

## 進捗報告

### 1 進捗

- コマ画像データの切り出し
- コマ画像データのみでの感情推定
- セリフの遷移情報を加味する
- 発話者情報を加味する
- i2v の fine-tuning

### 6 タスク

情報処理学会関西支部 資料作り

### 2 実験設定の変更

全体的な実験設定変更として、エポック数を 200 から 50 に変更した。精度への影響を確かめるため、セリフ 1 文の感情推定をやり直した。識別器としては 3 層 MLP を用いた。入力次元 768, 隠れ層次元 30, 出力次元 2 とした。その結果を表 1 に示す。

### 3 コマ画像データのみでの感情推定

識別器としては 3 層 MLP を用いた。入力次元 4096, 隠れ層次元 768, 出力次元 2 とした。その結果を表 2 に示す。

### 4 セリフ 1 文のマルチモーダル感情推定

識別器としては 3 層 MLP を用いた。入力次元 768+4096, 隠れ層次元 300, 出力次元 2 とした。BERT は最終層のみをチューニングした。その結果を表 3 に示す。

### 5 裏でやっていること

- 過去のセリフを考慮した感情推定 やり直し (セリフの遷移情報を加味)
- コマベクトルを PCA で圧縮したものを用いたマルチモーダル感情推定

表 1: セリフ 1 文の感情推定

	ギャグ			少女漫画			少年漫画			青年漫画			萌え			5 タッチ総合		
	Acc	P-Recall	P-F1	Acc	P-Recall	P-F1	Acc	P-Recall	P-F1	Acc	P-Recall	P-F1	Acc	P-Recall	P-F1	Acc	P-Recall	P-F1
BERT fixed	0.712	0.200	0.174	0.567	0.632	0.623	0.766	0.083	0.118	0.692	0.643	0.474	0.594	0.591	0.500	0.666	0.510	0.473
BERT fine tuning	<u>0.818</u>	0.200	<u>0.250</u>	<u>0.612</u>	0.711	<u>0.675</u>	0.766	0.083	0.118	<u>0.862</u>	0.500	<u>0.609</u>	<u>0.609</u>	0.591	<u>0.510</u>	<u>0.733</u>	0.521	<u>0.535</u>
ベースライン	0.848	0	0	0.432	0	0	0.812	0	0	0.784	0	0	0.656	0	0	0.705	0	0

表 2: コマ画像データのみでの感情推定

	ギャグ			少女漫画			少年漫画			青年漫画			萌え			5 タッチ総合		
	Acc	P-Recall	P-F1	Acc	P-Recall	P-F1	Acc	P-Recall	P-F1	Acc	P-Recall	P-F1	Acc	P-Recall	P-F1	Acc	P-Recall	P-F1
koma only	0.697	0.100	0.091	0.612	0.526	0.606	0.438	0.750	0.333	0.477	0.571	0.320	0.594	0.364	0.381	0.564	0.479	0.393
ベースライン	0.848	0	0	0.432	0	0	0.812	0	0	0.784	0	0	0.656	0	0	0.705	0	0

表 3: コマ画像データのみでのマルチモーダル感情推定

	ギャグ			少女漫画			少年漫画			青年漫画			萌え			5 タッチ総合		
	Acc	P-Recall	P-F1	Acc	P-Recall	P-F1	Acc	P-Recall	P-F1	Acc	P-Recall	P-F1	Acc	P-Recall	P-F1	Acc	P-Recall	P-F1
multi	0.773	0.200	0.211	0.687	0.763	0.734	0.703	0.417	0.345	0.769	0.643	0.545	0.641	0.500	0.489	0.715	0.583	0.546
ベースライン	0.848	0	0	0.432	0	0	0.812	0	0	0.784	0	0	0.656	0	0	0.705	0	0