

進捗報告

1 今週やったこと

- 画像と台詞の分散表現を取得
- 画像の分散表現のサイズを可変に

1.1 画像と台詞の分散表現を取得

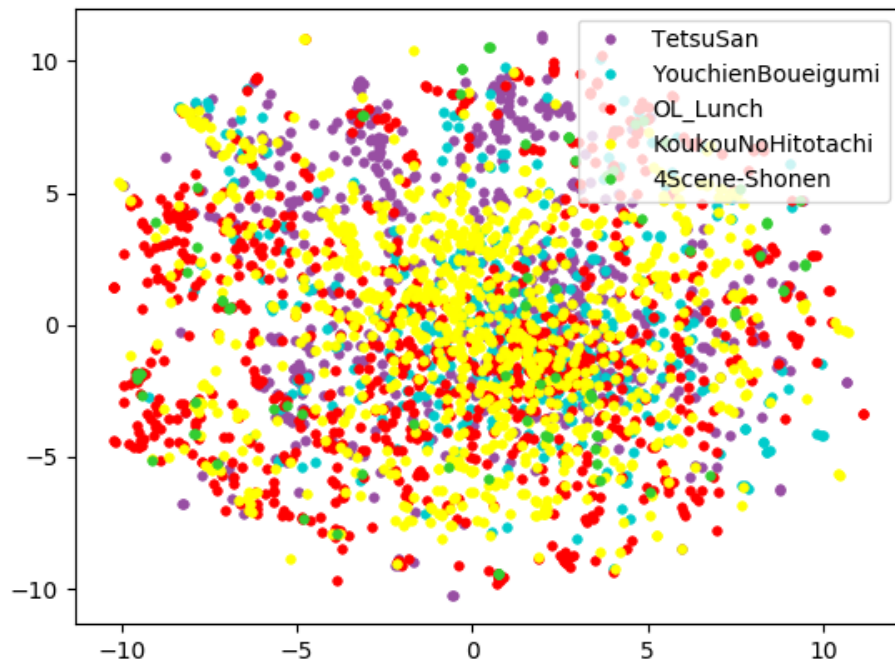
Metric Learning に取り組むにあたって画像と台詞の分散表現がどれくらい分離できているのか調べた．表 1 に t-SNE によるプロットを示す．いずれも 768 次元から 2 次元への圧縮をしている．例によって CAE で得られた分散表現はタイトル毎の分離は見られないが，台詞は少し分離できているような気もする...？ OL Lunch だけやたら分離できている領域があるのも気になった．解析用のコードを書いたので見せながら．

1.2 画像の分散表現のサイズを可変に

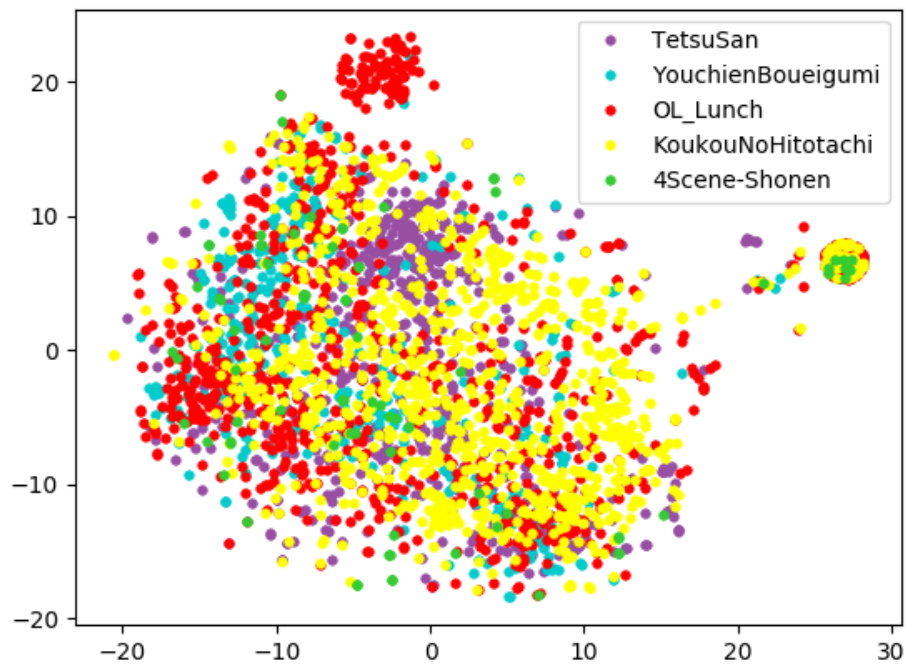
画像 台詞と台詞 画像のマッチングの識別率に違いがあったため，画像と台詞の分散表現の次元を揃えて再実験した．実装的には CAE のエンコーダの後ろに線形層を挟み中間層の次元を可変にしている (BERT の埋め込み次元 768 と合わせるため)．ちなみにこの処理は keras 時代にもしようとしていたがロスが下がらず画像もほとんど復元できていなかったため断念していた．PyTorch でこの実装を書き，ロスが正常に下がり，解像度は落ちているがしっかりと元画像が復元されていることを確認した．図 2 に再学習したマッチング実験のロスの遷移を示す．画像と台詞の次元を一致させたが，やはり台詞 画像のマッチング問題が難しいらしい (これはなぜ?) ．

2 来週の予定

Metric Learning の知見をためる．かつできれば簡単なモデルを実装．

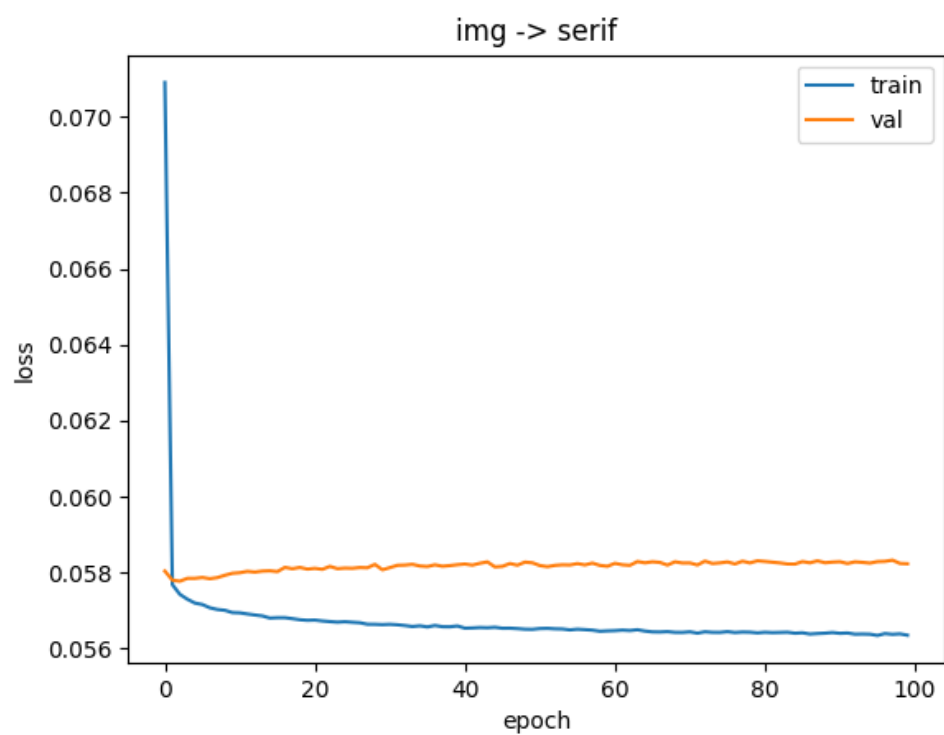


(a) 画像の分散表現プロット

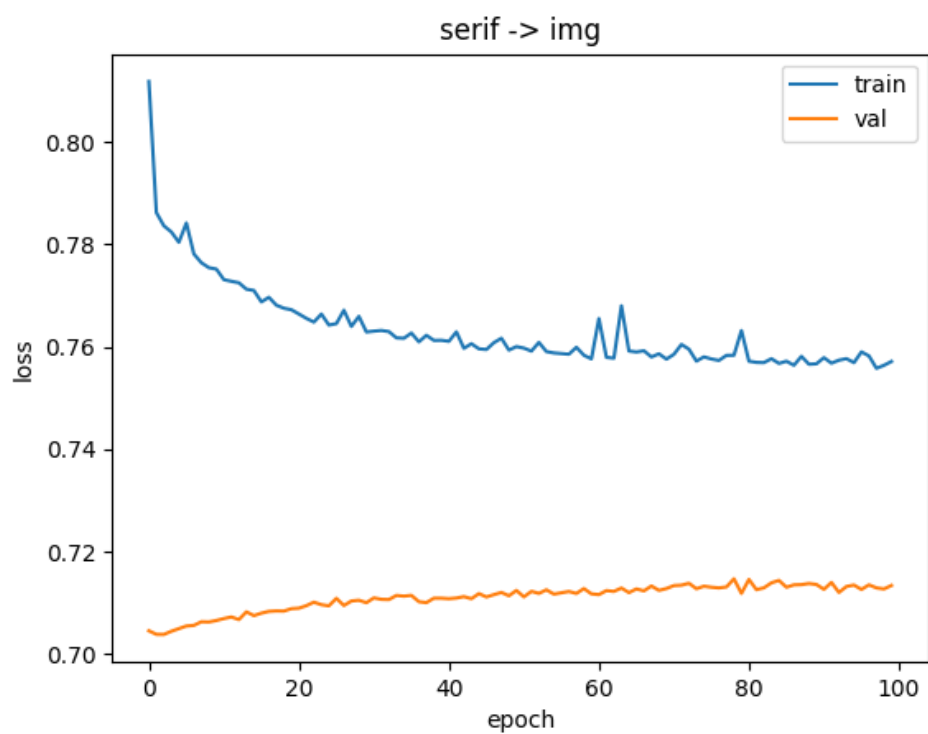


(b) 台詞の分散表現プロット

図 1: t-SNE



(a) 画像 台詞



(b) 台詞 画像

図 2: ロスの遷移