

マルチモーダルAIを活用した 漫画作品の感情推定に関する研究

香川大学 創造工学部 / 創発科学研究科
情報システム・セキュリティコース



松尾雄伊



藤井美月



(喜田研究室)

※ 右上のコマから、漫画の順で読んで下さい

マルチモーダルAIによる 漫画の感情推定

目的 両方の情報（マルチモード）をAIで解析し感情を推定！

イラスト



セリフ

信じられない！



アプローチ：

ボトムアップ解析

事前のヒントなしで、「セリフ」と「イラスト」の両方から、
まずは独立してキャラクターの感情を予想

トップダウン解析

「セリフ」と「イラスト」の解析結果が食い違った場合、
「セリフ」をヒントに、「イラスト」を再度解析し、感情を予想
この解析にはマルチモーダルAIを使用

二段階解析

漫画の魅力を世界に！

漫画の翻訳は機械翻訳では
現状まだ難しく人手に依存！

▶ 90% 未翻訳

信じられない！

感情に
よって
翻訳が
変化

【喜び】 → "It's like a dream!"
【怒り】 → "That's unacceptable!"
【悲しみ】 → "I can't accept this..."
【恐怖】 → "This can't be happening!"
【驚き】 → "Unbelievable!"
【嫌悪】 → "I don't believe you."
【無感情】 → "That is incredible."

セリフとイラストの情報を組みわせる！

ボトムアップ解析の結果である

セリフとイラストの感情が一致しているかチェックする

一致時：

両方の結果が同じならその感情を正しい感情として決定

不一致時：

一致しないときは、トップダウン解析に移行

ボトムアップ解析の結果を事前情報にし、再度解析！！

マルチモーダルAI

再度、解析を行う！

トップダウン解析



VQA

7つの感情
Anger, Disgust,
Fear, Happiness
Sadness, Surprise, Neutral
1つの感情を決定し出力

ボトムアップ解析
の結果

ボトムアップ解析の結果を
事前情報として与え、再度解析！

イラストから感情推定！

プロンプトエンジニアリング



* VQA (Visual Question Answering)
画像と質問を統合的に理解し、回答を出力する技術

VQA

7つの感情
Anger, Disgust,
Fear, Happiness
Sadness, Surprise, Neutral
1つの感情を決定し出力

内容を改良し、正しい感情の推定を目指す！

ボトムアップ解析

セリフから感情推定！

自然言語処理

セリフ抽出

単語分割

感情決定

STEP1:セリフ抽出 OCRを使ってコマ画像からセリフを抽出
STEP2:単語分割 抽出したセリフを最小単位の単語に切り分け
STEP3:感情決定 単語と感情の「近さ」を計算し、感情を決定

実行結果



今後の課題

回数	解析モジュール	出力された感情
1回目	イラスト解析	Neutral
	テキスト解析	Happiness
	整合性判定	×
2回目	イラスト解析	Happiness
	最終結果	Happiness
正解ラベル		Happiness

- 感情以外の情報
- 漫画特有の要素への対応
- 文脈を考慮しての処理 etc...

システム構成図

