

# JIATAO GU

Ph.D.

@ jgu@fb.com    ☎ +1 (650) 885-1592 (US)    📍 770 Broadway, New York, NY, 10003    🌐 http://jiataogu.me  
🐦 @thoma\_gu    in jiatao-gu-204b2672    🔄 MultiPath    🆔 orcid.org/0000-0003-3578-2711

## RESEARCH INTEREST

---

*Natural Language Processing, Efficient Models, Multi-model Learning, Representation Learning*

## EDUCATION

---

Ph.D. in Electrical and Electronic Engineering

**The University of Hong Kong**

📅 Sept. 2014 – Aug. 2018

📍 Hong Kong, China

- Supervised by Prof. Victor O.K. Li, Chair Professor of Information Engineering, HKU
- Hong Kong PhD Fellowship (HKPF) holder (Top 100 PhD students in Hong Kong).
- Dissertation: *Efficient Neural Machine Translation*.

---

B.S. in Electronic Engineering

**Tsinghua University**

📅 Sept. 2010 – Jul. 2014

📍 Beijing, China

- GPA: 3.76, Rank: 18/243
- Dissertation: *A Compositional Semantic Distributed Model and its application on Text Summarization*.

---

Exchange Student Program

**The University of Tokyo**

📅 Oct. 2012 – Mar. 2013

📍 Tokyo, Japan

## EXPERIENCE

---

Research Scientist

**Facebook AI Research (FAIR), Meta Platforms, Inc.,**

📅 Aug. 2018 – Now

📍 New York City, NY, USA

- Manager: Y-Lan Boureau, Research Scientist Leader
- Research on advanced techniques for improving NLP and other deep learning applications

---

Research Intern

**Salesforce Research (MetaMind), Salesforce Inc.**

📅 Sept. 2017 – Dec. 2017

📍 Palo Alto, CA, USA

- Adviser: Richard Socher, Chief Scientist
- Research on Non-autoregressive decoding of neural machine translation

Research Intern

**Microsoft AI & Research, Microsoft Corporation**

📅 Apr. 2017 – Aug. 2017

📍 Redmond, WA, USA

- Adviser: Awadalla Hany Hassan, Principal NLP Scientist
- Research on transfer learning for Zero-Resource neural machine translation

## Visiting Scholar

**CILVR Group, Courant Institute of Mathematical Sciences, New York University**

📅 Jun. 2016 – Jan. 2017

📍 New York, NY, USA

- Adviser: *Kyunghyun Cho*, Assistant Professor, CIFAR Azrieli Global Scholar
- Research on trainable decoding for neural machine translation with reinforcement learning.

## Research Intern (Collaborative Research)

**Noah's Ark Lab, Huawei Technologies**

📅 Sept. 2015 – Jun. 2016

📍 Hong Kong, China

- Adviser: *Zhengdong Lu*, Senior Researcher (now CTO of deeplycurious.ai)
- Research on neural symbolic behaviors in NLP.

---

## Student Research Training (SRT)

**Tsinghua-iFlyTek Joint Laboratory for Speech Technologies, Tsinghua University & iFlyTek China.**

📅 Apr. 2013 – Sept. 2013

📍 Beijing, China

- Adviser: *Ji Wu*, Associate Professor
- Research on extractive summarization and language modeling.

## Exchange Student Lab Research

**IBA Lab, Graduate School of Information Science and Technology, The University of Tokyo**

📅 Oct. 2012 – Mar. 2013

📍 Tokyo, Japan

- Adviser: *Hitoshi Iba*, Professor
- Research on simple compositional methods for distributed and co-occurrence lexical patterns.

# PROFESSIONAL ACTIVITIES

---

## Teaching

- Teaching Assistant, ELEC 2501/3541 S/W Engineering and O.S., HKU Spring, 2015
- Teaching Assistant, ELEC 3226/3442 Embedded System, HKU Spring, 2016
- Teaching Assistant, ELEC 3641 Human computer interaction, HKU Spring, 2018

## Organizer

IWSLT 2020 (Simultaneous Speech Translation Task)

## Area Chair

ACL Rolling Review   ACL 2021   EMNLP 2020   NLPCC 2020

## Reviewer (# of paper reviewed)

ACL 2017 (1)   IJCNLP 2017 (2)   AACL 2018 (8)   NAACL 2018 (1)   WNMT 2018 (2)  
EMNLP 2018 (5) Best Reviewer Award   NeurIPS 2019 (5)   ICML 2019 (1)   ICLR 2019 (4)  
ACL 2019 (5)   EMNLP 2019 (3)   NeurIPS 2019 (3)   ICLR 2020 (4)   NeurIPS 2021 (8)  
CVPR 2022 (5)

## Invited Talk

- **StyleNeRF: A style-based 3d-aware generator for high-resolution image synthesis**
    - VALSE Webinar, virtual talk Dec. 29, 2021
  - **Multilingual Denoising Pre-training for Neural Machine Translation**
    - Google Translate, virtual talk Jul. 17, 2020
    - ACL2020 WNGT Workshop, virtual talk Jul. 10, 2020
  - **Understanding Knowledge Distillation in Neural Sequence Generation**
    - Tsinghua University, Beijing, China Jan. 03, 2020
    - Microsoft Translate, Redmond, WA, USA Dec. 04, 2019
  - **Non-autoregressive Neural Machine Translation**
    - New York University, New York, NY, USA Dec. 19, 2017
    - Google Brain, Mountain View, CA, USA Nov. 20, 2017
  - **Learning to Share: towards Zero-Resource Neural Machine Translation without Pivots**
    - Microsoft Research, Redmond, WA, USA Jul. 27, 2017
  - **Trainable Decoding for Neural Machine Translation**
    - NYU Shanghai, Shanghai, China Apr. 14, 2017
  - **Learning to Translate in Real-time with Neural Machine Translation**
    - Carnegie Mellon University, Pittsburgh, PA, USA Oct. 27, 2016
    - New York University, New York, NY, USA Oct. 19, 2016
- 

## MEDIA COVERAGE

---

- Reports mentioned **Data2vec**
  - <https://bit.ly/3Gr3Vhq> (MIT Technology Review)
- Reports mentioned **StyleNeRF: A 3D-Aware Generator for High-Resolution Image Synthesis**
  - <https://bit.ly/350eT49> (SyncedReview)
- Reports mentioned **Multilingual Fine-tuning of Extensible Multilingual Pretraining**
  - <https://bit.ly/3gpt01I> (Towards data science)
- Reports mentioned **Multilingual Denoising Pre-training for Neural Machine Translation**
  - <https://bit.ly/3L8VpHp> (SyncedReview)
- Reports mentioned **Levenshtein Transformer**
  - <https://zd.net/3sdVEIN> (ZDNET)
- Reports on **Non-Autoregressive Neural Machine Translation**
  - <https://cnb.cx/3siXyHV> (CNBC)
  - <https://bit.ly/3GqC4hw> (VentureBeat)
- Reports on **Learning to Translate in Real-time with Neural Machine Translation**
  - <https://bit.ly/3sgr5SL> (Slator)

## HONORS & AWARDS

---

- IEEE CI HK FYP Competition and Graduate Student Paper Competition 2015/16, 2nd Runner-up 2016
- HKPFS 2014/15 (Hong Kong PhD Fellowship Scheme) Award 2014
- CSC Scholarship for Excellent Undergraduate Exchange Program 2012
- National Endeavor Scholarship Award 2013, 2012, 2011

## LANGUAGES

---

English  
Chinese  
Japanese



# LIST OF PUBLICATIONS

---

\*=equal contribution

## CONFERENCES

- [1] **Jiatao Gu**, Lingjie Liu, Peng Wang, and Christian Theobalt. "StyleNeRF: A Style-based 3D-Aware Generator for High-resolution Image Synthesis". In: ICLR. 2022.
- [2] **Jiatao Gu\*** and Xiang Kong\*. "Fully Non-autoregressive Neural Machine Translation: Tricks of the Trade". In: ACL Findings. 2021.
- [3] Xiang Kong, Adithya Renduchintala, James Cross, Yuqing Tang, **Jiatao Gu**, and Xian Li. "Multilingual Neural Machine Translation with Deep Encoder and Multiple Shallow Decoders". In: EACL. 2021, pp. 1613–1624.
- [4] Lingjie Liu, Marc Habermann, Viktor Rudnev, Kripasindhu Sarkar, **Jiatao Gu**, and Christian Theobalt. "Neural Actor: Neural Free-view Synthesis of Human Actors with Pose Control". In: SIGGRAPH ASIA. 2021.
- [5] Jiajun Shen, Peng-Jen Chen, Matt Le, Junxian He, **Jiatao Gu**, Myle Ott, Michael Auli, and Marc'Aurelio Ranzato. "The source-target domain mismatch problem in machine translation". In: EACL. 2021.
- [6] Yuqing Tang, Chau Tran, Xian Li, Peng-Jen Chen, Naman Goyal, Vishrav Chaudhary, **Jiatao Gu**, and Angela Fan. "Multilingual translation with extensible multilingual pretraining and finetuning". In: ACL Findings. 2021.
- [7] Changan Wang, Wei-Ning Hsu, Yossi Adi, Adam Polyak, Ann Lee, Peng-Jen Chen, **Jiatao Gu**, and Juan Pino. "fairseq S<sup>2</sup>: A Scalable and Integrable Speech Synthesis Toolkit". In: EMNLP Demo track. 2021.
- [8] Lior Yariv, **Jiatao Gu**, Yoni Kasten, and Yaron Lipman. "Volume rendering of neural implicit surfaces". In: *NeurIPS*. (oral). 2021.
- [9] Chunting Zhou, **Jiatao Gu**, Mona Diab, Paco Guzman, Luke Zettlemoyer, and Marjan Ghazvininejad. "Detecting Hallucinated Content in Conditional Neural Sequence Generation". In: ACL Findings. 2021.
- [10] Maha Elbayad, **Jiatao Gu**, Edouard Grave, and Michael Auli. "Depth-adaptive transformer". In: 2020.
- [11] Junxian He\*, **Jiatao Gu\***, Jiajun Shen, and Marc'Aurelio Ranzato. "Revisiting self-training for neural sequence generation". In: ICLR. 2020.
- [12] Jungo Kasai, James Cross, Marjan Ghazvininejad, and **Jiatao Gu**. "Non-autoregressive machine translation with disentangled context transformer". In: ICML. PMLR. 2020, pp. 5144–5155.
- [13] Hang Le, Juan Pino, Changan Wang, **Jiatao Gu**, Didier Schwab, and Laurent Besacier. "Dual-decoder Transformer for Joint Automatic Speech Recognition and Multilingual Speech Translation". In: COLING. 2020.
- [14] Lingjie Liu\*, **Jiatao Gu\***, Kyaw Zaw Lin, Tat-Seng Chua, and Christian Theobalt. "Neural Sparse Voxel Fields". In: *NeurIPS*. vol. 33. (spotlight). 2020.
- [15] Xutai Ma, Mohammad Javad Dousti, Changan Wang, **Jiatao Gu**, and Juan Pino. "Simuleval: An evaluation toolkit for simultaneous translation". In: EMNLP Demo Track. 2020.
- [16] Xutai Ma, Juan Pino, James Cross, Liezl Puzon, and **Jiatao Gu**. "Monotonic multihead attention". In: ICLR. 2020.
- [17] Arya D McCarthy, Xian Li, **Jiatao Gu**, and Ning Dong. "Addressing Posterior Collapse with Mutual Information for Improved Variational Neural Machine Translation". In: ACL. 2020, pp. 8512–8525.
- [18] Chau Tran, Yuqing Tang, Xian Li, and **Jiatao Gu**. "Cross-lingual retrieval for iterative self-supervised training". In: *NeurIPS*. 2020.

- [19] Changhan Wang, Kyunghyun Cho, and **Jiatao Gu**. “Neural machine translation with byte-level subwords”. In: *AAAI*. vol. 34. 05. 2020, pp. 9154–9160.
- [20] Changhan Wang, Juan Pino, and **Jiatao Gu**. “Improving Cross-Lingual Transfer Learning for End-to-End Speech Recognition with Speech Translation”. In: *Interspeech*. 2020.
- [21] Changhan Wang, Juan Pino, Anne Wu, and **Jiatao Gu**. “Covost: A diverse multilingual speech-to-text translation corpus”. In: 2020.
- [22] Anne Wu, Changhan Wang, Juan Pino, and **Jiatao Gu**. “Self-supervised representations improve end-to-end speech translation”. In: *Interspeech*. 2020.
- [23] Saining Xie, **Jiatao Gu**, Demi Guo, Charles R Qi, Leonidas Guibas, and Or Litany. “PointContrast: Unsupervised pre-training for 3D point cloud understanding”. In: *ECCV*. Springer, Cham. 2020, pp. 574–591.
- [24] Chunting Zhou\*, **Jiatao Gu\***, and Graham Neubig. “Understanding knowledge distillation in non-autoregressive machine translation”. In: *ICLR*. 2020.
- [25] **Jiatao Gu**, Changhan Wang, and Jake Zhao. “Levenshtein Transformer”. In: *NeurIPS*. 2019.
- [26] **Jiatao Gu\***, Yong Wang\*, Kyunghyun Cho, and Victor OK Li. “Improved Zero-shot Neural Machine Translation via Ignoring Spurious Correlations”. In: *ACL*. 2019.
- [27] Changhan Wang, Anirudh Jain, Danlu Chen, and **Jiatao Gu**. “VizSeq: A Visual Analysis Toolkit for Text Generation Tasks”. In: 2019.
- [28] **Jiatao Gu**, James Bradbury, Caiming Xiong, Victor OK Li, and Richard Socher. “Non-Autoregressive Neural Machine Translation”. In: *ICLR*. 2018.
- [29] **Jiatao Gu**, Hany Hassan, Jacob Devlin, and Victor OK Li. “Universal Neural Machine Translation for Extremely Low Resource Languages”. In: *NAACL*. 2018.
- [30] **Jiatao Gu**, Daniel Jiwoong Im, and Victor OK Li. “Neural machine translation with gumbel-greedy decoding”. In: *AAAI*. 2018.
- [31] **Jiatao Gu**, Yong Wang, Yun Chen, Kyunghyun Cho, and Victor OK Li. “Meta-learning for low-resource neural machine translation”. In: 2018.
- [32] **Jiatao Gu**, Yong Wang, Kyunghyun Cho, and Victor OK Li. “Search Engine Guided Non-Parametric Neural Machine Translation”. In: *AAAI*. 2018.
- [33] **Jiatao Gu**, Kyunghyun Cho, and Victor OK Li. “Trainable greedy decoding for neural machine translation”. In: 2017.
- [34] **Jiatao Gu**, Graham Neubig, Kyunghyun Cho, and Victor OK Li. “Learning to translate in real-time with neural machine translation”. In: *EACL*. 2017.
- [35] Victor O. K. Li, Jacqueline C. K. Lam, Yun Chen, and **Jiatao Gu**. “Deep Learning Model to Estimate Air Pollution Using M-BP to Fill in Missing Proxy Urban Data”. In: *GLOBECOM*. 2017.
- [36] **Jiatao Gu**, Zhengdong Lu, Hang Li, and Victor OK Li. “Incorporating copying mechanism in sequence-to-sequence learning”. In: *ACL*. 2016.
- [37] **Jiatao Gu** and Victor OK Li. “Efficient Learning for Undirected Topic Models”. In: *ACL Short Paper*. 2015.

## JOURNALS

- [1] Yinhan Liu\*, **Jiatao Gu\***, Naman Goyal\*, Xian Li, Sergey Edunov, Marjan Ghazvininejad, Mike Lewis, and Luke Zettlemoyer. “Multilingual denoising pre-training for neural machine translation”. In: *Transactions of the Association for Computational Linguistics* (2020).
- [2] **Jiatao Gu**, Qi Liu, and Kyunghyun Cho. “Insertion-based decoding with automatically inferred generation order”. In: *Transactions of the Association for Computational Linguistics* (2019).

- [3] JQ James, Wen Yu, and **Jiatao Gu**. "Online vehicle routing with neural combinatorial optimization and deep reinforcement learning". In: *IEEE Transactions on Intelligent Transportation Systems* 20.10 (2019), pp. 3806–3817.
- [4] James Jian Qiao Yu and **Jiatao Gu**. "Real-time traffic speed estimation with graph convolutional generative autoencoder". In: *IEEE Transactions on Intelligent Transportation Systems* 20.10 (2019), pp. 3940–3951.
- [5] JQ James, David J Hill, Albert YS Lam, **Jiatao Gu**, and Victor OK Li. "Intelligent time-adaptive transient stability assessment system". In: *IEEE Transactions on Power Systems* 33.1 (2017), pp. 1049–1058.

## WORKSHOP & PREPRINTS

- [1] Alexei Baevski, Wei-Ning Hsu, Qiantong Xu, Arun Babu, **Jiatao Gu**, and Michael Auli. "data2vec: A General Framework for Self-supervised Learning in Speech, Vision and Language". <https://ai.facebook.com/research/data2vec-a-general-framework-for-self-supervised-learning-in-speech-vision-and-language/>. 2022.
- [2] Ann Lee, Hongyu Gong, Paul-Ambroise Duquenne, Holger Schwenk, Peng-Jen Chen, Changhan Wang, Sravya Popuri, Juan Pino, **Jiatao Gu**, and Wei-Ning Hsu. "Textless Speech-to-Speech Translation on Real Data". 2021.
- [3] Ann Lee, Peng-Jen Chen, Changhan Wang, **Jiatao Gu**, Xutai Ma, Adam Polyak, Yossi Adi, Qing He, Yun Tang, Juan Pino, et al. "Direct speech-to-speech translation with discrete units". arXiv preprint arXiv:2107.05604. 2021.
- [4] Ebrahim Ansari, Nguyen Bach, Ondřej Bojar, Roldano Cattoni, Fahim Dalvi, Nadir Durrani, Marcello Federico, Christian Federmann, **Jiatao Gu**, Fei Huang, et al. "Findings of the IWSLT 2020 Evaluation Campaign". Proceedings of the 17th International Conference on Spoken Language Translation. 2020.
- [5] Peng-Jen Chen, Ann Lee, Changhan Wang, Naman Goyal, Angela Fan, Mary Williamson, and **Jiatao Gu**. "Facebook AI's WMT20 News Translation Task Submission". arXiv preprint arXiv:2011.08298. 2020.
- [6] Zhuofeng Wu, Sinong Wang, **Jiatao Gu**, Madian Khabsa, Fei Sun, and Hao Ma. "CLEAR: Contrastive Learning for Sentence Representation". arXiv preprint arXiv:2012.15466. 2020.
- [7] Juan Pino, Liezl Puzon, **Jiatao Gu**, Xutai Ma, Arya D McCarthy, and Deepak Gopinath. "Harnessing indirect training data for end-to-end automatic speech translation: Tricks of the trade". IWSLT. 2019.
- [8] Yong Wang, Xiao-Ming Wu, Qimai Li, **Jiatao Gu**, Wangmeng Xiang, Lei Zhang, and Victor OK Li. "Large Margin Meta-Learning for Few-Shot Classification". NeurIPS Meta-learning Workshop. 2018.
- [9] **Jiatao Gu**, Baotian Hu, Zhengdong Lu, Hang Li, and Victor OK Li. "Guided Sequence-to-Sequence Learning with External Rule Memory". ICLR Workshop. 2016.