

進捗報告

1 進捗

- データセット再構築
- i2v 事前学習済みモデルの考察

練・テストデータの切り分けには注意しなければならない。拡張後について、UNK ラベルのものは拡張しないとすれば、拡張後について UNK ラベルの付いたデータの割合は 0.05% にまで落とせる。

2 データセット再構築

表 1 に拡張前の 4 コマ漫画ストーリーデータセットの感情ラベル数、表 2 には日本語 WordNet のシソーラスを用いて拡張した拡張後の感情ラベル数を示す。書き文字はいずれの場合も除いている。UNK ラベルのデータの分布は各タッチで異なっており、訓

3 i2v 事前学習済みモデルの考察

事前学習済みモデルにおいて、画像の分散表現にばらつきがある問題について、各層からの出力を取って、10 回試行の分散を計算した。その結果が表 3 である。また、モデルのネットワーク図を図 1 に示す。分散を求める際は各出力を 1 次元ベクトルに平坦

表 1: 拡張前データ

emotion	ギャグ	少女	少年	青年	萌え	合計
ニュートラル	43	8	55	33	30	169 (25.6%)
喜楽	25	77	27	33	47	209 (31.7%)
驚愕	19	16	17	29	20	101 (15.3%)
悲哀	25	12	13	16	13	79 (11.9%)
恐怖	6	11	8	8	9	42 (6.3%)
憤怒	4	5	2	7	2	20 (3.0%)
嫌悪	2	4	3	3	4	16 (2.4%)
UNK	7	3	5	2	6	23 (3.4%)
合計	131	136	130	131	131	659

表 2: 拡張後データ

emotion	ギャグ	少女	少年	青年	萌え	合計
ニュートラル	3065	186	4180	3123	2032	12586 (31.5%)
喜楽	1697	5121	1861	2018	3663	14360 (36.0%)
驚愕	593	563	484	1149	693	3482 (8.7%)
悲哀	1769	689	770	696	739	4663 (11.7%)
恐怖	241	609	583	438	610	2481 (6.2%)
憤怒	80	593	9	681	9	1372 (3.4%)
嫌悪	2	244	5	8	184	443 (1.1%)
UNK	147	3	175	2	180	507 (1.3%)
合計	7594	8008	8067	8115	8110	39894

化してから計算した.

表 3: 分散

10 回試行	分散
input	0.0
conv1_1	1.1e-07
conv2_1	1.5e-07
conv3_1	8.7e-08
conv3_2	5.3e-08
conv4_1	2.3e-08
conv4_2	9.3e-09
conv5_1	5.8e-09
conv5_2	3.1e-09
conv6_1	1.4e-09
conv6_2	3.3e-10
conv6_3	4731.3
encode1	36.4

表 3 より, 明らかに最後の畳み込み層からの出力が原因であると考えられる. 図 1 から, 最後の畳み込み層の後には Dropout 層が付随しているが, ソースコード上では train モードが False になっているため大丈夫だと思っていたが, 実際には Dropout が働いていたらしい. (chainerrusing_config 内で設定することで分散はほぼ 0 に近づいた.)

この設定で, 先週と同じ入力として初音ミクの画像を用いた時の PyTorch 移植版とのコサイン類似度は 1 になった.

4 次やること

データの切り分け方を決めて hottoSNS-BERT を動かす.

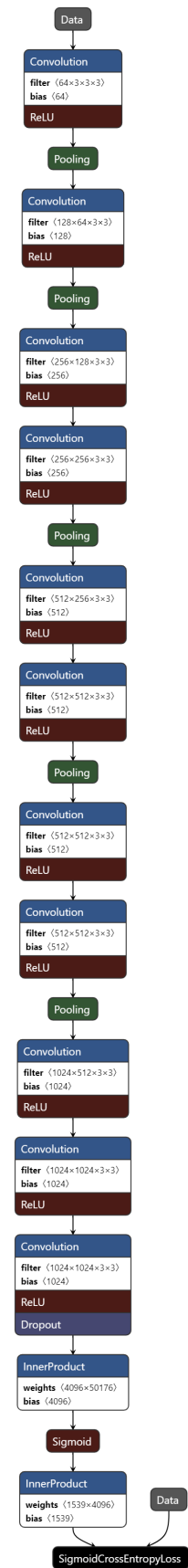


図 1: i2v モデル図