

Handwriting Recognition for Dyslexics

Student

Conradin Kleinstein

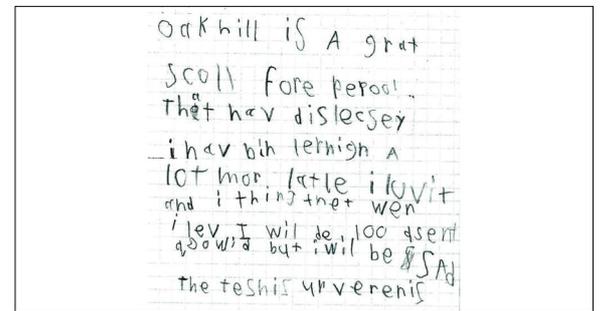
Introduction: This project thesis is about the creation of a prototype for a handwritten text recognition for dyslexic handwriting. The often chaotic and erroneous handwriting makes the recognition task challenging.

Approach: To achieve this goal the following actions were taken. At first a literary analysis was conducted which built the basis to choose a dataset and a recognition model. The chosen model was trained on the data and then evaluated. The next step consisted of processing of input text images where the secluded characters were separated and tested with the recognition model.

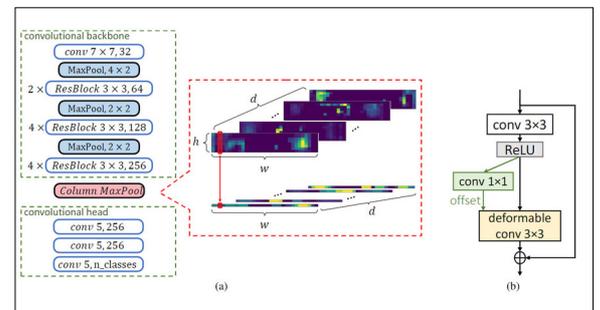
Conclusion: The chosen recognition method achieved an accuracy of 91.52% and is language independent due to working on the character-level and not on the word-level. It can distinguish between uppercase and lowercase letters and digits. The chosen recognition model performs best on uppercase letters and worse on the lowercase letter and numbers. Common mistakes include the confusion between similar upper and lowercase letters and similar digits. The preprocessing of the input data could be conducted on a simplified test text, but failed when confronted with realistic text images where a grid lines were present. The attempted removal of the grid lines with morphological and frequency filtering failed. The simple approach was able to segment all characters and yields a case sensitive character error rate of 29.76% and a word error rate of 75%. The case insensitive error rate is at 5.95% and has a word error rate of 0%, which confirms that the majority of errors is the confusion between upper and lower case characters. This project thesis proposes a prototype for handwritten text recognition for dyslexic handwritings. Due to time constrictions the prototype

could not be adapted to dyslexic handwriting. The main problem is the image preprocessing in which it fails to successfully segment the handwritten text into the individual characters. The use of transformer networks structure and a German dataset to update this prototype is recommended.

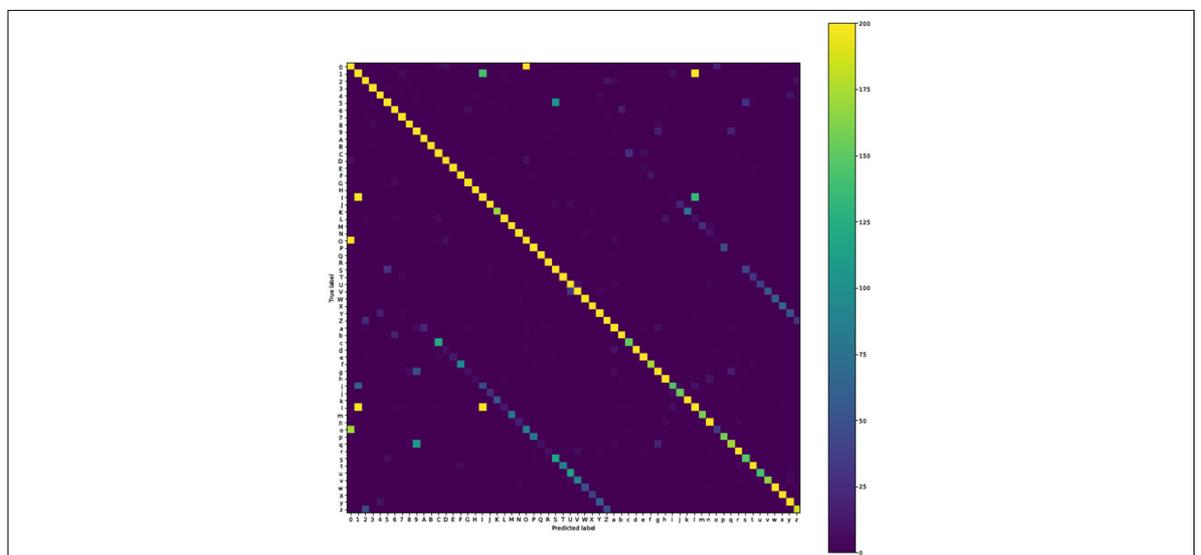
Example Text of Dyslexic Handwriting
www.internationalschoolparent.com



Adapted Deformable Convolutional Network
G. Retsinas, Deformation-invariant networks for HTR



Confusion Matrix of the Recognition Model
Own presentment



Advisor
Hannes Badertscher

Subject Area
Data Science

Project Partner
toz GmbH, Frauenfeld,
TG