HTR+ in Transkribus

Gundram Leifert
Tobias Grüning

Computational Intelligence Technology Lab
University of Rostock
Topics

- Technical Comparison HTR vs. HTR+
- Performance Comparison HTR vs. HTR+
- Best Practice
- Discussion
Technical Comparison HTR vs. HTR+

First: What stays unchanged?

- **input:** image
- **output:** text and ConfMat
- trained with Connectionist Temporal Classification (CTC)
  ⇒ The existing workflow stays untouched

Main technical between HTR and HTR+

- **change Software:** from proprietary to TensorFlow
- **change Hardware:** from CPU-training to GPU-training
- **change Neural Network:** deeper and larger
- **change Proprocess:** simpler and faster
Technical Comparison HTR vs. HTR+

First: What stays unchanged?

- **input**: image
- **output**: text and ConfMat
- trained with Connectionist Temporal Classification (CTC)

⇒ The existing workflow stays untouched

Main technical between HTR and HTR+

- **change Software**: from proprietary to TensorFlow
- **change Hardware**: from CPU-training to GPU-training
- **change Neural Network**: deeper and larger
- **change Proprocess**: simpler and faster
The Architecture

Input-Output-Mapping by HTR+

Figure: An example of the input (image) and output (ConfMat)

Hierarchical approach

Figure: The layers are stacked to guarantee a meaningful step-wise mapping between input and output.
## Performance Comparison HTR vs. HTR+

### Edelfeld (642 pages)

<table>
<thead>
<tr>
<th></th>
<th>CER train</th>
<th>CER val</th>
<th>training time</th>
<th>training lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTR</td>
<td>16.83%</td>
<td>21.11%</td>
<td>20h 57m</td>
<td>200,000</td>
</tr>
<tr>
<td>HTR+</td>
<td>4.09%</td>
<td>9.86%</td>
<td>5h 41m</td>
<td>1,638,400</td>
</tr>
</tbody>
</table>

### Diagrams

- **HTR**
  - Graph showing performance over epochs.

- **HTR+**
  - Graph showing performance over epochs.
### ABP OA Protokolle (48 pages)

<table>
<thead>
<tr>
<th></th>
<th>CER train</th>
<th>CER val</th>
<th>training time</th>
<th>training lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTR</td>
<td>16.59%</td>
<td>19.84%</td>
<td>9h 54m</td>
<td>400,000</td>
</tr>
<tr>
<td>HTR+</td>
<td>0.72%</td>
<td>6.35%</td>
<td>3h 59m</td>
<td>1,638,400</td>
</tr>
</tbody>
</table>

**Notes:**
- HTR: HTR
- HTR+: HTR+
- CER: Character Error Rate
- OA Protokolle: Original Accuracy Protokolle
- Training time and lines data may vary due to different methods and time constraints.
Performance Comparison HTR vs. HTR+

Konzilsprotokolle (2135 pages)

<table>
<thead>
<tr>
<th></th>
<th>CER train</th>
<th>CER val</th>
<th>training time</th>
<th>training lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTR</td>
<td>10.51%</td>
<td>10.08%</td>
<td>20h 06m</td>
<td>200,000</td>
</tr>
<tr>
<td>HTR+</td>
<td>2.30%</td>
<td>2.62%</td>
<td>8h 44m</td>
<td>1,638,400</td>
</tr>
</tbody>
</table>
Conclusion

- training speed acceleration by factor 10 to 100
- relative CER-reduction by 50% to 75%
Best Practice

GT quality

- The more GT the better
- **The better GT the better**

⇒ define rules, how to transcribe text
Same transcription for same character

\[ \text{sehen, seineswegs} \]

\[ \text{fehen, seineswegs} \]
Best Practice

Transcribe diplomatic, not modernized

Conclusion Wird dem Wallfahrtsprecher Engelbert

⇔

Concl. Wird dem Wallfahrtspr. Engelbert
Best Practice

Do not train with wrong transcripts

- set **Status = Ground Truth** of transcibed pages and use checkbox **use Groundtruth versions** in training setup

- check that the latest version of your transcription is **NOT an HTR result**
Best Practice

When HTR has bad performance

Sometimes HTR is good, but polygon creation fails

but...

We will improve that in the next month!
Best Practice

When HTR has bad performance

Sometimes HTR is good, but polygon creation fails

but...

We will improve that in the next month!
Thank you!

Questions?