

進捗報告

1 今週あったこと

- i2v の PyTorch 移植完了
- データセット引継ぎとアノテーション作業
- データセットの再構築

2 i2v の PyTorch 移植完了

- 結局, 自分でラッパーを書いて動作確認をした
- .caffemodel を .npz にしてから .pth にした

図1に, 自作ラッパー版とオリジナルの出力比較を載せる. 入力 is オリジナルの GitHub でも紹介されているデモ用画像を用い, 確率のしきい値は 0.5 とした. 図1より, おおよその結果は一致したが所々出力が異なった. また, 4096 次元の分散表現同士のコサイン類似度は 0.979 であった. なぜ違いが生まれるのかについては要検証だが, ファインチューニングもできるようになった.

3 データセット引継ぎとアノテーション作業

寺内さんから改めて4コマ漫画ストーリーデータセットのオリジナルと, 寺内さんが用意してくれた各コマについてレイヤー分けされたものと結合するスクリプトを頂いた. これを利用することでほとんどのデータについてはセリフの白抜きだけでなく任意のレイヤー結合が可能. ごく一部のデータは全レイヤーが結合された状態のものしかないので手作業で切らなければならないという現状. 後世のためにもアノテート作業(レイヤー分けされたものがどの種類か, という作業)を完成しておいた方がよい.

4 データセットの再構築

先週示したデータ不備をもとに再構築中です.

5 usagi

一応, cuda が機能するようになっていたが遅い気がする?


	PyTorch 移植版	Original (caffemodel)
 Input Image	<pre>{ 'general': [('thighhighs', 0.99), ('1girl', 0.98), ('twintails', 0.98), ('solo', 0.96), ('aqua hair', 0.91), ('long hair', 0.88), ('very long hair', 0.83), ('detached sleeves', 0.74), ('skirt', 0.67), ('necktie', 0.56), ('aqua eyes', 0.55)], 'character': [('hatsune miku', 0.99)], 'copyright': [('vocaloid', 0.99)], 'rating': [('safe', 0.97), ('questionable', 0.02), ('explicit', 0.00)] }</pre>	<pre>{ 'general': [('thighhighs', 0.99), ('twintails', 0.98), ('1girl', 0.98), ('solo', 0.96), ('aqua hair', 0.91), ('long hair', 0.89), ('very long hair', 0.83), ('detached sleeves', 0.72), ('skirt', 0.72), ('necktie', 0.67), ('aqua eyes', 0.65), ('boots', 0.58), ('thigh boots', 0.56)], 'character': [('hatsune miku', 0.99)], 'copyright': [('vocaloid', 0.99)], 'rating': [('safe', 0.98), ('questionable', 0.01), ('explicit', 0.00)] }</pre>

図1: i2v 比較