GUAN CONG

✉ bkannsou@gmail.com · ☎ (+81) 08080083038

EDUCATION

Waseda University, Japan   2023.10 – 2026.10
Ph.D. candidate Computer Science

Nanjing University   2020.9 – 2023.6
Master's Software Engineering - GPA 3.6/4

Southast University   2016.9 – 2020.6
Bachlor's Software Engineering - GPA 3.8/4

ACADEMIC PUBLICATION

A real-time and accurate obstacle detection framework based on enhanced YOLOv8 network
Under Review

First author

- Designed and implemented a real-time railway obstacle detection system, focusing on enhancing feature representation, dynamic adaptability, and computational efficiency. Leveraged a multi-scale feature enhancement strategy to improve small object detection under challenging scenarios, such as low-light and occluded environments. Achieved a significant boost in detection accuracy and processing speed, ensuring reliable performance in diverse and complex railway conditions.

ARM : nnU-Net with Arena Mechanism for Medical Image Segmentation   ICASSP 2025

Co-first author

- Designed a three-stage pipeline to automate module selection and optimize feature representation, leveraging SCConv and ConvNeXt for improved model adaptability. Developed a Hybrid Weighting Module integrating Depthwise Separable Convolution, Channel Attention Mechanism, and Singular Value Decomposition to enhance representation learning. Introduced cooperative loss functions to foster branch collaboration, significantly improving overall performance. This approach achieved top results on the ACDC, BraTS21, and KiTS datasets, outperforming resource-intensive models while maintaining efficiency.

MetaCert: Metabolic Attention Network Utilizing Uncertainty Estimation for Multimodal Aspect-Category-Sentiment Triple Extraction   ICASSP 2025

Co-first author

- Designed a Metabolic Attention Mechanism (MAM) inspired by biochemical metabolic networks, utilizing cross-attention and iterative information flow to enhance feature interactions across multiple layers. Developed an Uncertainty Estimation Network (UEN) combining self-attention and multilayer perceptrons to dynamically balance modality contributions and mitigate noise. Introduced advanced loss functions, including KL divergence, CRF loss, contrastive loss, and text loss, to achieve optimal feature alignment, further enriched by clipcap-generated text data. This approach delivered state-of-the-art performance on TWITTER-15 and TWITTER-17 datasets, setting a new benchmark for multimodal aspect-category-sentiment triple extraction tasks.

Full-time experience

Amazon - AWS   2023.9 – 2024.9

Machine Learning Engineer

1. **Developed a semantic embedding model based on sparse coding**, supporting both *doc-only* and *bi-encoder* pipelines. This enables users to perform semantic search on local machines without requiring powerful GPUs. Link: Neural sparse model v2 optimize
2. **Designed and implemented a two-stage accelerated search pipeline** for multi-linear additive search models. This improved the search speed of sparse semantic models by $5-8$ times, allowing users to achieve speeds comparable to BM25. Link:Neural sparse two phase algorithm
3. **Developed AI pipelines and related GitHub CI/CD workflows**, facilitating the release of pre-trained models available for cloud download.
4. **Optimized joint queries in Lucene columnar indexes** through research on bit set and quantization technologies, paving the way for future quantized storage of models.

Intern Experince

Amazon   2022.6 – 2023.4

- Developed a scalable data lake platform for heterogeneous data sources, enabling seamless integration of distributed data services.
- Leveraged elastic serverless computing and distributed data storage technologies to process massive expense data efficiently.
- Designed and implemented pipelines for data ingestion, cleaning, and loading, ensuring high reliability and performance.
- Utilized technologies including Java, AWS, and Flink to build robust fee data backend processing systems.

D5 Render

Sole developer of the prototype and the first release version

1. Designed and implemented a system to load architectural meshes in Rhino using RhinoCDK and synchronize them in real-time with a renderer.
2. Enabled support for incremental updates, including fine-grained modifications to geometry, lighting, and material properties.
3. Delivered a robust solution that has been successfully adopted by consumers globally for commercial applications.

Others

Skills

Language	Level		Deep learning	Level		Engineer	Level
English	Business		Computer Vision	▬▬▬		DataBase	▬▬▬
Japanese	Conversational		Diffusion Model	▬▬		Backend	▬▬▬
Chinese	Native		Pytorch	▬▬▬		Infra	▬▬▬