

# コミックの探索的情報アクセスを目的とした 統計モデルに基づく特徴解析とその活用に関する研究

朴 炳宣

関西大学大学院 総合情報学研究科

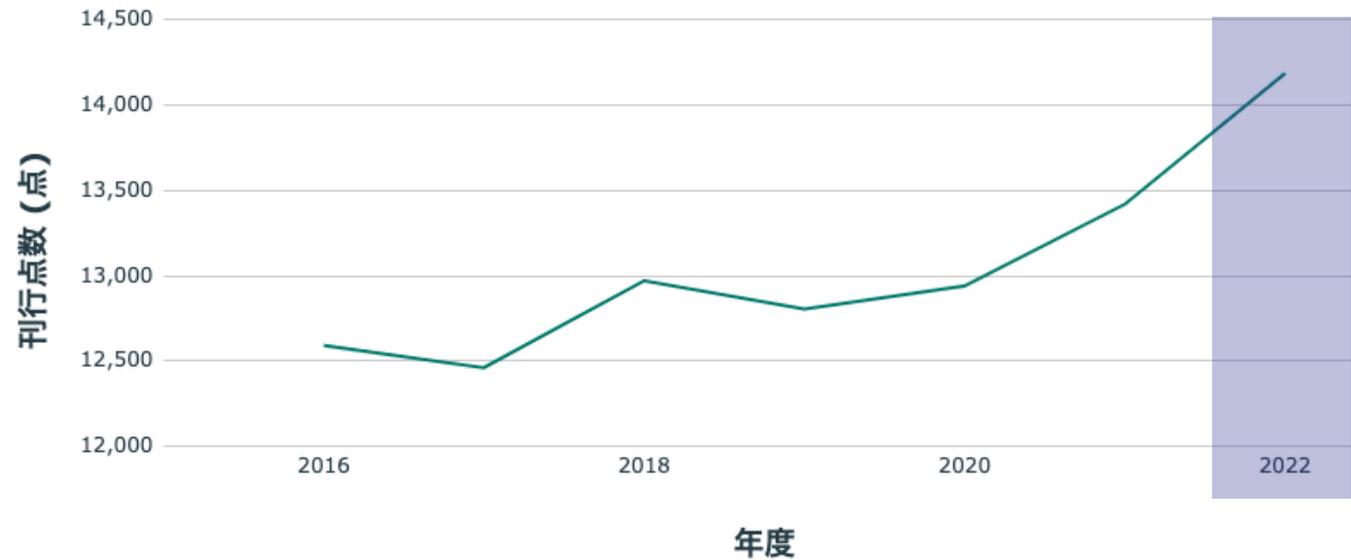


# 目次

- 本研究の背景と目的
- コミックを対象とする情報検索
- 本研究のアプローチ
- 提案手法 1 (本論第 4 章)
  - 固有表現抽出を活用したレビューからの内容情報の抽出
- 提案手法 2 (本論第 5 章)
  - コミックの表紙からの内容推定
- 提案手法 3 (本論第 6 章)
  - 画風に基づく深層クラスタリング
- 議論
- まとめ

# 本研究の背景と目的

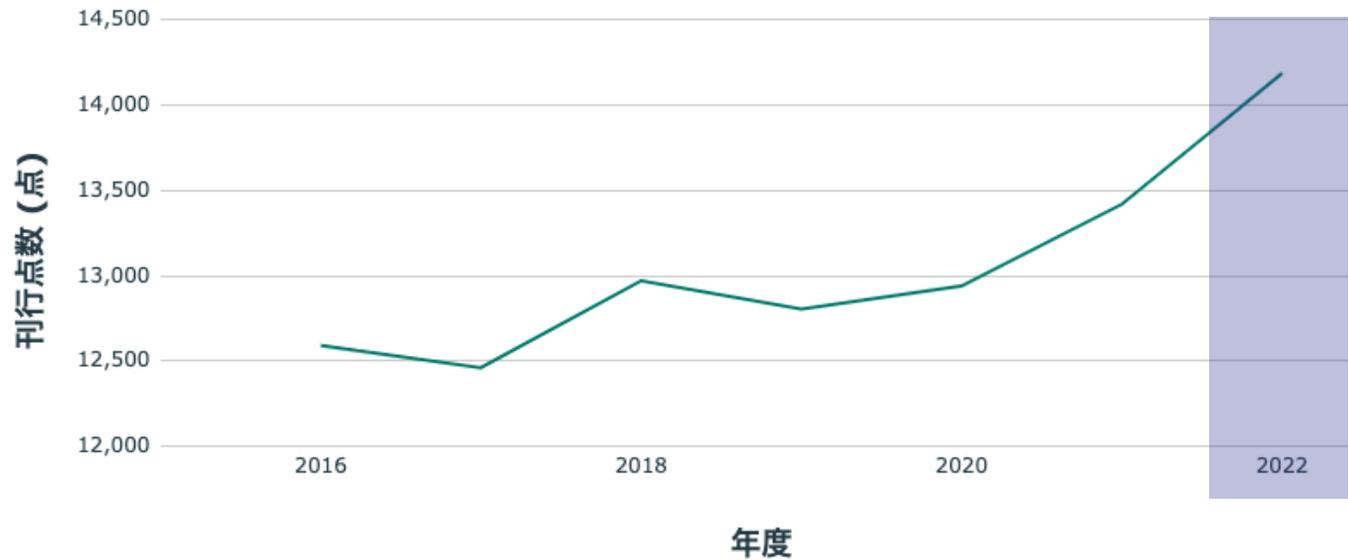
# 本研究の背景



2022年度の新刊点数だけで  
14,000 点以上

コミックの新刊単行本の発行点数の推移 [123]

# 本研究の背景



2022年度の新刊点数だけで  
14,000 点以上

→ 読者が新しい作品を発見することが  
日々困難になっている

コミックの新刊単行本の発行点数の推移 [123]

# 本研究の目的

ユーザの嗜好が反映された円滑な **コミック検索** システムの実現

**コミック検索**

読者が自身の嗜好に適した新しいコミック作品を発見する行為

# コミックを対象とする情報検索

# WEBサービスでのコミック検索

LINEマンガ 毎日無料 ストア インディーズ ログイン 検索

## LINEマンガ インディーズ 作品投稿受付中!

LINEマンガユーザーに、  
あなたの作品を読んでもらおう!

© Suzuko Yuki © haruko © Motoki Takeuchi © Toriko Sakurai

ランキングから探す ◯ 配信曜日から探す ◯ ジャンルから探す ◯ インディーズ作品投稿



コミック検索を行う代表的な空間

物理的制約を持つ書店は異なり  
膨大な数の作品を扱うことができる

→ 1つずつ作品を確認することは  
ユーザの負担になる

## 総合ランキング

もっと見る >



俺だけレベルMAX  
なビギナー  
ファンタジー・SF |  
毎週金曜更新



悪役陛下、執着する  
相手を間違ってます  
ファンタジー・SF |  
毎週日曜更新



入学傭兵  
バトル・アクション |  
毎週木曜更新



略奪された花嫁  
恋愛 | 毎週日曜更新



アカデミーの天才剣士  
ファンタジー・SF |  
毎週火曜更新



若き王の蜜夜  
恋愛 | 毎週土曜更新

# WEBサービスでのコミック検索の特徴: 検索機能

LINEマンガ 毎日無料 ストア インディーズ ログイン 検索

## LINEマンガ インディーズ 作品投稿受付中!

LINEマンガユーザーに、あなたの作品を読んでもらおう!

© Suzuko Yuki © haruko © Motoki Takeuchi © Torako Sakurai

ランキングから探す ◯ 配信曜日から探す ◯ ジャンルから探す ◯ インディーズ作品投稿



キーワードやカテゴリを用いた作品検索機能を提供することでユーザの負担を軽減

## 総合ランキング

もっと見る >



俺だけレベルMAXなビギナー  
ファンタジー・SF | 毎週金曜更新



悪役陛下、執着する相手を間違ってます  
ファンタジー・SF | 毎週日曜更新



入学 傭兵  
バトル・アクション | 毎週木曜更新



略奪された花嫁  
恋愛 | 毎週日曜更新



アカデミーの天才剣士  
ファンタジー・SF | 毎週火曜更新



若き王の蜜夜  
恋愛 | 毎週土曜更新

# WEBサービスでのコミック検索の特徴: 作品推薦機能

## 購入履歴からのおすすめマンガ



アフターゴッド  
(1) (裏少年サン…



終末のワルキューレ  
1巻 (ゼノンコミッ…



LILY-MEN  
(1) (ヤングマガ…



バウンサー 1 (ヤン  
グチャンピオン…

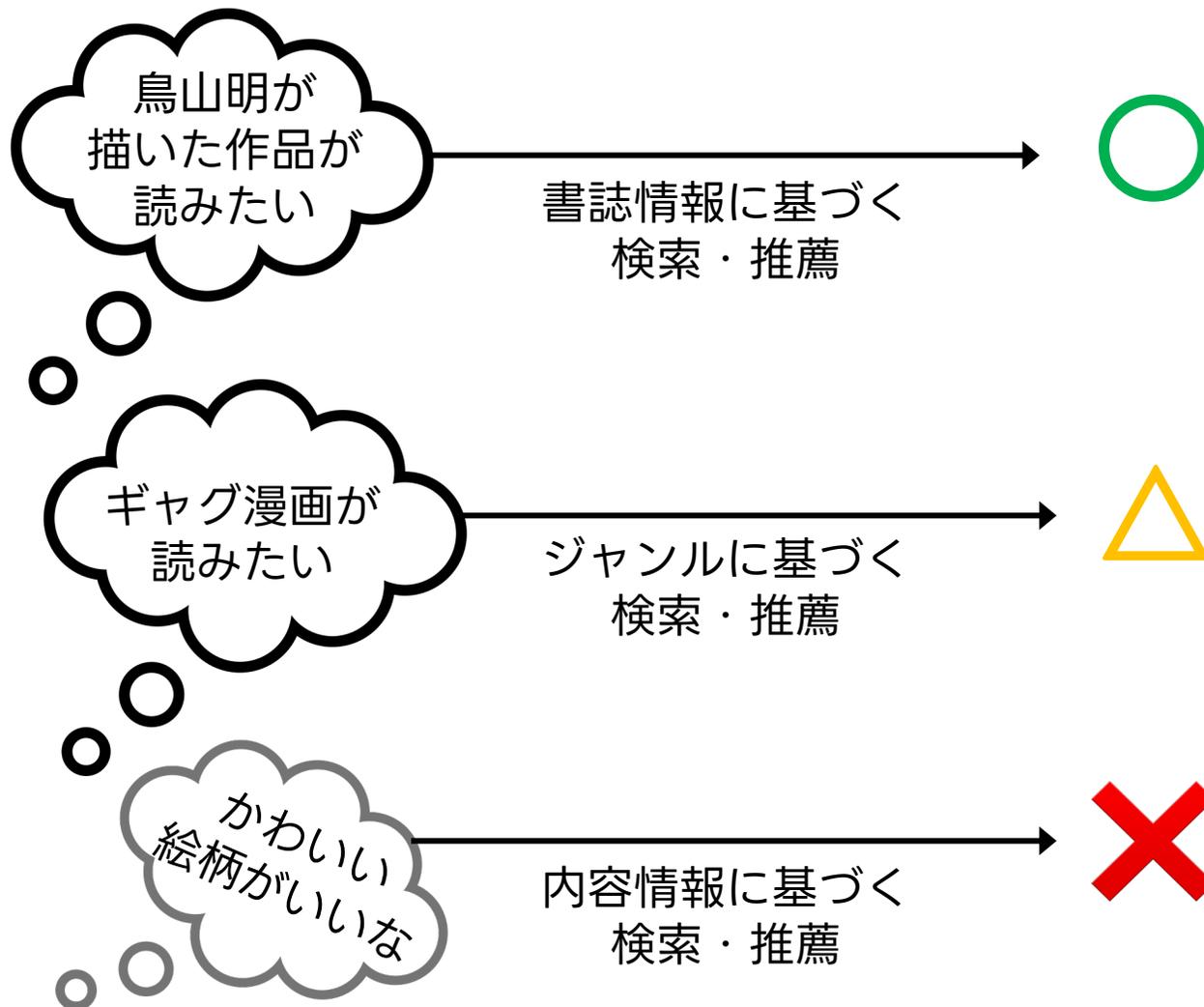


僕の呪いの吸血姫 1  
巻 (デジタル版ガン…

作品の情報やユーザの行動を考慮した  
作品推薦機能を提供することで  
ユーザの負担を軽減

# 本研究のアプローチ: コミック検索のための Document representation

# 現状の WEB サービスによるコミック検索の限界



書誌情報や大まかなジャンル情報に基づく検索・推薦を対応

ストーリーや画風といった  
内容情報に基づく検索・推薦は対応できない

# 現状の WEB サービスによるコミック検索の限界

絵柄に着目した場合  
各作品の表紙を見比べて把握



俺だけレベルMAXなビギナー

お気に入り登録



ファンタジー・SF WAN.Z(redice studio)(脚色)・Maslow(原作)・swingbat(作画) LINEマンガ  
LINE WEBTOON

ゲーム実況ストリーマーとして活動している主人公・鈴木裕人(すずきひろと)。彼は、万人があきらめたクソゲー『試練の塔』を挑戦しクリアした唯一のプレイヤーであった。念願のエンディングを見ることができたものの、もう誰も興味をもたないこのゲームはサービス終了が確定視されていた…。11年もかけクソゲーをクリアした裕人はこれを機に現実と向き合い、ゲーム実況者としての生活を終わらせることを決めた。が次の日、現実世界に異変が起きる。街中に突如現れた不吉な塔、そして人類に課される生死をかけた試練。それはまさに裕人が11年の月日を費やしクリアした『試練の塔』そのものであった。塔の攻略法を把握している裕人はその経験を活かし、誰よりも強い最強のビギナーとしてあらたな試練に挑む！

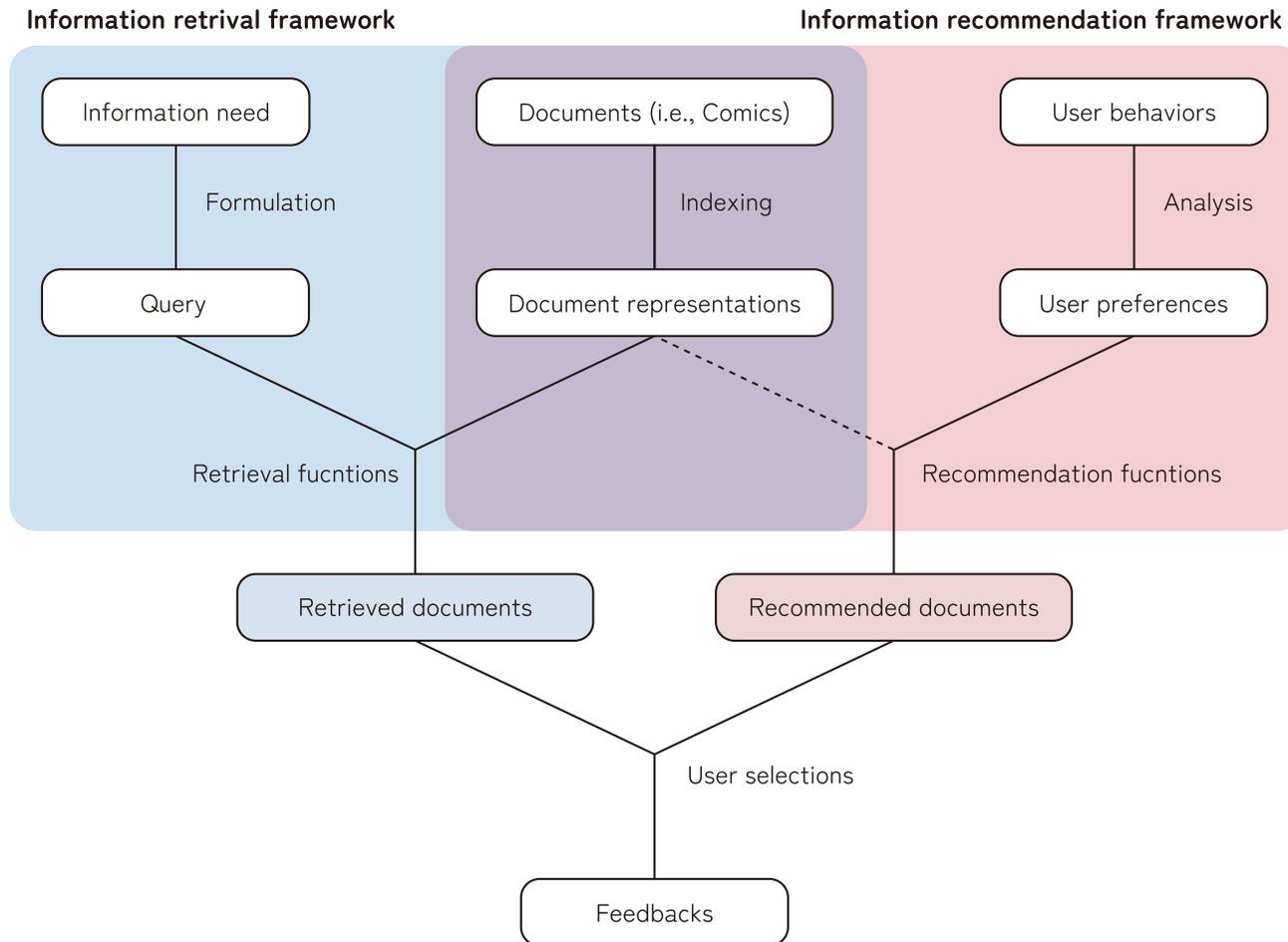
ストーリーに着目した場合  
各作品のあらすじを読んで把握

システムが内容情報に基づく  
検索・推薦を対応できない場合、  
ユーザは各情報について自身で  
情報収集を行わざるを得ない

→ 内容情報に基づくコミック検索は  
ユーザの負担が大きいまま

↑ 検索・推薦機能に用いられる  
内容情報が限定的であるため

# コミック検索における Document representation



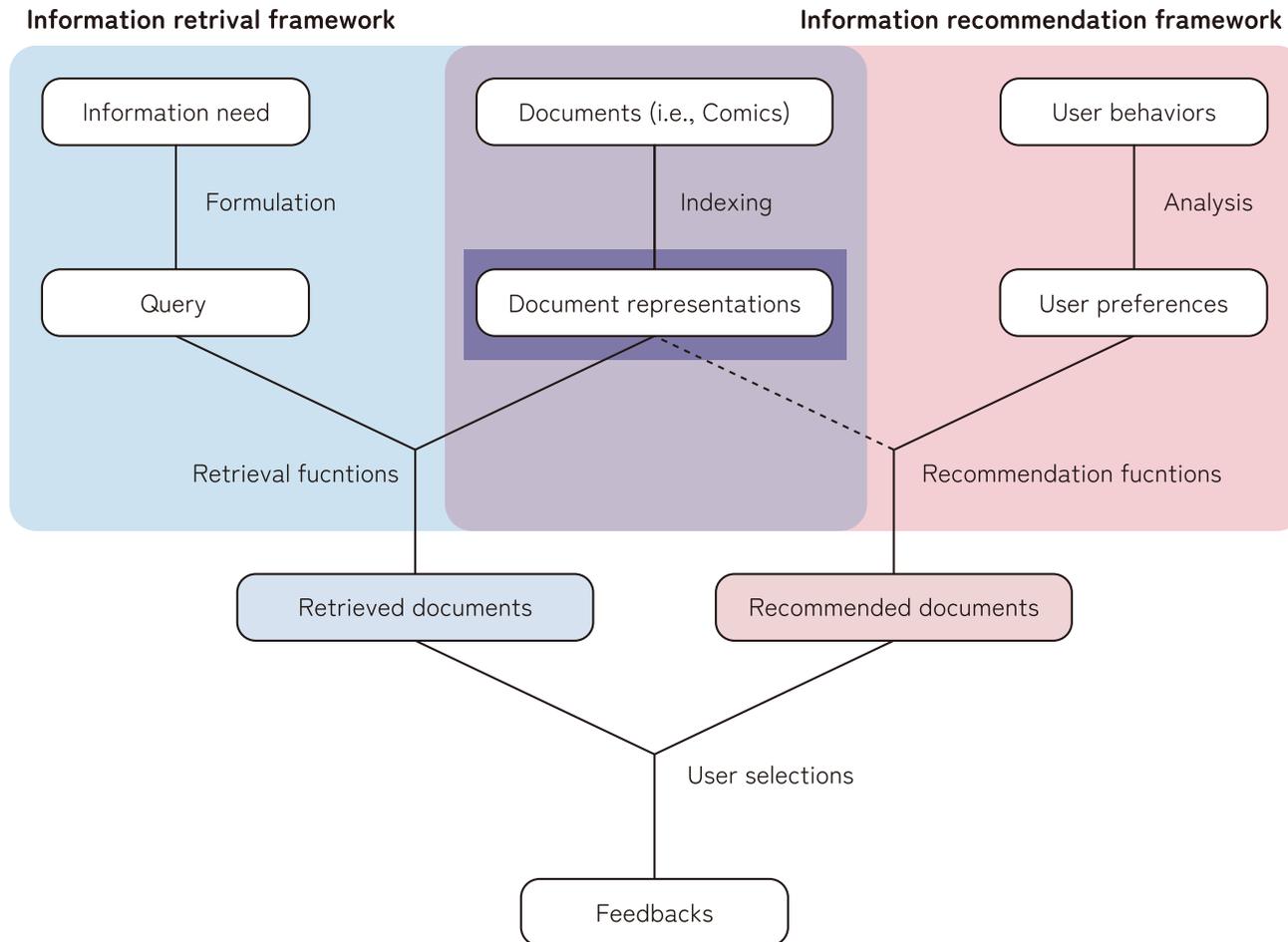
## 作品検索機能

ユーザの要求を反映した Query と  
作品の特徴を表現した  
Document representation を照合

## 作品推薦機能

ユーザの行動に基づいた嗜好情報と  
作品の Document representation に  
よって判断

# コミック検索における Document representation

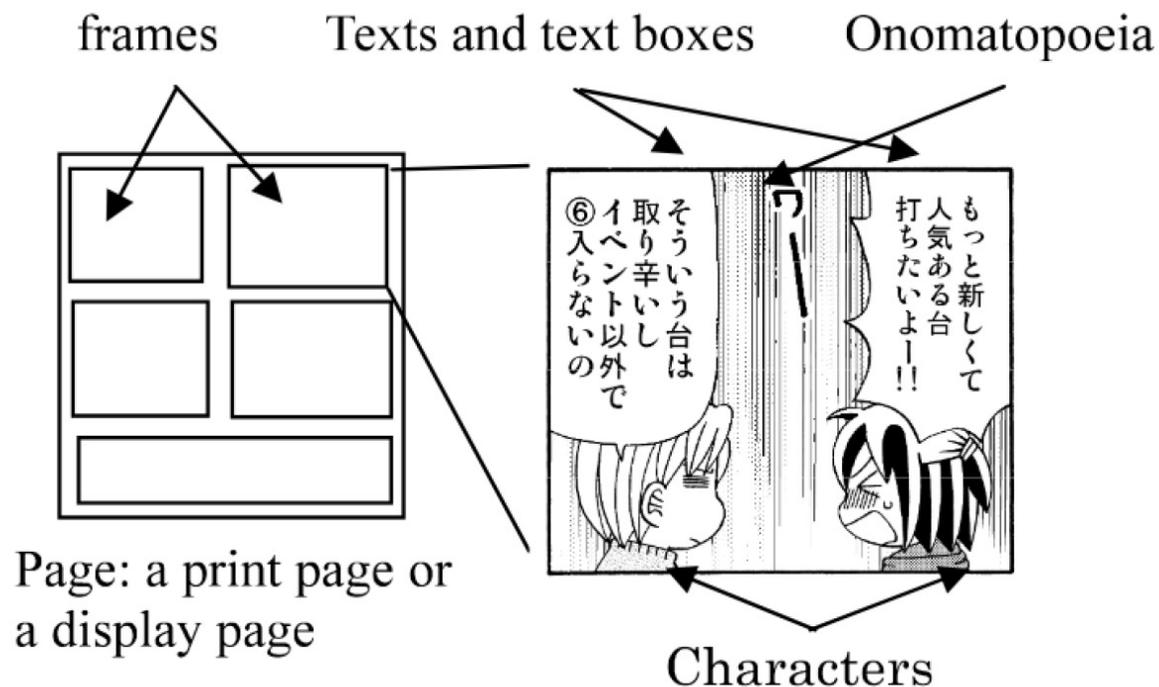


内容情報に基づく検索・推薦のためには  
コミックの内容情報を  
**Document representation** として  
蓄積する必要がある

→ そもそも利用されている内容情報が  
限定的な理由は？

↑ **コミックの内容情報の抽出が  
困難であるため**

# コミックの内容情報抽出の難しさ



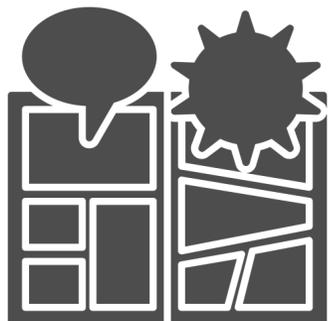
コミックの視覚的要素の一般的な構造 [118]

コミックのストーリーを解析するためには

- 各ページのコマの同定
- 各コマのキャラクター・セリフの同定
- キャラクター・セリフの対応づけ
- 各コマの意味の理解
- 連続したコマによる文脈の理解

など、さまざまなタスクの解決が必要

# コミックの内容情報の多様な情報源



## コミック本文

- 大量の絵・テキストが混在
- 複数のページ・コマが存在
- ストーリーや画風など



## 表紙

- 絵・少量のテキストが混在
- 1枚絵
- コンセプトや画風など



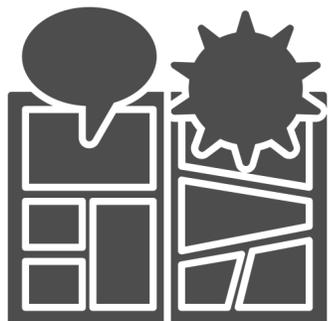
## レビュー

- テキストのみ
- 複数のユーザによって作成
- 作品の概要や感想など

各情報源によって得られる情報が異なり  
そのモダリティも多様

→ 各特性に応じた抽出手法が必要

# コミックの内容情報の多様な情報源



## コミック本文

- 大量の絵・テキストが混在
- 複数のページ・コマが存在
- **ストーリーや画風**など



## 表紙

- 絵・少量のテキストが混在
- 1枚絵
- コンセプトや画風など



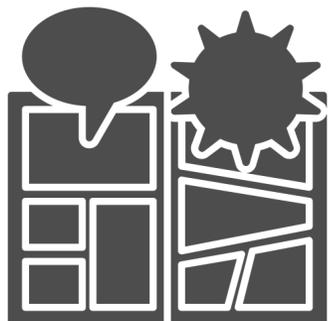
## レビュー

- テキストのみ
- 複数のユーザによって作成
- **作品の概要や感想**など

作品の感想を得られるのはレビューのみ

→ 情報源として有効に活用可能

# コミックの内容情報の多様な情報源



## コミック本文

- 大量の絵・テキストが混在
- 複数のページ・コマが存在
- ストーリーや画風など



## 表紙

- 絵・少量のテキストが混在
- 1枚絵
- コンセプトや画風など



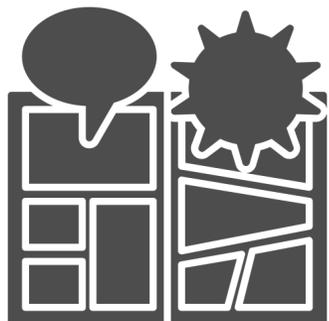
## レビュー

- テキストのみ
- 複数のユーザによって作成
- 作品の概要や感想など

レビューは不特定多数のユーザによって自由に記述される

→ 記述パターンの多様性・情報の品質における課題がある

# コミックの内容情報の多様な情報源



## コミック本文

- 大量の絵・テキストが混在
- 複数のページ・コマが存在
- ストーリーや画風など



## 表紙

- 絵・少量のテキストが混在
- 1枚絵
- コンセプトや画風など



## レビュー

- テキストのみ
- 複数のユーザによって作成
- 作品の概要や感想など

多様な内容情報によって  
Document representation を  
拡張するためには  
**単一の情報源による情報抽出では不十分**

各情報源から  
適切に内容情報を抽出するためには  
**各情報源とそのモダリティの特性による  
固有の課題が存在**

# 本研究の取り組み

## 本研究の目的

ユーザの嗜好が反映された円滑なコミック検索システムの実現

## 課題

- 従来のコミック検索システムで活用されるコミックの内容情報が限定的
- コミックの内容情報は各情報源や情報の特性によって抽出が困難

## 本研究のアプローチ

Document representation の拡張のための内容情報抽出手法

# 本稿で扱う情報源と内容情報に関する課題と提案手法

## 課題 1 (本論第 4 章)

読者によって作品の概略や感想が記述されたレビューを情報源とした内容情報抽出に関する課題

## 課題 2 (本論第 5 章)

作品の概要を一枚絵として表現した表紙の画像を情報源とした内容情報抽出に関する課題

## 課題 3 (本論第 6 章)

コミックの画像を情報源にした際の画風などの抽象的な嗜好に紐づく内容情報抽出に関する課題

# レビューを情報源とした内容情報抽出に関する課題とアプローチ

## 情報源としての有効性

作品の概要や感想が複数のユーザの視点によって記述されている  
→ 多角的な観点による作品への解釈・評価情報を獲得可能

## 課題

レビューに含まれる固有表現はユーザの作品の認知度によって得られる情報量にばらつきがある

## アプローチ

レビューに含まれる固有表現の抽出手法の提案と固有表現がユーザに及ぼす影響の調査

# 表紙の画像を情報源とした内容情報抽出に関する課題とアプローチ

## 情報源としての有効性

作品のコンセプトや画風といった特徴を1枚絵によって表現している  
→ コミック本文の情報を低コストで獲得できる

## 課題

表紙の画像に表現されたコンセプトは複数の要素の組み合わせによって構成されている

## アプローチ

表紙画像内のオブジェクト抽出とその組み合わせによってコンセプトを表現する仕組みの提案

# コミックの画風を対象とした情報抽出に関する課題とアプローチ

## 情報源としての有効性

コミックは絵画によってストーリーや感情を表現するコンテンツ

→ 画風は作品の演出に強く影響するためユーザの嗜好にも大きく関与する

## 課題

画風は複数の視覚的特徴によって複雑に構成され、主に抽象的な認知に基づく

## アプローチ

画風を表現する特徴量としてクラスタリングモデルによる潜在的特徴量を活用する手法

# コミックの画風を対象とした情報抽出に関する課題とアプローチ

## 情報源としての有効性

コミックは絵画によってストーリーや感情を表現するコンテンツ

→ 画風は作品の演出に強く影響するためユーザの嗜好にも大きく関与する

## 課題

**各手法により Document representation の拡張が可能となり  
本研究で指向するコミック検索システムの実現に寄与できる**

画風は複数の視覚的特徴によって複雑に構成され、主に抽象的な認知に基づく

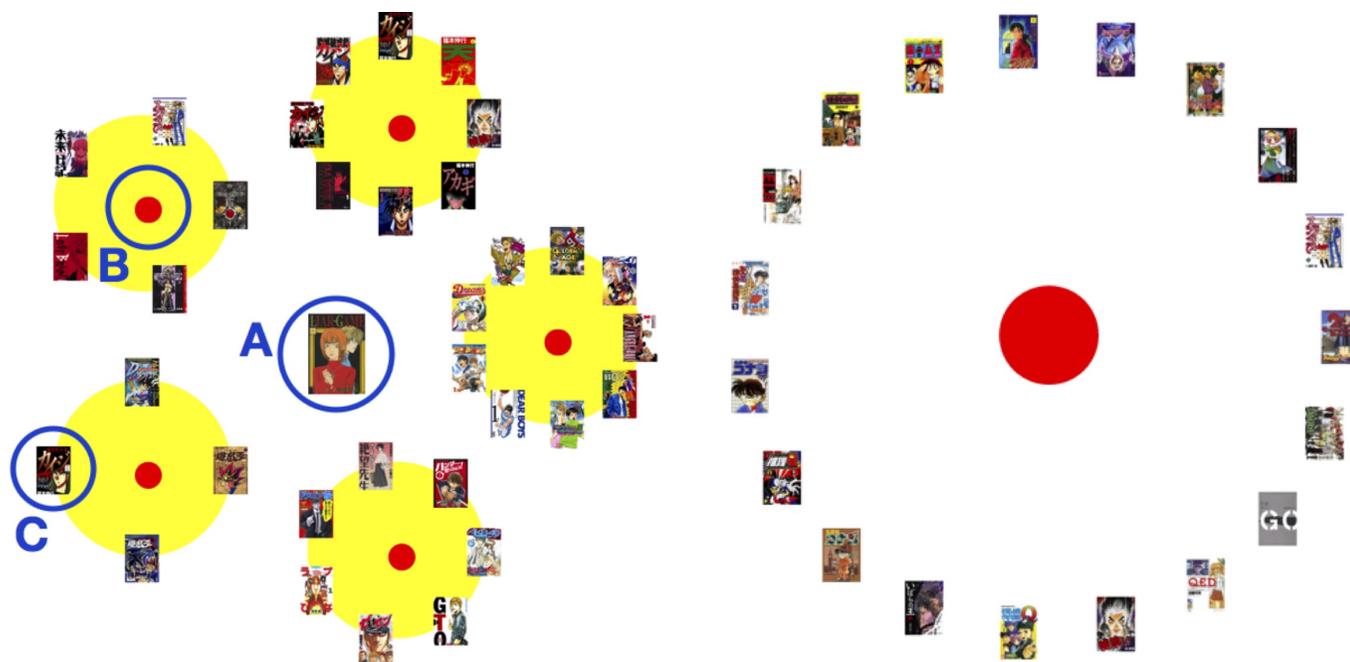
## アプローチ

画風を表現する特徴量としてクラスタリングモデルによる潜在的特徴量を活用する手法

## 提案手法 1:

固有表現抽出を活用したレビューからの内容情報の抽出

# レビューを用いたコミックの内容情報に基づく作品検索



## 山下らのコミック検索システム [135]

コミックのレビューから  
解説・感想などのコミックに関する  
内容情報の抽出とその情報を活用した  
検索インタフェースを提案

# コミックの特徴語群



コミックのレビューから  
コミックの特徴を表す語群を抽出

検索インタフェースで表示することで  
ユーザが各作品の内容を推定する  
判断材料として活用

山下らのシステムによって提示される  
コミックの特徴語群の例 [135]

# コミックの特徴語群に含まれる固有表現



作品名や登場人物といった  
作品内で定義がなされた固有表現が  
多く含まれている

→ 作品に対する認知度によって  
得られる情報量が大きく異なる

山下らのシステムによって提示される  
コミックの特徴語群の例 [135]

# 提案手法の概要

## 課題

(A) レビューに含まれている固有表現は作品の認知度によって得られる情報量が異なる

## アプローチ

### (A) 固有表現抽出モデルの構築

- 文の構造や記述パターンに基づいた規則を用いる手法 [115, 137]
- 統計モデルを用いる手法 [114, 136]
- コミックの固有表現は日々生まれ続けるため未知データへの頑健性が求められる  
→ 条件付き確立場 (Conditional random fields; CRF) [52] を活用

# 提案手法の概要

## 課題

- (A) レビューに含まれている固有表現は作品の認知度によって得られる情報量が異なる
- (B) 固有表現の知名度によって得られる効果が異なる  
(e.g., ドラゴンボールに似ている作品, ドラえもんの道具みたい)

## アプローチ

### (A) 固有表現抽出モデルの構築

- 文の構造や記述パターンに基づいた規則を用いる手法 [115, 137]
- 統計モデルを用いる手法 [114, 136]
- コミックの固有表現は日々生まれ続けるため未知データへの頑健性が求められる  
→ 条件付き確立場 (Conditional random fields; CRF) [52] を活用

### (B) レビューの特徴語群中の固有表現の知名度や割合によるユーザへの影響の調査

# コミックレビューの固有表現抽出: タスク定義

系列ラベリングモデルとして定式化

入力列 ( $N =$  系列の長さ)

$$\mathbf{x} = [x_1, x_2, \dots, x_N]$$

ラベル列

$$\mathbf{y} = [y_1, y_2, \dots, y_N]$$

例

$\mathbf{x} = [$  モンキー・D・ルフィ, は, 集英社, の, 作品, に, 登場, する, キャラクター, ……]

$\mathbf{y} = [$  “人物名”, -, “組織名”, -, -, ……]

# CRF に基づく固有表現抽出モデル [52]

入力列  $\mathbf{x} = [x_1, x_2, \dots, x_N]$    ラベル列  $\mathbf{y} = [y_1, y_2, \dots, y_N]$

## スコア関数

$$\text{score}(\mathbf{x}, \mathbf{y}) = P(\mathbf{y}|\mathbf{x}) = \frac{\exp(s(\mathbf{x}, \mathbf{y}))}{\sum_{\mathbf{y}' \in Y_{\mathbf{x}}} \exp(s(\mathbf{x}, \mathbf{y}'))} \quad (Y_{\mathbf{x}} : \text{すべてのラベル列の集合})$$

$$s(\mathbf{x}, \mathbf{y}) = \sum_{i=0}^N T_{y_i, y_{i+1}} + \sum_{i=0}^N E_{x_i, y_i} \quad (T_{y_i, y_{i+1}} : \text{遷移スコア}, E_{x_i, y_i} : \text{生成スコア})$$

## 損失関数

$$L(\mathbf{y}) = -\log \text{score}(\mathbf{x}, \mathbf{y}) = \log \sum_{\mathbf{y}' \in Y_{\mathbf{x}}} \exp(s(\mathbf{x}, \mathbf{y}')) - s(\mathbf{x}, \mathbf{y})$$

## 推論

$$\hat{\mathbf{y}} = \arg \max_{\mathbf{y}' \in Y_{\mathbf{x}}} \{\text{score}(\mathbf{x}, \mathbf{y}')\}$$

# 固有表現クラス

- Information Retrieval and Extraction Exercise (IREX) 固有表現クラス\*
- コミック特有の固有表現のための拡張クラス
  - IREX 固有表現クラスに含まれないパターンの固有表現への対応

表 4.1: IREX で定義した固有表現の種類と例

略称	種類	例
ART	固有物名	ノーベル文学賞, Windows7
LOC	地名	アメリカ, 千葉県
ORG	組織	自民党, NHK
PSN	人名	安倍晋三, メルケル
DAT	日付	1月29日, 2016/01/29
TIM	時間	午後三時, 10:30
MNY	金額	241円, 8ドル
PNT	割合	10%, 3割

表 4.2: コミックのための拡張固有表現クラス

略称	種類	例
STORY	物語	アラバスタ編, 開戦の章
ITEM	道具	どこでもドア, ゴムゴムの実
SKILL	技	北斗百烈拳, かめはめ波

# データ収集

- Wikipedia から 20 作品を選定し、該当作品の固有表現辞書を構築
- 該当作品のレビュー文を収集し、辞書と固有表現クラスに基づきアノテーション
- アノテーションの結果、5,052 文を確保

# ユーザの内容理解における固有表現が及ぼす影響に関する実験

- 実験目的
  - コミックの特徴語群に含まれる固有表現がコミックの内容推定に及ぼす影響の把握
    - 仮定: 固有表現の知名度によって得られる効果が異なる
- 比較条件の設定
  - 条件 (1) 固有表現をそのまま残したレビュー文から得られる特徴語群
  - 条件 (2) 固有表現を全て削除したレビュー文から得られる特徴語群
  - 条件 (3) 知名度の低い固有表現のみ削除したレビュー文から得られる特徴語群
- 特徴語群の抽出対象
  - 山下ら<sup>[134]</sup>の実験に用いられた 1,000 作品のレビュー

# ユーザの内容理解における固有表現が及ぼす影響に関する実験

## - 実験方法

- 各作品の特徴語群と全体のレビューに含まれるトピックとの関連有無の判定
  - トピック: トピックモデルによって生成された特徴語群間の共通の話題を表す語群
- 正解データはシステムによる関連有無の判定
  - 1 トピックあたり 4 つの候補作品を用意しその内 1~3 作品にシステム正解作品を配置
  - システムと実験参加者の回答の一致度を比較

提示トピック	
試合, 野球, スポーツ, バスケット, サッカー, チーム, 選手, 高校, 努力, 描写, 監督, スラムダンク, 成長, 中, 熱い, 練習, 柔道, リアル, 戦, 翼, ライバル, シュート, キャプテン, アメフト, 大会, 相手, プレー, 花道, ボクシング, プロ	
	設問 1
野球, 御幸, 試合, 強豪校, 青道, 甲子園, 谷, チーム, 選手, 高校, 投手, エース, 留学, スポーツ, 栄, 純, 名門校, ムービング, 稲実, 野球部, ダイヤ, 監督, 名門, 引退, 英才, 弱小, エース像, 練習, チームメイト, レギュラー, 捕手, ピッチャー, 先輩, 予選, キャッチャー, マガジン, 投手陣, 敷, A, 争い, 薬師, エリート, 弱小校, 轟, 選手達, フォーム, 魔球, スポコン, 市大, クセ球	

特徴語群からみて  
トピックと関係がある  
作品なのか？

↑ トピック及び  
特徴語群の内容の推測が必要

## ユーザの内容理解における固有表現が及ぼす影響に関する実験

- 実験データ: 36 設問 (3 条件 x 3 トピック x 4 作品)
- 実験協力者数: 10 名

## 実験結果

- t 検定により有意差を確認: 条件 (1) < 条件 (2) & 条件 (3) ( $p < 0.05$ )  
→ 固有表現の割合の減少により、特徴語群の内容の把握が容易になることが示唆された

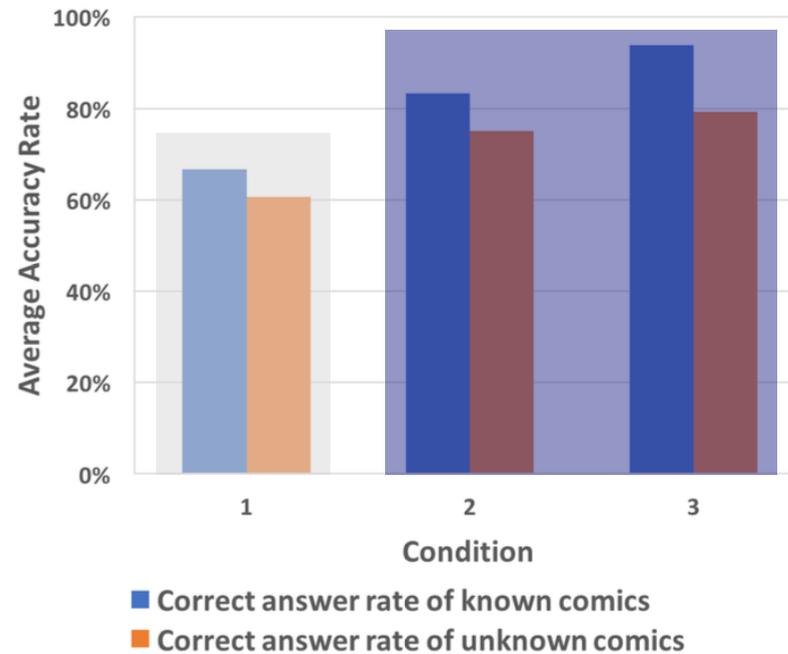


図 4.4: 特徴語群における固有表現の条件ごとの正解率

## 実験結果

- 知らない作品に対する正解率が最も高かったのは条件 (3), i.e., 79.5%  
→ 知名度の高い固有表現はユーザの内容推測に妨げにならないことが示唆された

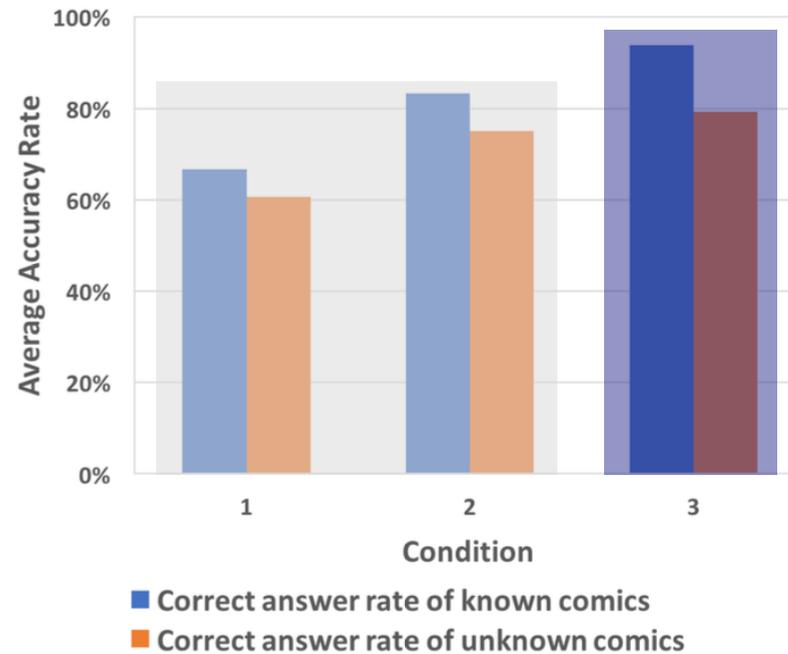


図 4.4: 特徴語群における固有表現の条件ごとの正解率

# まとめ

## 課題

- (A) レビューに含まれている固有表現は作品の認知度によって得られる情報量が異なる
- (B) 固有表現の知名度によって得られる効果が異なる

## アプローチ

- (A) CRF に基づく固有表現抽出モデルの構築
- (B) レビューの特徴語群中の固有表現の知名度や割合によるユーザへの影響の調査  
→ 固有表現の割合の減少により特徴語群の内容の把握が容易になる

## Document representation に関する貢献

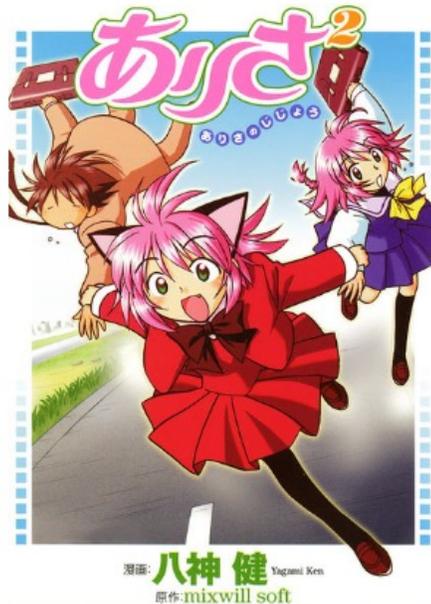
レビューから抽出された情報を Document representation に活用する際の  
固有表現に対する適切な前処理とその重要性

## 提案手法 2: コミックの表紙からの内容推定

# コミックの表紙



(a) 征神記ヴァルナス<sup>1</sup>



(b) ありさ<sup>2 2</sup>



(c) エヴリデイおさかなちゃん<sup>3</sup>

コミックの表紙の例

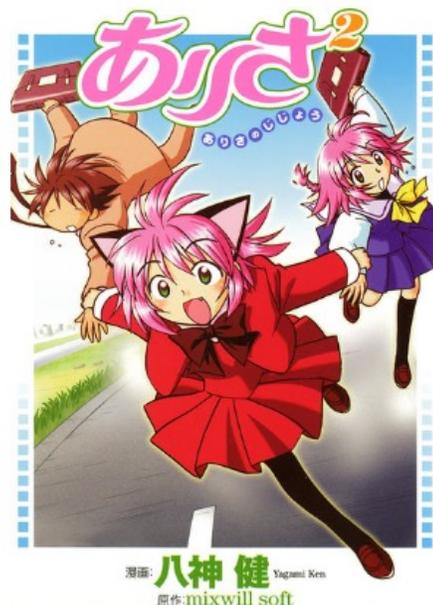
作品を広告する上で重要な役割を持つ

作品の特徴や魅力を伝えるために  
デザインされている

# コミックの表紙に含まれる作品コンセプト



(a) 征神記ヴァルナス<sup>1</sup>



(b) ありさ<sup>2 2</sup>



(c) エブリデイおさかなちゃん<sup>3</sup>

コミックの表紙の例

表紙に描かれた衣服や背景によって  
作品の題材や時代的・文化的背景に  
もとづく作品のコンセプトを把握できる

e.g., (b) と (c) は時代的背景が類似  
(a) はどちらとも異なる

# 提案手法の概要

## 課題

(A) コミクの絵画は作者の画風によって描写が一定でない特徴を持つため解析が困難



## アプローチ

(A) 事前学習された深層ニューラルネットワークモデルを活用した多ラベル分類モデルの構築

# 提案手法の概要

## 課題

- (A) コミックの絵画は作者の画風や演出によって描写が一定でない特徴を持つため解析が困難
- (B) 作品のコンセプトはオブジェクトの有無ではなく組み合わせによる文脈によって表現される
  - e.g., 「剣」と「ローブ」 vs. 「剣」と「着物」 → 異なるコンセプトが形成される

## アプローチ

- (A) 事前学習された深層ニューラルネットワークモデルを活用した多ラベル分類モデルの構築
- (B) トピックモデルによってオブジェクトを共起関係を考慮した関連付けの仕組みを提案

# 表紙のオブジェクト情報抽出モデル: タスク定義

多ラベル分類 (Multi-label classification) 問題として定式化

入力画像  $x$

正解ラベルセット  $y$

$y_1, y_2, y_3, \dots, y_N, y_i \in \{0, 1\}$  ( $N =$  ラベルの数)

各ラベルは出現有無のバイナリ状態によって表現される

分類モデル

$$\hat{y} = f(x)$$

例



$x =$

剣, ローブ, ドレス, T シャツ

,  $y = [1, 1, 1, 0, \dots], f(x) = [0.9, 0.8, 0.7, 0.05, \dots]$

# 表紙のオブジェクト情報抽出モデル: 事前学習モデルの活用

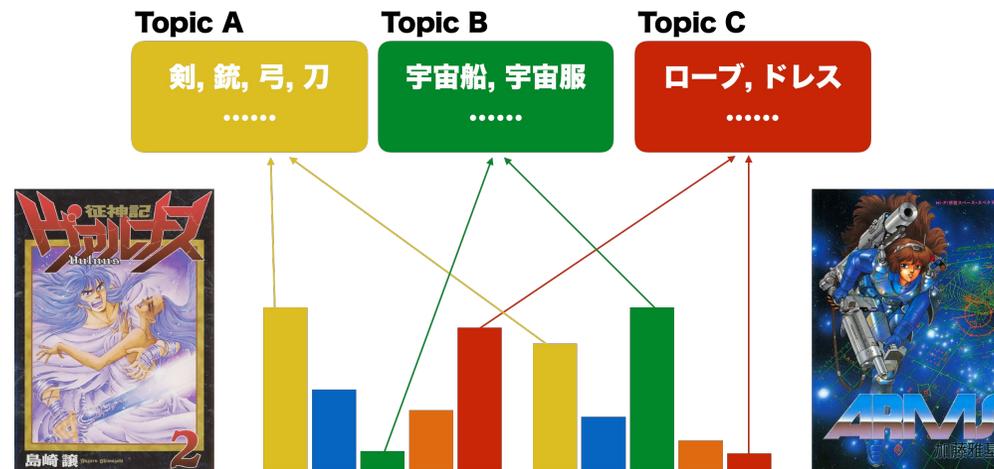
- ImageNet データセット [25] で学習された VGG-16 [87] を事前学習モデルとして活用
  - 本研究で必要な情報がアノテーションされたコミック表紙画像の数は限定的
  - 大量の画像データによって学習されたパラメータにより学習を安定化
- 学習時の詳細
  - VGG-16 の最終出力層を提案手法にタスクに必要なラベルセットに合わせて変更
  - 学習時は最終層から近い層以外はパラメータを凍結

# データ収集

- Amazon に販売されているコミックからランダムで選定された 100 作品
  - 各作品の最新刊までに画像を収集し、合計 536 枚の画像を収集
  - 429 枚と 107 枚をそれぞれ学習データ及び評価データで活用
- アノテーションによって 560 種類のタグを確保
  - 希少ラベルを削減するために 5 作品以上が共通的に出現するラベル 100 種類のみを利用

# トピックモデルを用いたコミックの関連付け

- 表紙内の作品のコンセプトは特定のラベルの有無ではなくその共起関係によって表現される
  - 「剣」と「ローブ」 vs. 「剣」と「着物」 → 異なるコンセプトが形成される
  - 「剣」「弓」「銃」は表記は異なるが共通の意味を持つ
- 文書における単語の共起関係に基づいて文書間の共通のトピックを推定するトピックモデルによって表紙内のオブジェクトの共起関係に基づいたコンセプトを表現
  - トピックモデルとして LDA [14] を活用
- 各作品ごとのトピックの分布についてバタチャリア係数 [15] によって類似度を算出

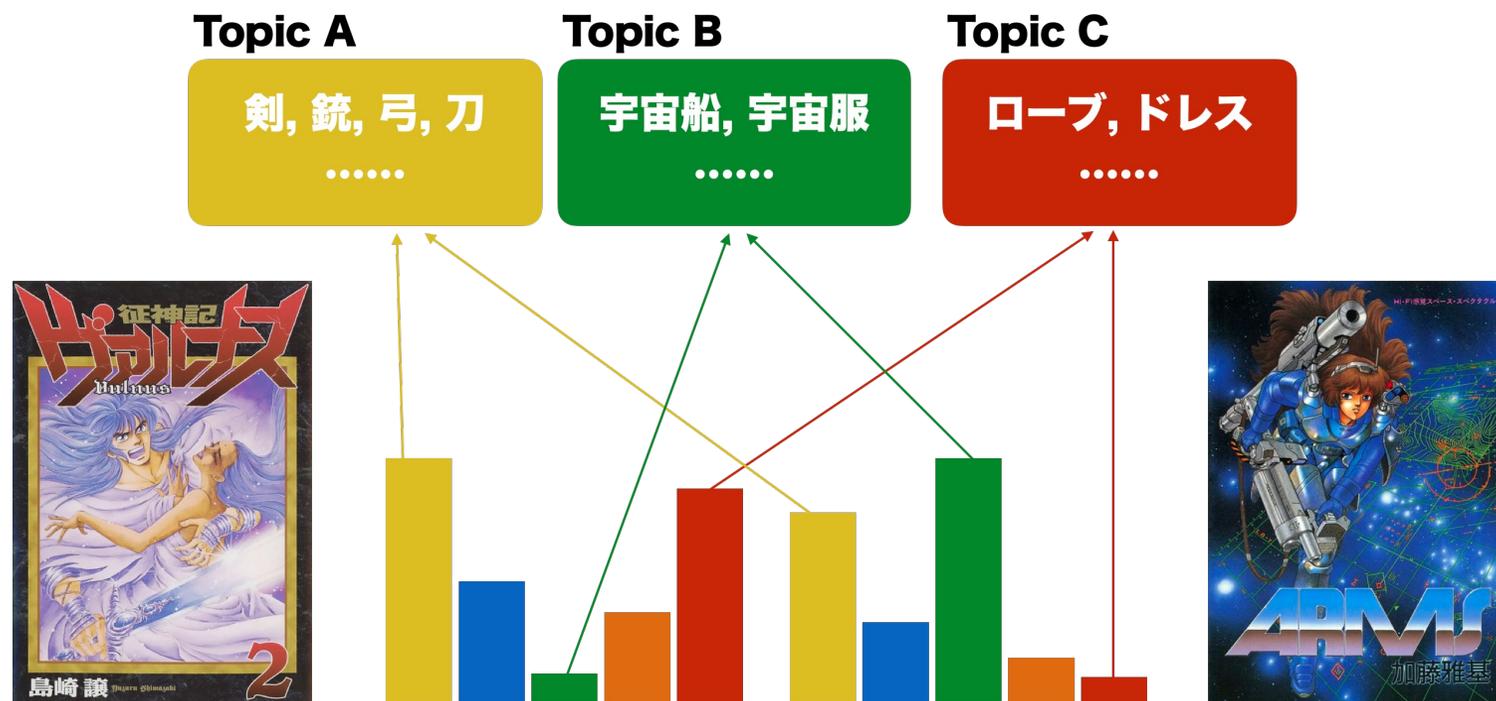


[14] Blei, D. M., Ng, A. Y. and Jordan, M. I.: Latent Dirichlet allocation, Journal of Machine Learning Research, Vol. 3, pp. 993–1022 (2003).

[15] Cha, S.: Comprehensive survey on distance/similarity measures between probability density functions, International Journal of Mathematical Models and Methods in Applied Sciences, Vol. 1, No. 4, pp. 300–307 (2007).

# コンセプト情報に基づく類似度算出に関する実験

- 実験目的
  - システムによるコミック間の関連性をユーザが測ることができるのかを明らかにする



# コンセプト情報に基づく類似度算出に関する実験

## - 実験方法

- クエリとなる作品の表紙画像 1 枚と、
- クエリ作品との関連の有無を判断する候補作品の 表紙画像 10 枚について、コンセプトにおける関連の有無を ○ か × で回答
  - システムによって関連があると判断した 3 作品 (類似度上位 3 作品, i.e., ○)
  - システムによって関連がないと判断した 7 作品 (類似度下位 7 作品)
  - 実験協力者には ○ は必ず 3 つつけるように指示 (i.e., ○)
  - システムによる正解作品と実験協力者の正解作品との一致度によって評価



クエリ作品



候補作品

# 実験結果

- Top-3 Accuracy は 80%, 平均正解作品数は  $1.24 \pm 0.16$  (95%信頼区間)
    - システム正解作品とユーザ正解作品は高い確率で1つ以上の作品が一致している
- 提案手法によるコンセプトに基づく作品間の類似度は実験協力者の認知と近似

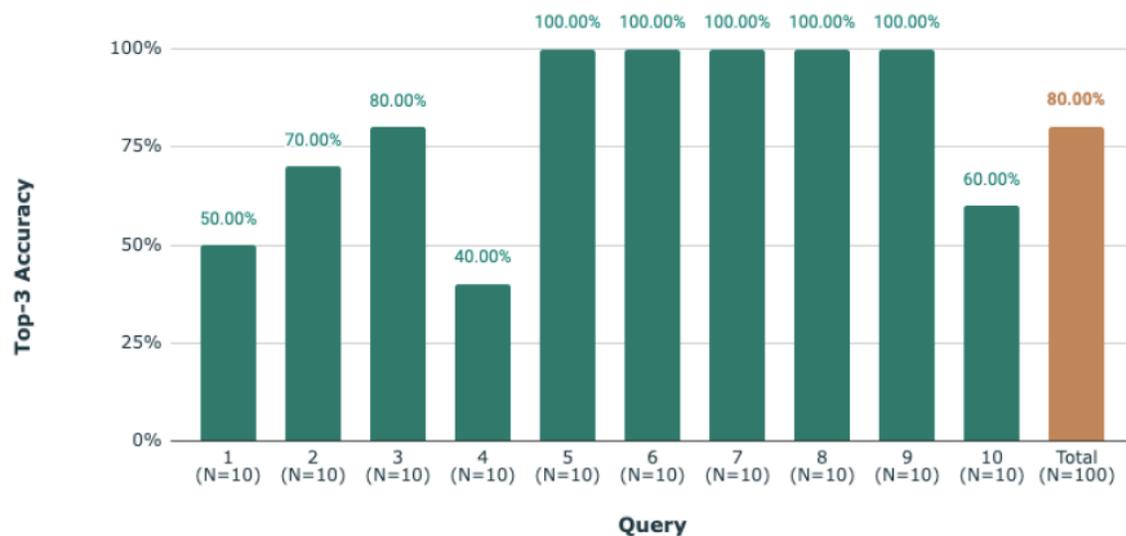


図 5.2: 各クエリごとの Top-3 Accuracy

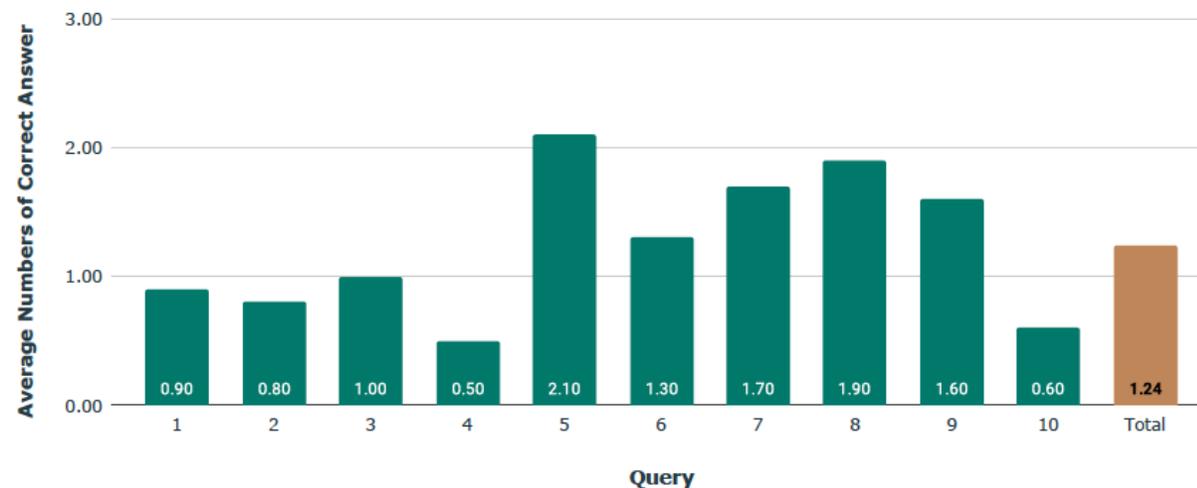


図 5.3: 各クエリごとの正解作品数

# まとめ

## 課題

- (A) コミックの絵画は作者の画風によって描写が一定でない特徴を持つため解析が困難
- (B) 作品のコンセプトはオブジェクトの有無ではなく組み合わせによる文脈によって表現される

## アプローチ

- (A) 事前学習された VGG16 モデルを活用した多ラベル分類モデルの構築
- (B) トピックモデルによってオブジェクトを共起関係を考慮した関連付けの仕組みを提案  
→ 提案手法によるコンセプト情報を用いて類似検索が可能であることが示唆

## Document representation に関する貢献

本研究で指向するコミック検索システムのコンセプトに基づく類似検索のための Document representation として有効な活用が可能

## 提案手法 3: 画風に基づく深層クラスタリング

# コミックの画風



© Shinpei Ito

© Ebifurai

© Kaasan

図 6.1: 画風による描写の違いの一例

絵画に表れた作者の特色 [54]  
(e.g., 線の太さ, 彩色方法,  
目や顔の輪郭などの描写方法)

ユーザの趣向に大きく関わる要素

視覚的特徴の組み合わせによって形成  
抽象的な認知に基づいて判断される

# 提案手法の概要

## 課題

画風情報は複雑な組み合わせによって形成され、抽象的な認知に基づいて判断される

## アプローチ

深層クラスタリングモデルによる潜在的特徴量を活用して画風を表現

# 提案モデル: 生成モデルと距離学習を用いた深層クラスタリングモデル

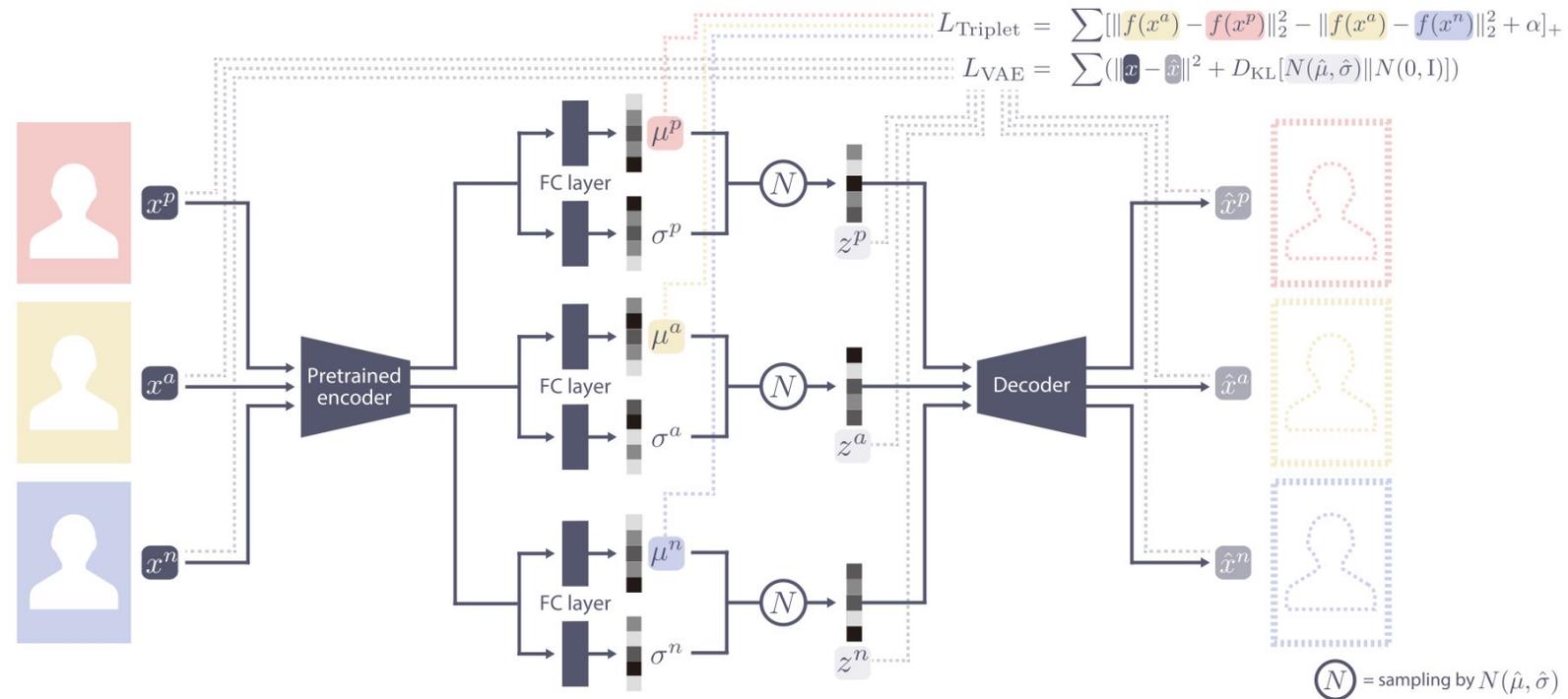
## - 先行研究

- 既定のラベルセットによって学習された分類モデルを用いる手法 [44, 128]
  - 未知データへの頑健性における懸念がある
- 生成モデルに基づくクラスタリング [28, 57]
  - 自己学習により特定の観点に着目した最適化が難しい
- 距離学習に基づくクラスタリング [81, 101]
  - モデルの精度が学習データのサンプリング方式に依存する

## - 本稿でのアプローチ

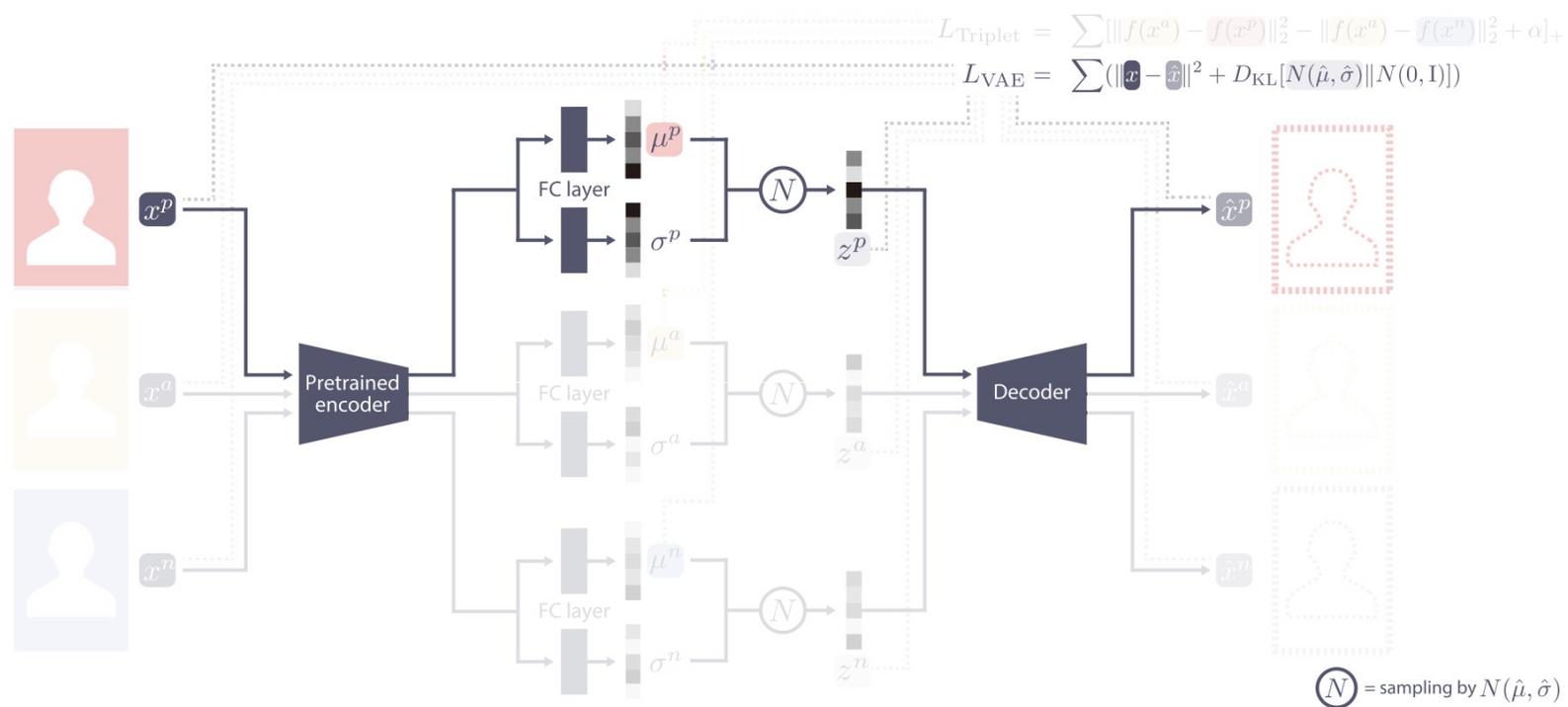
- 生成モデルである VAE モデル [46] を距離学習手法である Triplet loss [84] によって共同最適化することで、ラベルセットを用いず未知データに頑健なモデルを構築
- 事前学習モデルを用いてモデル学習を安定化する

# 提案モデル: 生成モデルと距離学習を用いた深層クラスタリングモデル



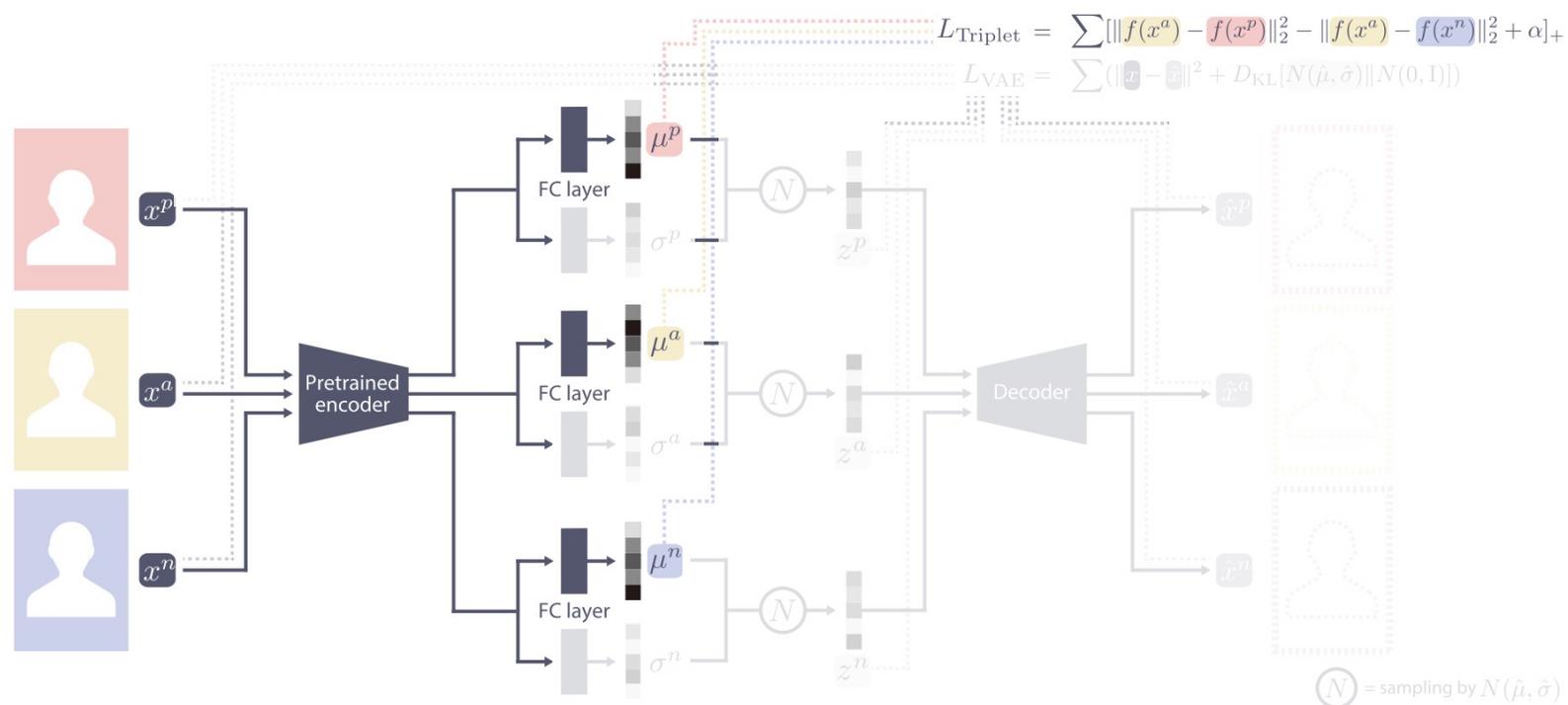
生成モデルである VAE モデル [46] と距離学習手法である Triplet loss [84] によって共同最適化

# 提案モデル: VAE に基づく生成過程



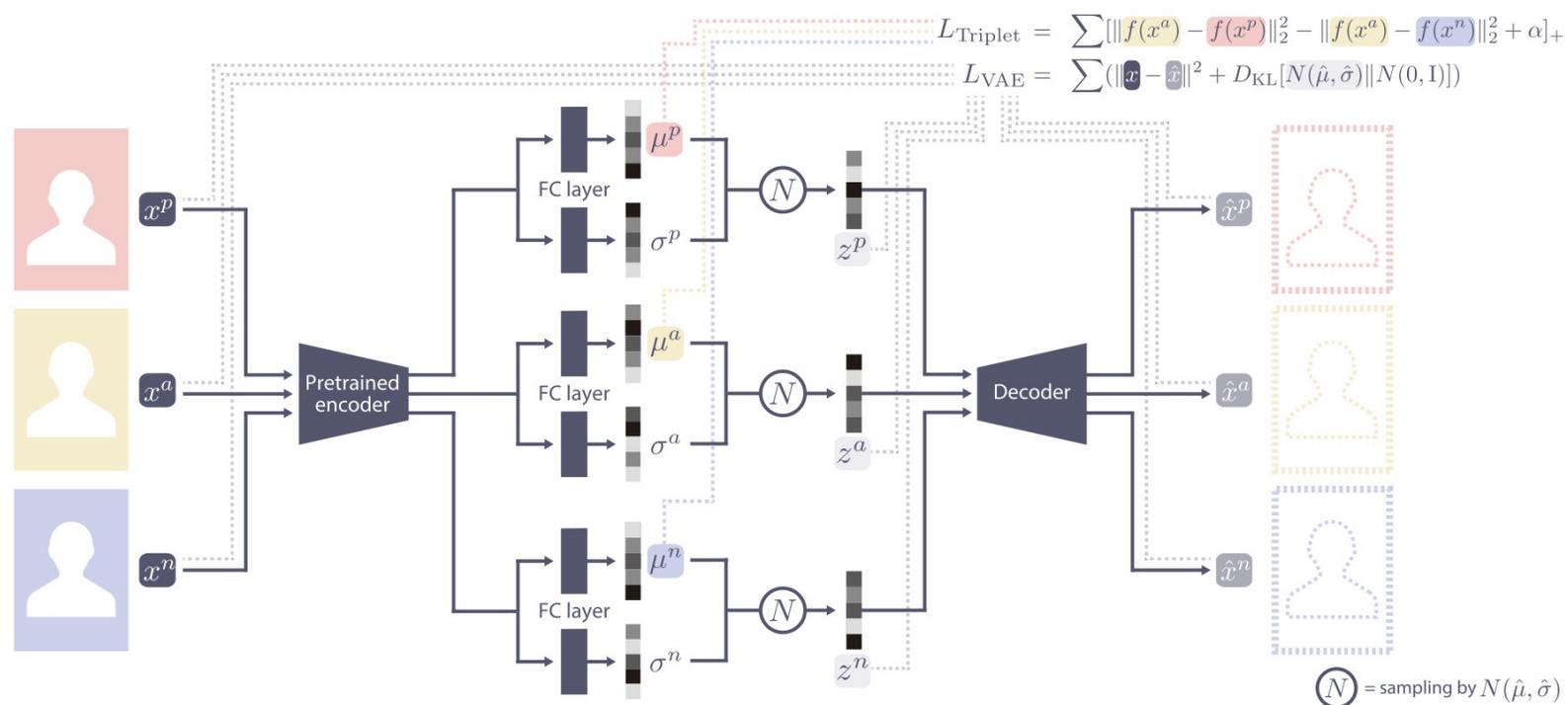
入力画像について Encoder によって得られる潜在変数と  
潜在変数を用いて Decoder によって再構成された出力画像について最適化

# 提案モデル: Triplet loss による距離学習



エンコーダーによって得られる潜在特徴量について  
同じ画風を持つ画像であれば距離を近づけ  
異なる画風を持つ画風であれば距離を遠ざけるように学習

# 提案モデル: 生成モデルと距離学習を用いた深層クラスタリングモデル



生成モデルである VAE モデル [46] と距離学習手法である Triplet loss [84] によって共同最適化  
→ 生成モデルの良質な潜在特徴量を画風に関する明確な指示によって最適化

# 画風に基づく深層クラスタリングモデルに関する評価

## - データセット

- 事前学習: ImageNet [80], Danbooru\*
- 学習データ: Danbooru
  - 事前学習に含まれない作者 1,000 名を選定し作者ごとに 300 枚まで収集
- 評価データ: Danbooru, Manga109 [1, 64]

用途	データセット	クラス数	画像数
事前学習	ImageNet	-	1,281,167
	Danbooru	-	1,000,000
学習	Danbooru	1,000	157,330
評価 (Seen)	Danbooru	1,000	50,000
評価 (Unseen)	Danbooru	100	5,000
評価 (Face)	Manga109	92	157,152
評価 (Body)	Manga109	92	118,715
評価 (Frame)	Manga109	92	103,900
評価 (Cover)	Manga109	91	107

[\*] <https://gwern.net/danbooru2021>

[1] Aizawa, K., Fujimoto, A., Otsubo, A., Ogawa, T., Matsui, Y., Tsubota, K. and Ikuta, H.: Building a Manga dataset “Manga109” with annotations for multimedia applications, IEEE MultiMedia, Vol. 27, No. 2, pp. 8–18 (2020)

[64] Matsui, Y., et al. : Sketch-based manga retrieval using manga109 dataset, Multimedia Tools and Applications, Vol. 76, No. 20, pp. 21811–21838 (2017)

[80] Russakovsky, O., et al. : ImageNet large scale visual recognition challenge, International Journal of Computer Vision, Vol. 115, No. 3, pp. 211– 252 (2015)

# 画風に基づく深層クラスタリングモデルに関する評価

## - 比較モデル

- 全 18 モデルを比較
- 損失関数の設定によるグループ

(1)  $L_{Arc}$  を用いる - (a), (b), (c) ラベルセットを用いた手法のベースライン

(2)  $L_{Triplet}$  を用いる - (d), (e), (f) 距離学習のみを用いた手法のベースライン

(3)  $L_{VAE}$  を用いる - (g), (h), (i)

(4)  $L_{GMVAE}$  を用いる - (j), (k), (l) 生成モデルのみを用いた手法のベースライン

(5)  $L_{Arc}$  と  $L_{VAE}$  を用いる - (m), (n), (o)

(6)  $L_{Triplet}$  と  $L_{VAE}$  を用いる - (p), (q), (r) 提案手法

## - 事前学習の設定によるグループ

(1) 事前学習を行わない - (a), (d), (g), (j), (m), (p)

(2) ImageNet を用いる - (b), (e), (h), (k), (n), (q)

(3) Danbooru を用いる - (c), (f), (i), (l), (o), (r)

# 実験結果: Danbooru データセットを用いた定量的評価

Model	Pretraining domain	Loss	Seen			Unseen		
			Accuracy	ARI	NMI	Accuracy	ARI	NMI
(a)	-	$L_{Arc}$	6.42%	1.08%	47.02%	14.82%	5.13%	35.60%
(b) [128]	ImageNet	$L_{Arc}$	8.45%	1.93%	48.93%	19.70%	8.17%	41.10%
(c)	Danbooru	$L_{Arc}$	9.80%	2.60%	50.30%	20.46%	8.93%	43.52%
(d)	-	$L_{Triplet}$	4.73%	0.45%	45.28%	10.42%	2.57%	29.98%
(e)	ImageNet	$L_{Triplet}$	11.69%	3.87%	51.41%	26.68%	13.90%	46.45%
(f)	Danbooru	$L_{Triplet}$	14.61%	5.57%	53.51%	30.22%	16.65%	50.02%
(g)	-	$L_{VAE}$	4.21%	0.46%	39.06%	7.06%	1.01%	19.93%
(h)	ImageNet	$L_{VAE}$	4.26%	0.47%	40.98%	7.10%	1.06%	21.22%
(i)	Danbooru	$L_{VAE}$	4.27%	0.47%	41.28%	7.10%	0.92%	21.65%
(j) [28]	-	$L_{GMVAE}$	4.42%	0.50%	41.85%	7.44%	1.01%	21.54%
(k)	ImageNet	$L_{GMVAE}$	4.37%	0.46%	42.76%	7.08%	1.01%	23.11%
(l)	Danbooru	$L_{GMVAE}$	4.48%	0.54%	42.86%	7.88%	1.20%	23.64%
(m)	-	$L_{Arc} + L_{VAE}$	6.69%	1.24%	47.07%	14.60%	5.11%	35.08%
(n)	ImageNet	$L_{Arc} + L_{VAE}$	8.45%	1.95%	48.68%	20.04%	8.83%	40.86%
(o)	Danbooru	$L_{Arc} + L_{VAE}$	10.53%	3.02%	50.60%	24.66%	12.28%	45.80%
(p)	-	$L_{Triplet} + L_{VAE}$	4.00%	0.17%	44.02%	7.62%	1.03%	25.74%
(q)	ImageNet	$L_{Triplet} + L_{VAE}$	13.06%	4.55%	52.12%	26.64%	14.52%	47.06%
(r) (i.e., ours)	Danbooru	$L_{Triplet} + L_{VAE}$	15.44%	6.06%	53.94%	32.70%	18.71%	51.71%

提案手法は  
ベースライン手法より  
高いクラスタリング性能を  
持つことを確認

# 実験結果: Danbooru データセットを用いた定量的評価

Model	Pretraining domain	Loss	Seen			Unseen		
			Accuracy	ARI	NMI	Accuracy	ARI	NMI
(a)	-	$L_{Arc}$	6.42%	1.08%	47.02%	14.82%	5.13%	35.60%
(b) [128]	ImageNet	$L_{Arc}$	8.45%	1.93%	48.93%	19.70%	8.17%	41.10%
(c)	Danbooru	$L_{Arc}$	9.80%	2.60%	50.30%	20.46%	8.93%	43.52%
(d)	-	$L_{Triplet}$	4.73%	0.45%	45.28%	10.42%	2.57%	29.98%
(e)	ImageNet	$L_{Triplet}$	11.69%	3.87%	51.41%	26.68%	13.90%	46.45%
(f)	Danbooru	$L_{Triplet}$	14.61%	5.57%	53.51%	30.22%	16.65%	50.02%
(g)	-	$L_{VAE}$	4.21%	0.46%	39.06%	7.06%	1.01%	19.93%
(h)	ImageNet	$L_{VAE}$	4.26%	0.47%	40.98%	7.10%	1.06%	21.22%
(i)	Danbooru	$L_{VAE}$	4.27%	0.47%	41.28%	7.10%	0.92%	21.65%
(j) [28]	-	$L_{GMVAE}$	4.42%	0.50%	41.85%	7.44%	1.01%	21.54%
(k)	ImageNet	$L_{GMVAE}$	4.37%	0.46%	42.76%	7.08%	1.01%	23.11%
(l)	Danbooru	$L_{GMVAE}$	4.48%	0.54%	42.86%	7.88%	1.20%	23.64%
(m)	-	$L_{Arc} + L_{VAE}$	6.69%	1.24%	47.07%	14.60%	5.11%	35.08%
(n)	ImageNet	$L_{Arc} + L_{VAE}$	8.45%	1.95%	48.68%	20.04%	8.83%	40.86%
(o)	Danbooru	$L_{Arc} + L_{VAE}$	10.53%	3.02%	50.60%	24.66%	12.28%	45.80%
(p)	-	$L_{Triplet} + L_{VAE}$	4.00%	0.17%	44.02%	7.62%	1.03%	25.74%
(q)	ImageNet	$L_{Triplet} + L_{VAE}$	13.06%	4.55%	52.12%	26.64%	14.52%	47.06%
(r) (i.e., ours)	Danbooru	$L_{Triplet} + L_{VAE}$	15.44%	6.06%	53.94%	32.70%	18.71%	51.71%

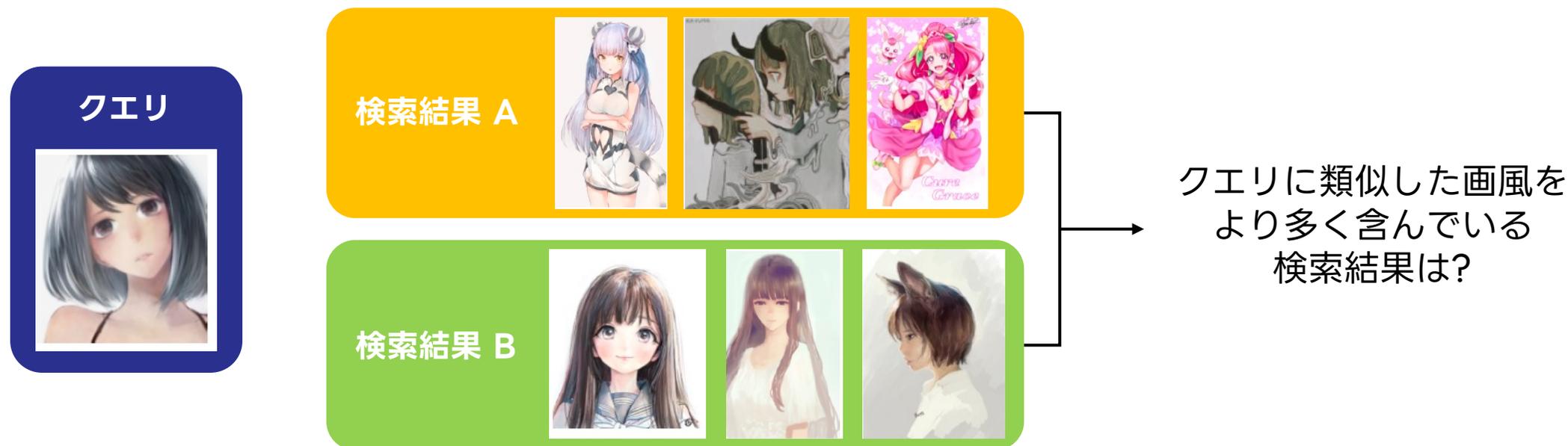
ベースラインは  
未知作者について性能が  
大きく劣化  
(-7.83%)

提案手法は劣化を  
抑えられたことが確認  
(-2.23%)

→ 提案手法は  
ベースライン手法より  
未知データへの高い頑健性を  
持つことを示唆

# 実験結果: AB テストによる定性的評価

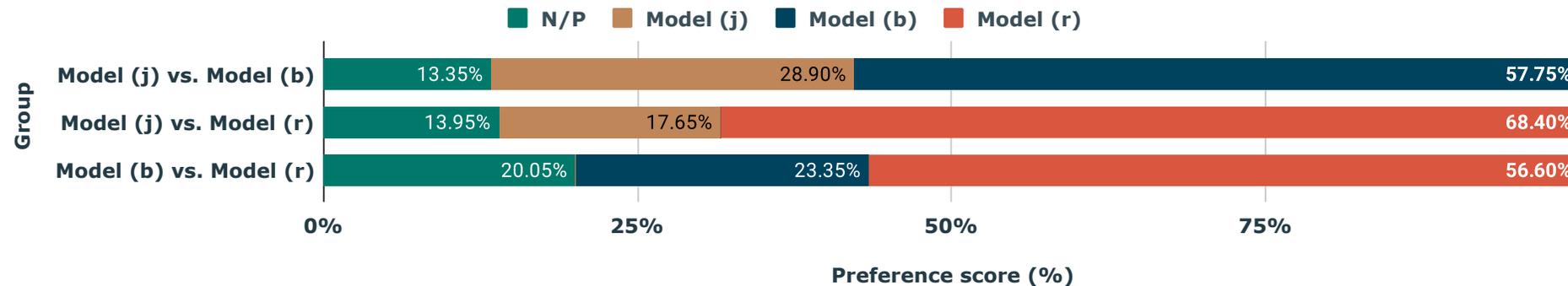
- 評価目的
  - 各モデルの画風特徴量に基づくクラスタリング結果が人間の知覚と一致するかを確認
- 評価方法
  - データセット: Danbooru Unseen サブセット
  - クエリ画像について各モデルの画風特徴量を用いた最近傍探索による Top-3 画像を収集
  - クエリ画像と 2 つの比較モデルの検索結果を提示しクエリ画像に類似した結果に投票
  - クエリ画像は 20 種類を活用
  - クラウドソーシングにより各比較グループごとに 100 人が参加



# 実験結果: AB テストによる定性的評価

## - 評価結果

- 定量的評価によって確認されたモデル間の優劣と比例した結果を確認
  - Model (j) < Model (b) < Model (r)
- 提案手法は人間による知覚と一致する結果をより多く提示したことを示唆



# まとめ

## 課題

画風情報は複雑な組み合わせによって形成され、抽象的な認知に基づく既存手法では未知のデータへの頑健性に懸念がある

## アプローチ

深層クラスタリングモデルによる潜在的特徴量を活用して画風を表現  
→ 定量的評価及び定性的評価にて提案手法の有効性を確認

## Document representation に関する貢献

本研究で指向するコミック検索システムのDocument representation として有効な活用が可能

# 議論

# 議論: 統計モデルを活用した Document representation の拡張

## 提案手法 1 (本論第 4 章)

CRF を用いたレビューに含まれる固有表現の抽出手法

## 提案手法 2 (本論第 5 章)

事前学習された畳み込みニューラルネットワークを用いた表紙画像内のオブジェクト抽出手法  
トピックモデルを用いたオブジェクトの共起関係に基づくコンセプト情報抽出手法

## 提案手法 3 (本論第 6 章)

深層クラスタリングモデルによる潜在的特徴量を用いた画風情報の抽出手法

# 議論: 統計モデルを活用した Document representation の拡張

## 提案手法 1 (本論第 4 章)

Document representation に必要なコミックの内容情報の抽出手法を獲得  
定量的評価・定性的評価により各手法による内容情報のコミック検索における有効性を確認

## 提案手法 2 (本論第 5 章)

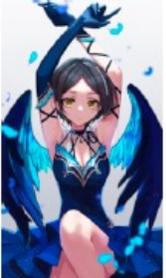
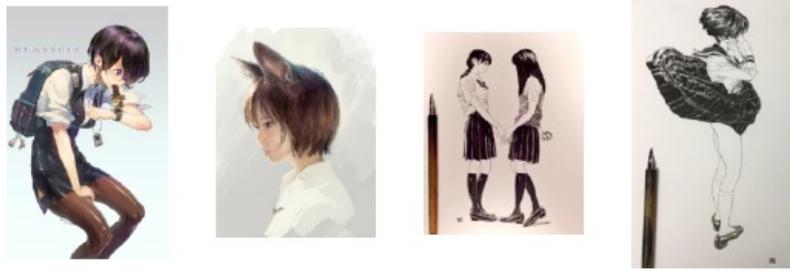
→ コミック検索システムにおける  
ユーザの多様な要求への対応性が向上されることが期待できる  
(e.g., 近未来な SF コンセプトでかわいい絵柄の作品が読みたい)

## 提案手法 3 (本論第 6 章)

→ ユーザの嗜好が反映された円滑なコミック検索に寄与

深層クラスタリングモデルによる潜在的特徴量を用いた画風情報の抽出手法

# 議論: より柔軟なコミック検索の可能性

Query A	Intermediate images	Query B
(1) 		
(2) 		
(3) 		

## Art style morphing

提案手法による画風特徴量を応用した画風に基づく作品検索

異なる画風を持つ任意のイラスト画像から両イラストの画風における特徴が混合された特徴を持つ画像を検索できる

→ 言語化が困難な画風についてより柔軟にユーザの嗜好を反映できる

# まとめ

# まとめ

## 本研究の目的

ユーザの嗜好が反映された円滑なコミック検索システムの実現

## 課題

従来のコミック検索システムの Document representation はコミックの内容情報が限定的

## 本研究のアプローチ

統計モデルを用いた内容情報抽出手法による Document representation の拡張

## 結果

コンセプト・画風などの内容情報を可能にし各情報のコミック検索における有効性を確認

各手法を用いた Document representation の拡張により本研究の目的の達成に貢献できる

# Appendix

# データ収集

- Wikipedia から 20 作品を選定し、該当作品の固有表現辞書を構築
- 該当作品のレビュー文を収集し、辞書と固有表現クラスに基づきアノテーション
- アノテーションの結果、5,052 文を確保

I/like/the/chapter/of/Alabasta

- “B-STORY” tag
- “I-STORY” tag
- “E-STORY” tag
- “O” tag

図 4.3: IREX と固有表現クラスを活用したアノテーション例

# モデルの学習

- データセットについて5分割交差検定を用いてハイパーパラメータ探索を実施
  - Elastic Net [42] における L1, L2 ペナルティ値を選定
  - 探索の結果、 $L1 = 1e-4$ ,  $L2 = 1e-4$  を採用

	Precision	Recall	F1-score
<b>B-ART</b>	0.676	0.646	0.580
<b>E-ART</b>	0.672	0.650	0.578
<b>I-ART</b>	0.000	0.000	0.000
<b>B-ITEM</b>	0.586	0.316	0.352
<b>E-ITEM</b>	0.586	0.316	0.352
<b>I-ITEM</b>	0.200	0.004	0.008
<b>B-LOC</b>	0.792	0.372	0.460
<b>E-LOC</b>	0.796	0.388	0.478
<b>I-LOC</b>	0.720	0.144	0.230
<b>B-ORG</b>	0.874	0.704	0.762
<b>E-ORG</b>	0.882	0.702	0.766
<b>I-ORG</b>	0.778	0.590	0.610
<b>B-PSN</b>	0.822	0.486	0.560
<b>E-PSN</b>	0.830	0.500	0.584
<b>I-PSN</b>	0.822	0.412	0.478
<b>B-SKILL</b>	0.650	0.280	0.358
<b>E-SKILL</b>	0.676	0.356	0.434
<b>I-SKILL</b>	0.972	0.468	0.568
<b>B-STORY</b>	0.806	0.566	0.628
<b>E-STORY</b>	0.838	0.578	0.648
<b>I-STORY</b>	0.684	0.644	0.644
<b>avg/total</b>	0.788	0.458	0.526

# WEBサービスでのコミック検索

いい  
キーワードが  
思いつかない



LINEマンガ 毎日無料 ストア インディーズ ログイン 検索

## LINEマンガ インディーズ 作品投稿受付中!

LINEマンガユーザーに、  
あなたの作品を読んでもらおう!

© Suzuko Yuki © haruko © Motoki Takauchi © Torako Sakurai

ランキングから探す ◯ 配信曜日から探す ◯ ジャンルから探す ◯ インディーズ作品投稿

検索機能を利用する

## 総合ランキング

もっと見る >



俺だけレベルMAX  
なピギナー  
ファンタジー・SF |  
毎週金曜更新



悪役陛下、執着する  
相手を間違ってます  
ファンタジー・SF |  
毎週日曜更新



入学傭兵  
バトル・アクション |  
毎週木曜更新



略奪された花嫁  
恋愛 | 毎週日曜更新

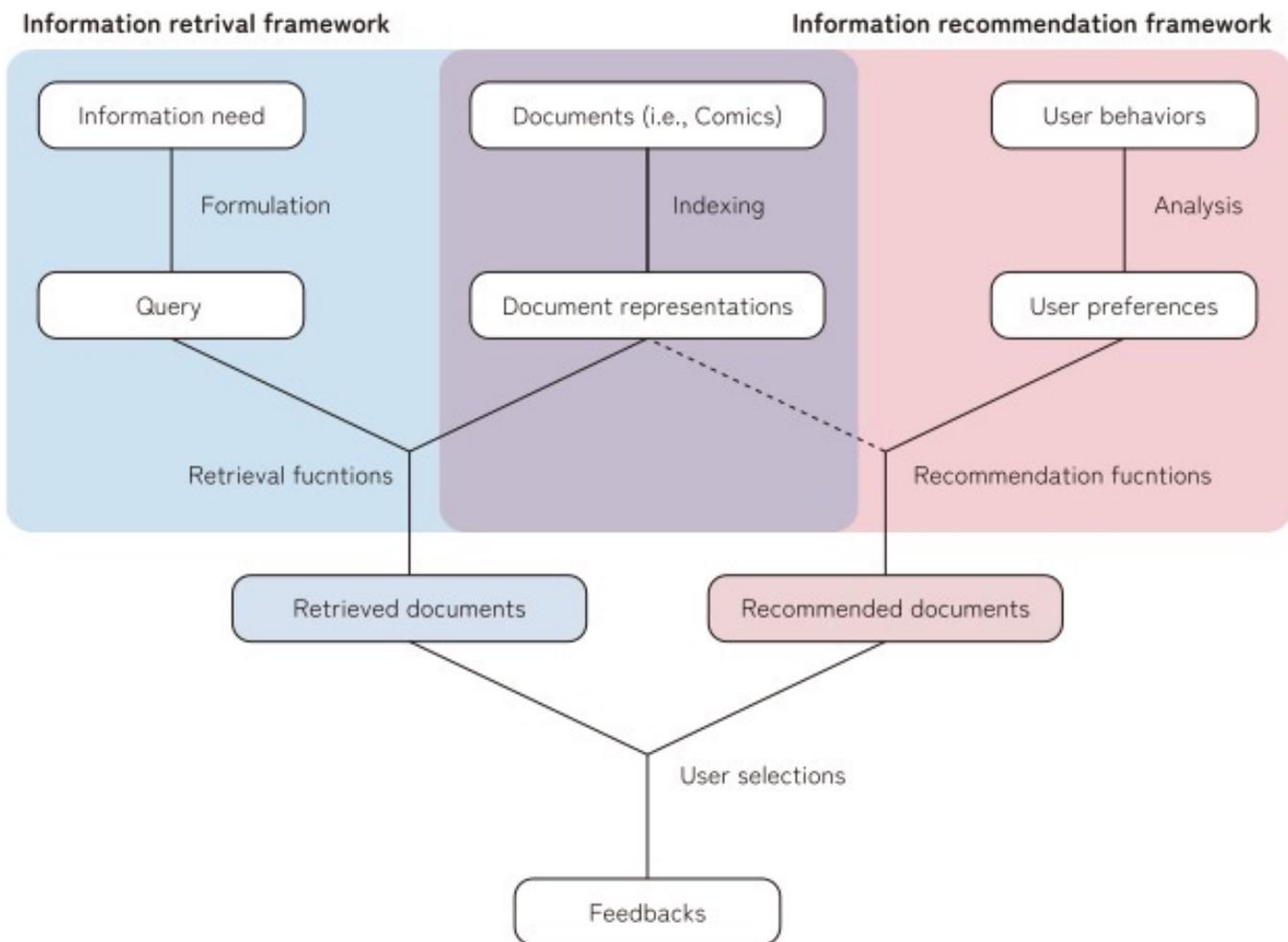


アカデミーの天才剣士  
ファンタジー・SF |  
毎週火曜更新



若き王との蜜夜  
恋愛 | 毎週土曜更新

# コミック検索モデル



## WEB サービスにおける ユーザのコミック検索の振る舞い

情報検索によって提示される情報と  
情報推薦によって提示される情報を  
横断しつつ探索的に検索を行う

# コミック検索モデル



**WEB サービスで扱われている  
大量のコミックへのアクセスを実現**

WEB サービスにおける  
ユーザのコミック検索の振る舞い

情報検索によって提示される情報と  
情報推薦によって提示される情報を  
横断しつつ探索的に検索を行う

# 実験結果: t-SNE を用いた画風特徴量の可視化

Group (e):  
Illustrations of Hatsune Miku  
by various artists



Group (d):  
deformed characters  
with detailed coloring



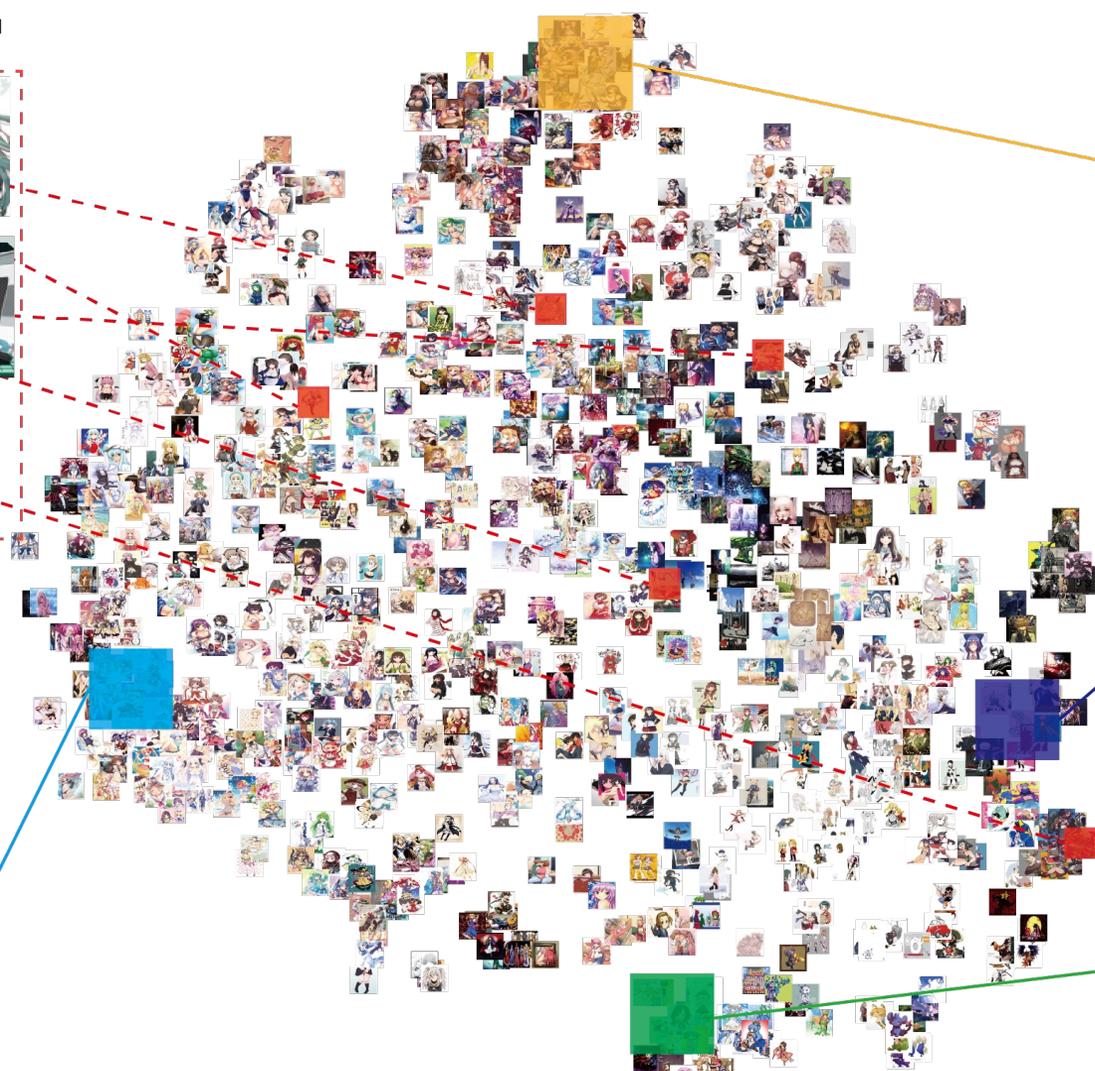
Group (a):  
realistic characters  
with detailed coloring



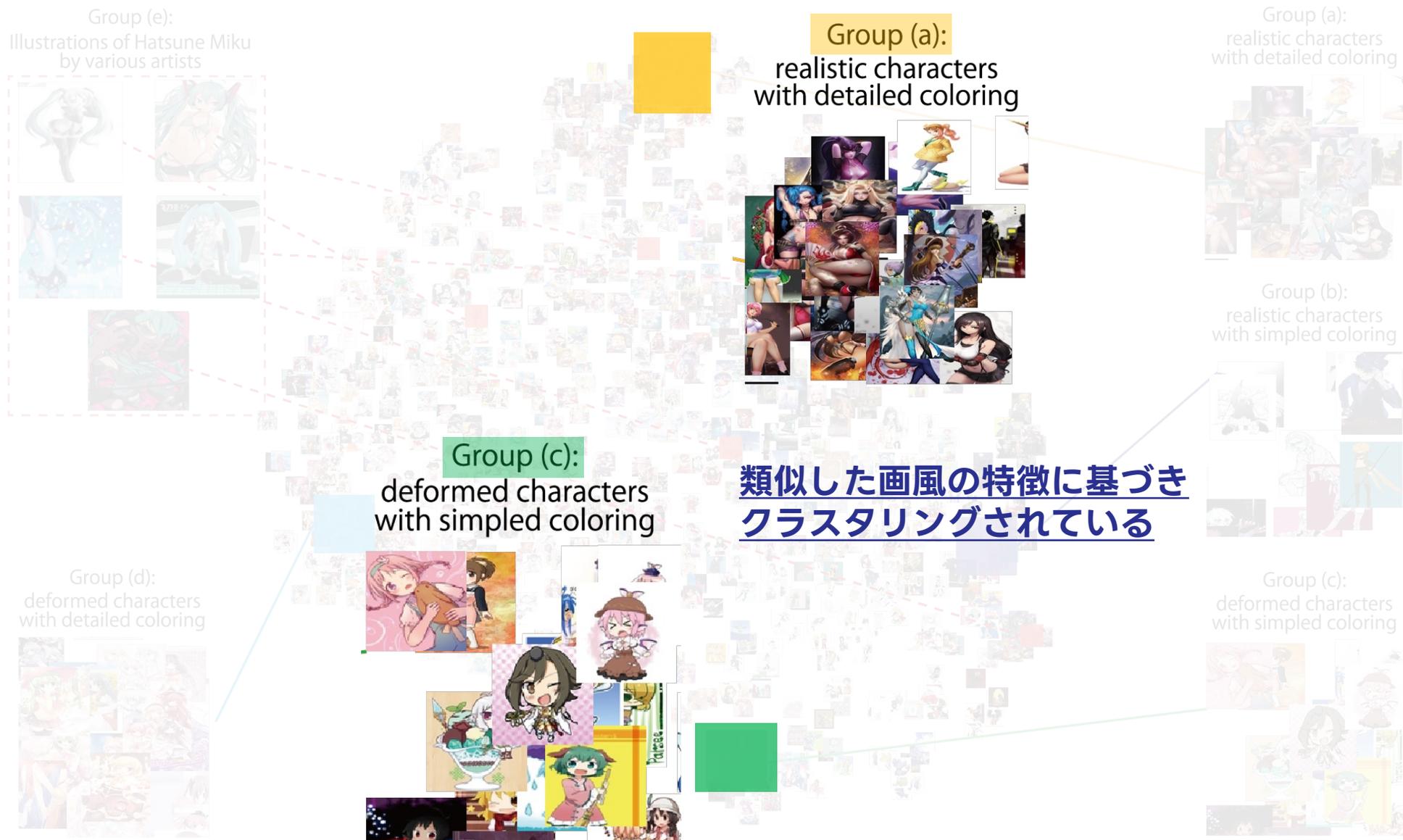
Group (b):  
realistic characters  
with simplified coloring



Group (c):  
deformed characters  
with simplified coloring

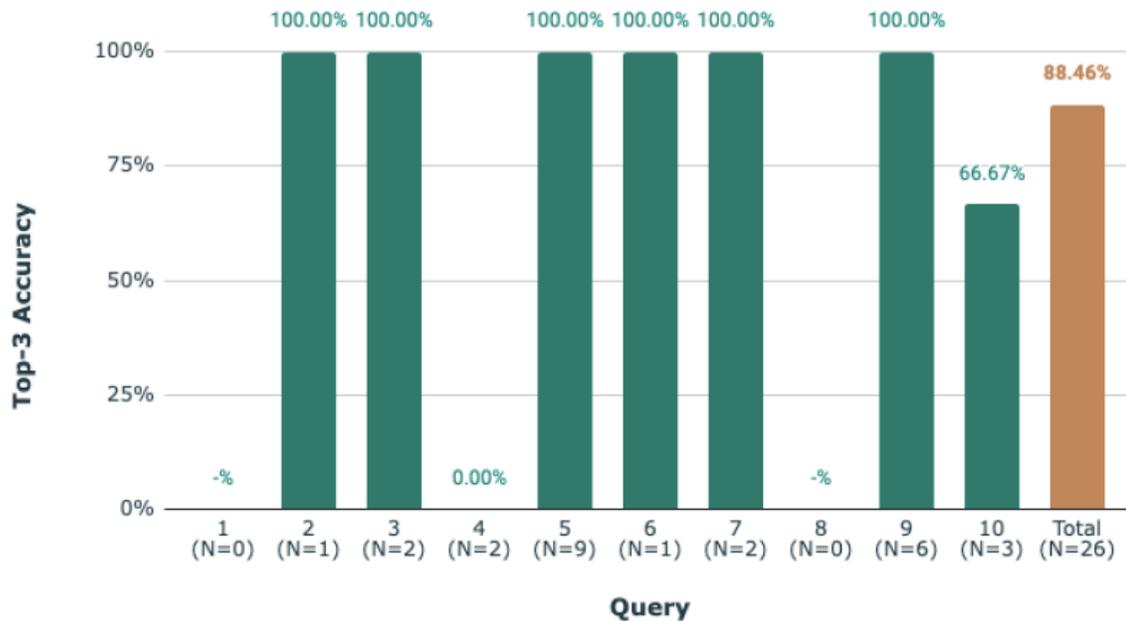


# 実験結果: t-SNE を用いた画風特徴量の可視化

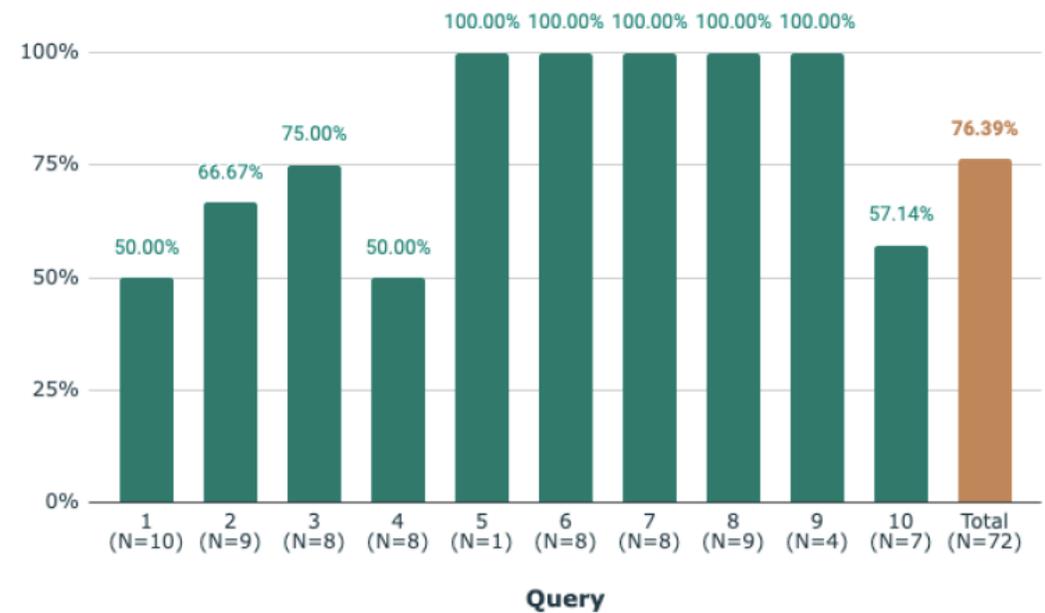


# 実験結果

- 実験協力者がどちらかのみを、またはどちらも知らない場合でも高い精度を達成
  - 提案手法によって推定されたコンセプト情報の未知の作品に関する有用性を示唆



どちらかの作品のみを知っている場合



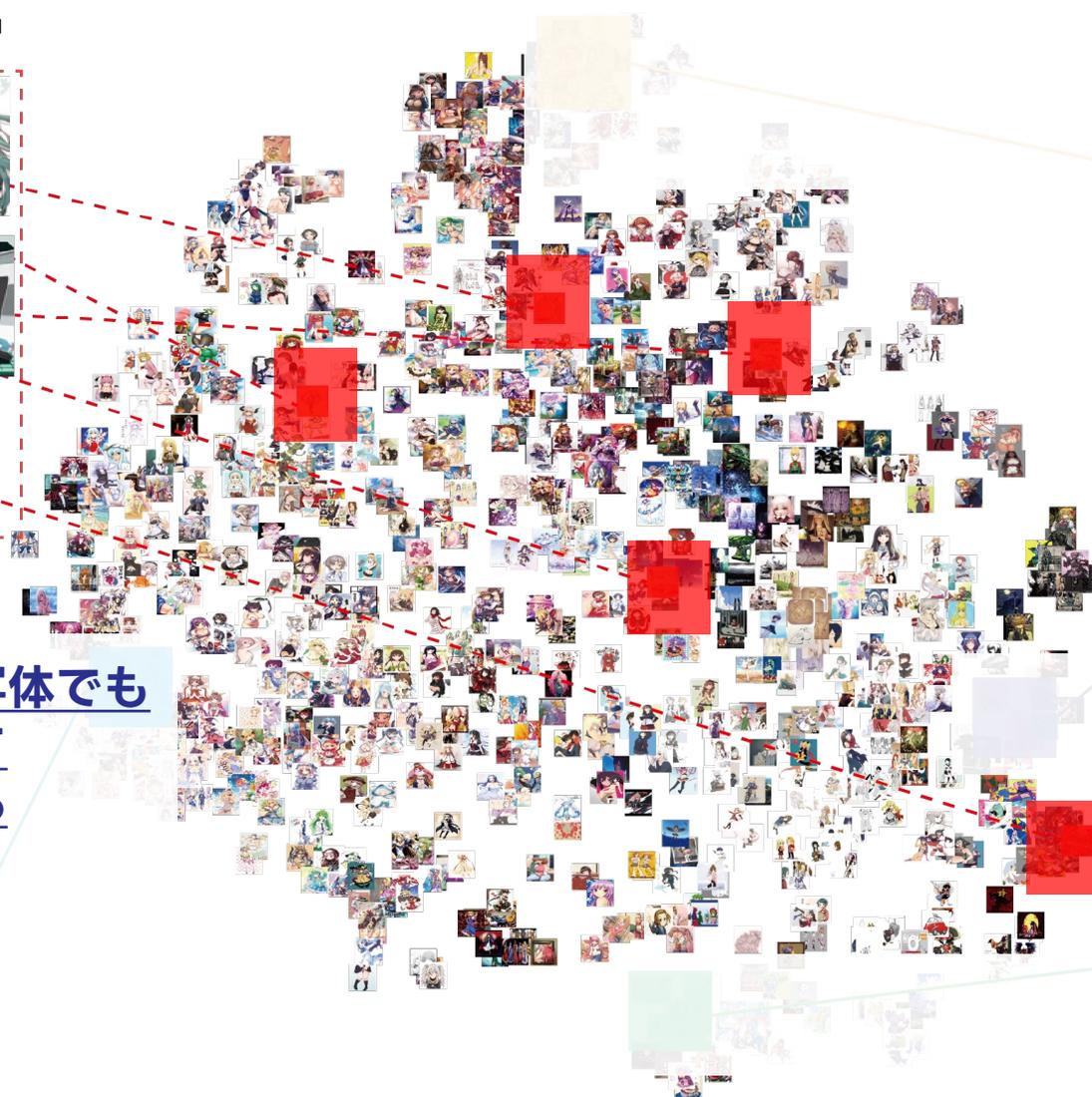
どちらかの作品も知らない場合

# 実験結果: t-SNE を用いた画風特徴量の可視化

Group (e):  
Illustrations of Hatsune Miku  
by various artists



同じ視覚的特徴を持つ被写体でも  
画風における特徴によって  
クラスタリングされている



Group (a):  
realistic characters  
with detailed coloring



Group (b):  
realistic characters  
with simplified coloring

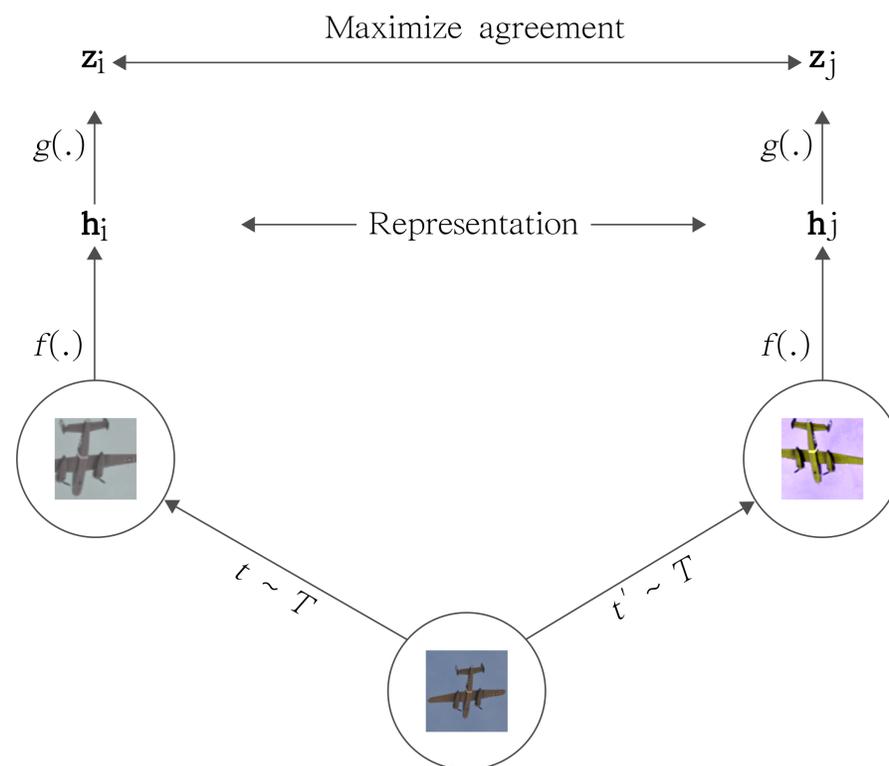


Group (c):  
deformed characters  
with simplified coloring



# 提案モデル: SimCLR [17] に基づく事前学習

- 対照学習 (Contrastive learning)による自己学習モデル
- 事前学習後の Encoder を提案モデルの Encoder として活用



# 実験結果: Manga 109 データセットを用いた定量的評価

Model	Face			Body			Frame			Cover		
	Accuracy	ARI	NMI	Accuracy	ARI	NMI	Accuracy	ARI	NMI	Accuracy	ARI	NMI
(b) [128]	4.92%	1.12%	9.98%	5.85%	1.51%	11.41%	6.33%	1.58%	12.21%	74.77%	-0.31%	95.15%
(j) [28]	4.38%	0.63%	5.54%	2.85%	0.28%	3.05%	2.86%	0.30%	2.93%	74.77%	-0.36%	94.88%
(f)	7.29%	2.14%	14.24%	8.66%	2.57%	15.77%	9.00%	2.56%	14.75%	79.44%	15.51%	95.51%
(r) (i.e., ours)	<b>7.56%</b>	<b>2.28%</b>	<b>14.89%</b>	<b>9.21%</b>	<b>2.79%</b>	<b>15.99%</b>	<b>9.21%</b>	<b>2.63%</b>	<b>15.14%</b>	<b>80.37%</b>	<b>15.97%</b>	<b>95.70%</b>



© Mai Asatsuki

Face: コマに描かれた人物の顔領域を切り取った画像のみ含まれるサブセット  
→ 学習データ (i.e., Danbooru) には希少な特徴を持つ

→ 提案手法はベースライン手法より  
未知ドメインへの高い頑健性を持つことを確認

# 実験結果: 各モデルを用いた最近傍探索結果の例

	(1)	(2)
Query		
Top-3 results	  	  
	(j) (b) (r)	(j) (b) (r)

提案手法は  
 ベースライン手法より  
 彩色や人物の描写といった観点から  
クエリ画像の特徴が類似した画像を  
 より多く提示