SAID: Spatial and Interaction-Aware Directed Heterogeneous Graph Neural Network for Gene Mutation Prediction from Histopathology Whole Slide Images

Yifei Wang[1], Jun Shi[2*], Shihao Wu[1], Jiyang Li[1], Zhiguo Jiang[3], Yushan Zheng[4*]

1 School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, China

2 School of Software, Hefei University of Technology, Hefei, China

3 Tianmushan Laboratory, Beihang University, Hangzhou, China

4 Beijing Advanced Innovation Center for Biomedical Engineering, School of Engineering Medicine, Beihang University, Beijing, China

*Corresponding author: juns@hfut.edu.cn, yszheng@buaa.edu.cn

Outline

01	Background	Pathology slides, mutations, and what a graph is
02	Method	Graph construction and network framework
03	Results	Two TCGA datasets, ablation and sensitivity
04	Interpretability	Which interactions carry the signal

01

PART ONE

Background

Pathology slides, mutations, and what a graph is

Yifei Wang et al. ICPR'26

Background

Mutation Status and Treatment

- A tumor's **genotype** decides which therapy is given.

> *"The detection of genetic mutations is critical for precision therapy."*
>
> Abstract, SAID

- **EGFR**, **RAS**, **BRAF** — tested before treatment is chosen.

- Two tumors can look alike under the microscope and still need *different drugs*.

Background

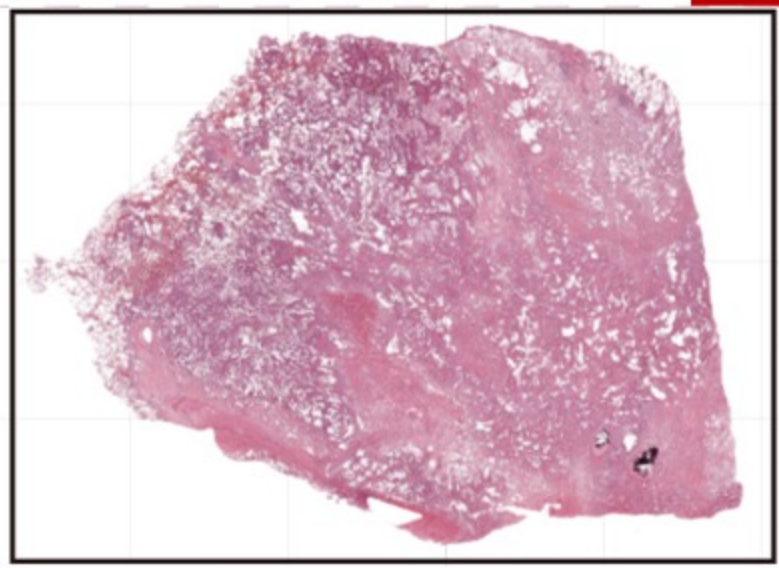
Next-generation sequencing

STILL THE GOLD STANDARD

- **High cost** per case.
- **Long turnaround** — the treatment decision waits.
- Together, these *motivate faster and cheaper alternatives*.

Background

H&E Slides in Routine Practice

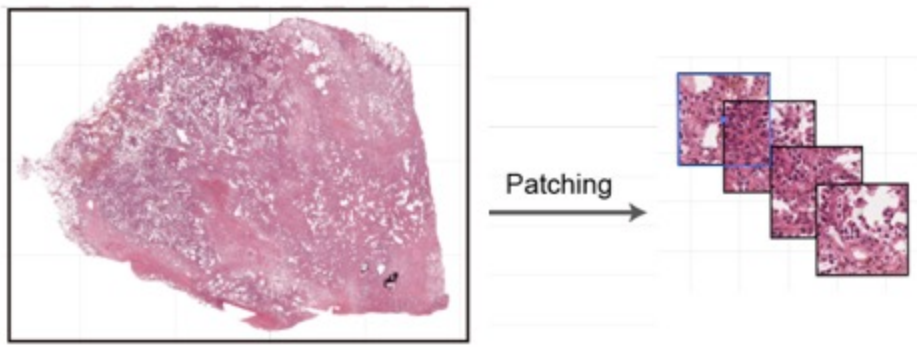


- H&E staining is **routine** — the slide already exists for nearly every case.
- Genetic mutations can lead to changes in the tumor microenvironment.
- Its rich morphology **reflects underlying genetic alterations**.
- So the task becomes: predict mutation status *directly from the image*.

Background

From Slide to Patches

- One slide is gigapixel-scale — *no model reads it whole*.
- Tessellate into **256 × 256 patches.**
- The label belongs to the slide, not the patch — **weak supervision**.



Background

Limits of the Bag-of-Patches View

- **Multiple instance learning** treats a slide as an unordered bag of patches.
- Attention decides *which* patches matter — never *where* they are.
- Tissue architecture and **interactions between regions** are lost.
- Those are exactly the cues tied to genetic alterations.

Background

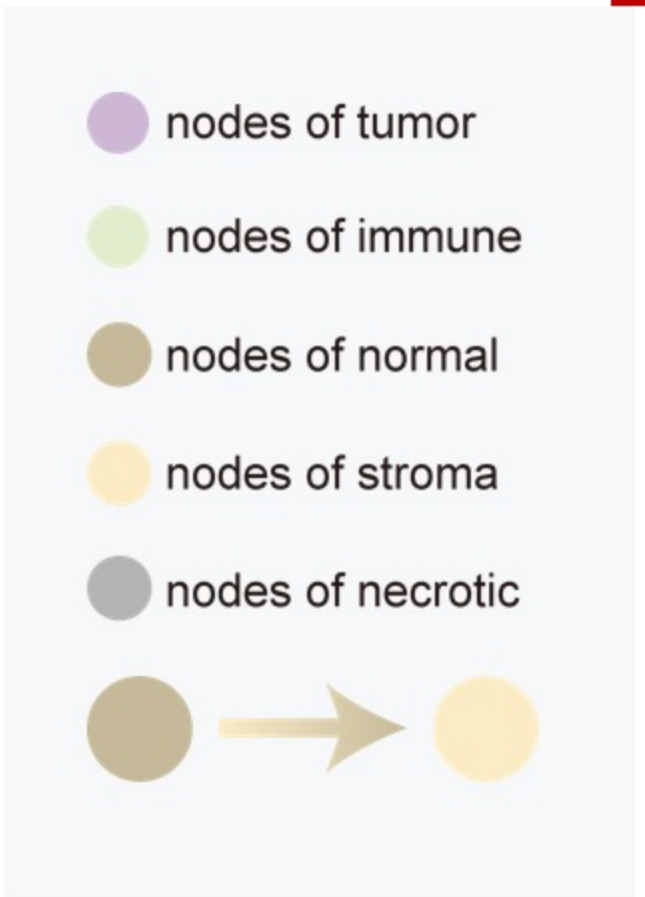
The Tumor Microenvironment

- Five tissue types share the same slide, each with a distinct biological function.
- They do more than sit next to each other — they **interact**.
- And interaction has a direction: **tumor → immune** is not the same event as **immune → tumor**.

Background

Graphs in One Minute

- **Node** = one patch.
- **Edge** = a relationship between two patches.
- **Message passing**: each node updates itself from its neighbors.
- A **directed** edge carries information one way only.



Background

Limits of Existing Graph Models

HOMOGENEOUS

One node type, one edge type

Spatial structure is modelled, but every tissue is treated as the same thing.

PatchGCN, SlideGraph+, LA-MIL

HETEROGENEOUS

Types kept, edges undirected

Edge types come from correlation coefficients — *no clear biological meaning* .

HEAT, ProtoSurv

DERECTED, HETEROGENEOUS

Typed *and* directed

Edges follow meta-paths defined by tissue type and direction, so each one names a real interaction.

SAID (This paper)

02

PART TWO

Method

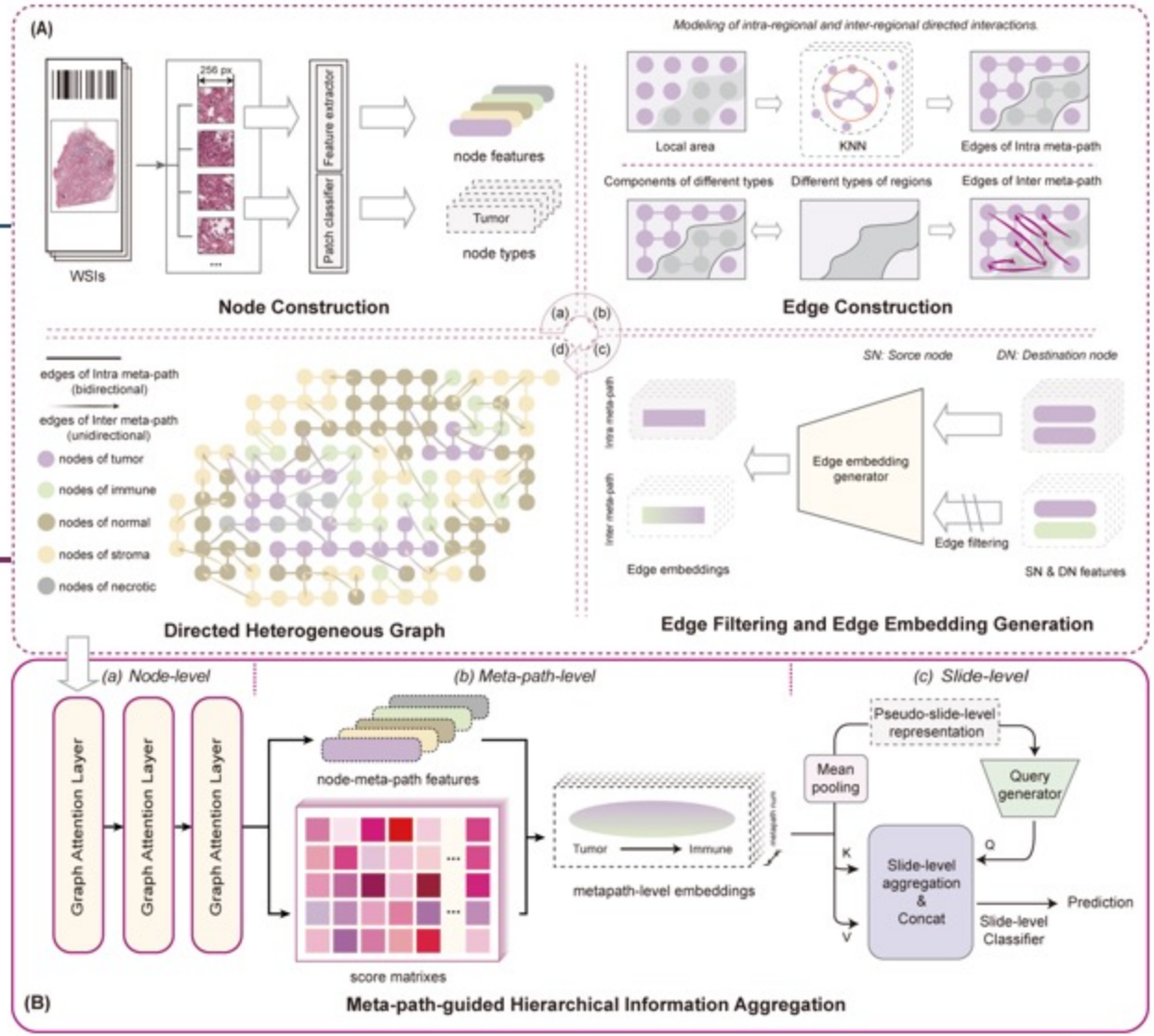
Graph construction and network framework

Method

Overall Framework

(A) Build the graph. Patches become typed nodes; edges are drawn along meta-paths, then embedded and filtered.

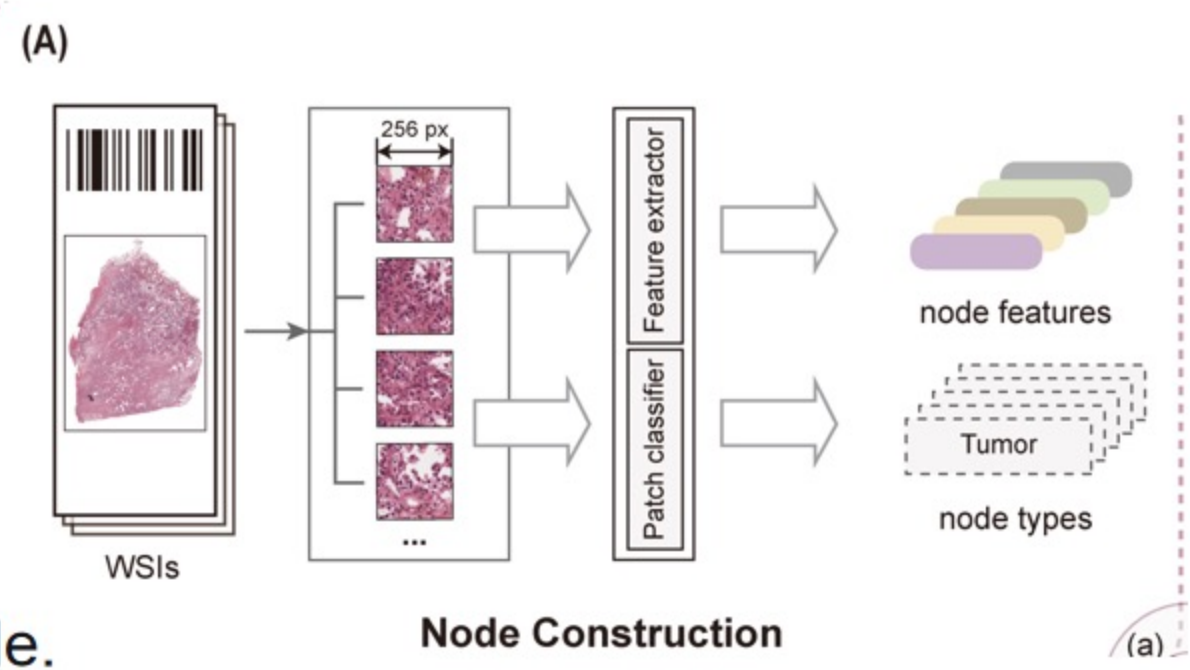
(B) Read the graph. Aggregate node → meta-path → slide, then classify the mutation status.



Method

Nodes: Patch, Type, Position

- **Feature** — frozen PLIP foundation model, 512-d.
- **Type** — HoVer-Net assigns one of the five tissue classes.
- **Position** — the patch coordinate on the slide.



Method

Meta-paths: 25 Typed Relations

A meta-path is a template — **(source type, relation, target type)**.

5

intra-meta-paths

Same tissue type — spatial proximity inside one region.

20

inter-meta-paths

Different tissue types — directed functional interaction.

25

relations in total

Each one gets its own attention and its own embedding.

(Tumor, inter, Immune) is *not* the same relation as (Immune, inter, Tumor).

Method

Edge Construction

INTRA-EDGES

KNN on coordinates

→ *spatial* relationships

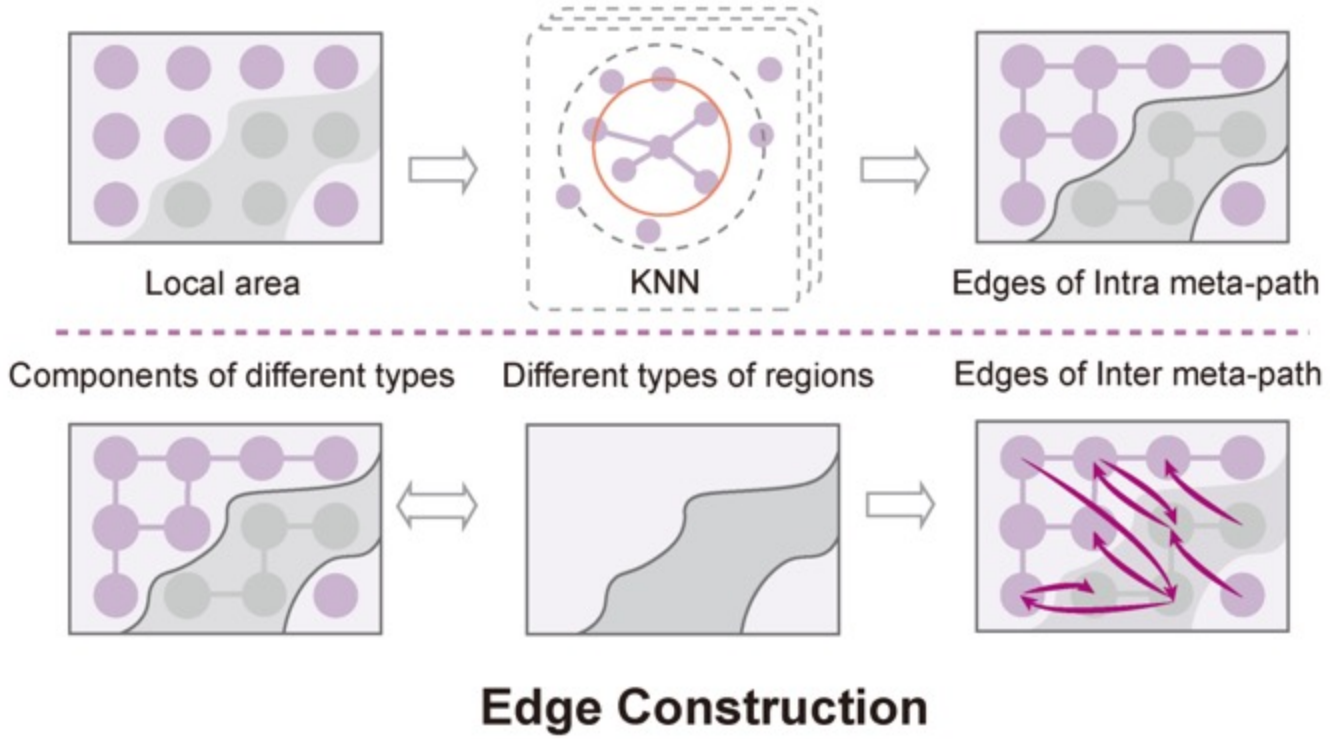
K = 4 | same type | distance-capped

INTER-EDGES

Cosine similarity

→ *functional* interactions

top pairs only | sparse by design



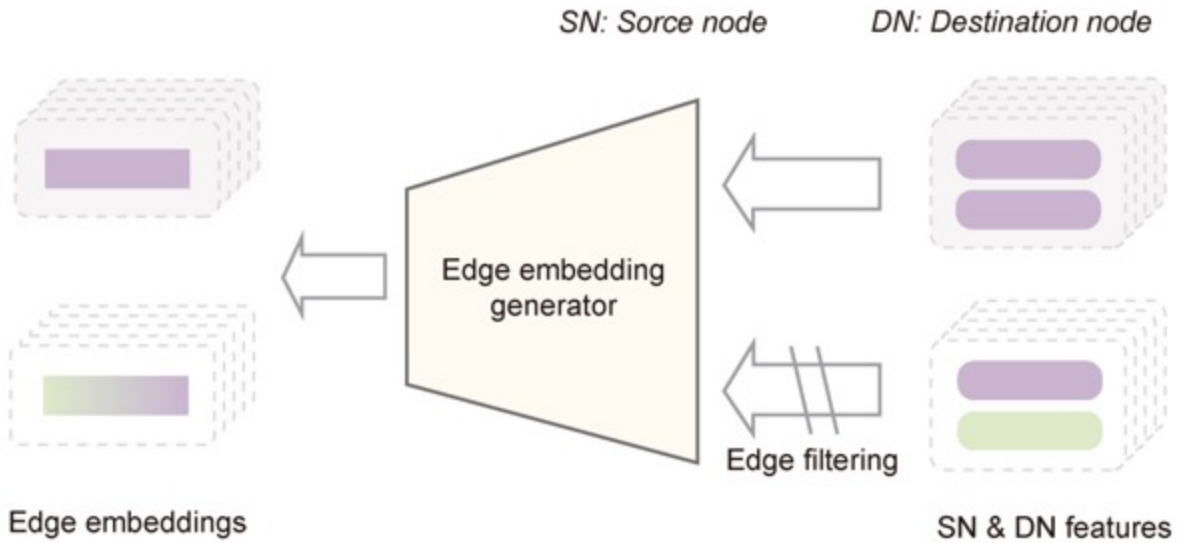
Edge Construction

Method

Edge Embedding and Filtering

- Edge embedding from both endpoints — **source first**
- Order encodes direction — **two directions, two embeddings**
- Learnable score keeps the top **ρ = 0.4** of inter-edges
- Intra-edges never filtered — local structure intact

$$r_{ij} \neq r_{ji}$$

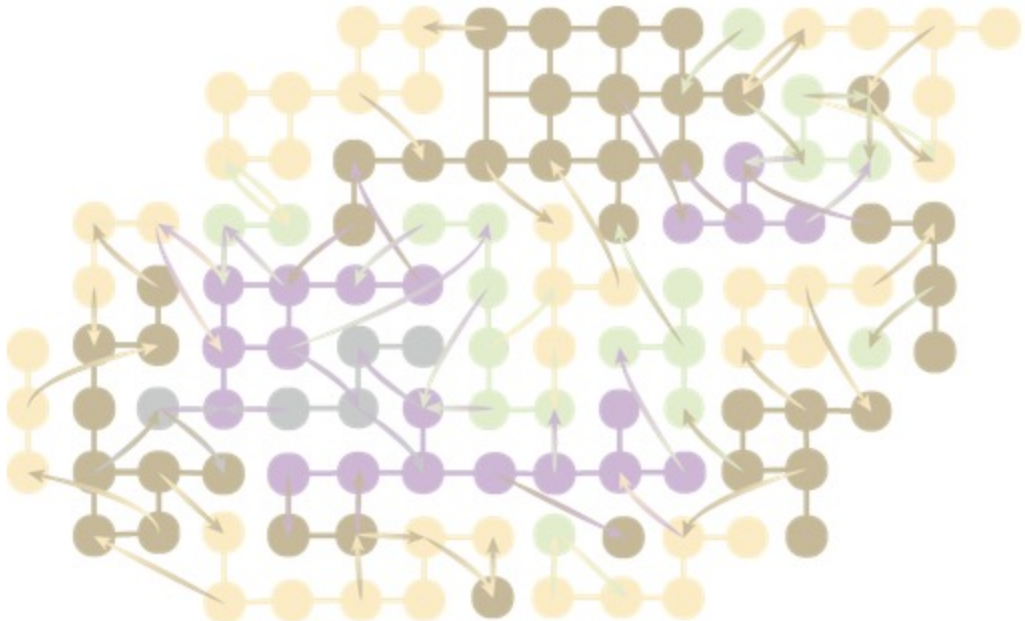


Edge Filtering and Edge Embedding Generation

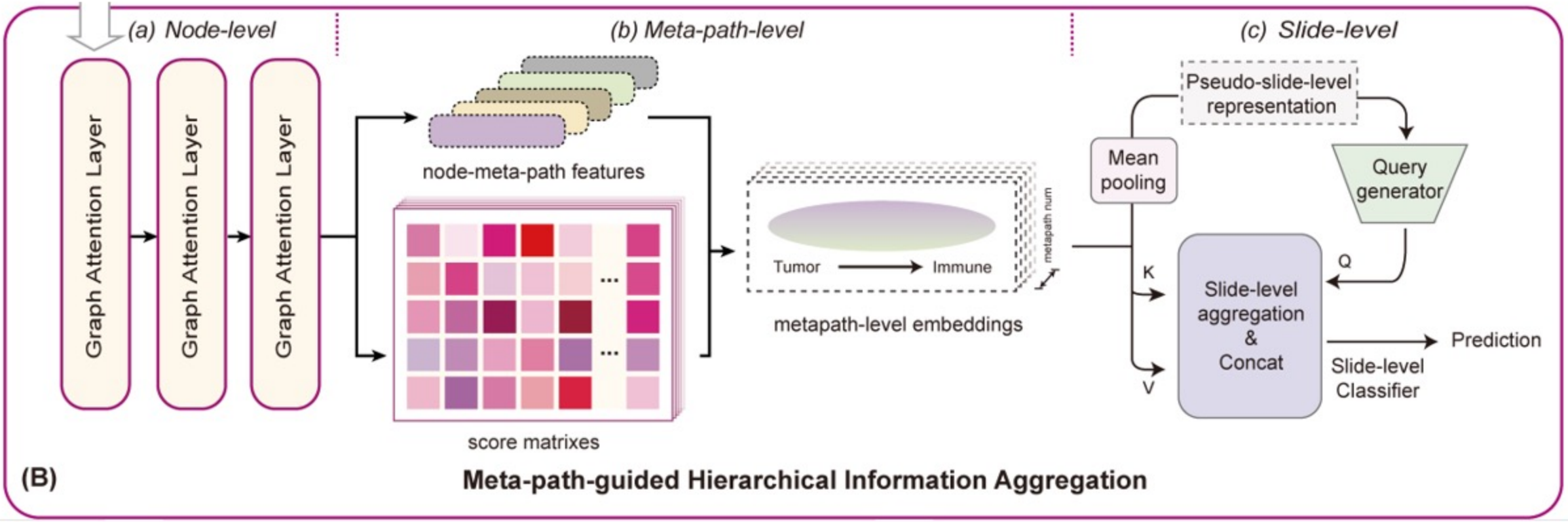
Method

Directed Heterogeneous Graph

- Undirected edges describe **spatial relationships.**
- Directed edges depict the **functional interaction**
- Nodes of different colors represent different types of patches.
- Connected components correspond to the organizational regions.



Method Hierarchical Aggregation



Node level

3 attention layers · weights from edge embeddings

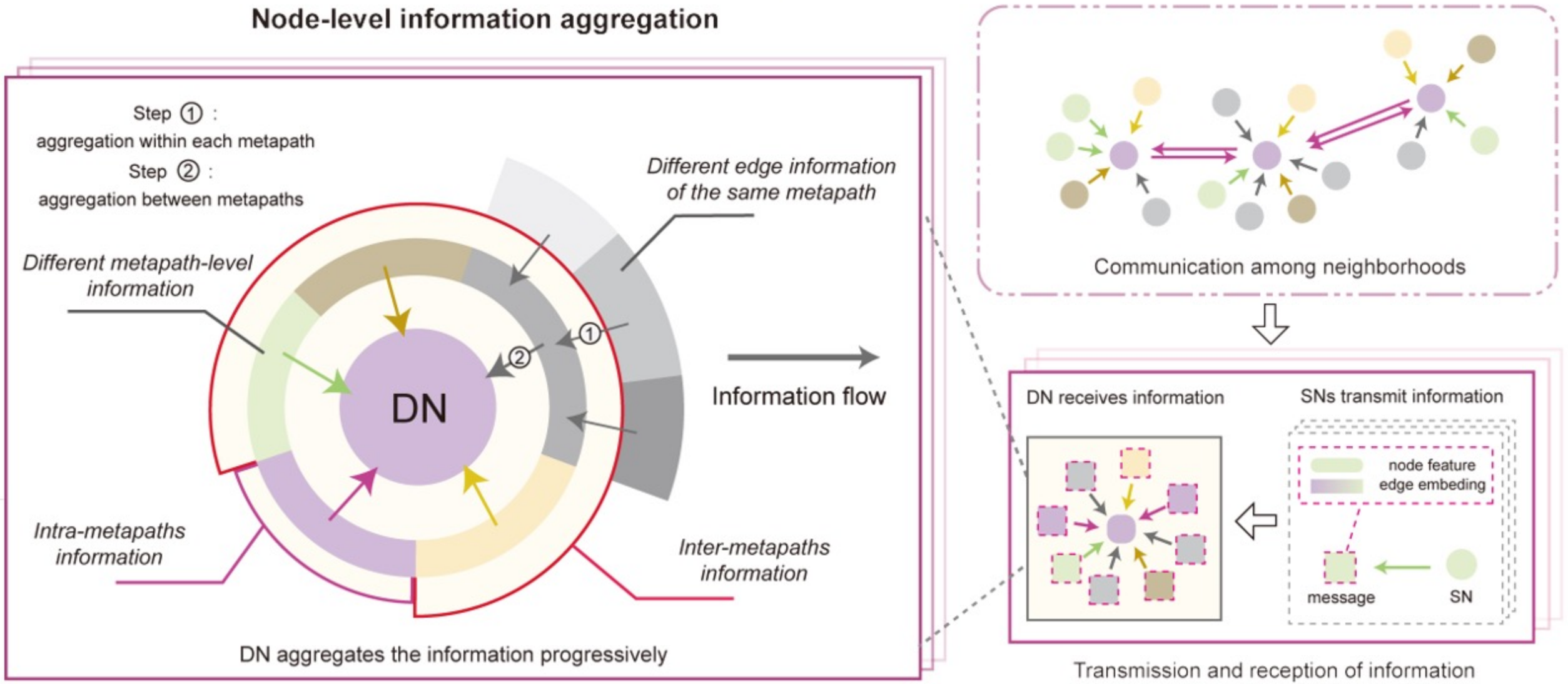
Meta-path level

Normalize across nodes → 1 embedding per meta-path

Slide level

Cross-attention fuses 25 → MLP predicts

Method Hierarchical Aggregation



03

PART THREE

Results

Two TCGA datasets, ablation and sensitivity

Results

Datasets and Setup

TCGA-EGFR

695

whole slide images

589 wild-type · 106 EGFR-mutant

2 classes

TCGA-RAS

222

whole slide images

91 wild-type · 110 RAS · 21 BRAF

3 classes, strongly imbalanced

Identical PLIP features for every method · patient-level 7:3 split · five-fold cross-validation · AUC, macro-F1, accuracy

Results

Performance Comparison

Method	TCGA-EGFR (2 classes)			TCGA-RAS (3 classes)		
	AUC(%)	F1(%)	ACC(%)	AUC(%)	F1(%)	ACC(%)
ABMIL [12]	71.84 ± 2.7	49.29 ± 5.1	84.40 ± 0.7	55.12 ± 6.5	23.78 ± 1.0	49.39 ± 1.4
CLAM [14]	76.91 ± 3.1	64.30 ± 4.6	84.78 ± 1.9	55.13 ± 7.0	35.60 ± 6.1	50.00 ± 5.4
TransMIL [21]	79.66 ± 2.9	61.89 ± 8.7	85.26 ± 1.0	55.51 ± 4.7	35.48 ± 3.8	47.88 ± 3.7
S4MIL [9]	77.93 ± 3.8	66.29 ± 6.2	86.41 ± 1.3	52.84 ± 4.7	30.69 ± 3.5	52.42 ± 3.0
DTFD-MIL [26]	72.33 ± 2.7	55.57 ± 9.0	82.39 ± 2.1	57.92 ± 7.6	35.08 ± 5.1	52.42 ± 5.2
PatchGCN [4]	79.20 ± 1.8	69.22 ± 2.0	84.88 ± 0.8	60.53 ± 7.9	28.90 ± 8.0	51.52 ± 4.3
SlideGraph+ [15]	72.84 ± 1.2	59.91 ± 4.6	83.25 ± 1.6	60.39 ± 4.5	39.93 ± 4.3	55.76 ± 2.2
LA-MIL [20]	75.52 ± 3.6	66.40 ± 3.7	84.59 ± 2.9	61.58 ± 9.5	40.65 ± 6.4	53.33 ± 6.2
HEAT [3]	63.06 ± 1.7	57.48 ± 3.8	81.33 ± 1.9	62.16 ± 1.5	33.55 ± 4.9	50.00 ± 2.1
ProtoSurv [24]	75.13 ± 2.4	58.41 ± 8.6	83.54 ± 2.0	59.79 ± 4.1	34.07 ± 8.1	50.30 ± 3.6
Ours	81.59 ± 1.3	69.67 ± 2.8	84.78 ± 1.4	67.85 ± 3.4	47.06 ± 9.4	57.58 ± 2.8

Results

Ablation and Sensitivity

Inter-edge directionality · TCGA-EGFR

Edge Type	Total	AUC(%)	F1(%)	ACC(%)
Bidirectional	N	78.44 ± 4.0	69.16 ± 4.1	84.68 ± 2.6
Bidirectional	$2N$	76.05 ± 3.9	66.11 ± 3.9	79.61 ± 4.2
Unidirectional	N	$\underline{81.59 \pm 1.3}$	$\underline{69.67 \pm 2.8}$	$\underline{84.78 \pm 1.4}$

Direction adds **+3.15** AUC at equal edges · **+5.54** at equal node pairs

03

PART FOUR

Interpretability

Which interactions carry the signal

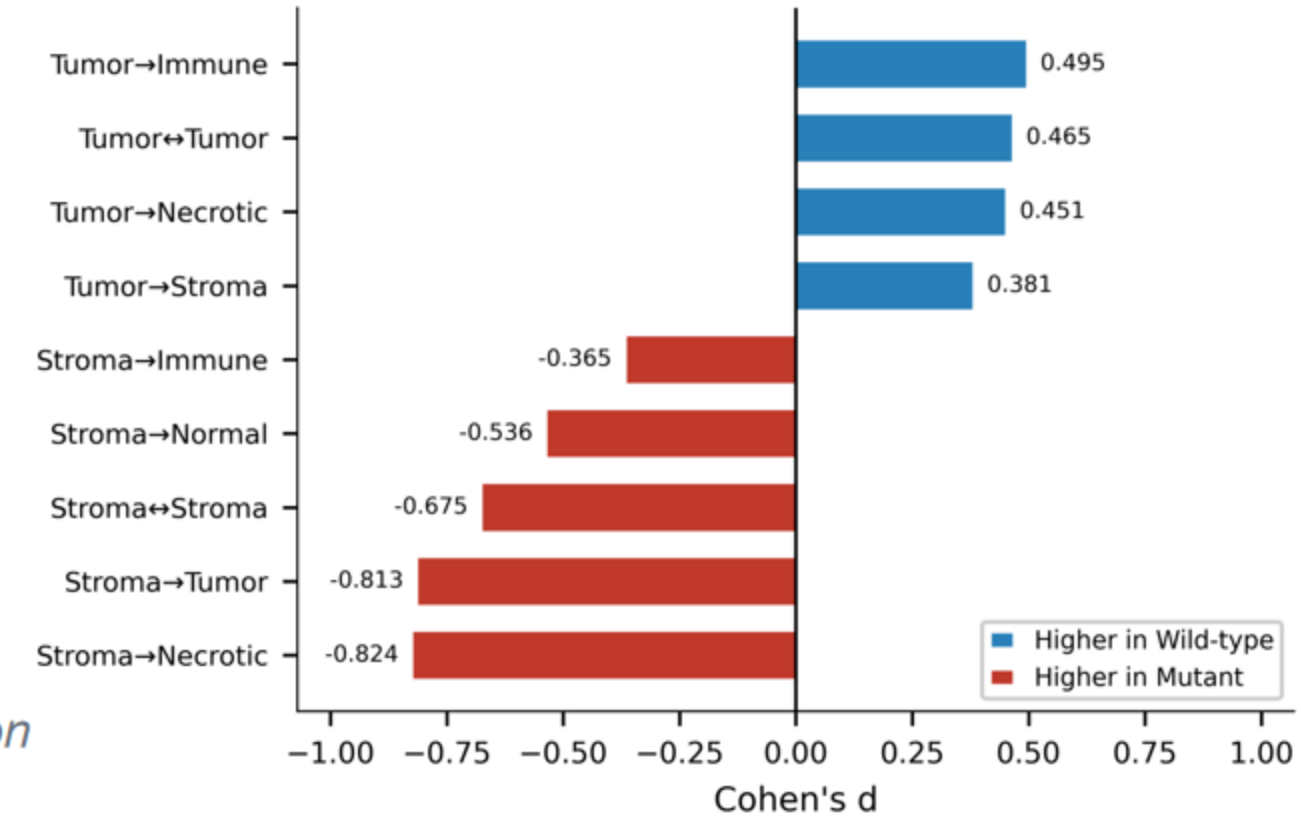
Yifei Wang et al. ICPR'26

Interpretability

Discriminative Interaction Patterns

- All 25 meta-paths tested — Welch's t-test, BH correction
- **9 survive** $P_{FDR} < 0.05$ and $|d| > 0.2$.
- **Higher in wild-type** — every path starts at **Tumor**
- **Higher in mutant** — every path starts at **Stroma**

Cohen's d — standardized difference in mean attention



Interpretability

Biology Behind the Pattern

Wild-type

Higher mutational burden · **more immune infiltration**

Tumor → Immune

$d = +0.495$

EGFR-mutant

Stromal hyperplasia · immunosuppressive microenvironment

Stroma → Tumor · Stroma → Immune

$d = -0.813 \cdot -0.365$

Interaction-level interpretability — matches known TME biology

Interpretability

Asymmetry Between Tumor and Stroma

Tumor → Stroma

+0.381

higher attention in wild-type

Stroma → Tumor

−0.813

higher attention in mutant

Same tissue pair, opposite directions, **opposite signals** . An undirected edge merges them — **|d| = 1.194** lost.

Thanks for your attentions!

SAID: Spatial and Interaction-Aware Directed Heterogeneous Graph Neural Network for Gene Mutation Prediction from Histopathology Whole Slide Images

Yifei Wang, Jun Shi*, Shihao Wu, Jiyang Li, Zhiguo Jiang, Yushan Zheng*

*Corresponding author: juns@hfut.edu.cn, yszheng@buaa.edu.cn