Ruiyang Zhang

✉ 17307130064@fudan.edu.cn | ⌗ Ruiyang-061X | 🌐 ruiyang-061x.github.io

RESEARCH INTERESTS

My research interests mainly lie in **computer vision** and related topics. I'm now focusing on **open-vocabulary 3D object detection** and its application in the autonomous driving scenario. I also have a background in **3D object understanding** and **vision-language model**.

EDUCATION

Fudan University   Shanghai, China
Bachelor's Degree in Computer Science and Technology   Sept. 2017 - Jun. 2021

WORK EXPERIENCE

Backend Development Engineer, Meituan, Shanghai   Jul. 2021 - Jun. 2023

SELECTED AWARDS

Second Prize, China Undergraduate Mathematical Contest in Modeling (CUMCM), 2019
Third Class Scholarship for Outstanding Students, Fudan University, 2018 & 2019
First Prize, Chinese Mathematical Olympiad in Jiangsu Province (CMO), 2016
First Prize, National Olympiad in Informatics in Jiangsu Province (NOIP), 2014 & 2015

RESEARCH EXPERIENCE

Open-vocabulary 3D Object Detection   Nov. 2023 - Present

– **Supervisor: Dr. Hu Zhang, The University of Queensland** and **Prof. Zhedong Zheng, University of Macau**
– Detect 3D objects from point clouds and images without human-annotated 3D labels.
– Generate pseudo labels for large objects and small objects from point clouds and images respectively. Train a class-agnostic 3D object detector in a self-training manner based on those pseudo labels.
– Use 2D open-vocabulary detector such as GroundingDINO to detect 2D objects from images and classify 3D objects by matching them with their 2D counterparts.

Fine-grained 3D Object Understanding   Aug. 2023 - Nov. 2023

– **Supervisor: Runsen Xu, MMLab, CUHK**
– Developed a dataset called ShapeNetPartTriplet based on ShapeNetPart, comprising part-level triplets of point cloud, image and text.
– Conducted contrastive learning between point cloud and image, as well as point cloud and text, to enhance the part-level understanding of objects of the 3D encoder.
– Performed part segmentation experiments to validate the model's capabilities.

Video Instance Segmentation   Mar. 2021 - Jun. 2021

– **Supervisor: Prof. Wenqiang Zhang, Fudan University**
– Developed a DNN which encodes the video into a spatio-temporal feature pyramid.
– Implemented channel attention and spatio-temporal attention mechanisms to enhance performance.
– Demonstrated superior results on the Youtube-VIS dataset compared to previous approaches.

COVID-19 Detection Mar. 2020 - Jun. 2020

– **Supervisor: Prof. Junping Zhang, Fudan University**
– Designed an advanced deep learning model based on DenseNet with CBAM(Convlution Block Attention Module).
– Enabled accurate classification of chest X-rays into Covid, No-Covid, and Normal categories.
– Developed a heatmap visualization to identify disease areas in X-ray images.

Speech Recognition Mar. 2019 - Jun. 2019

– **Supervisor: Prof. Xiangyang Xue, Fudan University**
– Collect and annotate a speech dataset of 20 Chinese words.
– Utilized FFT(Fast Fourier Transform) for feature extraction and trained a CNN for robust speech classification.
– Achieved an impressive recognition accuracy of 91.3% through rigorous experimentation.

WORK PROJECTS

Zebra System Mar. 2022 - Mar. 2023

– Spearheaded the development of a comprehensive marketing supply activity system.
– Implemented key features such as supply activity configuration, operation authority management, and supply data retrieval.
– Utilized a technology stack including Spring Boot, Pigeon, Thrift, MySQL, Redis, Kafka, Crane, Caffeine, and Lion.

Rainbow System Jul. 2021 - Feb. 2022

– Played a key role in supporting the daily iterations of the system.
– Developed supply list components and incorporated exposure labels for improved functionality.
– Designed and implemented an Elasticsearch index for efficient supply retrieval.
– Technology stack: Spring Boot, Pigeon, MySQL, Redis, Kafka, Crane, Lion, Elasticsearch.

SKILLS

Programming: Python (Pytorch), C/C++, MATLAB, Java
Language: English (IELTS 7.0/9.0), Chinese (Native)