Contents

1 Introduction

2 background

2.1 neural networks and deep learning

keywords to explain (maybe) from [1] - knowledge distillation - generalist models - large unsupervised training data sets - transformers - cnns - deep neural networks - self attention - convolution - transfer learning - encoder/decoder - model compression - loss functions - cross-entropy - data augmentation - training/validation/test data sets - learning rate - batch size - tokenizing - image patches - self-attention - multi-head self-attention - performance metrics - precision - recall - f1

2.2 paleoecology

3 data methods etc

things to try data augmentation

4 results

5 conclusion

References

[1] M. Li, T. Lv, L. Cui, Y. Lu, D. Florencio, C. Zhang, Z. Li, and F. Wei. *TrOCR: Transformer-based Optical Character Recognition with Pre-trained Models*. 2021. arXiv: 2109.10282 [cs.CL].