Segregating Offline and Online Handwriting for Conditional Classification Analysis

L C Wong and W P Loh*

School of Mechanical Engineering, Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Penang, Malaysia.

Corresponding author: meloh@usm.my

Abstract. Offline and online handwriting patterns vary when exposed to external factors. Recent researches on both type handwriting analysis were focused on the image feature extraction, pattern recognition, and classification approach. However, no studies have considered segregating the conditional effects of handwriting patterns on two categories: offline vs. online and normal vs. vibration. Hence, the main goal of the study was to investigate the effects of classifying induced vibration on the offline and online handwriting patterns. There were 25 experimental handwriting samples retrieved under four pre-defined class: FLN, FLV, NLN, and NLV. Extracted data attributes consisted of seven handwriting size metrics and two demographic data. JRip algorithm and Decision Stump algorithm of WEKA tool were employed on tenfold cross-validation mode for handwriting classification and a further segregation by offline and online handwriting. JRip and Decision Stump merely resulted in 54% and 44% classification accuracy respectively. On data categorization between the offline and online handwriting, both algorithms achieved classification accuracies of 56% and 76% respectively. Findings showed that the misclassifications between offline and online were reduced to 5 and 7 instances for JRip and Decision Stump algorithms respectively, whereas the total misclassification between normal and vibration were reduced to 22 and 13 instances.

Keywords. Classification, Handwriting, Offline, Online, Vibration

1. Introduction

Handwriting is the output of the writing process using simple writing tools such as a pen and pencil. Each subject’s handwriting gives an individualistic unique identity. The handwriting feature extraction and recognition have captured diverse attention from forensic science, computer science, psychology, or pattern recognition areas. There are two types of handwriting recognition: offline and online. The offline handwriting recognition involves the interpretation and conversion of handwritten characters from hard copy documents. Online handwriting recognition, whereas, involves digitized representation of a digital pen movement, where the data describes sequential information about a position, velocity, and acceleration of the handwriting. However, both handwriting data patterns are prone to external factors’ impact such as the writing surface, pen or pencil used, and the writing force.

Recent researches on offline handwriting analysis were reported in Liwicki et al. [1], Kumar et al. [2], and Mukherjee and De [3]. Offline handwriting works in Surinta et al. [4] concerned extracting high dimensional feature vector with local gradient feature descriptors for handwritten character recognition. In addition, the authors compared the classification accuracy performance between K-Nearest Neighbour (K-NN) and Support Vector Machine (SVM) classifiers. Meanwhile, Wakahara et al. [5], Mandal et al. [6], and Verma et al. [7] focused on the online handwriting. In Samanta et al. [8], the authors approached online handwriting recognition through the application of von Mises and Gaussian
distribution on circular and linear feature components respectively. The circular, linear and combined features were fed into a fully connected non-homogeneous hidden Markov model (HMM) for classification. Handwriting researches on both offline and online handwritings were considered in Liwicki and Bunke [9], Plamondon and Srihari [10], and Zhang et al. [11]. In Bunke [9], a multiple classifier systems (MCS), which was based on HMM and bidirectional long short-term memory (BLSTM) classifier for both offline and online handwritten text-line recognition. The output sequences of the MCS were combined with Recognizer Output Voting Error Reduction (ROVER) combination strategy. External factor’s effect on handwriting analysis was discussed in Schomaker and Plamondon [12], and Chang et al. [13]. Schomaker and Plamondon [12] studied the relationship between axial pen force and pen-point kinematics. Meanwhile, the effects of hand soreness in handwriting patterns evaluations were also a concern [13]. The soreness was induced through prolonged writing and measured with a visual analogue scale. In general, researches mainly focused on the efficiency of feature extraction and classification, while the studies on the effects of an external factor in the handwriting patterns were relatively scarce.

Hence, this study adds to the body of knowledge by investigating the effects of induced vibration on the offline and online handwriting patterns. The goal of the study was to identify the classification accuracies of offline and online handwriting under the influence of vibration and to improve the classification system through data segregation. Quantitative attributes (letter width, spacing, size, inclined angle, and slant) were extracted to facilitate information analysis. Study data involve offline and online experimental handwritings of 25 participants under the normal and vibration exposure condition. The Rule-based classifier: JRip and Tree-based classifier: Decision Stump algorithms were used for handwriting conditions classification on tenfold cross-validation mode. The pre-defined conditions are offline normal (FLN), offline with vibration (FLV), online normal (NLN), and online with vibration (NLV). Data segregation into offline and online handwriting was performed based on the confusion matrices evaluation.

Section 2 describes the research methodology that includes data collection approach, pre-processing, classification and knowledge discovery. The analysis and findings are discussed in Section 3. Section 4 concludes the main findings from the paper.

2. Methodology

2.1. Experimental data collection
A total of 25 right and left-handed university students (15 males, 10 females, 22 ± 2 years old) were involved on a voluntary basis with signed informed consents. The subjects were required to write the phrase “Sphinx of black quartz, judge my vow” under two writing conditions: normal and vibration exposed, using the Pilot G2 05 gel ink rollerball pen on a sheet of white paper followed by the stylus (S Pen) on Samsung Tablet A.

The experiment was performed on a desk at 0.74 m height (elbow height) for the normal condition. Subsequently, the Mondial Slim Beauty Fitness Massager (100V – 240AC50/60HZ) was strapped on the desk to simulate the vibrational effects of the vibration exposed condition.

Figure 1. The phrase “Sphinx of black quartz, judge my vow” written under vibrational effects in (a) offline and (b) online handwriting.
2.2. Pre-processing

The recorded data were essentially made up of 25 samples of handwritings with 9 attributes and 100 instances through the image to numeric transformation (Table 1). Handwriting size metrics include average letter width (W), average spacing (S), size (SZ1-SZ3), inclined angle (IA), and slant (SL) was extracted as shown in figure 3.

Table 1. Descriptive summary of data attributes.

| Attributes | Description                  | Offline | Online |
|------------|------------------------------|---------|--------|
| W          | Average letter width (cm)    | [0.16 – 0.32] | [0.26 – 0.54] |
| S          | Average spacing width (cm)   | [0.22 – 0.57] | [0.21 – 0.44] |
| SZ1        | Size of “tall” letter (cm)   | [0.30 – 0.78] | [0.53 – 1.20] |
| SZ2        | Size of “short” letter (cm)  | [0.20 – 0.55] | [0.28 – 0.75] |
| SZ3        | Size of “tail” letter (cm)   | [0.40 – 0.91] | [0.65 – 1.55] |
| IA         | Inclined angle (°)           | [-2.00 – 3.00] | [-1.50 – 1.25] |
| SL         | Slanting of writing          | {L, S, R}  | {L, S, R}  |
| H          | Handedness                   | {L, R}     | {L, R}     |
| G          | Gender of subject            | {M, F}     | {M, F}     |

Figure 3. Handwriting size metrics ($w_i, s_i, IA, SZ1-SZ3$) features.

\[
W = \frac{\sum_{i=1}^{n} w_i}{k}, \quad n = 7 \text{ (total number of words)}
\]

\[
S = \frac{\sum_{i=1}^{N} s_i}{N}, \quad N = 6 \text{ (total number of spacings)}
\]

where: $w_i =$ width word number $i$, and $s_i =$ spacing between word number $i$
The attributes were represented in interval scales, except for SL, handedness, and gender represented in the nominal scale: \{left (L), right (R), and straight (S)\}, \{left (L), and right (R)\}, and \{male (M), and female (F)\}. The W, S, and SZ1-SZ3 were measured with Endo Keiki 6”/15cm Steel Ruler (in cm), whereas, the IA was measured with a protractor (in °). The computation of W and S attributes were based on Equations (1) and (2).

The outlier and extreme value screenings were based on equations (3) and (4). The mean substitution was used to treat the missing values, whereas instances with outliers or extreme values present were removed.

\[
Q_3 + 3 \times IQR < \text{Outliers} \leq Q_3 + 6 \times IQR \quad \text{or} \quad (3)
\]

\[
Q_1 - 6 \times IQR \leq \text{Outliers} < Q_1 - 3 \times IQR
\]

\[
\text{Extreme Values} < Q_1 - 6 \times IQR \quad \text{or} \quad (4)
\]

\[
\text{Extreme Values} > Q_3 + 6 \times IQR
\]

where:

- \(Q_1\) = lower quartile (25% data)
- \(Q_3\) = upper quartile (75% data)
- \(IQR\) = interquartile range

2.3. Data Classification
This process involved recognizing handwriting features extracted into four predefined data classes; FLN, FLV, NLN, and NLV using WEKA tool. JRip algorithm from rule-based classifier and Decision Stump algorithm from tree-based classifier were adopted to recognize the handwriting data under tenfold cross-validation. This training method allows determination of robustness of general models in predicting the classes of new handwriting data. ZeroR algorithm was adopted as the benchmarking algorithm to evaluate the classification performance.

2.4. Knowledge Discovery
At this level, the confusion matrices generated by the JRip and Decision Stump algorithms could be mined by inspecting the misclassification between offline and online, normal and vibration. Based on the confusion matrix analysis, appropriate data segregation can be performed on the handwriting data to improve the classification performance of both algorithms.

3. Results and Discussion
The data contained neither missing values nor outliers or extreme values in the pre-processing stage. Hence, the handwriting data was confirmed clean.

The accuracy of the all conditions handwriting data categorized into four pre-defined classes: FLN, FLV, NLN, and NLV using JRip and Decision Stump algorithms on tenfold cross-validation were 54% and 44% respectively. Both classification accuracies were found reliable (above 20% accuracy found with ZeroR as a benchmark) as shown in figure 4. The confusion matrices generated from both algorithms are shown in table 2.
Figure 4. Classification accuracy on all conditions, offline and online, offline only, and online only handwriting with JRip and Decision Stump algorithm.

Table 2. Confusion matrix showing the number of instances classified into FLN, FLV, NLN and NLV on JRip / Decision Stump classification.

|          | FLN   | FLV   | NLN   | NLV   | Classified as |
|----------|-------|-------|-------|-------|---------------|
|          | 17/21 | 6*/2* | 1/1   | 1*/1* | FLN           |
|          | 11*/22* | 10/1 | 2*/1* | 2*/1 | FLV           |
|          | 4*/4  | 2*/0  | 7/9   | 12*/12* | NLN         |
|          | 1*/1* | 1/0   | 3*/11* | 20/13 | NLV          |

# correctly classified instances
#* misclassified between normal and vibration
# misclassified between offline and online

Based on table 2, the number of misclassified instances were located outside the matrix diagonal. It can be observed that there were 14 instances misclassified between offline and online, whereas 38 instances were misclassified between normal and vibration condition, amounting to 46 misclassified instances for the JRip algorithm, where there were 6 instances misclassified between all four conditions. As for the Decision Stump algorithm, the misclassifications between offline and online, and between normal and vibration conditions were 9 and 50 instances respectively, amounting to 56 misclassified instances, with 3 instances misclassified between all four conditions.

Findings show that there were fewer instances misclassified between offline and online handwriting as compared to between normal and vibration conditions. This led to the further refining classification analysis by categorizing handwriting data into two pre-defined conditional classes: offline and online. The results obtained as depicted in figure 4 showed that the JRip and Decision Stump algorithms were able to achieve the percentage classification accuracy of 95% and 93% respectively. The results showed that the JRip algorithm performed better than Decision Stump in classifying between offline and online handwriting data. Based on table 3, the number of misclassification instances between offline and online for both algorithm was reduced from 14 and 9 instances to 5 and 7 instances respectively on JRip and Decision Stump.
Table 3. Confusion matrix showing the number of instances classified into offline (FL) and online (NL) on JRip / Decision Stump classification.

|        | FL      | NL      |
|--------|---------|---------|
| 49/46  | 1/4     | FL      |
| 4/3    | 46/47   | NL      |

# correctly classified instances
# misclassified between offline and online

Tenfold cross-validation classification was employed by both JRip and Decision Stump algorithms to categorize the offline and online handwriting data separately into two pre-defined classes: N and V. Classification accuracy of JRip and Decision Stump achieved were 56% for offline handwriting, and 74% for online handwriting. All accuracies were found reliable (> 40% on ZeroR as a benchmark). The hierarchy sequence of the split classifications beginning with the type of handwriting (offline vs. online), followed by the presence of vibration (normal vs. vibration) were summarized in figure 5.

![Hierarchy breakdown of conditional handwriting classification by percentage using JRip and Decision Stump.](image)

Legend: Offline (49/1) = Type of handwriting (# correctly classified instances / # incorrectly classified instances)

Although the classification accuracies by JRip and Decision Stump algorithms were equivalent for the offline handwriting data, JRip algorithm had categorised more normal instances into vibration condition correctly as compared to the Decision Stump (figure 5). Meanwhile, more than half of vibration instances were categorised into normal instances by the Decision Stump algorithm (figure 5). For the online handwriting, both algorithms gave similar misclassification confusion between normal and vibration condition; 12 normal instances were misclassified into vibration, and only one vibration instances were misclassified into normal.

Through the further classification split of handwriting data into offline and online, JRip and Decision Stump algorithms were able to improve the classification accuracy by 2% to 30%. The number of misclassification between normal and vibration were reduced from 38 and 50 instances to 22 and 13 instances respectively on JRip and Decision Stump. This suggests that both algorithms provide better results for conditional classification of offline and online handwriting data.

Conclusion

This study considered experimental offline and online handwriting patterns analysis that were exposed to normal and vibration conditions. A total of nine attributes: seven handwriting (W, S, SZ1, SZ2, SZ3, IA, and SL), and two demographics (gender and handedness) data were extracted from the handwriting. The tenfold cross-validation was employed with JRip and Decision Stump algorithms used to classify the handwriting into four pre-defined conditional classes (FLN, FLV, NLN, and NLV). The classification accuracy of both algorithms though low but was proven reliable (above ZeroR benchmark). An examination of instances classified into the confusion matrix analysis from the JRip
and Decision Stump algorithms gave hint that the offline and online were better distinguished compared to between normal and vibration conditions.

Hence, handwriting data were segregated into offline and online to exhibit improved handwriting classification into N and V classes. In particular, both types of handwriting under vibrational stress was tested and proven better for conditional classification. The approach of conditional classification reveals that through segregation of handwriting data between offline and online, the percentage classification accuracy was improved. Future works could be extended to include compatible public handwriting case study data to validate the findings. Other handwriting segregation such as between gender or handedness can be adapted as added factors into the handwriting conditions for analysis.

References

[1] M Liwicki, S Ebert and A Dengel 2014 Bridging the Gap between Handwriting Recognition and Knowledge Management Pattern Recognit. Lett. Vol 35 No 1 pp 204–213
[2] G Kumar, P K Bhatia and Indu 2013 Analytical Review of Preprocessing Techniques for Offline Handwritten Character Recognition Vol 3 No 3
[3] S Mukherjee and I De 2017 Feature Extraction from Handwritten Documents for Personality Analysis Int. Conf. Comput. Electr. Commun. Eng. ICCECE 2016
[4] O Surinta, M F Karaaba, L R B Schomaker and M A Wiering 2015 Recognition of Handwritten Characters using Local Gradient Feature Descriptors Eng. Appl. Artif. Intell. Vol 45 pp 405–414
[5] T Wakahara, H Murase and K. Odaka 1992 On-line handwriting recognition Proc. IEEE Vol 80 No 7 pp 1181–1194
[6] S Mandal, S R M Prasanna and S Sundaram 2018 GMM Posterior Features for Improving Online Handwriting Recognition Expert Syst. Appl. Vol 97 pp 421–433
[7] B Verma, J Lu, M Ghosh and R Ghosh 2004 A Feature Extraction Technique for Online Handwriting Recognition Neural Networks 2004 Proceedings 2004 IEEE Int. Jt. Conf. Vol 2 pp 1337–1341
[8] O Samanta, U Bhattacharya and S K Parui 2014 Smoothing of HMM Parameters for Efficient Recognition of Online Handwriting Pattern Recognit. Vol 47 No 11 pp 3614–3629
[9] M Liwicki and H Bunke 2009 Combining Diverse On-line and Off-line Systems for Handwritten Text Line Recognition Pattern Recognit. Vol 42 No 12 pp 3254–3263
[10] R Plamondon and S N Srihari 2000 On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey IEEE Trans. Pattern Anal. Mach. Intell. Vol 22 No 1 pp 63–84
[11] X Y Zhang, Y Bengio and C L Liu 2017 Online and Offline Handwritten Chinese Character Recognition: A Comprehensive Study and New Benchmark Pattern Recognit. Vol 61 pp 348-360
[12] L R B Schomaker and R Plamondon 1990 The Relation between Pen Force and Pen-point Kinematics in Handwriting Biol. Cybern. Vol 63 No 4 pp 277–289
[13] S H Chang, C L Chen and N Y Yu 2015 Biomechanical Analyses of Prolonged Handwriting in Subjects with and without Perceived Discomfort Hum. Mov. Sci. Vol 43 No 8 pp 1–8