Error-corrected deep learning approach to handwritten text recognition of Gregg shorthand

Alexander Weimer
alexander.weimer@a0.ax

under the direction of Minnetonka Research
18301 MN-7, Minnetonka, MN 55345

May 20 2025

Shorthand, also known as pen stenography, is a family of writing systems for English and other languages that emerged out of a need for a fast and efficient writing system in a pre-digital age. Of the many English shorthand systems, Gregg shorthand is the most prevalent (Zhai et al., 2018). While largely made obsolete by general-purpose computers, the cultural and legal value within old shorthand documents means that being able to efficiently scan shorthand documents into modern computer systems holds significant value. This investigation explored the implementation of a model built around a Gated Convolutional Neural network for purposes of handwritten text recognition of Gregg shorthand. An accuracy of 0.04 was achieved after minimal training. The finalized model is freely licensed and made available online for public access.

Keywords: *handwritten text recognition, pen stenography, shorthand*

Introduction

Shorthand, also known as pen stenography, is a family of writing systems for English and other languages that emerged out of a need for a fast and efficient writing system in a pre-digital age.

Of the many English shorthand systems, Gregg shorthand—first developed in 1888—is the most prevalent (Zhai et al., 2018). With the advent of digital text input and storage, shorthand has largely fallen out of use in favor of standard typing and digital stenography (Rajasekaran & Ramar, 2012).

Optical Character Recognition (OCR) is the transliteration of (often handwritten) text samples from photographs or scans of physical material into digitally encoded text.

Handwritten text recognition (HTR) is the transliteration of full handwritten words or sentences, from both digital (i.e. touchscreen or stylus) and physical (i.e. pen and paper, captured through photographs or scans) input methods into digitally encoded text.

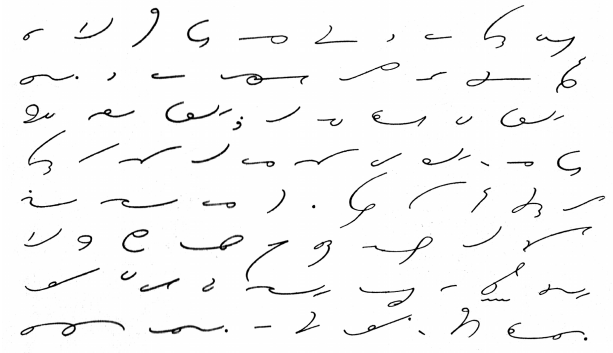

Figure 1: Gregg shorthand sample (Gregg, 2001)

The transliteration of short-hand scripts into regular text presents a unique challenge for several reasons. As can be seen in the Gregg shorthand text sample provided in Figure 1, Shorthand characters often lack distinct features, sometimes varying only in length or degree of curvature. Furthermore, shorthand lexicons are often simplified, often missing vowels or other defining features of words. While the human mind can accommodate for these kinds of omissions, creating a digital system that can do the same poses a challenge.

Reading of manuscripts written in shorthand is vital to understanding documents in a wide variety of fields where time-efficient handwriting was necessary in the past, such as law and medicine. The digitization and thereby preservation of shorthand documents, therefore, presents possible benefits in preservation of history and culture. Further, the development of extensible HTR and OCR systems, in this case with a focus on English shorthand, feasibly opens avenues for the creation of HTR and OCR systems for other written languages—thus presenting possible benefits for the preservation of world languages and cultures.

Digital applications of Gregg shorthand (2) have been investigated in the literature as far back as 1990 by Agarwal1 (Agarwal, 1990). In 2004, HTR of Pitman shorthand—a script related to Gregg—was investigated as a means of rapid text entry into mobile devices by Higgins, who also presents important data on the feasibility of transliteration based on the simplified lexicon of many shorthand systems (Higgins, 2004). More recently, publications by Rajasekaran and Ramar in 2012 as well as Zhai et al. in 2018 have focused on machine learning approaches to HTR of Gregg shorthand (Rajasekaran & Ramar, 2012) (Zhai et al., 2018). Zhai et al. release a dataset Gregg-1916, consisting of 16,000 common English words written in Gregg shorthand as part of their investigation (Zhai et al., 2018).

In June 2024, Heil and Nauwerck present a deep-learning approach to HTR of the Swedish-language Melin shorthand with fairly impressive results. As part of their publication, they include the LION dataset of shorthand manuscripts written in the Melin system (Heil & Nauwerck, 2024). Recent shorthand recognition models still struggle with accurately identifying entire words correctly, with error rates near or above the 50% mark (Heil & Nauwerck, 2024). This presents an opportunity to develop shorthand text recognition models further. Connecting Heil and Nauwerck's modern deep learning

approach to shorthand HTR in the Swedish Melin system with past research into HTR of the English-language Gregg system, as well as other modern deep learning techniques would enable broader, more effective HTR within an English-language context. This would be beneficial for preservation, understanding and accessibility of historical texts and records written in shorthand.

Methods

A Gregg-shorthand handwritten-text-recognition machine learning model was developed using the `PyTorch` deep learning framework, itself based on the `torch` ML library.

The PyTorch ecosystem was chosen for its advantages in flexibility and extensibility compared to popular alternative models within the deep learning field, most significantly TensorFlow. To allow processing of image data, `torchvision`, a package that extends PyTorch with common image transformations for data augmentation and preprocessing was implemented in the realization of a Gregg shorthand recognition model. PIL, the Python Imaging Library, is one of many other libraries that was used in the training of the model.

To allow for the effective processing of sequentially ordered Gregg shorthand characters, the aforementioned PyTorch model was based on the Gated Recurrent Unit (GRU) neural network architecture. It was trained on the Gregg-1916 dataset introduced by Zhai et. al. in 2018, comprising greater than 16,000 Gregg shorthand words.

The `matplotlib` package was used to create graphical representations of accuracy data.

Results

The results of the investigation were generally promising.

An accuracy of 0.04 was achieved after training the model under certain conditions as described. The model was trained on a limited subset of the available data provided by Zhai et. al. via the following method.

A sample of a quarter of the available image-caption pairs was first selected. This sample was then further split into training, validation and testing subsets, yielding a training dataset of roughly 3,000 image-caption pairs. This smaller size provides benefits in the form of reduced training time, but reduces the maximally feasible model performance.

The model was trained on this dataset for 10 epochs, corresponding to roughly 24 hours of compute time.

The finalized model is made available at `github.com/a0a7/GreggRecognition` and licensed under the permissive MIT license.

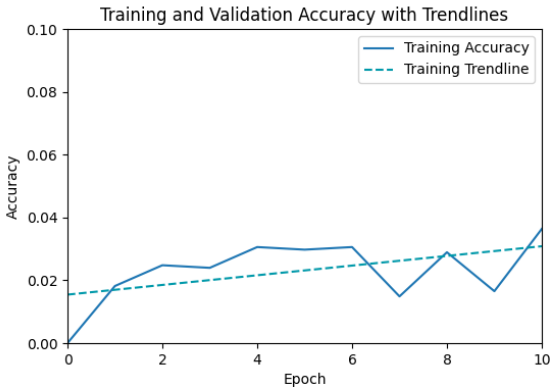

Figure 2: Training and Validation Accuracy

Figure 2 shows the progression in accuracy of the training and validation stages of the training process.

Discussion

Results could be stronger but are not trivial, especially given the very limited quantity of data used as well as the limited duration of training.

Training on a more complete dataset, maybe even larger than the one created by Zhai et al. would mean much better results with the model, as would training for a larger number of epochs. An epoch already taking several hours, this would require either significant model optimizations or significantly more hardware resources. Future research could also explore the implementation of different advanced machine learning techniques, especially more emergent tools.

References

Agarwal, A. (1990, January). Off-line Shorthand Recognition System [[Online; accessed 9. Sep. 2024]]. https://www.academia.edu/10139084/Off_line_Shorthand_Recognition_System

Gregg, J. R. (2001). Gregg Shorthand A Light-Line Phonography for the Million. *The Gregg Publishing Company*. https://www.amazon.com/Gregg-Shorthand-Light-Line-Phonography-Million/dp/B001V0MUP6

Heil, R., & Nauwerck, M. (2024). Handwritten stenography recognition and the LION dataset. *IJDAR*, 1–16. https://doi.org/10.1007/s10032-024-00479-6

Higgins, C. (2004). Knowledge based transcription of handwritten Pitman's Shorthand using word frequency and context. *... on Development and ...* https://www.academia.edu/12601630/Knowledge_based_transcription_of_handwritten_Pitmans_Shorthand_using_word_frequency_and_context

Rajasekaran, R., & Ramar, K. (2012). Handwritten Gregg Shorthand Recognition. *International Journal of Computer Applications*. https://www.semanticscholar.org/paper/Handwritten-Gregg-Shorthand-Recognition-Rajasekaran-Ramar/94a46afa25a456bf11862347225a55c74d52b517?p2df

Zhai, F., Fan, Y., Verma, T., Sinha, R., & Klakow, D. (2018, September). A Dataset and a Novel Neural Approach for Optical Gregg Shorthand Recognition. In *Text, Speech, and Dialogue* (pp. 222–230). Springer. https://doi.org/10.1007/978-3-030-00794-2_24