[Q3-3]

SciGA-Vec：学術論文におけるベクタ画像形式の Graphical Abstract データセット

川田 拓朗, 北田 俊輔, 彌冨 仁 法政大学大学院 理工学研究科 {takuro.kawada.3g@stu.,iyatomi@}hosei.ac.jp

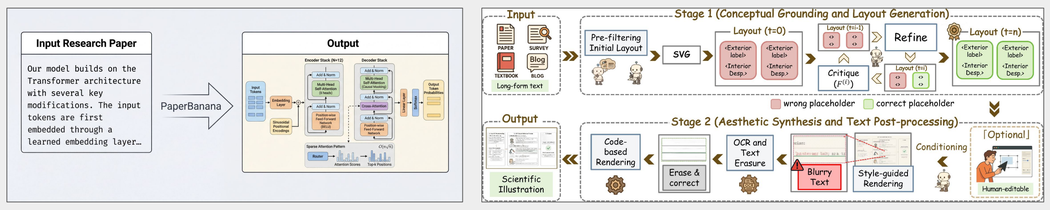
NLP 2026 UTSUNOMIYA

本研究のポイント

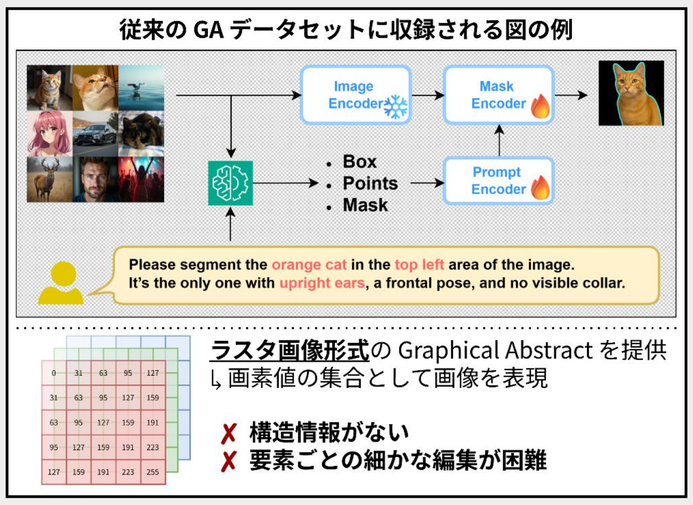
- PowerPoint 等で**編集可能なポンチ絵の自動生成の向けたデータ基盤**を構築
- **ラスタ画像形式の Graphical Abstract をベクタ画像形式に変換**するパイプライン

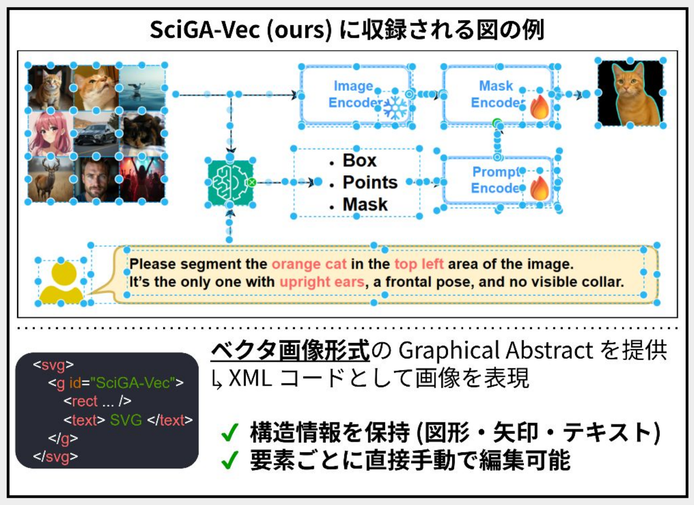
Background | 学術論文の顔となるポンチ絵 Graphical Abstract (GA)

🔑GA = 論文の核心的な内容を伝える図 (所謂 Fig.1)
😄論文の注目度を高める [Bennett+, Scientometrics'23]
➜ 自動生成の試みが活発化 [Zhu+, ICLR'25; Zhu+, arXiv'26]



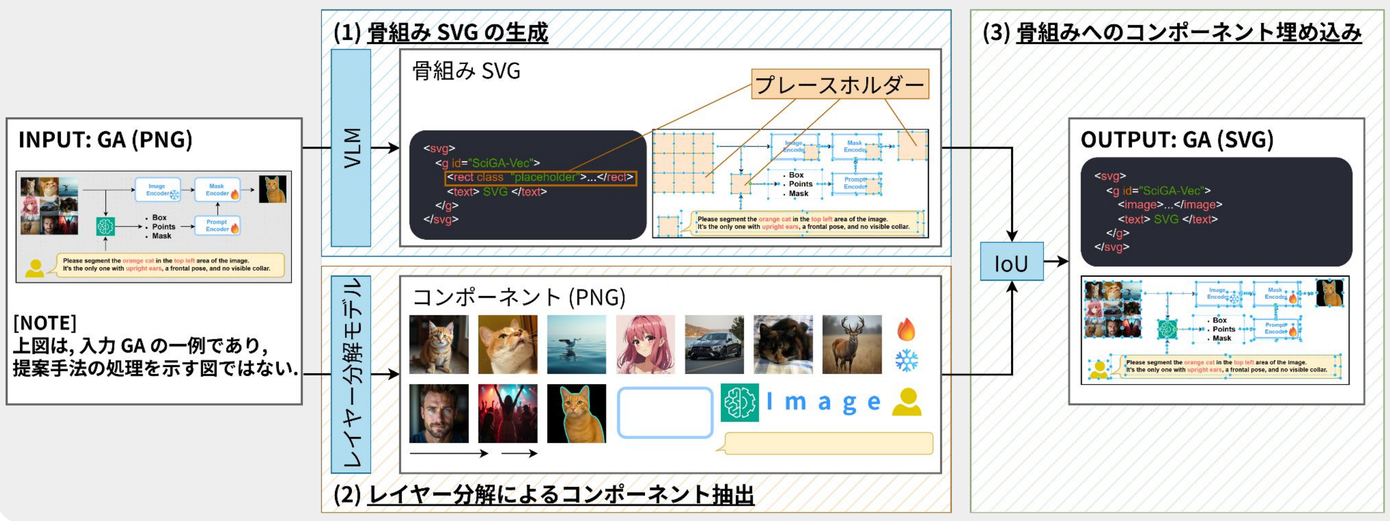
😭従来モデルの出力はラスタ画像で編集不可能
➜ **ベクタ画像形式の GA のデータ基盤が求められる**





SciGA-Vec | 編集可能なポンチ絵生成のデータ基盤

- GA データセット SciGA-145k [kawada+, CVPR'26 Findings] に含まれるラスタ画像形式の GA をベクタ画像 (SVG) として再構成
- 写真等はプレースホルダーを置き, ラスタ画像のまま埋め込む



バックボーンモデル

- VLM: Gemini-3-Pro [Google, 2025]
- レイヤー分解モデル: LayerD [suzuki+, ICCV'25]

従来データセットとの比較

データセット	画像形式	Infographic を対象とするか？
ChartGalaxy [Li+, arXiv'25]	Raster	✓ (チャート図, 表)
ArxivCap [Li+, ACL'24]	Raster	✓ (科学論文の図)
SciGA-145k [Kawada+, CVPR'26 Findings]	Raster	✓ (GA)
OmniSVG [Yang+, NeurIPS'25]	Vector	✗ (アイコン, イラスト)
UniSVG [Li+, ACMMM'25 Dataset Track]	Vector	✗ (アイコン)
SciGA-Vec (ours)	**Vector**	**✓ (GA)**

GA 1 枚あたりに含まれる各要素の統計情報

# 矩形	#矢印	# テキスト	# 画像
0.26 ± 14.77	19.84 ± 20.55	2.40 ± 3.73	4.38 ± 6.72

Results | ベクタライズされた GA の品質評価実験

(i) 視覚的情報保持性

- 元画像 x と再構成画像 x′ の視覚的類似度を図り, ベクタライズ前後で外観が保持されているかを評価

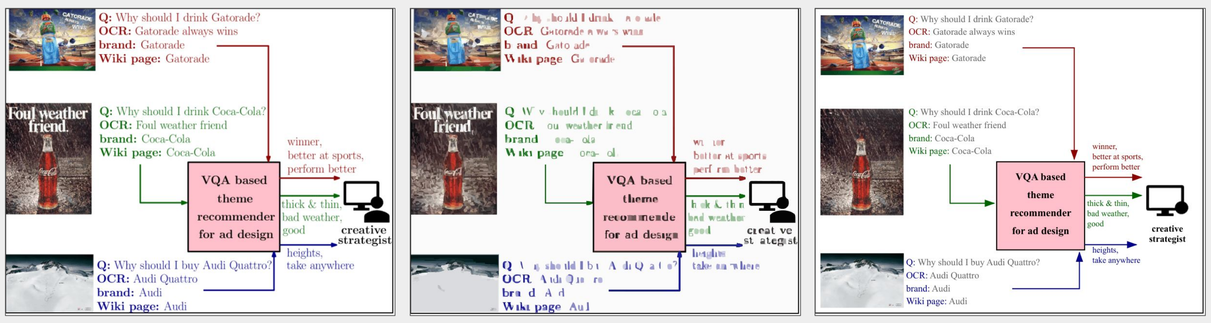
(ii) 意味的情報保持性

- GA だけを見て論文の貢献を当てる 4 択 VQA を解き, 主要貢献の伝達能力が保持されているかを評価

$$SemanticRetentionScore(SRS) = \frac{(\text{再構成画像}x'\text{入力時の}Acc.)}{(\text{元画像}x\text{入力時の}Acc.)}$$

(iii) コードのトークン長

- 再構成画像 x′ の SVG コードの表現効率を評価



(a) 元画像 (b) VTracer (c) Ours

➜ 各要素 (テキスト・図形) を保持したベクタ GA を再構成し, 従来手法の**ストロークで写真や文字を近似する問題を解消**

ベクタライズ手法	(i) 視覚的情報保持性					(ii) 意味的情報保持性		(iii) コードのトークン長	
	FID ↓	CLIP-I ↑	PSNR ↑	SSIM ↑	LPIPS ↓	Gemini-3 ↑	GPT-5.2 ↑	gemma-2b ↓	o200k_base ↓
VTracer [VisionCortex, 2020]	**31.654**	0.876	**33.912**	**0.878**	**0.172**	**0.936**	0.934	520,645	687,300
Our Method	44.210	**0.903**	33.324	0.573	0.348	0.926	**0.946**	**421,214**	**403,273**

➜ **視覚・意味保持性能は従来手法と同等**で, フォントの違いや座標のズレからSSIM/LPIPS は低下するものの**表現効率は向上**

Application | 編集可能な GA の自動生成

- 論文＋**SVG 参照例**から, Gemini-3-Pro で GA (SVG) 生成
  arXivID: 2401.14185 の GA 生成例

(a) PaperBanana [Zhu+, arXiv'26] (b) Few-shot w/ SciGA-Vec

➜ 情報粒度の整った編集可能な GA を安定して生成可能に

Conclusion

- 構造要素を保持したベクタ画像に変換するベクタ化手法, ベクタ化された GA データセット SciGA-Vec を提案
- ストローク近似に依存しない構造保持型ベクタ表現を実現
- 意味情報を保持しつつ編集可能性と表現効率を向上

今後の展望

- より魅力的な GA を自動生成する手法の開発
- 科学図に限らない Infographic への拡充