

uplatex を使った 2 段組文書作成

Haruka *

2021 年 11 月 26 日

1 はじめに

最近では Over Leaf 等のオンラインの L^AT_EX サービスがあるので、それらを使うといいと思うが画像のコンパイル等が重くなるのでやっぱりローカルで環境を作ったほうがいいかもしれない。upLaTeX は Unicode が使えるので記号等の問題が少ないのがメリットかもしれない。

2 関連研究

既存研究として、Vinyals ら [1] は、畳み込みニューラルネットワーク (CNN) と再起型ニューラルネットワーク (RNN) を用いたエンコーダー・デコーダー構造により、画像特徴とキャプションから、文の生成を学習する手法を提案している。

3 提案手法

Vinyals らの提案する Image Captioning のエンコーダー・デコーダー構造を参考に、人間を用いた人力キャプション生成を提案する (図 1)。



図 1 絵を見て内容を言語にて回答

4 生成実験

友達と協力させる人力キャプション生成を実装する (図 2)。既存手法として Vinyals らの手法 (NIC) と、人力キャプション生成モデル、それに友人を導入したと友

表 1 評価用データセットに対するキャプション生成の品質。1,2,3,4 は n -grams を使った Bleu スコア。数値が高いほど Ground Truth との類似度が高い。Ours+F はモデルと友人 (Friend) を相談させるネットワークとする。

	Bleu				Bert		
	1	2	3	4	P	R	F1
NIC[1]	0.113	0.018	0.001	0.000	0.856	0.878	0.867
Ours	1.00	0.990	0.980	0.900	0.990	0.990	0.990
Ours+Fr	1.00	1.00	0.999	0.990	0.999	0.999	0.999



図 2 友人と協力モデル (Ours+Fr) の図

人協力モデルを比較した。表 1 に示すように、人力キャプション生成モデルの方が著しく高いスコアとなった。また、友人協力モデルは最も高い評価となったものの、人力キャプション生成モデルと比較すると差はわずかである。

5 まとめ・今後の課題

完璧すぎてありません。

参考文献

- [1] Oriol Vinyals, A. Toshev, S. Bengio, and D. Erhan. Show and tell: A neural image caption generator. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3156–3164, 2015.

* 国立高等専門学校 全知全能科