Automated rating of Multiple Sclerosis test results using a convolutional neural network

Martin Birchmeier and Tobias Studer
The project
Multiple Sclerosis (MS)

Demyelinating disease

▶ Insulating cover of nerve cells in the brain and spinal cords are damaged

Symptoms

▶ Double vision
▶ Muscle weakness
▶ Coordination problems
▶ Cognitive deficits
  ▶ Processing speed slower
  ▶ Attention function impaired
  ▶ Learning / memory function inferior
  ▶ Executive functions impaired
Brief Visuospatial Memory Test Revised (BVMT-R) I

Assessment to observe the memory function of MS patients

**BVMT-R**
- Patient inspect 2 x 3 geometric figures
- 10 seconds time to memorize the figures
- Draw the figures from memory in correct shapes and on the correct position
Brief Visuospatial Memory Test Revised (BVMT-R) II

Test execution
### Brief Visuospatial Memory Test Revised (BVMT-R) IV

<table>
<thead>
<tr>
<th>Score</th>
<th>Precision</th>
<th>Placement</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Precisely drawn</td>
<td>Correctly placed</td>
</tr>
<tr>
<td>1</td>
<td>Precisely drawn</td>
<td>Not placed correctly</td>
</tr>
<tr>
<td>1</td>
<td>Not drawn correctly, but recognizable as a given figure</td>
<td>Correctly placed</td>
</tr>
<tr>
<td>0</td>
<td>Missing or not drawn correctly and not recognizable as a given figure</td>
<td>Not placed correctly</td>
</tr>
</tbody>
</table>

**Full Credit Example**

![Full Credit Example](image)

**Partial Credit Examples**

![Partial Credit Examples](image)
From paper to machine learning result

Label 1: (99%)
Label 2: (<1%)
Label 0: (<1%)
Implementation of Machine Learning
100% Raw data with label

Convolutional Neural Network

Training experiment

Azure

Browser Drag & Drop
ML Workflow Designer
(Custom Vision)
Live demo
Live-Demo I

Procedure

► Fill in your PatID

► Memorize the six figures (10sec)

► Draw the six figures

► Upload the drawn figures
Live-Demo II

Procedure

- Rate the figures through machine learning
- Present every figure with its proposed probability (Label 0,1 or 2)
Evaluation of machine learning
Evaluation of machine learning I

Accuracy of the rating of the algorithm (Average of all 6 algorithms)

- 69% conformity of drawings with physician’s assessment 0
- 57% conformity of drawings with physician’s assessment 1
- 76% conformity of drawings with physician’s assessment 2
Insights about algorithms

▶ If the physician assigns labels of 1 or 2, the algorithm will in very few cases assign a score of 0.

▶ The difference between the scores 1 and 2 still causes problems.
Evaluation of machine learning III

Possible reasons for deviations in comparison to a physician

- Only 135 training examples to train the algorithm
- Better and more precise results in case of more training and test data
- Different hit rates for the individual figures
  - Not the same amount of training data per algorithm
- Physician’s evaluation considered as given
- Certain subjectivity
  - Daily constitution
  - Algorithms are trained with wrong training data
    - Be better than a physician is not possible
- Not good enough for full automatism but good enough for support the physician
Failure example for expected 0, but 2 given

Original figure  

Machine learning rated with 2
Outlook
Outlook

Digitalizing the BVMT-R
- Full workflow implementation

Study
- Compare paper based BVMT-R results with the digital based BVMT-R results

Improve the machine learning algorithm
- With more training data

Preclassification
- Implement the digital workflow in the clinical environment to support the physician
Thank you for your attention!

Many thanks to our project partners and people who supported us:
Prof. Serge Bignens
Prof. Dr. med. Andreas Lutterotti, University Hospital Zurich
Team I4MI
## Evaluation of individual algorithms

<table>
<thead>
<tr>
<th>Number</th>
<th>Figure</th>
<th>Rating 0</th>
<th>Rating 1</th>
<th>Rating 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image1.png" alt="Figure 1" /> ((n=26, m=101))</td>
<td>0% ((n=2, m=5))</td>
<td>67% ((n=6, m=23))</td>
<td>83% ((n=18, m=73))</td>
</tr>
<tr>
<td>2</td>
<td><img src="image2.png" alt="Figure 2" /> ((n=25, m=102))</td>
<td>67% ((n=6, m=24))</td>
<td>91% ((n=11, m=45))</td>
<td>63% ((n=8, m=33))</td>
</tr>
<tr>
<td>3</td>
<td><img src="image3.png" alt="Figure 3" /> ((n=26, m=104))</td>
<td>67% ((n=4, m=26))</td>
<td>63% ((n=10, m=31))</td>
<td>67% ((n=12, m=47))</td>
</tr>
<tr>
<td>4</td>
<td><img src="image4.png" alt="Figure 4" /> ((n=26, m=102))</td>
<td>100% ((n=10, m=38))</td>
<td>50% ((n=4, m=18))</td>
<td>67% ((n=12, m=46))</td>
</tr>
<tr>
<td>5</td>
<td><img src="image5.png" alt="Figure 5" /> ((n=26, m=107))</td>
<td>88% ((n=8, m=34))</td>
<td>50% ((n=6, m=23))</td>
<td>92% ((n=12, m=50))</td>
</tr>
<tr>
<td>6</td>
<td><img src="image6.png" alt="Figure 6" /> ((n=26, m=108))</td>
<td>93% ((n=14, m=56))</td>
<td>20% ((n=5, m=22))</td>
<td>86% ((n=7, m=30))</td>
</tr>
</tbody>
</table>

### Average

| \(n, m = \text{total}\) | 69% \((n=44, m=183)\) | 57% \((n=42, m=162)\) | 76% \((n=69, m=279)\) |
Link to Azure Custom Vision