

# [S3-P27] 学術論文におけるGraphical Abstract自動生成の初期検討

川田 拓朗<sup>1</sup>, 根本 颯汰<sup>2</sup>, 北田 俊輔<sup>2</sup>, 彌富 仁<sup>1,2</sup> <sup>1</sup>法政大学 理工学部, <sup>2</sup>法政大学大学院 理工学研究科

{takuro.kawada.3g@stu.,iyatomi@}hosei.ac.jp



## Summary

### 分野間の特徴を考慮したGraphical Abstract (GA) 自動生成手法の探索

- GA, 本文, 図, メタデータを含む論文データセットを構築
- 論文内からGAの構成要素となり得る図を検索

## Background

- 😊 GAは論文の要点を視覚的に表し、読者の迅速な理解を促す
- 😊 GAにはImpact Factor増加効果がある [Lee+, Sci. Editing'23]
- 😓 効果的なGA作成には手間とデザインスキルが求められる
  - より効果的なGAの分析, および提案システムが望まれる
- GAの多くは論文内の図をそのまま利用するか, 複数の図を組み合わせて作成される
  - 初期検討として, 論文内からGAの構成要素となる図を検索

## Dataset

- ar5iv:04:2024 dataset [Ginev, 2024] を基に28,935件の論文に関するTextデータ (Abstract, 本文, 図のcaption) を収集
- 別途, メタデータ (題目, 著者, 投稿日, 分野, コメント, DOI), 256,212件のImageデータ (GA, 図) を収集

## Methods

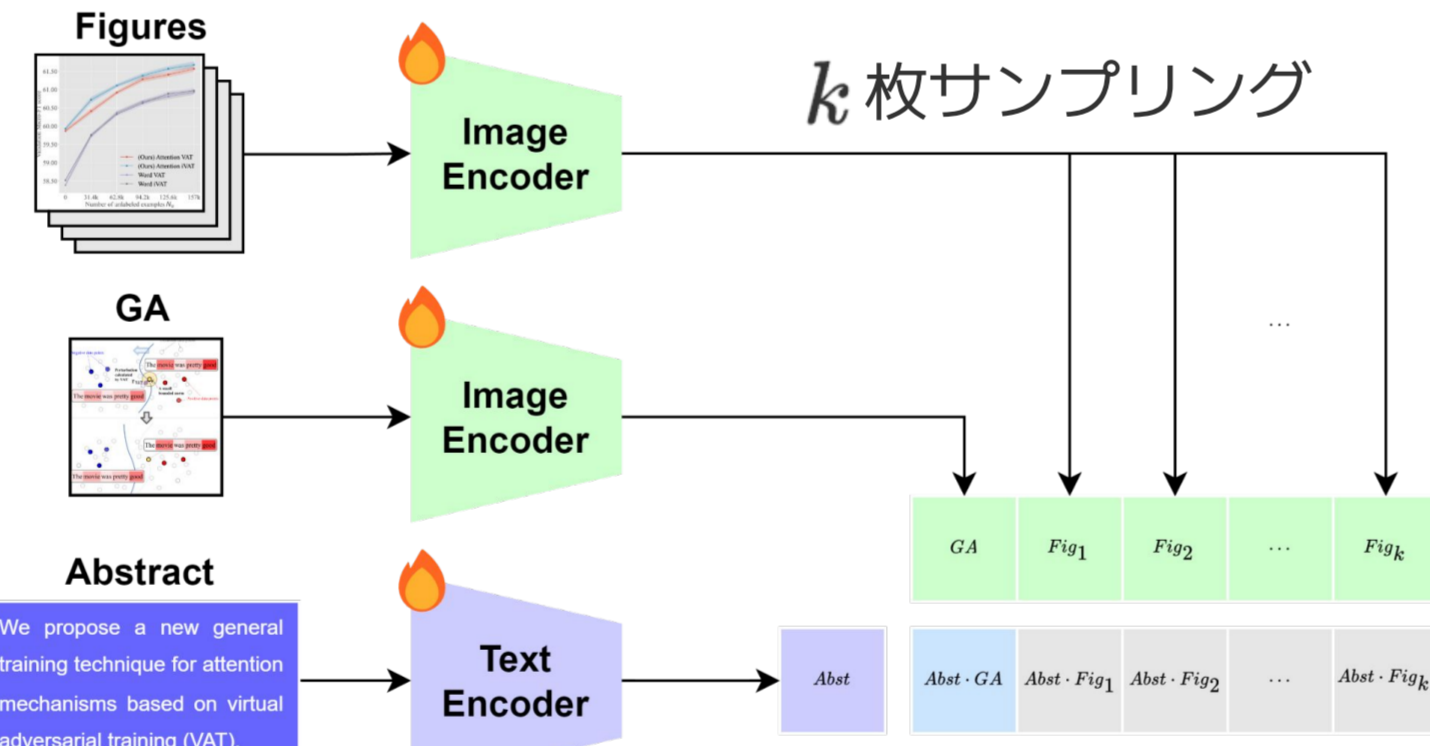
### 対照学習によるAbst2GA

Abstractと最も類似度の高い画像をGAとして検索するタスク Long-CLIP [Zhang+, ECCV'24] で, AbstractとGA, 図を対照学習 → 248トークンのTextを入力でき, Abstract全体を処理可能

#### Intra Loss $\mathcal{L}_I$

同一論文内で, GAと他の図を識別できるように

- $k$  枚の図をサンプリング 不足分は0パディング
- 図の枚数の平均値: 8.74 枚 中央値: 6 枚
- $k = 7$ を採用



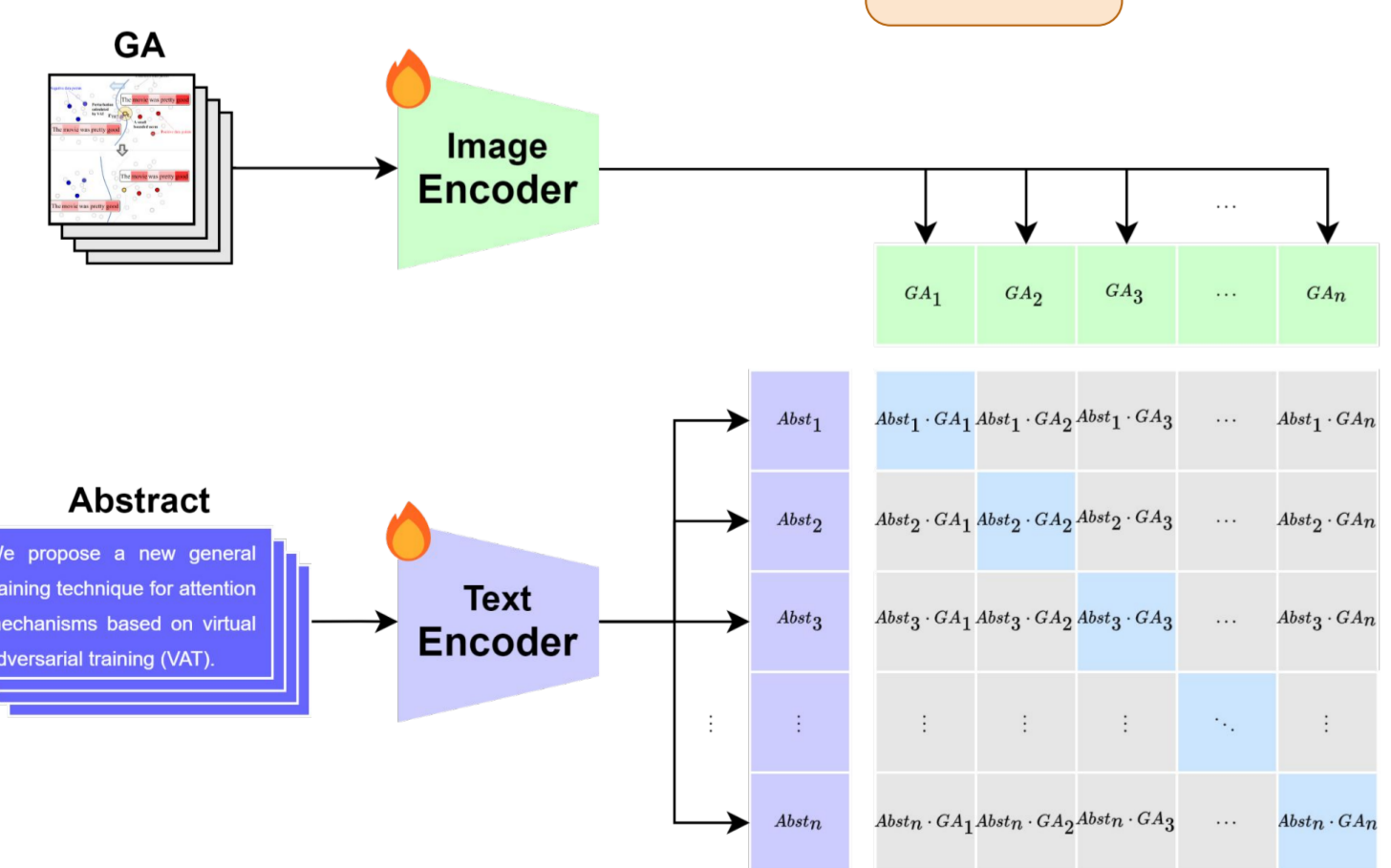
#### 1. $\mathcal{L}_I$ に異なるLossを加算

##### Inter Loss $\mathcal{L}_{In}$

論文間で, GAの違いを識別 "できる" ように

- AbstractとGAで対照学習

$$\mathcal{L}_{I+In} = \mathcal{L}_I + \lambda \mathcal{L}_{In}$$

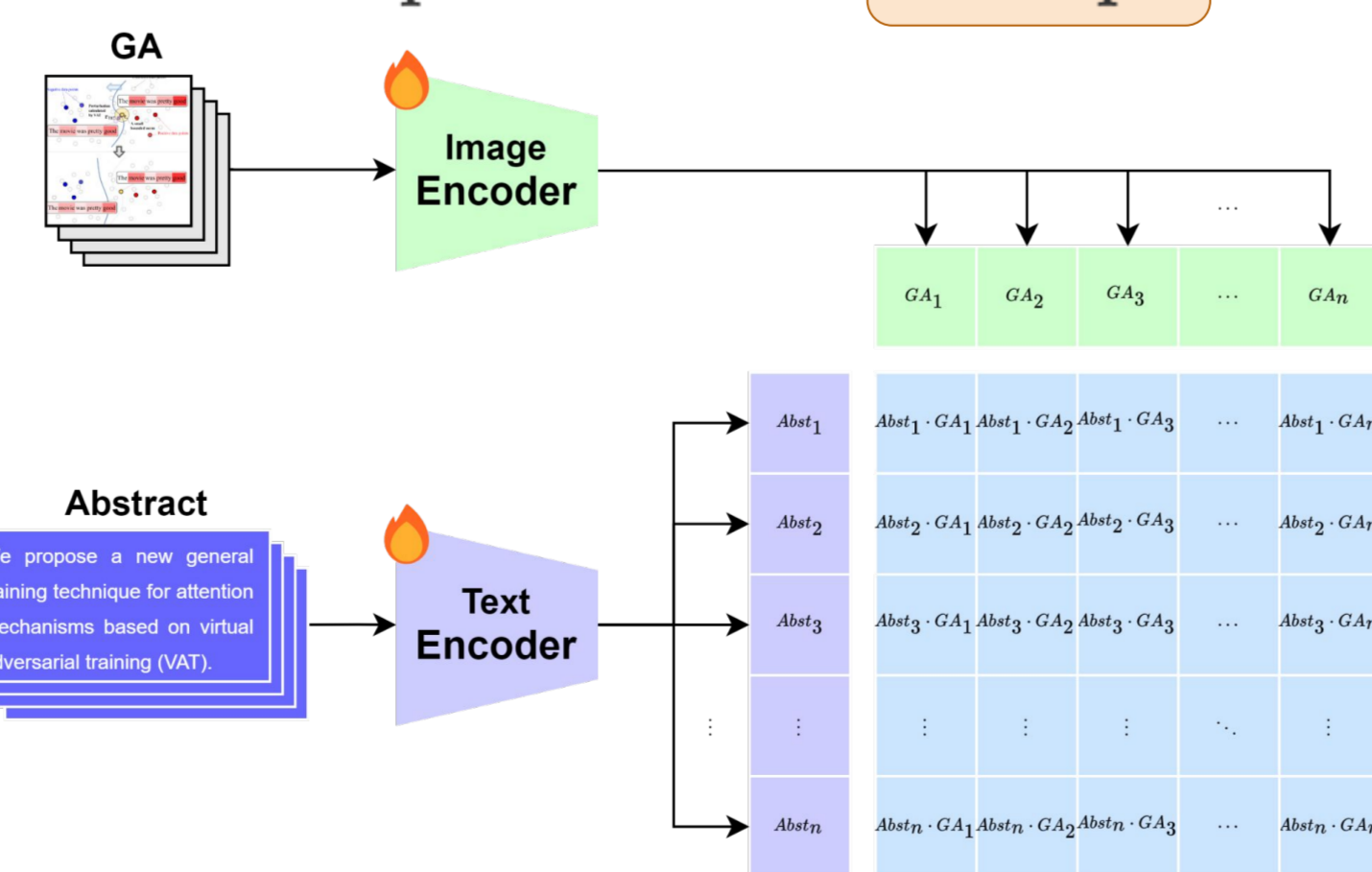


##### Equalized-Inter Loss $\mathcal{L}_{EqIn}$

論文間で, GAの違いを識別 "できない" ように

- 正解ラベルを  $1/n$  に平滑化

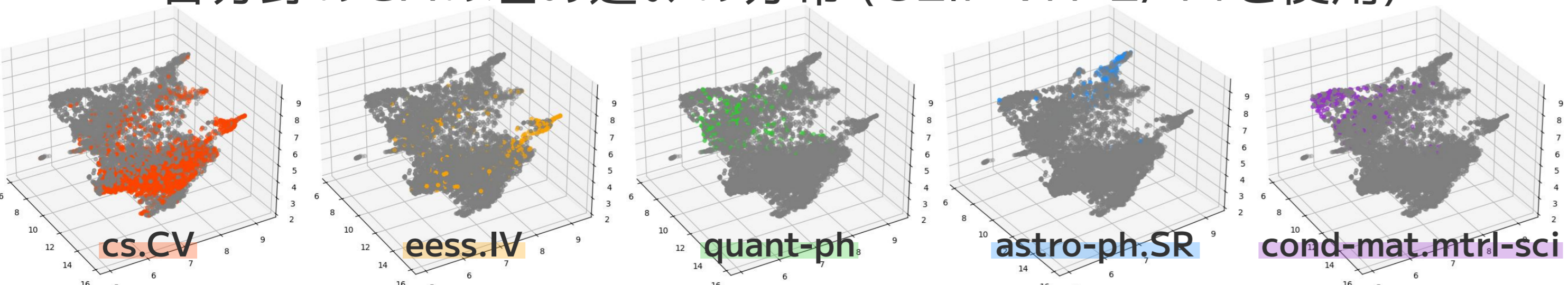
$$\mathcal{L}_{I+EqIn} = \mathcal{L}_I + \lambda \mathcal{L}_{EqIn}$$



#### 2. Domain-Consistent Batch Sampling (DCSamp)

GAは分野ごとに特徴がある (e.g., CVでは画像を格子状に並べたGAが多い)

各分野のGAの埋め込みの分布 (CLIP ViT-L/14を使用)



- 各バッチで同じ分野の論文をサンプリングし, 対照学習で類似度を測るデータに一貫性を持たせる

## Experiments & Results

### 実験方法

- GA, Teaser, Intro. 内のFig. 1をGround Truth (GT) とし, GTが1枚になるように前処理 (データ数: 6,951)
- データセットを学習: 検証: 評価 = 16: 4: 5に分割
- 各提案手法を用い, Long-CLIP-LをFine-tuning
- 評価データのAbstractとGA, 図をエンコード
- Abstractとcos類似度が高い画像をGAとして推論

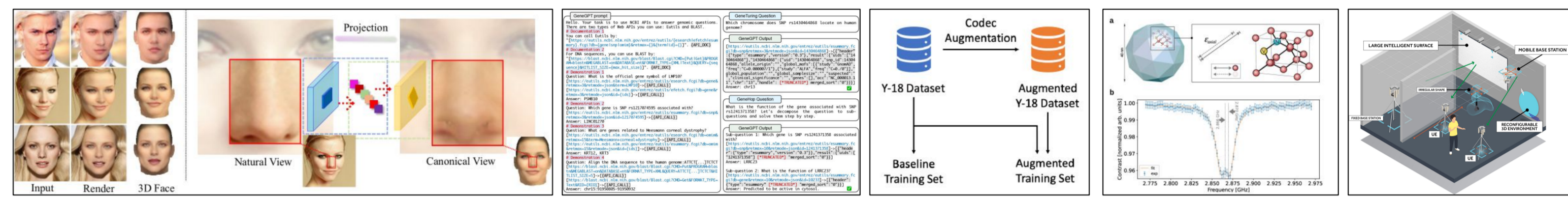
### 実験結果

各提案手法の損失で学習したモデルでの推論結果

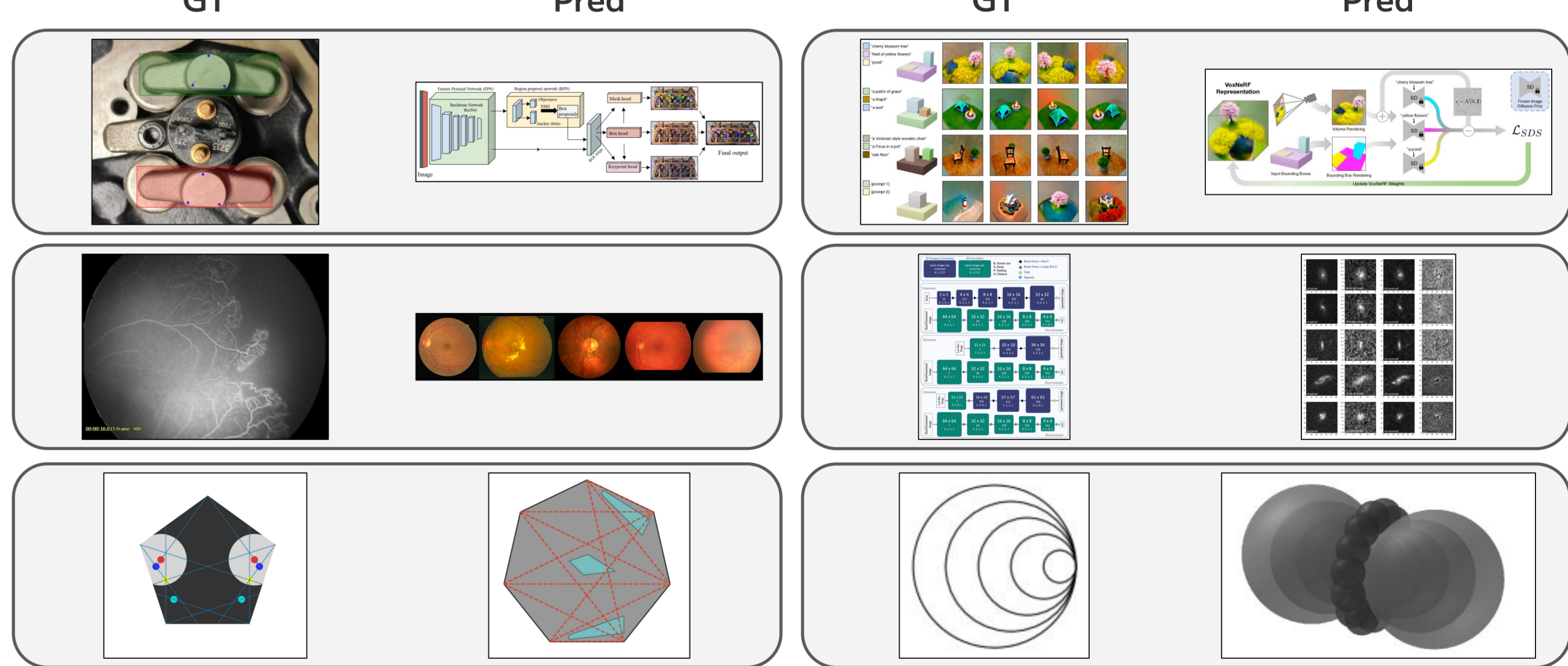
Method	Recall@1	Recall@3	Recall@5
Intra	0.521	0.814	0.920
Intra (DCSamp)	0.499	0.809	0.916
Intra + Inter	0.503	0.811	0.912
Intra + Inter (DCSamp)	0.521	0.807	0.917
Intra + Equalized-Inter	0.526	0.829	0.922
Intra + Equalized-Inter (DCSamp)	0.516	0.826	0.920

- Equalized-Inter Lossの加算により全指標でスコアが向上 DCSampはInter Lossに対して効果的

成功例



失敗例



## Discussion & Future Work

### 各手法の効果

- Inter Loss
  - 同一論文内でのGA検索のボトルネックに...
  - バッチ内のGAの分布が大きく異なる場合, 効果が薄い
- Equalized-Inter Loss
  - 分野を超えたGA共通の特徴の抽出を可能に
- DCSamp
  - 各バッチで分野を横断した多様性が失われる
  - 一方, Inter Lossの安定性と分野ごとの適切な学習を促進

### 定性評価から見える推論の誤りパターン

- GTとは異なるが, GAの構成要素としては妥当な図
- データ数の多い分野 (e.g., cs.CV) の特徴を受けた推論
- 論文内にGAの構成要素として適切な図がない論文
  - GA構成要素の取得という目的において, 十分な検索能力
  - 一方, 論文中の図を用いたGA生成が困難な論文も存在

### Future Work

- Vision Language Model, レイアウト生成技術の活用
- より質の高いデータセットの構築
- 分野ごとの特徴を効果的に活用した手法の開発

