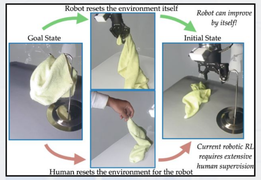
SciGA: A Comprehensive Dataset for Designing Graphical Abstracts in Academic Papers

Takuro Kawada, Shunsuke Kitada, Sota Nemoto, Hitoshi Iyatomi
Hosei University, Tokyo, Japan

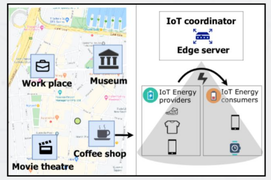
CVPR JUNE 3-7, 2026 DENVER COLORADO

Motivation

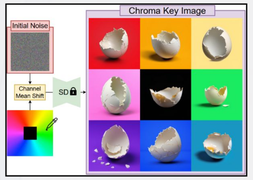
- **GAs are journal-hosted visual summaries** that highlight a paper's key findings
- They can increase paper visibility and social media engagement [Bennett+, Scientometrics'23]
- In-paper teasers, often shown as *Figure 1*, serve as de facto GAs
- Creating compelling GAs requires advanced visualization skills [Jeyaraman+, Cureus'23]
- ➜ **We support accessible GA design to enhance scientific communication**



[Sharma+, CoRL'23]



[Lakhdari+, ICSOC'20]



[Morita+, CVPR'25]

SciGA-145k Dataset

145,080 papers	1,148,882 figures	309 GAs	30,724 Teasers

- **The first large-scale dataset to support GA design**
  - Textual data: full-text / sections / footnotes / tags
  - Visual data: figures / captions / subfigures / GA labels
  - Metadata: research fields / comments / venues / etc.
- Source: arXiv.org Jan. 2021 – Nov. 2025
- License: C-UDA 1.0

▼ CLIP embed. of GAs/Teasers

▼ CLIP embed. of Abstracts

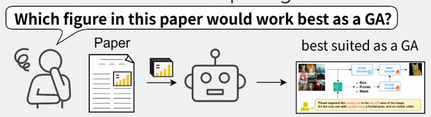
- CV
- Cond. Matter
- NLP
- Math
- Network Architecture
- Others

▼ Comparison with Conventional Datasets

Datasets	Support Contents: Full-text	Support Contents: Figures	#Papers	#Figures	Annotations: #GA	Annotations: #Teaser
S2ORC [Lo+, ACL'20]	✓	✗	81,100k	N/A	N/A	N/A
unarXive 2022 [Rodriguez+, WACV'23]	✓	✗	1,900k	N/A	N/A	N/A
Paper2fig 100k [Rodriguez+, WACV'23]	✗	✓	69k	102k	N/A	N/A
ArxivCap [Rodriguez+, WACV'23]	✗	✓	572k	6,400k	N/A	N/A
CentralFigure [Yang+, ICDAR'19]	✗	✗	7k	N/A	N/A	7,295
MMSci [Li+, arXiv'24]	✓	✓	131k	742k	N/A	N/A
SciGA-145k (ours)	✓	✓	**145k**	**1,100k**	**309**	**30,724**

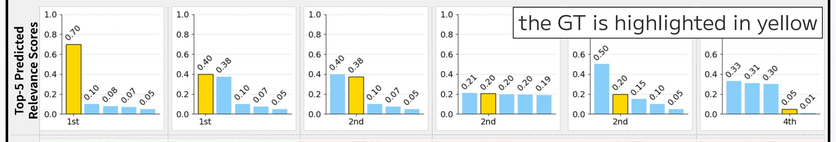
Intra-GA Recommendation

- A task of **identifying GA-suitable figures** within a paper
- ➜ Helps platforms present GAs without requiring authors to submit them



Confidence Adjusted top-1 GT Ratio (CAR)

- Conventional metrics rely on the ground-truth (GT) rank
- Multiple figures may plausibly serve as GAs in Intra-GA Recommendation



R@1	1.000	1.000	0.000	0.000	0.000	0.000
nDCG	1.000	1.000	0.631	0.631	0.631	0.431
CAR	0.868	0.687	0.644	0.488	0.269	0.106
	Clearly Correct	Uncertain Success	Understandable Error		Clearly Incorrect	

$$\mathrm{CAR@}k = \underbrace{\frac{p_{\mathrm{GT}}}{p_{\text{top-1}}}}_{\text{Top-1 GT Ratio}} \times \underbrace{\left(1 - \frac{1}{2}\max\left(0, \frac{H(P) - \log k/2}{H_{\max}(P) - \log k/2}\right)\right)}_{\text{Confidence}}$$

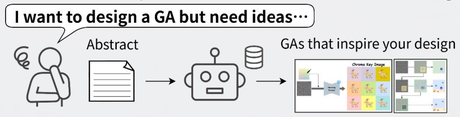
- P : normalized top-k relevance scores
- $p_{\text{top-1}}, p_{\mathrm{GT}} \in P$: top-ranked / labeled-GA scores
- $H(P)$: entropy of P
- $H_{\max}(P) = \log k$: max entropy

GA near top? [0.0, 1.0] — Ranking certainty. [0.5, 1.0]

➜ **CAR provides an interpretable, instance-level continuous score in [0, 1]**

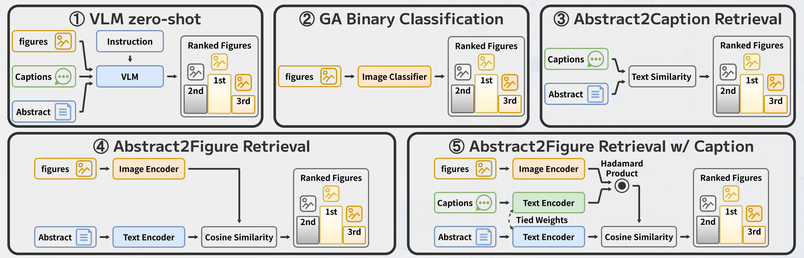
Inter-GA Recommendation

- A task of **retrieving GAs as design inspirations** from other papers
- ➜ Helps authors create more impactful GAs with concrete design references



Experiments

Benchmark Methods



Intra-GA Recommendation

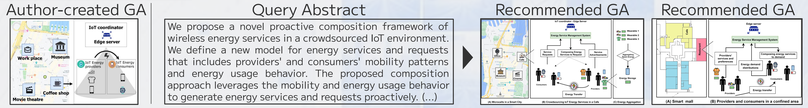
Qualitative Results [Brunnbauer+, ICRA'25]

Quantitative Results

Backbone	Recall@1	Recall@3	CAR@5
① GPT-5.2 [OpenAI, 2025]	0.438	0.871	0.526
② CLIP [Radford+, ICML'22]	0.493	0.826	0.518
③ BM25 [Robertson+, TREC-3'94]	0.508	0.849	0.528
④ SigLIP2 [Tschannen+, arXiv'25]	0.558	0.869	0.588
⑤ LongCLIP [Zhang+, ECCV'24]	**0.637**	**0.914**	**0.615**

Inter-GA Recommendation

Qualitative Results [Lakhdari+, ICWS'21]



Selecting a Silver-Standard Signal via User Study

- No explicit GT
- ➜ We compare human 2AFC choices with four axes to select a silver-standard relevance signal.

- **Field Match** — same arXiv category
- **Semantic Coherence** — query – retrieved abstract similarity
- **Visual Coherence** (Human-aligned!!) — author-created – retrieved GA similarity
- **Aesthetic Quality** — visual appeal / readability

Quantitative Results

Backbone	nDCG@5	nDCG@10	nDCG@30
Rand. Samp.	0.770	0.794	0.862
③ CIDEr [Vedantam+, CVPR'15]	0.820	0.839	0.892
④ CLIP [Radford+, ICML'22]	**0.855**	**0.870**	**0.913**
⑤ OpenCLIP [Cherti+, CVPR'23]	0.824	0.843	0.894

Ongoing Work

We are extending SciGA from GA recommendation toward automatic editable GA generation!!