

管 聰 (GUAN CONG)

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

学歴

早稲田大学 博士課程（コンピュータサイエンス専攻）	2023年10月 – 2026年10月
南京大学, QS 103 修士（ソフトウェア工学）	2021年9月 – 2023年6月

学術論文

Dual Camera Super-Resolution via Domain Modulation and Multi-scale Efficient Matching
AAAI 2025 査読中

筆頭著者

- 望遠カメラを参照として広角画像を高画質化するデュアルカメラ超解像（SR）のため、マルチスケールブロックマッチング戦略とドメインモジュレーションモジュールを提案。推論効率を高めるプルーニング技術および軽量なドメイン適応手法により、現実世界のデータセットにおいて最先端の性能を実現。

CLIP-driven rain perception: Adaptive deraining with pattern-aware network routing and mask-guided cross-attention

Pattern Recognition 査読中

筆頭著者

- CLIP を用いた雨パターン認識およびサブネットワーク選択に基づく新しい除雨フレームワーク CLIP-RPN を提案。雨と背景の相互作用を捉えるクロスアテンション機構 MGCA を設計し、動的損失スケジューリング (DLS) を導入。複雑な混合雨シナリオを含む複数のベンチマークにおいて SOTA 性能を達成。

ROD-YOLO: A Fast and Accurate Obstacle Detection Framework for Railways based on Feature Enhancement and Context Aggregation

accepted by IEICE Transactions on Information and Systems

筆頭著者

- YOLOv8 をベースに、特徴強調モジュール (FEM)、改良型空間ピラミッドプーリング (SPPCSPC-F)、および LSK アテンションを導入し、鉄道環境における小規模・暗所・複雑シーンの障害物検出性能を向上。自作実環境データセットにて、YOLOv8n 比で mAP50 が 7.93% 向上、推論速度が 72.42% 改善、パラメータ数は 36.19% の増加に抑制。COCO や KITTI ベンチマークにおいても、小物体・天候変化に対して高いロバスト性を実証。

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

ICASSP 2025

共同第一著者

- SCConv と ConvNeXt を活用し、三段階構成による医療画像セグメンテーションフレームワークを設計。複数の損失関数と注意機構で性能を大幅に向上。

インターンシップ経験

ソニーセミコンダクタソリューションズグループ

2025年9月 - 2025年9月

リサーチインターン（コンピュータビジョン）

- 3週間の職場密着型インターンシップに参加し、カメラ ISP アルゴリズムにおけるデモザイキングおよびノイズ除去をテーマに研究・実験を実施
- 可変ノイズ環境に対してロバストな学習手法を探査し、モデルの汎化性能向上
- 生成モデルを活用し、RAW から SR RGB へのワンステップ変換を試み、デモザイキング・ノイズ除去・超解像を統合的に実現

ソニー R&D センター China Laboratory

2025年5月 - 2025年8月

リサーチインターン（コンピュータビジョン）

- 画像生成技術に基づく画像・動画の超解像アルゴリズムを研究。
- フローマッチングおよび拡散モデルに関する先行研究を調査し、実験検証を通じて技術的知見を蓄積。
- モデルの精度・速度・適用性を分析し、研究テーマの方向性決定と進捗に貢献。

アマゾン・ウェブ・サービス（AWS）

2023年9月 - 2024年9月

機械学習エンジニア

- スペースコーディングに基づいた意味埋め込みモデルを開発し、GPU を使用せずにローカル検索を実現。
- 検索速度を 5-8 倍高速化する二段階の検索パイプラインを設計。
- CI/CD ワークフローと事前学習モデルの配信パイプラインを構築。
- Lucene のカラムナインデックスにおいて、量子化およびビットセットの最適化を実施。

D5 Render

2021年6月 - 2021年9月

3D エンジニア

- RhinoCDK を用いて Rhino 上の建築メッシュを読み込み、レンダラーとリアルタイムで同期するシステムを設計・実装。
- ジオメトリ、ライティング、マテリアルプロパティなどの細部にわたる変更を含む、インクリメンタルな更新処理に対応。
- 商用アプリケーションとしてグローバルに採用されており、堅牢なソリューションを提供。

スキル

- 言語スキル:** 中国語を母語とし、英語および日本語においてもビジネスレベルでのコミュニケーションが可能です。
- ディープラーニング:** コンピュータビジョンや拡散モデルなどに関する研究・実装経験があり、PyTorch を用いたモデル構築とトレーニングに熟練しています。
- エンジニアリングスキル:** インフラ構築、バックエンド開発、データベース設計・運用など、システム全体の設計・実装を一貫して担当した経験があります。
- プログラミング言語:** C++、Python、Java に熟練しており、研究開発および実装業務で日常的に使用しています。