

# GenGA：学術論文における編集可能な Graphical Abstract の自動生成に関する初期検討

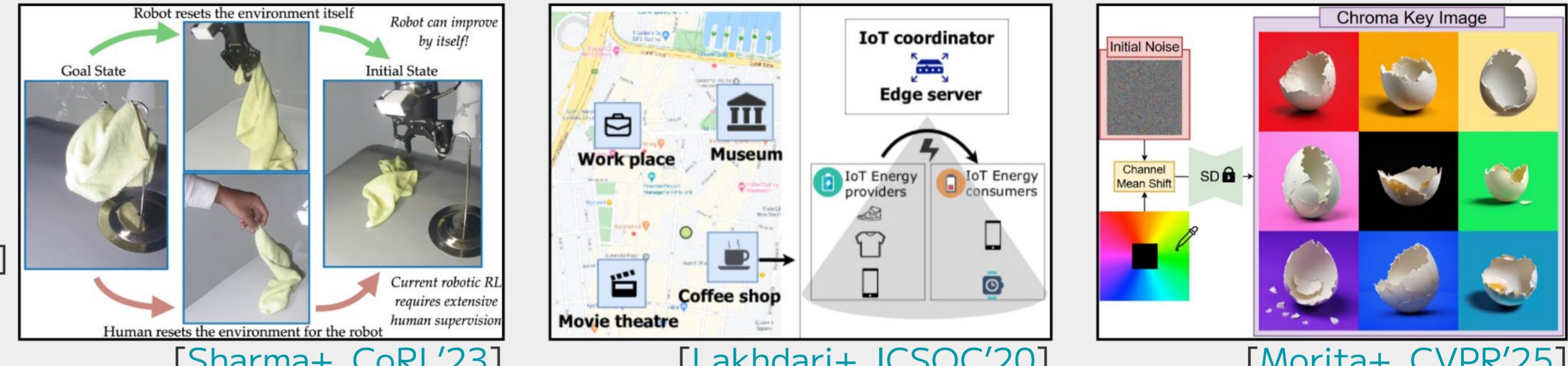
川田 拓朗, 北田 俊輔, 彌富 仁 法政大学大学院 理工学研究科 {takuro.kawada.3g@stu., iyatomi@}hosei.ac.jp

## 本研究のポイント

- Vision Language Model (VLM) を用いたポンチ絵生成フレームワーク GenGA を提案
- PowerPoint 等で編集可能なベクタ形式での出力を実現し、その有効性と限界を示す

## 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 (いわゆる Figure 1)
- 論文の注目度・SNS での拡散力を高める [Bennett+, Scientometrics'23]
- 読者に論文を読むきっかけを与える [Jeyaraman+, J. Orthop.'23]
- 効果的な GA の作成にはデザインスキルが必要 [Jeyaraman+, Cureus'23]
- GA 作成支援技術は科学コミュニケーションを効率化



## GenGA | 論文から「編集可能なポンチ絵」を描くフレームワーク

### ① 参照 GA の検索 (Reference GA RAG)

Abst. をクエリとし、デザインアイデアとなる既存論文の GA を検索

### ② GA の生成

Abst. + Intro. (+ 参照 GA) を入力とし、VLM で SVG 形式の GA をテキストコードとして生成

### ③ ラスタライズ

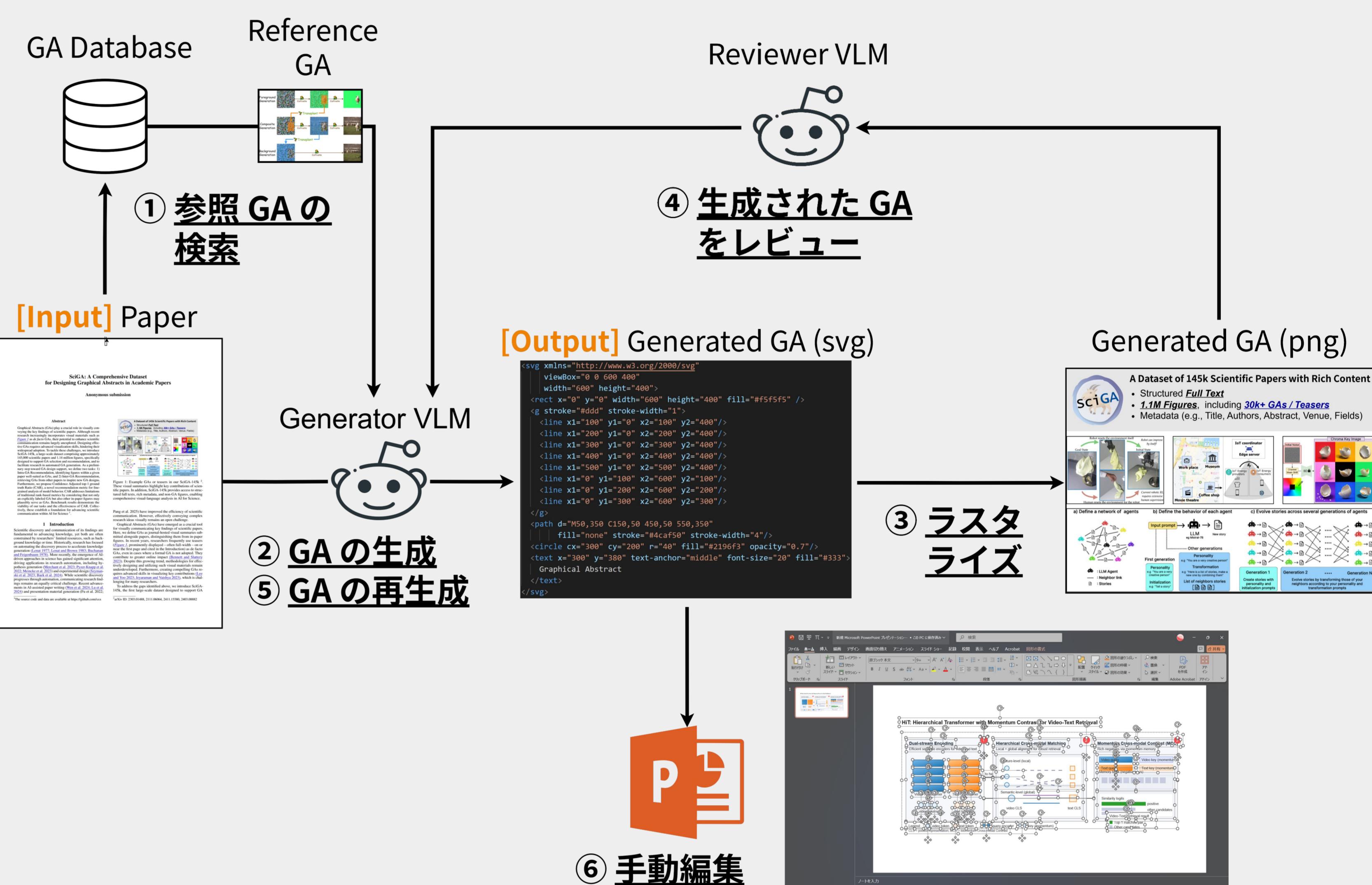
生成された SVG を PNG に変換し、GA を可視化

### ④ 生成された GA をレビュー

別の VLM で論文本文との整合性や視覚的可読性を評価

### ⑤ GA の再生成 (Self-Correction)

フィードバックをプロンプトに追加し、再度 GA を生成

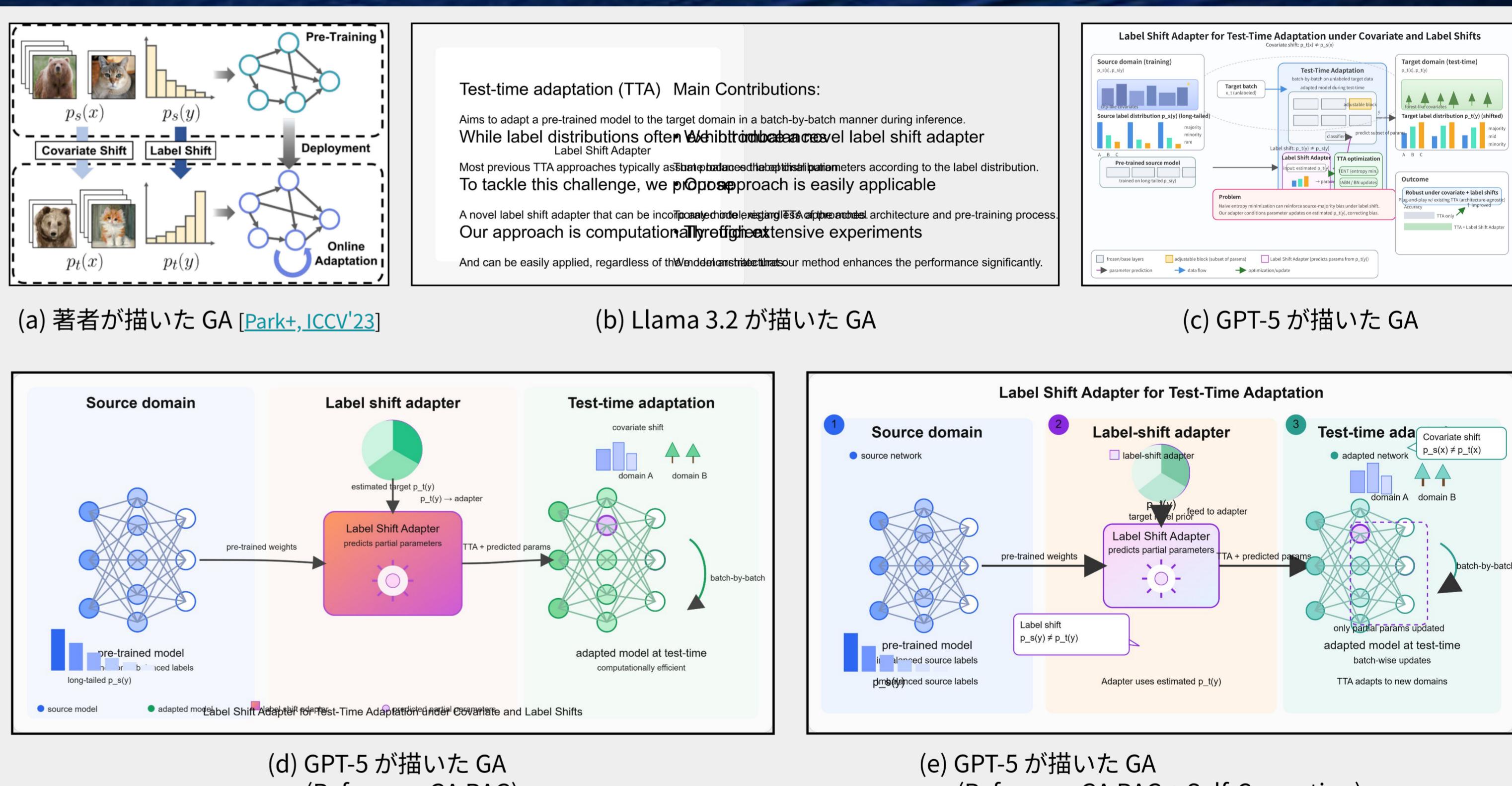


## Results | 生成された GA は一定の意味的整合性を持つが、構文安定性とレイアウト品質はまだ不十分

- データセット: SciGA-145k [Kawada+, arXiv'25] (CS 分野 500 枚)

Models	Reference GA RAG	Self-Correction	有効 SVG 生成率 (%) ↑	要素重なり面積率 (%) ↓	Abst. と生成 GA の CLIPScore ↑	実際の GA と生成 GA の CLIPScore ↑
Llama 3.2 11B Vision Instruct [Meta AI'25]	—	—	92.2	33.9	0.259±0.143	0.249±0.106
	✓	—	00.0	—	—	—
	—	✓	90.3	34.5	0.252±0.142	0.235±0.111
Gemini 2.5 Pro [Comaniciu+, arXiv'25]	✓	—	68.3	26.6	0.325±0.148	0.263±0.107
	—	✓	72.6	17.7	0.313±0.188	0.268±0.127
	✓	✓	46.6	30.6	0.308±0.165	0.256±0.113
GPT-5 [OpenAI'25]	✓	—	52.7	23.9	0.344±0.187	0.257±0.139
	—	—	58.8	49.2	0.311±0.148	0.284±0.097
	✓	—	62.5	44.1	0.327±0.145	0.291±0.097
	—	✓	46.3	50.3	0.301±0.145	0.284±0.092
	✓	✓	43.9	45.2	0.337±0.153	0.290±0.105

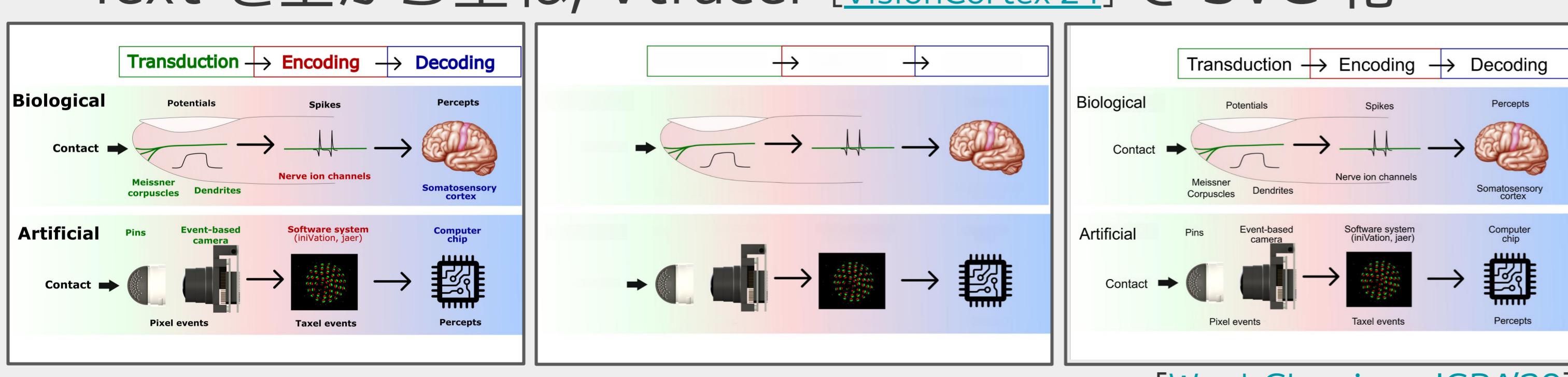
→ Reference GA RAG により、視覚的にすっきり簡潔に👍



→ Self-Correction により、デザイン性 UP 👍 / 情報過多 🙏

## SciGA-Vec | SVG 形式の GA データセット構築

- PNG 形式の実 GA を SVG 化し、学習・評価に使いたい
- OCR + LaMa [Suvorov+, WACV'22] で Inpaint Text を上から重ね、Vtracer [VisionCortex'21] で SVG 化



写真、イラストがストロークで描かれてしまう  
→ SVG コードのトークン長が数十万に...

## Conclusion

- 編集可能な Infographic 自動生成の可能性を実証
- 生成の安定性と品質（簡潔さ・美しさ）に課題
- LLM 間で SVG 生成力とデザインセンスに明確な

## Future Work

- 評価指標の拡充: 良い GA とは何か?
  - 主要な貢献のみを簡潔に抽出できているか
  - 視線誘導がスムーズか
- User Study による妥当性検証
- データセットの整備
- 生成モデルの改良