2012.

In 2011:

- Artem Solokov, Tanguy Urvoy, and **Hai-Son Le**. Low-dimensional feature learning with kernel construction. In *Proceedings of the NIPS 2011 Workshop on Deep Learning and Unsupervised Feature Learning*, Granada, Spain, 2011.
- Thomas Lavergne, Alexandre Allauzen, **Hai-Son Le**, and François Yvon. LIMSI's experiments in domain adaptation for IWSLT11. In *Proceedings of the 8th International Workshop on Spoken Language Translation (IWSLT)*, San Francisco, CA, 2011.
- Alexandre Allauzen, Gilles Adda, Hélène Bonneau-Maynard, Josep M. Crego, **Hai-Son Le**, Aurélien Max, Adrien Lardilleux, Thomas Lavergne, Artem Sokolov, Guillaume Wisniewski, and François Yvon. LIMSI @ WMT11. In *Proceedings of the 6th Workshop on Statistical Machine Translation (WMT)*, Edinburgh, Scotland, 2011.
- **Hai-Son Le**, Ilya Oparin, Abdel. Messaoudi, Alexandre Allauzen, Jean-Luc Gauvain, and François Yvon. Large vocabulary SOUL neural network language models. In *Proceedings of the 12th Annual Conference of the International Speech Communication Association (Interspeech)*, Florence, Italy, 2011.
- **Hai-Son Le**, Ilya Oparin, Alexandre Allauzen, Jean-Luc Gauvain, and François Yvon. Structured output layer neural network language model. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Prague, Czech, 2011.

In 2010:

- Alexandre Allauzen, Josep M. Crego, Ilknur Durgar El-Kahlout, **Hai-Son Le**, Guillaume Wisniewski, and François Yvon. LIMSI @ IWSLT 2010. In *Proceedings of the seventh International Workshop on Spoken Language Translation (IWSLT)*, Paris, France, 2010.
- **Hai-Son Le**, Alexandre Allauzen, Guillaume Wisniewski, and François Yvon. Training continuous space language models: Some practical issues. In *Proceedings of the 2010 Conference on Empirical Methods on Natural Language Processing (EMNLP)*, Cambridge, MA, 2010.

In 2008:

- Tat-Thang Vu, Tang-Khanh Nguyen, **Hai-Son Le**, Chi-Mai Luong. Vietnamese tone recognition based on Multi-layer Perceptron network. In *Proceedings of the 2008 Conference of Oriental Chapter of the International Coordinating Committee on Speech Database and Speech I/O system (Oriental COCOSDA)*, Kyoto, Japan, 2008.
- Quang-Trung Bui, Chi-Mai Luong, **Hai-Son Le**. A lexical tree based search network in continuous Vietnamese Speech Recognition. In *Proceedings of the 2008 Conference of Oriental Chapter of the International Coordinating Committee on Speech Database and Speech I/O system (Oriental COCOSDA)*, Kyoto, Japan, 2008.

Hobbies

Sport Football

Reading and Novel, poem

Writing