BERT’s Handwriting
Testing BERT Word Prediction Effect on Accuracy in Handwritten Text Recognition
Anna Widder : awidder@stanford.edu

Introduction

• Handwritten text recognition is essential to digitization of data and modern workflow
• The BERT encodings have pushed the boundaries of many NLP tasks—could it improve HTR accuracy?
• Can BERT can predict a word from context that normal HTR might miss?
• Relevant metrics: Character Error Rate (CER) Word Error Rate (WER)

Data

• IAM Handwriting Database
• 13,353 isolated and labeled text lines from 657 different people

CTC Loss and HTR

• Connectionist Temporal Classification is the fastest and most accurate method of training BLSTM output
• Current research being done to best decode CTC output
• Current methods are beam search, token passing and word beam search

Models and Method

Results

<table>
<thead>
<tr>
<th></th>
<th>Train (CER:WER)</th>
<th>TEST (CER : WER)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple HTR</td>
<td>6.0336 : 55.6923</td>
<td>5.2512 : 53.0769</td>
</tr>
<tr>
<td>HTR + BERT (avg. top score)</td>
<td>6.0903 : 55.5385</td>
<td>5.2832 : 52.7692</td>
</tr>
<tr>
<td>HTR + BERT (avg. char score)</td>
<td>6.2773 : 57.0769</td>
<td>5.5491 : 54.0000</td>
</tr>
</tbody>
</table>

Discussion

• Disappointed BERT did not significantly increase performance of HTR models
• Results seem to be skewed by data irregularity from the train to test set results.

Future Work

• Line segmentation for paragraph input data
• Increase dataset with more variety of handwriting
• Develop metric of certainty at word level
• Expand image processing to handle real-time data

Key repositories and model examples: