Forged Signature Recognition using Hotelling's T^2 Analysis

Emily Allen, Alicia Carriquiry, Danica Ommen

Iowa State University

Center for Statistics and Applications in Forensic Evidence

Introduction

As our world becomes more technologically centered, the increased use of digital signatures will ramp up the demand for digital signature analysis and forgery recognition. Digital signatures are commonly written on signature pads at various places such as banks and pharmacies. In addition to digitally written signatures, a software called MovAlyzeR can be used to collect signatures written on paper from a genuine writer and suspect.

This project aims to develop a statistical approach to differentiate a genuine signature from a forged one. The dataset used was obtained using MovAlyzeR. The MovAlyzeR software extracts dynamic information from a signature as it is written. The signatures are segmented into strokes, and each stroke's attributes are measured and recorded (Fig. 1). The variables collected for each stroke include size, duration, velocity, jerk, and pressure. We used the Hotelling's T^2 statistic as a similarity metric between pairs of signatures. These metrics provide a quantitative means to compare signatures. This analysis helps identify forgery and can supplement court cases that involve the crime of forgery.

Data & Methods

Forgery Types: Two test statistics were calculated separately for two forgery types: disguised with model (DWM) and disguised no model (DNM). The DWM signatures were written by forgers who were provided a model signature of the signature they are forging. The DNM signatures were written by forgers who were only provided a name to forge without a model (Fig. 3).

Test Statistics: To compute one Hotelling's T^2 Test statistic, we compared one questioned signature to five known genuine signatures. The questioned signatures fall into two categories: known forgeries from one of the two forgery groups and known genuine signatures. Ten genuine versus genuine statistics and ten genuine versus forged statistics were computed for 90 subjects. This resulted in 1800 test statistics for the DWM forgery metric and 1800 test statistics for the DNM forgery metric.

Metric: The 900 genuine versus genuine statistics and the 900 genuine versus forged statistics are plotted into two separate histograms that overlap in the center and are shown in Figs. 3 and 4. The Hotelling's similarity metric is shown on the x-axis. The light blue histogram shows the distribution of scores for pairs of genuine signatures, and the pink histogram shows the scores obtained when comparing a genuine and a forged signature. Since the Hotelling's T^2 statistic is a measure of the distance between the two signatures, we observe higher values of the statistic when comparing genuine and forged signatures.

Receiver Operating Characteristic Curve: A receiver operating characteristics curve was used to determine the threshold where sensitivity (true positive) and specificity (true negative) are approximately equal. This threshold is also where the Type I and Type II error rates are equal, meaning the rate of false conviction is equal to the rate of false acquittal (Figs. 6 and 7).

Data & Results

Figure 1

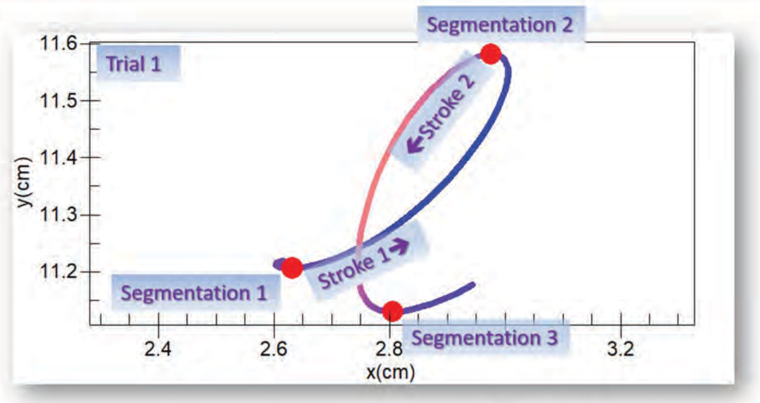
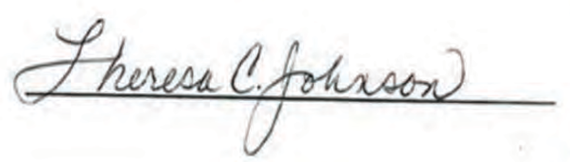


Figure 2

Signature Types

Text



Mixed

Stylized

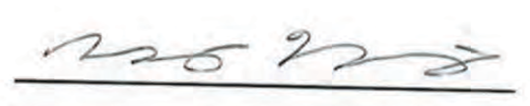
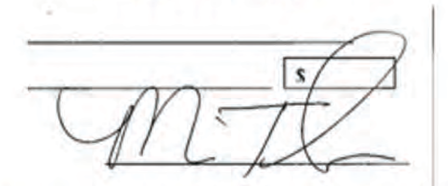


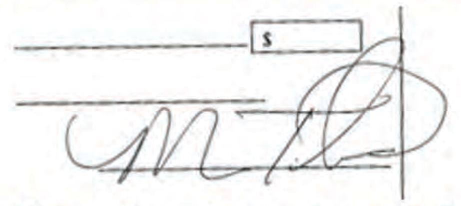
Figure 3

Writer Groups

Genuine



Disguised with Model



Disguised no Model

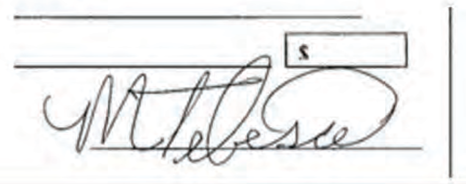


Figure 4

Disguised with Model Forgery Metric

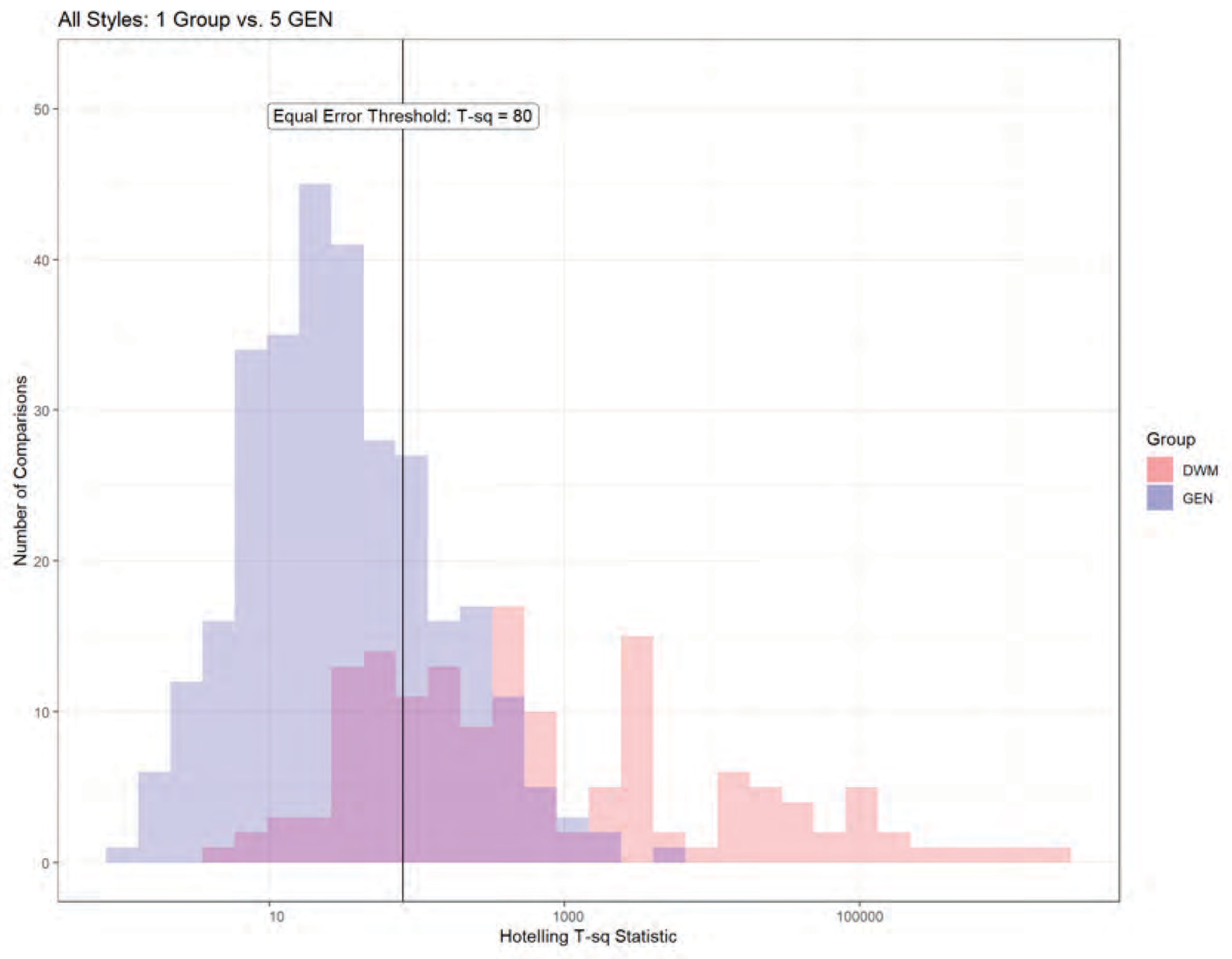


Figure 5

Disguised no Model Forgery Metric

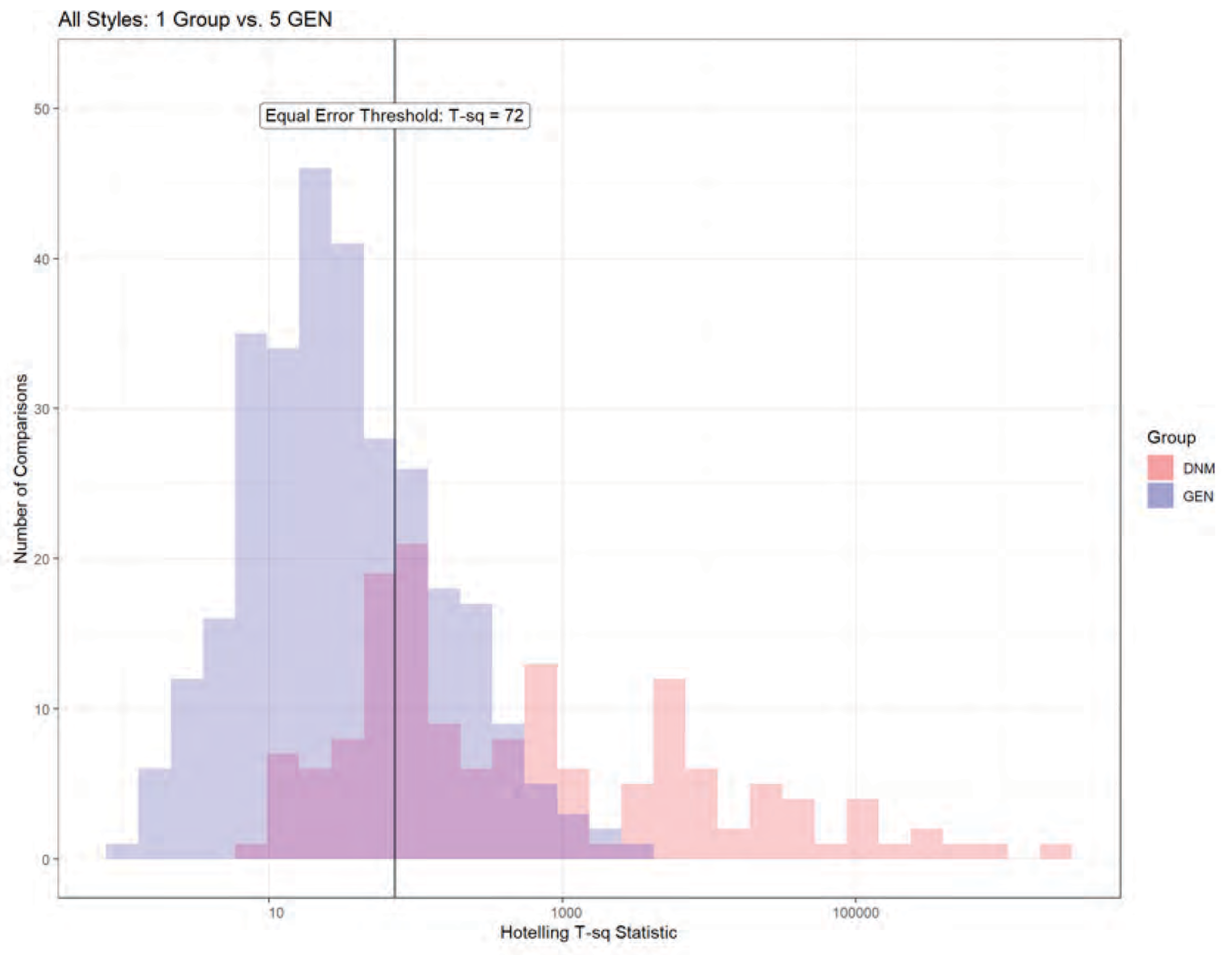


Figure 6

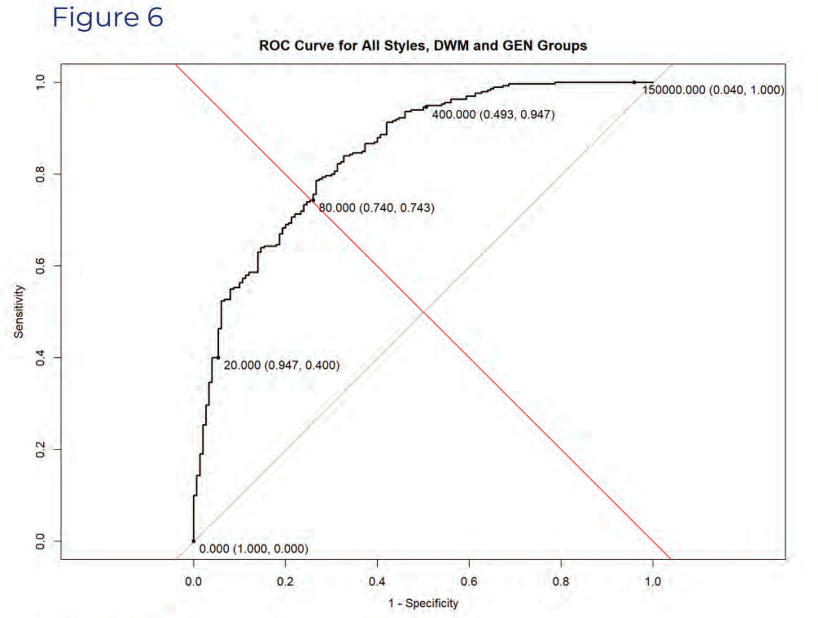
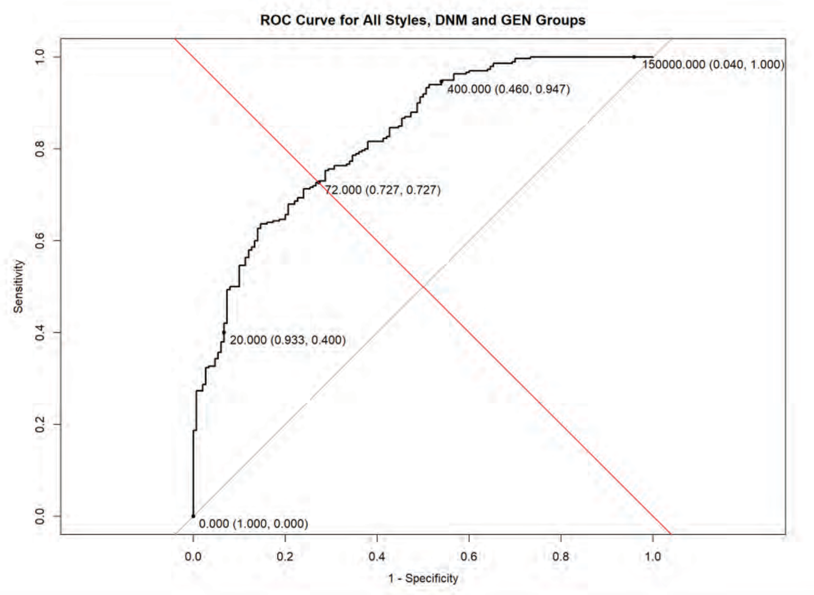


Figure 7



Results & Discussion

Bridging the Gap: This project reimagines the way pattern-based evidence is examined. The MovAlyzeR system allows for the examination of an individual writer's signature pattern by digitally collecting and quantifying features of each stroke.

Figure 1: The data is collected and analyzed for each stroke of writing. A stroke is indicated by a change in direction.

Figure 2: Signatures were categorized into three types: text, mixed, and stylized. Forgery metrics were created for each signature type to provide a more specific metric; however, these metrics were very similar, and we found it most useful to combine all signature types. Using a metric for all signature types prevents the misclassification of signature type.

Figure 3: The data is categorized into writer groups: genuine, disguised with model, and disguised no model. "Model" refers to an image of the genuine signature.

Figure 4-5: As expected, scores shown in Figs. 4 and 5 were higher for pairs including forged and genuine signatures. This was true for the two types of forgery DWM (Fig. 4) and DNM (Fig. 5). Findings are promising, in that the overlap between the two sets of scores in each case is moderate. Thus, the Hotelling's T^2 metric can help discriminate between genuine and forged signatures. The thresholds that result in the smallest classification error are shown in the figures.

Figure 6-7: The ROC curves (Figs. 6 and 7) show how the equal error threshold was selected. The red line indicates where the classification error is equal. Other points are labeled for reference. The area under the ROC curve for the DWM plot is 0.8456, and 0.8256 for the DNM plot.

Future Directions

To continue my research, I will apply a score-based likelihood ratio to the test statistic metric to provide a likelihood of forgery to supplement evidence in court. I also plan to implement other statistical methods to this data such as a random forest. I aim to combine these results with current document analysis to better objectify the field.

Resources & References

Thanks to Dr. Linton Mohammad of Forensic Science Consultants, Inc. for donating signature samples and supplying the data used in this project.

This work was funded (or partially funded) by the Center for Statistics and Applications in Forensic Evidence (CSAFE) through Cooperative Agreements 70NANB15H176 and 70NANB20H019 between NIST and Iowa State University, which includes activities carried out at Carnegie Mellon University, Duke University, University of California Irvine, University of Virginia, West Virginia University, University of Pennsylvania, Swarthmore College and University of Nebraska, Lincoln.