Guangkai Ren

kevin.r5009@gmail.com +86 184 0124 5270 Beijing, China

Research Interests Multimodal world models for embodied AI, 3D perception and reconstruction, and generative spatiotemporal modeling.

Publications & Research Outputs

Beyond Vision: Holistic World Models 2026

Qianwen Yang, Yanni Kang, **Guangkai Ren**, Qingwen Yang, and Nan Li.
Proceedings of the IEEE/CVF CVPR Workshops, pp. 7654–7662.
CVF Open Access

World Models: Principles and Engineering of the Next-Generation Embodied AI Engine 2026

Guangkai Ren, Qianwen Yang, and Yanni Kang.
Tsinghua University Press, Nov 2026. In Chinese. ISBN 9787302729242. Wrote Chapters 1–4 and 9–10.
Publisher Preface

An Agricultural E-Commerce Platform Based on Rough Sets and Semantics (CNKI) 2014

Registered research output, Henan Provincial Key Science and Technology Project No. 102102310058.
Implemented ontology-based semantic search with rough concept lattices and recommendation with association rules.

Education

University of Science and Technology Beijing Sep 2020 – Jan 2023

Master's degree in Software Engineering

Luoyang Normal University Sep 2009 – Jul 2013

Bachelor's degree in Computer Science and Technology

Applied Research Experience

Beijing Block Imagine Digital Technology Co., Ltd.
Senior Algorithm Engineer

Multimodal Video Generation and Editing for Digital Humans Jan – Apr 2025

- Developed a Diffusion Transformer with flow matching for audio-, text-, and image-conditioned video generation and editing.
- Designed region masking and attention controls for audio–visual alignment and implemented multi-step samplers for inference.

Sparse-View 3D Human Reconstruction Jun – Dec 2024

- Led development of a two-stage 3D Gaussian Splatting pipeline that fused sparse multi-view features using epipolar attention and an SMPL body-shape prior.

Real-Time Lip Synchronization for Digital Humans Mar – Jun 2024

- Led development of a GAN–VAE model with two-stage training and a lip-sync loss, achieving real-time generation at 30 FPS and 256 × 256 resolution.

Applied Research Experience

Geovis Technology Co., Ltd.
AI Algorithm Engineer

3D Severe Convective Weather Nowcasting

Jan – Apr 2026

Primary project contributor

- Developed a volumetric radar-echo nowcasting workflow combining physics-constrained advection–diffusion–source extrapolation with conditional diffusion for probabilistic forecasting.

Block-Level High-Resolution Weather Forecasting

May – Jul 2026

Primary project contributor

- Used large-eddy simulation (LES) to generate high-resolution training data for urban wind fields, including flows around buildings and in street canyons.
- Developed a Swin Transformer–based wind-field prediction model. Shifted-window attention captures local flow patterns and connects neighboring regions across spatial scales.

Additional Professional Experience

Agent-Based Product Development

Aug 2026 – Present

- Media network platform for topic selection, drafting, and publishing.
- Interstellar Travel online game.
- Video lessons and review games for New Concept English Books 1–4.
- MTV production and publishing workflow.
- Audience-growth automation tools for X.

China National Knowledge Infrastructure (CNKI)

Nov 2016 – Jun 2023

R&D Engineer and Senior Project Manager

- Designed and implemented academic literature retrieval components combining multiple retrieval routes, vector search, and learning to rank.
- Transformer-based machine translation.

Beijing Gome Online E-Commerce Co., Ltd.

Dec 2014 – Aug 2016

R&D Engineer

- Developed back-end search and recommendation algorithms for products, stores, and users, with Elasticsearch for indexing and retrieval.

Beijing Dawn Infinite Technology Co., Ltd.

Sep – Dec 2014

R&D Engineer

- Developed filters and image-blending algorithms for mobile photo editing and collage apps.

Languages & Technical Skills

English	IELTS 6.5 overall (all bands ≥ 6.0), CET-6 466
Model development & post-training	Python, PyTorch, Transformers, Diffusers, PEFT (LoRA), verl
Inference & deployment	SGLang, vLLM, CUDA, Docker
Agentic coding tools	Codex, Claude Code, DeepSeek Harness

任广凯

kevin.r5009@gmail.com　+86 184 0124 5270　中国 北京

研究兴趣　面向具身智能的多模态世界模型、三维感知与重建、生成式时空建模。

论文、著作与科研成果

Beyond Vision: Holistic World Models 2026

Qianwen Yang、Yanni Kang、**Guangkai Ren**、Qingwen Yang、Nan Li。
IEEE/CVF 计算机视觉与模式识别会议研讨会论文集（CVPR Workshops），第 7654–7662 页。
论文全文

世界模型：下一代具身智能引擎的原理与工程实践 2026

任广凯、杨倩文、康延妮。
清华大学出版社，2026 年 11 月。ISBN 9787302729242。负责撰写第 1–4 章及第 9–10 章。
出版社书目　前言

基于粗集结合语义的农产品电子商务交易平台的研究（CNKI） 2014

河南省重点科技攻关项目入库成果，项目编号 102102310058。
负责算法落地实现，结合粗概念格和农产品本体实现语义检索，利用关联规则挖掘实现个性化推荐。

教育背景

北京科技大学 2020.09 – 2023.01
软件工程硕士

洛阳师范学院 2009.09 – 2013.07
计算机科学与技术学士

应用研究经历

北京区块畅想数字科技有限公司
资深算法工程师

数字人多模态视频生成与编辑 2025.01 – 2025.04

- 研发基于扩散 Transformer（DiT）和流匹配的多模态视频生成框架，支持音频、文本和图像条件下的视频生成与编辑。
- 设计区域掩码和注意力控制机制，实现音视频特征对齐，并实现用于推理的多步采样器。

稀疏视角三维人体重建 2024.06 – 2024.12

- 牵头研发两阶段 3D 高斯泼溅（3DGS）人体重建流程，利用极线注意力融合稀疏多视角特征，并引入 SMPL 人体形状先验。

实时数字人口型同步 2024.03 – 2024.06

- 主导研发 GAN–VAE 口型同步模型，通过两阶段训练和同步损失，以 30 FPS 实时生成 256 × 256 视频。

应用研究经历

中科星图股份有限公司
AI 算法工程师

三维强对流临近预报 2026.01 – 2026.04

第一完成人

- 研发基于三维雷达反射率的对流回波临近预报流程，结合物理约束下的平流–扩散–源项外推与条件扩散模型，实现概率预报。

街区级精细化预报 2026.05 – 2026.07

第一完成人

- 利用大涡模拟（LES）生成高分辨率街区风场训练数据，涵盖建筑绕流与街谷内的局地气流。
- 研发基于 Swin Transformer 的风场预报模型，利用移位窗口注意力提取局地流动特征，在不同空间尺度上融合相邻区域的信息。

其他专业经历

基于智能体的产品开发 2026.08 – 至今

- 媒体矩阵选题、编撰与发布平台。
- 《星际旅行》在线游戏。
- 《新概念英语》1–4 册视频版及课后复习游戏。
- MTV 制作与发布工作流。
- X 平台受众增长自动化工具。

中国知网技术有限公司 2016.11 – 2023.06

研发工程师兼高级项目经理

- 设计并实现学术文献检索模块，结合多路召回、向量检索和学习排序方法。
- 基于 Transformer 架构的机器翻译。

北京国美在线电子商务有限公司 2014.12 – 2016.08

研发工程师

- 开发面向商品、店铺和用户的后端搜索与推荐算法，使用 Elasticsearch 构建索引与检索模块。

北京曙光无限技术有限公司 2014.09 – 2014.12

研发工程师

- 为移动端美图与拼图应用开发滤镜和图像融合算法。

语言与技术技能

英语	IELTS 总分 6.5，单项均不低于 6.0。大学英语六级 466 分。
模型开发与后训练	Python、PyTorch、Transformers、Diffusers、PEFT（LoRA）、verl
推理与部署	SGLang、vLLM、CUDA、Docker
智能体编程工具	Codex、Claude Code、DeepSeek Harness