

## 進捗報告

### 1 進捗

- 単文のマルチモーダル感情推定 (発表資料より抜粋しています)

#### 1.1 illustration2vec

illustration2vec[1] は, Masaki Saito と Yusuke Matsui らが提案した画像のベクトル化手法で, Danbooru と Safebooru から 100 万枚のイラストを用いて学習した事前学習済みモデルが公開されている. illustration2vec で扱った問題としてイラストに対する画像認識の難しさがあり, 既存の画像認識モデルのほとんどが ImageNet などを評価対象にしており, アニメや漫画といったイラストに対して評価を行っていなかったため, イラストのベクトル化手法に関してより良い期待が持てる手法である. 本稿では事前学習済みモデルを使って 4096 次元のコマ画像のベクトルを獲得する. 概要図を図 1 に示す.

### 2 実験 1 : 時系列を考慮しないセリフの感情推定

この実験では 1 つのセリフを入力し, 対応する感情ラベルを出力するような感情推定を行う. JUMAN++<sup>1</sup> によって分かち書きされたセリフを BERT への入力とし, 識別器としては 3 層 MLP を用いた. 表 1 に MLP で用いたパラメータ, そして表 2 に学習で用いたパラメータを示す. 学習率は optuna によって最適なパラメータを探索した.

<sup>1</sup><http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN++>

BERT の事前学習済みモデルの全ての重みを固定した場合 (BERT fixed) と最終層のパラメータだけをチューニングした場合 (BERT last layer) の 2 つの手法について感情推定を行い, 各タッチ・手法の結果について比較し, fine tuning の有用性を確かめることを目的とした.

表 3 は各タッチについて, 評価用データの結果をまとめたものである. 表 3 より, Accuracy に関しては差異はないが共にベースラインを超え, Recall, F1 値共に BERT fixed より BERT last layer の方が上回ったことから, fine tuning の有用性を確かめることができた.

### 3 実験 3 : 単文のマルチモーダル感情推定

実験 1 において, BERT から得た 768 次元のベクトルと, 入力したセリフが含まれているコマ全体の画像を illustration2vec に入力し, 得た 4096 次元のベクトルを concat した 4864 次元のベクトルを, 同様に 3 層 MLP に入力することでマルチモーダルな感情推定を行った. データセット整備の都合上, 少年漫画タッチ, 青年漫画タッチ, 萌えタッチについてのみ実験を行った.

表 4 は 3 タッチについて, 評価用データの結果をまとめたものである. 比較として, 実験 1 における 3 タッチの結果も載せている. 表 4 より, 実験 1 と比べ, Recall 以外の評価指標は下回ったが, 少年漫画タッチにおいては今回の実験を通して最も高い Recall と F1 値が出ており, マルチモーダルな感情推定を行う意味はあると推測できる.

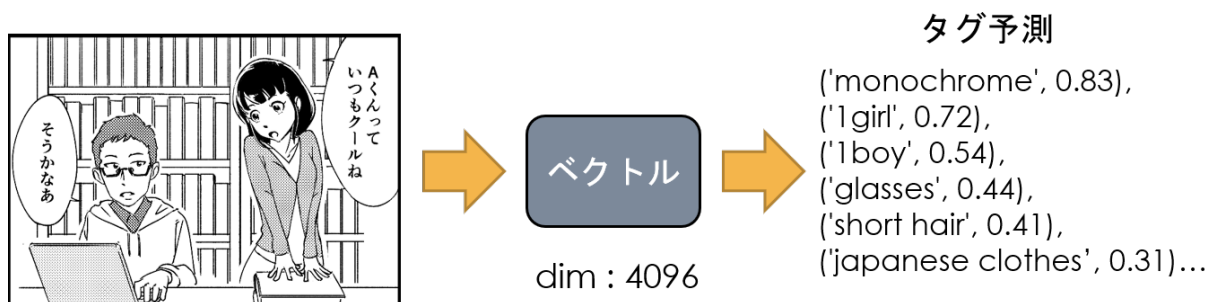


図 1: i2v 概要図

## 参考文献

- [1] Yusuke Matsui Masaki Saito. Illustration2vec:  
A semantic vector representation of illustrations.  
2015.

表 1: 実験 1 MLP パラメータ

|                     |            |
|---------------------|------------|
|                     | MLP        |
| (in,hidden,out)     | (768,30,2) |
| activation function | tanh       |
| dropout rate        | 0.5        |

表 2: 学習パラメータ

|               |                    |
|---------------|--------------------|
|               | 実験 1・2・3           |
| epoch         | 200                |
| batch size    | 16                 |
| loss function | Cross Entropy Loss |
| optimizer     | Adam               |

表 3: 実験 1 結果 (評価用データ)

|                 | ギャグ   |        |       | 少女漫画  |        |       | 少年漫画  |        |       | 青年漫画  |        |       | 萌え系   |        |       | 5 タッチ総合 |        |       |
|-----------------|-------|--------|-------|-------|--------|-------|-------|--------|-------|-------|--------|-------|-------|--------|-------|---------|--------|-------|
|                 | Acc   | Recall | F1    | Acc   | Recall | F1    | Acc   | Recall | F1    | Acc   | Recall | F1    | Acc   | Recall | F1    | Acc     | Recall | F1    |
| BERT fixed      | 0.757 | 0.200  | 0.200 | 0.552 | 0.684  | 0.634 | 0.796 | 0.000  | 0.000 | 0.784 | 0.357  | 0.416 | 0.718 | 0.454  | 0.526 | 0.720   | 0.447  | 0.485 |
| BERT last layer | 0.818 | 0.200  | 0.250 | 0.641 | 0.710  | 0.692 | 0.781 | 0.000  | 0.000 | 0.800 | 0.500  | 0.518 | 0.578 | 0.500  | 0.448 | 0.723   | 0.489  | 0.510 |
| ベースライン          | 0.848 | 0.000  | 0.000 | 0.432 | 0.000  | 0.000 | 0.812 | 0.000  | 0.000 | 0.784 | 0.000  | 0.000 | 0.656 | 0.000  | 0.000 | 0.705   | 0.000  | 0.000 |

表 4: 実験 3 結果 (評価用データ)

|                 | ギャグ |        |    | 少女漫画 |        |    | 少年漫画  |        |       | 青年漫画  |        |       | 萌え系   |        |       | 3 タッチ平均 |        |       |
|-----------------|-----|--------|----|------|--------|----|-------|--------|-------|-------|--------|-------|-------|--------|-------|---------|--------|-------|
|                 | Acc | Recall | F1 | Acc  | Recall | F1 | Acc   | Recall | F1    | Acc   | Recall | F1    | Acc   | Recall | F1    | Acc     | Recall | F1    |
| 実験 1 fixed      | -   | -      | -  | -    | -      | -  | 0.796 | 0.000  | 0.000 | 0.784 | 0.357  | 0.416 | 0.718 | 0.454  | 0.526 | 0.766   | 0.312  | 0.400 |
| 実験 1 last layer | -   | -      | -  | -    | -      | -  | 0.781 | 0.000  | 0.000 | 0.800 | 0.500  | 0.518 | 0.578 | 0.500  | 0.448 | 0.720   | 0.375  | 0.400 |
| 実験 3            | -   | -      | -  | -    | -      | -  | 0.562 | 0.500  | 0.300 | 0.784 | 0.500  | 0.500 | 0.562 | 0.363  | 0.363 | 0.637   | 0.437  | 0.375 |
| ベースライン          | -   | -      | -  | -    | -      | -  | 0.812 | 0.000  | 0.000 | 0.784 | 0.000  | 0.000 | 0.656 | 0.000  | 0.000 | 0.751   | 0.000  | 0.000 |