第二十八届中国科协年会
The 28th Annual Meeting of the China Association for Science and Technology
抢占制高点 催生新动能 为中国式现代化提供科技支撑

AI 赋能的生物医学大数据研究

Data- and knowledge-driven multimodal learning in Computational Pathology: A comprehensive survey

Mingxin Liu, Chengfei Cai, Depin Chen, Jun Li, Jinze Li, Wenlong Ming, Jun Xu

Jiangsu Key Laboratory of Intelligent Medical Image Computing, School of Artificial Intelligence,

Nanjing University of Information Science and Technology, Nanjing 21004, China

Link to Paper
EngMedicine (2026)

Visit IMIC Lab

1. Motivation & Scope

Why multimodal CPath matters?

Histology alone captures morphology, but clinical decisions combine tissue context with genomic, radiologic, and longitudinal patient evidence.

WSI | Radiology | EHR & reports | Omics profiles

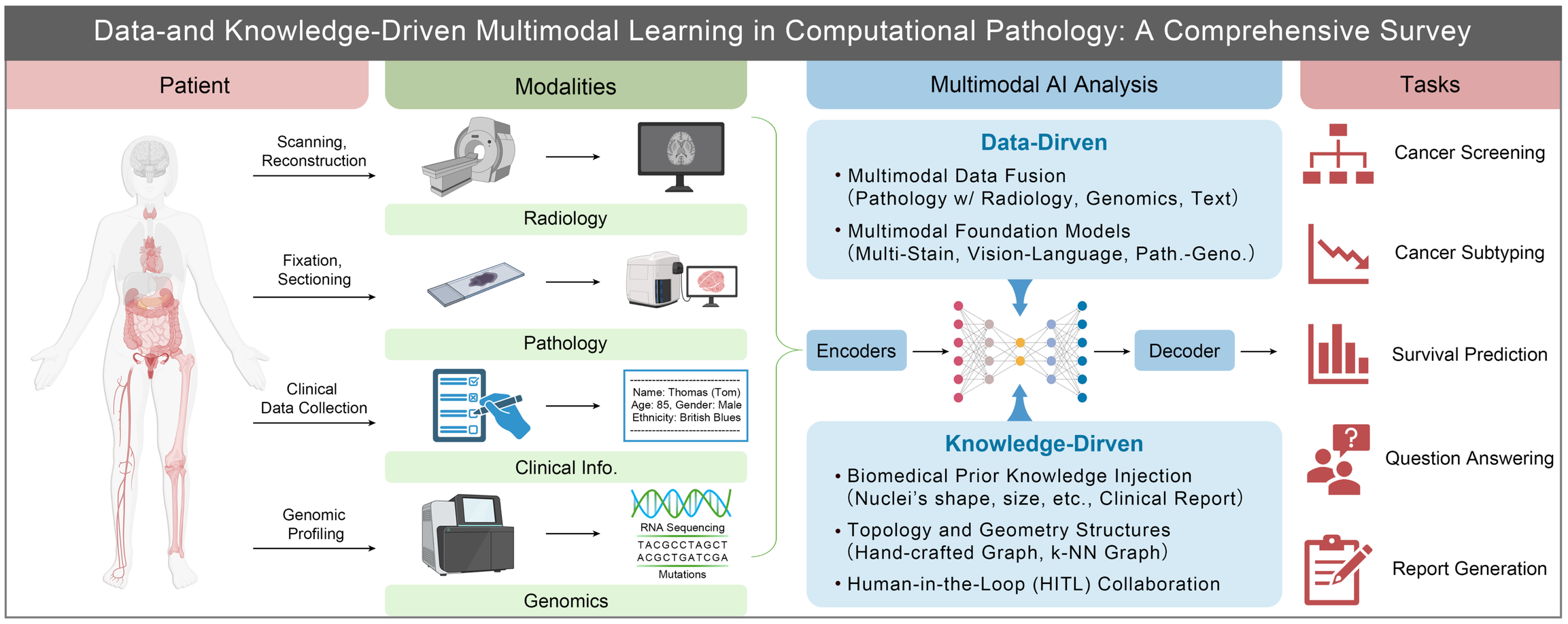
What this survey organizes?

1. Heterogeneous clinical evidence
WSIs describe tissue morphology; Radiology captures organ-level context; Genomics reveal functional perturbations; Reports encode clinical trajectories and expert reasoning.

2. Two methodological engines
Data-driven models learn alignment, fusion, and scalable representations. Knowledge-driven methods inject topology, prompts, and human feedback for interpretability and trust.

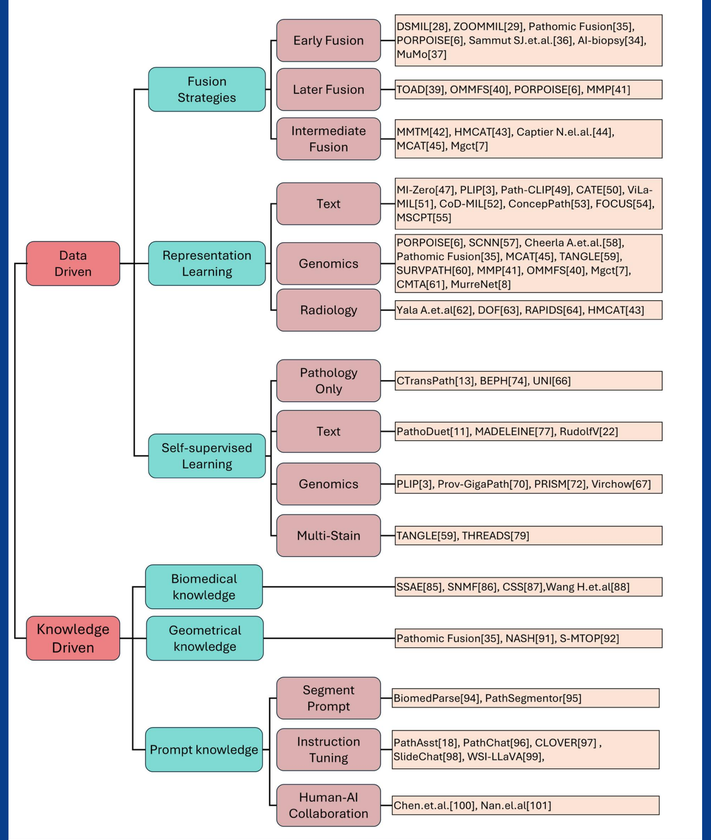
3. Clinical translation lens
Links technical advances to deployment barriers: missing modalities, scanner and stain variation, and the need for pathologist-valid explanations.

2. Graphical Abstract: multimodal pathology AI at a glance

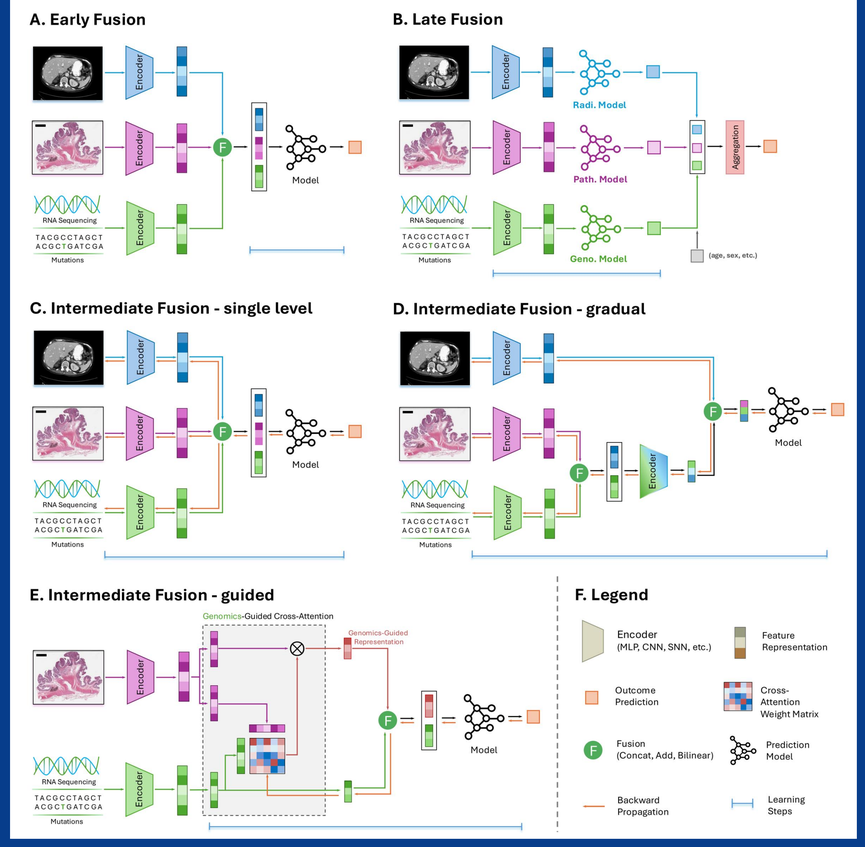


Multimodal CPath connects patient-level data with data- and knowledge-driven AI to support diverse downstream tasks in clinical practice.

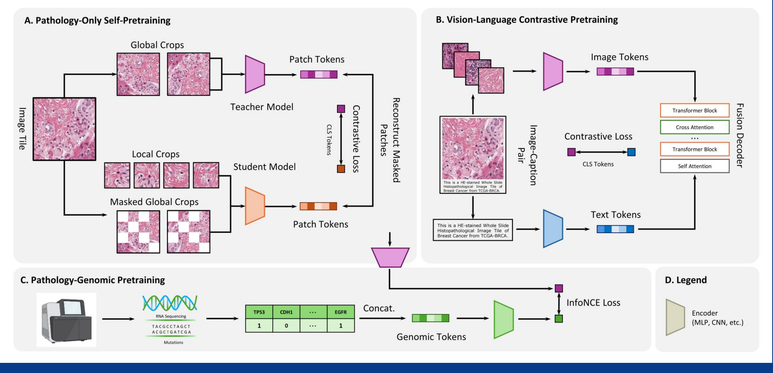
3. Hierarchical design taxonomy



4. Multimodal fusion taxonomy



5. Multimodal alignment taxonomy



6. Foundation models taxonomy

- **Modality & Knowledge:**
  - H&E / Multi-Stain / Vision-Language
- **Pretraining framework:**
  - Contrastive / Distillation / Masked Modeling
- **Architecture & Data scale:**
  - ViT-S / B / H / G & LongNet for Prov-GigaPath

7. Future directions for trustworthy multimodal CPath

Unified representations
Align WSIs, omics, radiology, reports, mIF, 3D pathology, and longitudinal slides

Dynamic fusion
Adapt to missing modalities, noisy data inputs, latency, and unequal data reliability

Knowledge-driven AI
Encode diagnostic criteria, biomedical priors, and pathologist-verifiable reasoning

Trustworthy evaluation
Avoid public-data leakage and build clinically meaningful, validated benchmarks

Robust generalization
Handle scanner, stain, center, specimen-quality, and rare-case shifts