



郭亮

政治面貌：预备党员

个人主页：<https://guoliang0925.me>

(+86) 19854272590

1834482296@qq.com

教育背景

2024.09 - 2027.07	中国石油大学(华东)	211/双一流	通信工程(含宽带网络、移动通信等)	硕士 (Rank 1/100)
2020.09 - 2024.07	中国石油大学(华东)	211/双一流	通信工程	本科

技术方向：时序预测(气象/交通)，视觉感知(2D分割/3D检测)，多模态大模型

成果与奖励：时序领域SCI论文2篇：Renewable Energy(一区TOP, IF=9.1)、Ocean Engineering(二区TOP, IF=5.5), 累计被引42次；国家奖学金、研究生一等学业奖学金×2、综合优秀奖学金、优秀学生。

专业技能

深度学习与模型开发：熟练使用Python与PyTorch框架，具备从数据构建到模型搭建/训练/评估的研发闭环；熟悉GRU、Transformer、ViT、CLIP、YOLO、BEVFormer等架构，项目经验覆盖时序(时空)预测、3D检测，语义分割与VLM等；熟练使用LoRA等方法训练LLM，有GRPO经验。

工程化与工具链：熟练掌握Linux环境搭建与Git协作，具备ODPS/SQL大规模数据开发经验；掌握ONNX导出与跨平台部署、vLLM加速推理；善于借助AI Coding Agent提升研发效率。

实习经历

高德地图 — 交通算法组 — 算法工程师实习生 (2026.01 - 2026.05)

项目：AI领航场景下道路车速语义信息预测

- ①基于ODPS处理分钟级千万级交通流数据，构建GRU-Seq2Seq时序预测基模，递归预测未来15分钟车流速度；
- ②设计数值到语义的评估体系，将预测数值映射为稳定和趋势（上升/下降）语义；导出ONNX封装Java接口接入MCP支撑线上调用，实现实时判断异常拥堵事件。

项目成果：稳定场景准确率与召回率均超90%，趋势场景准确率70%，显著优于Google TimesFM 2.5。

项目：基于多模态大模型的导航POI空间方位感知

- ①参与导航场景POI方位识别VLM研发，将目标空间方位建模为细粒度多分类问题（如“左前方”“右侧”等方位）；
- ②基于内部AI平台构建CoT+Answer训练数据，通过LoRA将大型VLM的方位识别能力蒸馏Qwen3-VL轻量模型，在降低推理成本的同时保持较高精度；

项目成果：方位判断任务正确率达85%+，验证了轻量VLM在导航空间理解任务上的可行性。

项目经历

◆ MeteoTime — 大规模预训练气象时序基座大模型

▪ 技术栈：时序大模型、因果Transformer、RoPE、RMSNorm、SwiGLU、SDPA、DDP、分位数预测

- ①构建Patch因果Transformer的时序大模型，支持最长1024点上下文、64点多步预测、分位数输出和自回归预测；
- ②搭建覆盖ERA5、BTS机场气象的百亿级数据管道，支持均衡采样、缺失掩码、随机掩码增强与长度分桶；
- ③建立覆盖Weather、中国近海海洋观测站与KDD Cup 2022风电场的Zero-Shot评测体系，对比持续性基线、Google TimesFM 2.5与Amazon Chronos-2。

▪ 项目成果：模型参数量仅约33.2M，支持最长1024时间步上下文的零样本时序预测。在输入长度为512下，在3个benchmark共8项气象预测任务上取得5项最优、2项次优；覆盖128/256/512/1024四类上下文长度、8项任务的共计32组对比实验中，该模型取得18项最优、6项次优，性能表现突出。GitHub：<https://github.com/GL0925/MeteoTime>

◆ MultiBEV — 面向空间感知的LiDAR-相机多模态融合3D目标检测

- **技术栈:** 视觉感知、多模态BEV表征、ViT、稀疏3D卷积、体素特征嵌入、DDP/AMP
- ①构建稀疏3D卷积骨干网络，设计高度注意力池化压缩Z维度聚合高度层的特征，更精细保留3D空间高度信息；
- ②基于Swin-Tiny预训练权重迁移学习，通过FPN多尺度融合提取图像特征，通过可学习深度头预测逐像素深度分布，配合深度投影将2D图像特征提升至3D BEV空间；
- ③设计门控残差融合模块自适应融合LiDAR与相机BEV特征，配合三支解耦检测头实现3D目标检测。
- **项目成果:** 模型38.2M参数量，实现多模态有效融合，在nuScenes验证集上较LiDAR-only基线提升约50%以上；Car类别mAP@1m达0.74，Bus类别mAP@1m达0.55。GitHub: <https://github.com/GL0925/MultiBEV>

◆ OilDetectionVLM — 多模态大模型驱动SAR海面溢油智能感知

- **技术栈:** 多模态大模型、Qwen3-VL、LoRA/PEFT、VLM SFT、多任务视觉问答
- ①将SAR溢油检测重构为视觉问答任务，设计“检测-计数-定位”三阶段输出框架；构建自动化VQA数据流水线，由Qwen3.7-Max 从分割掩码生成CoT结构化标签数据；
- ②对Qwen3-VL-8B-Instruct的ViT与语言侧同步施加LoRA微调，实现双侧联合优化、任务能力高效迁移。
- **项目成果:** 检测任务F1由0.77提升至0.96，计数任务MAE由2.02降至0.54，定位任务Hit Rate从0.59上升至0.88，全面超过原生Qwen3.7-Max(F1:0.92 / MAE:0.83 / HR:0.61)。GitHub: <https://github.com/GL0925/OilDetectionVLM>

◆ CLIP-ITMatch — 基于CLIP与ONNX量化的轻量图文检索模型

- **技术栈:** 多模态、CLIP、图文检索、LoRA/PEFT、对比学习、ONNX Runtime优化、权重量化
- ①基于 CLIP-ViT-B/16 构建图文检索模型，在图像/文本塔 q/k/v/out 层引入LoRA，完成Flickr30k图文对齐优化；
- ②构建 I2T/T2I 双向评估流程，以Multi-positive Recall评估一图多文本 TopK 检索；导出ONNX，对MatMul权重做INT8、token embedding做INT4量化，使用ONNX图优化加速推理。
- **项目成果:** Mean R@1/R@3达到0.7390/0.8784，较Zero-Shot相对提升24.4%/15.2%，W8E4A32量化后体积由571.2MB降至139.0MB，压缩75.7%；图优化后速度提升21.5%。GitHub: <https://github.com/GL0925/CLIP-ITMatch>

◆ 基于迁移学习与多尺度融合解码的SAR溢油图像语义分割

- **技术栈:** 视觉感知、Vision Transformer、FPN多尺度融合、迁移学习、DDP/AMP
- ①结合SAM与人工修正生成标签，在MiT骨干注入轻量Adapter增强任务适配，结合预训练权重实现迁移学习；
- ②以FPN替代U-Net解码器，构建语义-细节双路径精修模块，并引入CE-Lovász混合损失直接优化IoU指标。
- **项目成果:** 最终模型参数量仅为基线的57.5%，mIoU达0.734，溢油核心类别IoU达0.662，超过OffSeg、OneFormer、Mask2Former等主流分割模型。

◆ 其他项目:

- ①基于空间增强与预训练LLM的多站点风速预测 (<https://github.com/GL0925/GraphTimeLLM>，时序预测)
- ②基于BERT的海洋领域NER命名实体识别 (<https://github.com/GL0925/OceanNER>，自然语言处理)

学术成果

《Short-term multi-step wind speed forecasting with multi-feature inputs using Variational Mode Decomposition, a novel artificial intelligence network, and the Polar Lights Optimizer》

时序预测, Renewable Energy, IF=9.1, SCI, 中科院一区TOP

- ①采用皮尔逊相关系数完成特征筛选，结合 VMD 模态分解与时序对齐构建多维特征输入；
- ②设计三阶段融合神经网络，依托GCN与通道注意力挖掘特征关联、LSTM 捕捉时序依赖，通过显著性检验佐证模型有效性，相较 7 类基线模型实现性能全面提升。

《Medium-term offshore wind speed multi-step forecasting based on VMD and GRU-MATNet model》

时序预测, Ocean Engineering, IF=5.5, SCI, 中科院二区TOP

- ①提出分解-深度学习混合框架，以VMD分解信号，规避EMD系列方法的噪声累积，降低频率复杂度；
- ②据风速固有特征，提出GRU-MATNet，融合多头注意力，增强高频细节，中远期预测全面优于主流基线。