Yabo Zhang

21S003020@stu.hit.edu.cn +86-188-0042-0960
github.com/YBYBZhang Harbin, China

Education

- **Harbin Institute of Technology** Harbin, China
  Master in Computer Science *Sept. 2021 - June. 2023 (expected)*
  Supervised by Prof. Wangmeng Zuo.
- **Harbin Institute of Technology** Harbin, China
  B.Eng. in Computer Science; GPA: 92.86/100; Ranking: 8/261 (Top 3%) *Sept. 2017 - June. 2021*
  Graduated with the honor of "Excellent Graduate of Harbin institute of technology".

Work Experience

- **ByteDance Ltd.** China
  Research Intern (Full-time) *June. 2022 - Jan. 2023*
  - **Text-supervised Semantic Segmentation**: Excellent performance of semantic segmentation highly depends on labor-intensive labels. We attempt to alleviate this issue under a data-sufficient setting, *i.e.*, text supervision only.
- **ByteDance Ltd.** China
  Algorithm Engineer (Full-time) *Apr. 2021 - Aug. 2021*
  - **Arbitrary Shape Scene Text Recognition**: We adapt an efficient transformer-based module, *i.e.*, conformer used in 1D audio recognition, to 2D images and outperform all of prior methods with 87.51% and 94.93% accuracy on ArT19 and ReCTS datasets, respectively.
  - **Table Detection and Recognition in Photo Scene**: For table detection, we improve FCOS model so that it can detect arbitrary quadrilaterals and reduce noise area. For table recognition, we classify tables into borderless and bordered, where the former is implemented with Seq2Seq framework and the latter is processed with table-lines segmentation.

Research Experience

- **Y. Zhang**, M. Yao, Y. Wei, Z. Ji, J. Bai, W. Zuo. *Towards Diverse and Faithful One-shot Adaption of Generative Adversarial Networks.* **NeurIPS 2022.** *Dec. 2021 - May. 2022*

  Overview: One-shot generative domain adaption aims to transfer a pre-trained generator on one domain to a new domain using one reference image only. However, it remains very challenging for the adapted generator (i) to generate diverse images inherited from the pre-trained generator while (ii) faithfully acquiring the domain-specific attributes and styles of the reference image. In this paper, we present a novel one-shot generative domain adaption method, *i.e.*, DiFa, for diverse generation and faithful adaptation.

- **Y. Zhang**, B. Dong, M. Ni, Q. Wang, W. Zuo. *Boosting Visual Grounding by Enforcing Alignment and Reasoning in Vision-and-Language Pretraining.* *July. 2021 - Dec. 2021*

  Overview: Visual grounding aims to localize an object described by a query expression, which is mainly achieved by three core sub-tasks, *i.e.*, (1) the vision-language alignment between objects and query expressions based on (2) reasoning of relationships between objects, followed by (3) the basic object localization. To jointly optimize the core sub-tasks in a unified architecture, this paper proposes a novel Reasoning DETR (namely RDETR) along with two task-specific pre-training tasks for visual grounding.

Honors and Awards

- **Scholarships**
  - China National Scholarship *Oct. 2022*
  - Excellent Graduate, Harbin Institute of Technology *Jun. 2021*
  - First-Class People's Scholarship **Top 3%** *Fall 2017 & Spring 2019*
  - Huawei Enterprise Scholarship **Top 3%** *Oct. 2019*
- **Academic Competitions Awards**
  - Meritorious Winner in Mathematical Contest in Modeling (Top 7%) *Apr. 2020*
  - Provincial Second Prize in iCAN International Contest of innovAtioN *Oct. 2019*

Teaching Assistant

- CS32262: Pattern Recognition Deep Learning *Mar. 2022 - present*
- CS32131: Data Structures and Algorithms *Oct. 2021 - Jan. 2022*

Skills Summary

- **Languages**: Python, C++, SQL, Bash, JAVA
- **Frameworks**: Pytorch, TensorFlow, SpaCy