

Is There a Difference between Paper and Electronic Chinese Signatures?

Ji-Feng Luo, Yun-Zhu Pu, Jie-Yang Yin, Xiaohong Liu, Tao Tan, Yudong Zhang,*
and Menghan Hu*

The purpose of this study is to investigate whether there are differences in handwritten Chinese signatures on different media including paper and electronic devices. Participants were asked to sign specified names on various types of media and the signatures were scanned or saved digitally for subsequent analysis. In this study, using convolutional neural networks and Siamese neural networks as classifiers and comparators, the performance plunge is revealed and thus considerable dissimilarity between the signatures on different media is implied. To further explore this, cubic Bézier curves are fitted to the signatures using the least square method for quantitative statistical analysis. By analyzing the visual changes in the morphology of strokes, several features of signatures are selected and computed, and the paired *t*-test and the Wilcoxon signed-rank test are implemented, which provides a deeper substantiation and explanation of the findings.

also a popular biometric hallmark used in various areas, such as legal documents, forensic sciences, and banking validation.^[1,3]

Traditionally, signatures are written on paper and verified by signature identification specialists to identify the genuine signatory. With the rapid development of information technology, signatures can now be signed directly on touch panels. The increasing usage of the internet and social demands in many fields, including electronic commerce and electronic banking systems, has driven the growth of this usage. Nevertheless, these new input media raise the issue of the similarity between signatures from these media and those from paper: when switching to an electronic system, though existing stud-

1. Introduction

Biometric features are unique and nontransferable, making them intrinsic properties to every individual. As a result, they are widely used for identification and authentication purposes. An individual's handwriting signature, although controversial for susceptibility to be forged,^[1,2] is a widely used method for verifying their claimed identity, based on the proof provided by their previously registered signature. Therefore, signatures are

ies have shown that it is credible to establish a robust verification system on electronic media, the potential dissimilarity from those on paper may indicate the necessity to recollect signatures on these media. Otherwise, if they are similar, the system manager could scan the previously collected signatures on paper.

Advancements in technology have made it practical to automatically verify biometric features, which undoubtedly maintain its position in modern identity authentication. Signature recognition by machine can be generally classified into two types:

J.-F. Luo, M. Hu
Shanghai Key Laboratory of Multidimensional Information Processing
School of Communication and Electronic Engineering
East China Normal University
Shanghai 200241, China
E-mail: mhhu@ce.ecnu.edu.cn

J.-F. Luo
School of Mathematical Sciences
East China Normal University
Shanghai 200241, China


Y.-Z. Pu, J.-Y. Yin
School of Software Engineering
East China Normal University
Shanghai 200062, China

X. Liu
John Hopcroft Center for Computer Science
Shanghai Jiao Tong University
Shanghai 200240, China

T. Tan
Faculty of Applied Sciences
Macao Polytechnic University
Macao 999078, China

Y. Zhang
School of Computing and Mathematical Sciences
University of Leicester
Leicester LE1 7RH, UK
E-mail: yudongzhang@ieee.org

M. Hu
Key Laboratory of Artificial Intelligence
Ministry of Education
Shanghai 200240, China

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/aisy.202300439>.

© 2023 The Authors. Advanced Intelligent Systems published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/aisy.202300439

online recognition and offline recognition. Online recognition captures dynamic signature parameters such as speed, pressure, and stroke order, whereas offline recognition analyzes only the image of the signature.^[1,2] Marcus Liwicki et al. conducted a machine signature verification competition on Dutch and Chinese signatures, collected in dataset SigComp2011 and containing online and offline ones and their skilled forgeries and compared submitted 9 systems' performance.^[4] Although offline signature recognition is a challenging task, as some forgeries can even look more similar to the original signature than other genuine ones,^[2] and one's signature may vary due to illness, age, and psychological reasons,^[5] it continues to be widely used owing to the cost-effectiveness of not requiring specialized equipment and software procurement. Cui Wencheng et al. applied AlexNet, a classical convolutional neural network (CNN), to offline Chinese signature authentication where the signatures are from 20 volunteers and SigComp2011, and arrived at an accuracy rate of 87.5% and an error rate of 12.5%.^[6] Ping Wei et al. proposed an inverse discriminative network and deployed it on their collected approximately 29 000 Chinese signatures (Challenging Chinese Signature Dataset, CSD) and datasets in other languages including CEDAR, BHSig-B, and BHSig-H, and reached an accuracy rate of 90.17% and an error rate of 10.83%.^[7] Teressa Longjam et al. utilized a CNN to CEDAR and BH-Sig260, multi-language signature datasets, and their study can be applied to multicultural countries such as India.^[8]

Feature selection and extraction are indispensable part of offline signature verification, and in addition to utilizing convolutional kernels for feature extraction, other methods such as morphological techniques, geometric methods, grid analysis, wavelet transforms, curvelet transforms, principle component analysis (PCA), etc., can also serve as feature extraction approaches.^[5] A fusion of existing features is a wiser approach to combine the advantages of different extant features. Manabu Okawa proposed a new feature extraction approach based on PCA over the Fisher vector with fused features on a support vector machine (SVM), which is effective for offline verification and robust under skillfully forged datasets.^[9] Faiza Eba Batoool et al. exploited gray level co-occurrences matrices and geometric features to come up with an SVM for verification, suitable for small sample size.^[10] Muhammad Sharif proposed a strategy using genetic algorithm for feature selection and gained improved performance compared with existing approaches.^[5]

Though extant studies have gained excellent performance and been proved to be robust in many scenarios, they are focused on the signatures from only single input method. Our study is focused on offline signature recognition to uncover the extent of dissimilarity of signatures from paper versus those from electronic media, and thus it features the absence of anti-counterfeiting requirements and the focus solely on studying differences. Because existing datasets lack cross-media signatures, to implement the investigation, we invited 100 participants to write specific samples of signatures. For possible doubts on selection of samples and verification systems or any other systematic errors, we established a control group to compare the disparities more clearly. We then preprocessed the data using image processing methods, to filter out extraneous variables and make it more suitable for machine learning and other further analysis. Then, we employed popular neural network models to demonstrate a significant

performance decrease between signatures on different media. Finally, graphical features were extracted to compare their differences, and a Bézier curve fitting model was established to analyze the morphology of each batch of signatures to provide a deeper explanation for the observed discrepancy.

Our contributions primarily lie in 1) collected signatures from both paper and electronic media, 2) applied state-of-the-art neural network models to demonstrate the differences, 3) established a method of representing the morphology of Chinese characters using Bézier curves, and 4) statistically proved the differences between signatures from paper and those from other electronic media, and found out the features contributing to the dissimilarity.

The flow of data process and result overview are shown in **Figure 1**.

2. Data Collection

We have searched for existing datasets but were unable to find a suitable one for our study. Therefore, we decided to create our own dataset focusing on Chinese signatures using multiple input media. To collect data, we followed these steps: 1) invited 100 volunteers to participate in our study. 2) Randomly selected 10 popular Chinese names, consisting of five 2-character names and five 3-character names, as most names in China are either two or three characters long. 3) Recorded the entire signature signing process and asked each volunteer to sign with official fonts on three different media: paper, iPad using an Apple Pencil with Notability app, and a smartphone with OneNote app held by hand or placed on a desk, as shown in **Figure 2**. To simplify the model, here we have turned off the pressure sensitivity. 4) Scanned the paper signatures and compiled them with the digital signatures collected from the touch screen devices.

We also established a control group consisting of 11 volunteers who wrote the same names on paper at least twice, and the data were processed similarly.

3. Data Preprocess

The purpose of data preprocessing is mainly to filter out unnecessary factors, including background noise, stroke width, and stroke color. Figure 2 demonstrates the whole workflow of data preprocessing.

3.1. Squaring

Machine-learning models always require regular input, featuring the same size for all the images, so we make all the images square first, 224×224 each.

To minimize unnecessary parts and highlight the signature, we cropped the scanned pages of signatures. Then, we made all the cropped pictures with dark strokes and light backgrounds and inserted them into the center of a square background.

3.2. Binarization

Binarization is a commonly used method for filtering out background noise and simplifying subsequent computational

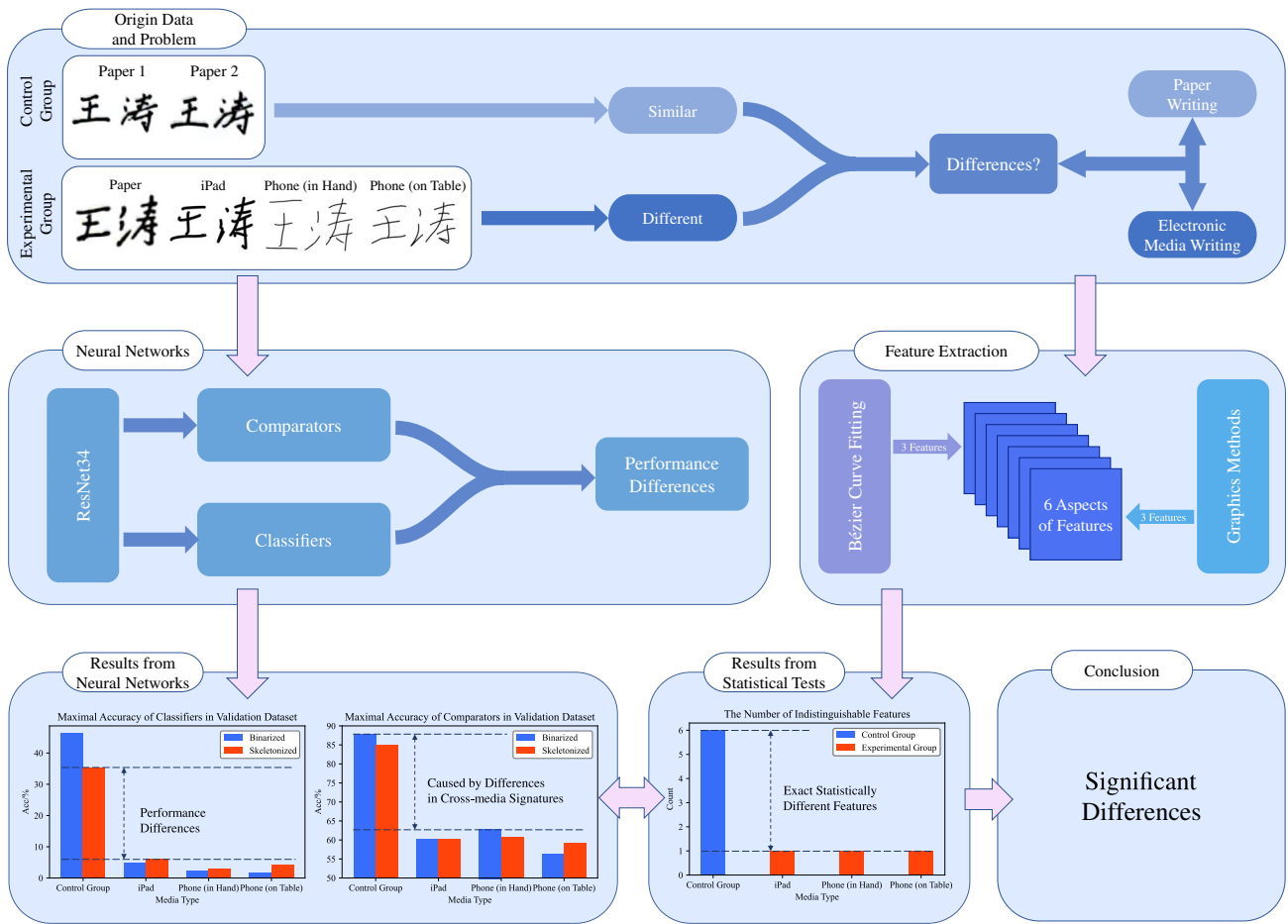


Figure 1. Data process and result overview.

complexity, and more importantly, extracting the skeleton.^[11] It is often necessary to set appropriate binarization thresholds manually, especially for certain inputs. For example, a paper-scanned input requires a lower threshold than an iPad input to clearly distinguish close strokes.

3.3. Skeletonization

Noise may occur due to accidental touch on the screen or the pen tip touching the paper, leaving small dots in the samples. Given that we have disabled the pressure sensor of the stylus, the thickness of all strokes is the same. To avoid these dots, and reduce the influence of extraneous variables from variations in stroke size due to different sources or writing settings, we should adjust the thickness of all strokes to the smallest size through morphology methods, keeping only skeletons of characters.^[11]

3.4. Data Cleansing

Some volunteers chose to write with a pen stroke that was too thick to identify, causing great difficulty in processing, which exceeds the ability of the algorithm applied.

Other volunteers forgot to write certain signatures, and only some of them later supplemented those missing data.

A few volunteers wrote an incorrect character within a signature but only corrected that particular character instead of the entire signature, we shifted the corrected character to its original position of the incorrect character, while attempting to maintain the spacing and balance between the characters as originally.

Finally, we abandoned the data from seven volunteers.

4. Neural Network Model

Neural networks are now more widely used than ever before, especially in the fields of computer vision, voice recognition, and natural language processing. Among these, CNNs are particularly suitable for image processing due to their distinctive convolutional and pooling layers. The convolutional layer extracts features from images through image convolution, improving the precision of image classification and other tasks. The pooling layer compresses data and reduces the dimensionality of the feature maps generated by the convolutional layer, mimicking the human vision system, and thereby reducing computational complexity and preventing overfitting.^[12]

ResNet is a well-known CNN model, originally proposed by Kaiming He and his team at Microsoft Research. They introduced a residual network to connect the layers of ResNet, which effectively prevents the problem of gradient vanishing and

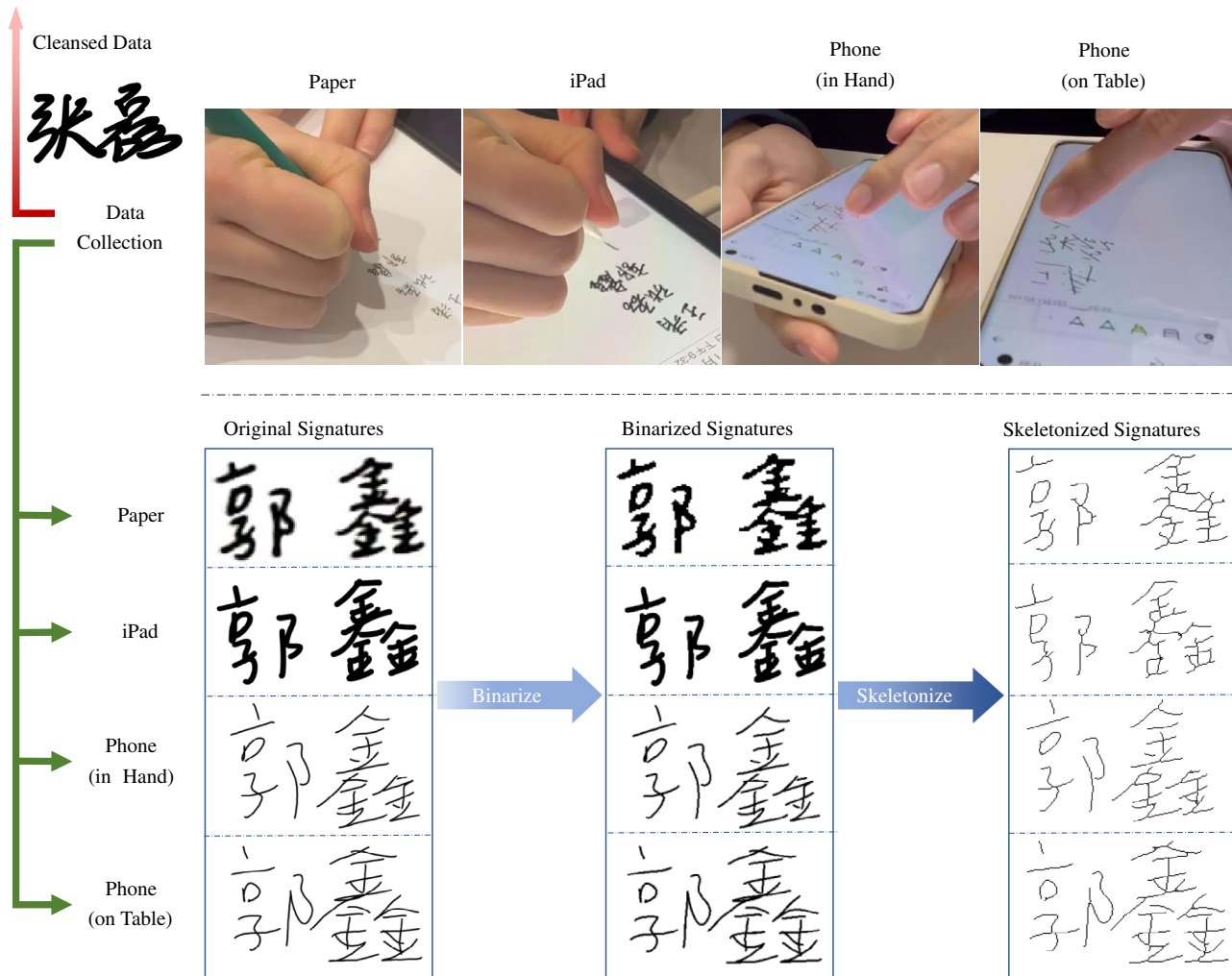


Figure 2. Flow of data collection and preprocess.

exploding. ResNet has fewer parameters, faster training speed, and better image classification performance compared to other network architectures.^[13] ResNet is also a state-of-the-art model for signature recognition.^[1] For this reason, we used the predefined ResNet34 model (a 34-layer ResNet) in PyTorch as a base model, and we also modified some of its structures to better suit our specific tasks.

4.1. Classifier

To use ResNet34 as a multi-label classifier, we modified the fully connected layer output vector dimension to match the number of volunteers, and typically used log-softmax function as the activation function and cross-entropy loss as the loss function.^[12] When using the model for classification, we considered the class with the maximum value of the log-softmax function to be the model's classification result. Figure 3 shows the model.

In the experimental group, we used signatures from paper for training and those from other media for validation. In the control group, pairs of signatures with identical characters were used, with one signature used for training and the other for validation.

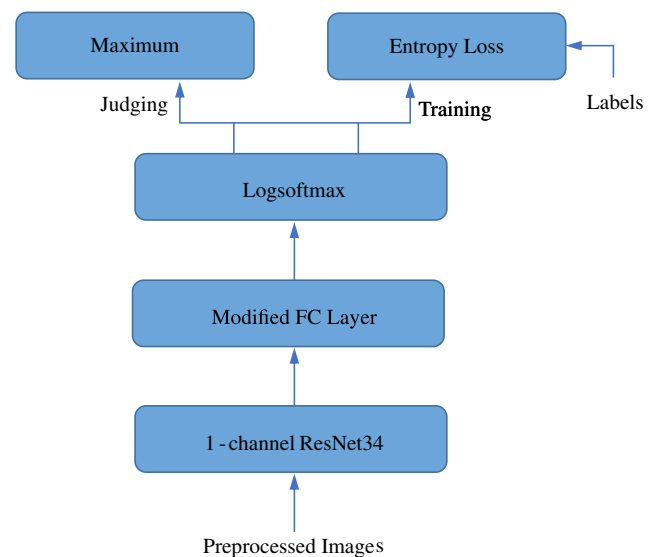


Figure 3. Convolutional neural network classifier.

4.2. Comparator

To implement a comparator, we introduced Siamese neural networks (SNNs). SNNs are composed of two neural networks with shared weights, allowing them to compute the dissimilarity between two similar inputs, and it is extremely suitable for signature recognition.^[2,3]

We modified the fully connected layer of ResNet34 as follows

$$\begin{aligned} Y_1 &= \text{ReLU}(W_1 X + B_1), X \text{ is the input}, Y_1 \in \mathbb{R}^{256} \\ Y_2 &= W_2 Y_1 + B_2, Y_2 \text{ is the output}, Y_2 \in \mathbb{R}^2 \end{aligned} \quad (1)$$

Contrastive loss function is commonly used for SNNs as a metric for image matching including signatures, and its objective is to reduce the distance between similar pairs in feature space while increasing the distance between dissimilar pairs.^[14,15] It is defined as follows, letting the label be 0 if inputs are the same and be 1 if not and setting the margin to 2. **Figure 4** shows the model.

$$\begin{aligned} \text{dis} &= \|Y_1 - Y_2\|_2, p = \max(\text{margin} - \text{dis}, 0) \\ \text{loss} &= (1 - \text{label}) * \text{dis}^2 + \text{label} * p^2, \text{label} \in \{0, 1\} \end{aligned} \quad (2)$$

When comparing signatures using the model, we consider them to be similar if the Euclidean distance between the output pair by the model is below the threshold of 0.9.

For the experimental group, we randomly selected seven signatures out of the ten as the training dataset and used the remaining three as the validation dataset for every volunteer. To compare each pair of signatures signing on two different media, we generated two sets of data pairs, one for a person versus himself and the other for the same versus another person. We shuffled these two sets of pairs and selected all pairs from the first set and an equal number of pairs from the second set. For each data pair, we used the same training and validating signatures. In this operation, we expected the accuracy of a randomly initialized network to be 50%. The same was applied to the control group, except that both signatures were from paper.

5. Stroke Morphology

Regular Chinese characters consist of 36 kinds of strokes, which are classified into 8 basic strokes as shown in **Figure 5**, including

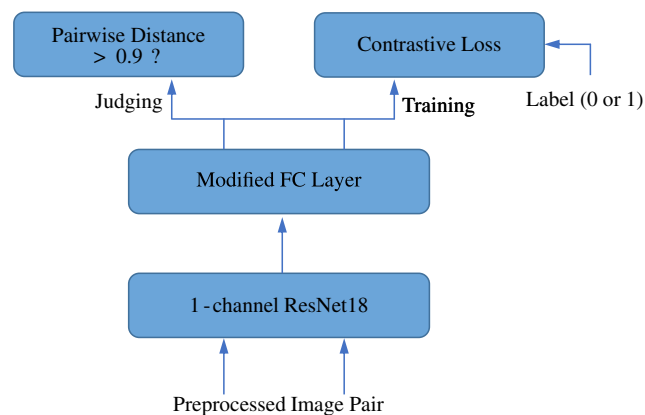


Figure 4. Siamese neural network comparator.

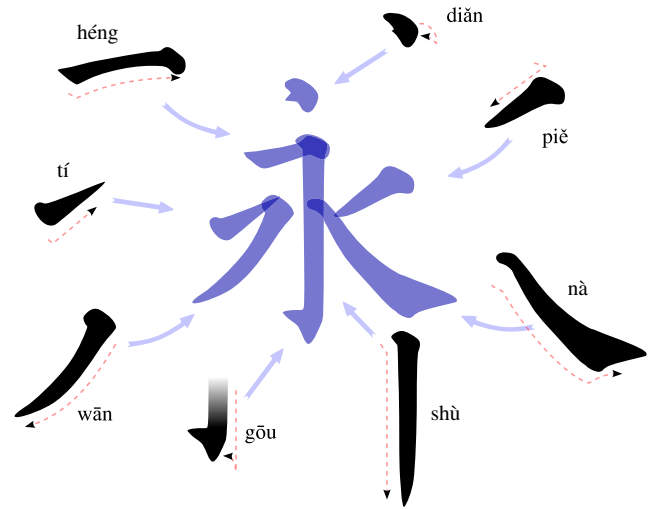


Figure 5. Basic strokes.^[39]

horizontal (héng), vertical (shù), left-falling (piě), right-falling (nà), dot (diǎn), hook (gōu), rising (tí), and turning (wǎn).^[16,17] Our analysis is based on these basic strokes, combined with some variations commonly used in daily writing. From the data collection process, the observed differences can primarily be attributed to the volunteers' level of familiarity with digital writing and their correction behaviors. We have specifically chosen several representative morphological changes to illustrate the extent of dissimilarity.

5.1. Stroke Continuity

In regular writing, one typically writes only one stroke each time the pen is put down. However, many people usually write two or even more strokes at the same time to speed up their handwriting as Chinese characters usually contain numerous strokes, which is called running style.^[17] This writing method can create additional continuity between strokes in handwritten text, which to some extent becomes a distinguishing feature of one's signature.

However, digital writing often differs in terms of continuity when compared to paper writing. For instance, when using an iPad, the discrepancy can be attributed to the less friction of writing on a screen. Additionally, when using a mobile device, the larger size of the fingertip as compared to a stylus or pen can lead to blocking more part of the contact area, contributing to difficulties in pinpointing the accurate position of the fingertip. In some cases, participants who were not familiar with using a stylus made an effort to write formally, where stroke-by-stroke writing leads to a significant reduction in their writing speed and, consequently, less continuity than that in their paper writing. Conversely, other participants wrote faster than they did on paper as the stylus was slippery so that they only needed to locate the trajectory of their fingertip on the screen once when writing a series of strokes or even a whole character, leading to greater continuity than paper writing.

The first two rows of **Figure 6** and **7** exhibit reverse altering of continuity.



Figure 6. Differences of original signatures between paper and other media.



Figure 7. Differences of skeletonized signatures between paper and other media.

5.2. Stroke Direction

Horizontal and vertical strokes are the most important to a character, determining the whole structure and balance of it.^[17] The standard for regular Chinese characters requires that some of the horizontal strokes should be parallel, and all vertical strokes should be perpendicular to the bottom/top edge of the paper.^[16,17] Individuals who have received school-based hand-writing training are usually able to closely meet this standard when writing on paper. However, maintaining this style of writing became more difficult when writing on a digital screen. Additionally, the orientation of other strokes also changed.

The third rows of Figure 6 and 7 illustrate the breaking of parallelism, where the two horizontal strokes in the left first character are parallel to each other, which is broken in the right. Moreover, the parallelism of horizontal strokes is also disrupted. We can also see many strokes alter their direction in the sample.

5.3. Stroke Size

In regular Chinese characters, dots are usually with shapes and often written as sesame-like marks instead of using the actual “.” symbol.^[16] While it was not difficult to achieve this shape when writing on paper or an iPad, some volunteers struggled to replicate it on mobile devices. Instead, they wrote small and even illegible “.” symbols, or just short straight lines, or wrote two symmetric dots together like a tilde symbol “~”. Although the “~” style is common to see in the running style of Chinese calligraphy,^[17] it was different from their paper writing. There are similar phenomena in hook strokes, volunteers often wrote the hook too small or too large, since such tiny strokes are harder to draw on a digital screen precisely in proper size.

The fourth rows of Figure 6 and 7 show how dots’ size shrinks to a minuscule “.”, and the fifth shows how the shape of a pair of symmetric dots changed to be “~”.

5.4. Stroke Misalignment

Misaligned strokes frequently occurred in complex characters written on mobile devices, often due to inaccurately judging the position of one’s fingertip on the screen. If left uncorrected, such misalignment can severely distort the structure of the character.

The last rows of Figure 6 and 7 demonstrate how some strokes were misaligned in mobile writing.

6. Curve Fitting Model

6.1. Bézier Curve Fitting

Bézier curves are widely used for plane curve designing on computers, including font descriptions in font libraries such as TrueType or OpenType, which use quadratic and cubic Bézier curves. Thus, we think it is suitable for describing our signatures and then enabling us to further quantitatively calculate their features for statistical analysis. We adopted a Bézier curve fitting method based on least squares.^[18]

In curve theory, a plane parametric curve is described as

$$C: r = r(t) = (x(t), y(t)), t \in [a, b] \quad (3)$$

It is typical to use lower-order Bézier curves to fit the local regions of a plane figure, especially a font,^[19,20] rather than using a single high-order curve, and here we choose cubic Bézier curve, whose parametric description is

$$B(t) = \sum_{i=0}^3 \binom{3}{i} P_i (1-t)^{3-i} t^i, t \in [0, 1] \quad (4)$$

where P_0, P_1, P_2, P_3 are four control points of it, leading the curve from P_0 to P_3 .^[18]

6.2. Partition of Signatures

To generate the proper set of neighbor points to fit, we should partition each connected component of a signature into some parts so that we can fit them with piecewise cubic Bézier curves.

We employ a straightforward depth-first search (DFS) algorithm to partition signatures into contiguous parts, which can be outlined as follows: 1) begin at a signature pixel and perform DFS to identify the neighboring 6 points that form a partitioned set. Fit these points with a cubic Bézier curve. 2) Commence from the last pixel of the previous segmented set and repeat the aforementioned process. 3) Iterate the previous step until there are insufficient pixels to form a partitioned set. We also fit the insufficient set more than or equal to 5 points with a cubic Bézier curve (4-point fitting always leads to overfitting). 4) Switch to another unexplored pixel and repeat the first step.

Figure 8 illustrates how a part of a signature is partitioned and fitted by cubic Bézier curves using least square.

7. Feature Extraction

Assume that we have fitted a signature with some Bézier curves $C_i: r_i = r_i(t), t \in [0, 1], i \in \{1, 2, \dots, n\}$, with which we implement the following computation to work out some features.

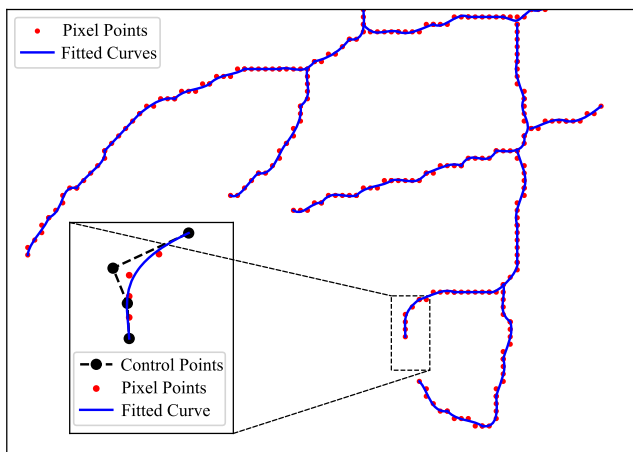


Figure 8. Cubic Bézier curve fitting of a part signature.

7.1. Connected Components

Here connected components are defined as in graph theory^[21] that we take a point and its eight neighboring points in a 3×3 grid as a whole component. We are interested in the number of connected components, which to some extent illustrates the degree of stroke discontinuity. Figure 9 shows the four connected components of a sample. We can easily find such connected components by using a DFS algorithm.

7.2. Minimum Area Rectangle

The minimum area rectangle (MAR) of a given set of points is defined as the rectangle that contains the point set with the smallest possible area. This problem was solved by Michael Shamos in 1978 using the rotating calipers algorithm on the convex hull of the set.^[22]

The features we are interested in here are the overall tendency angle of character writing and standardized length measurement, since the length can be affected by zooming.

To determine the overall tendency angle (overall angle) of the characters, one possible solution is to draw a line at the bottom or top of the characters. However, this approach presents the challenge of balancing the angle difference between these two lines. Alternatively, the angle of the bottom edge of the MAR of the signature can be a suitable choice, considering its close adherence to the written Chinese characters. For the same reason, the perimeter of the MAR can serve as a suitable standardized length measurement. Figure 10 shows how the red MAR adheres to a signature and the overall angle in blue.

7.3. Writing Density

Writing density refers to the proximity of strokes within a character. In the case of a Chinese character, it is typically written within an imaginary square,^[16,17] where higher density indicates a larger proportion of the square's area occupied by the character. Since we have retained only the skeleton of the signatures, we can represent the area using the arc length of the skeleton. To calculate the writing density, we divide the character arc length by the perimeter of the outer MAR. As is defined, this feature can also be considered as the relative length of the handwriting after excluding the length differences caused by scaling.

The arc length can be computed using a DFS algorithm, where a walk in the cardinal direction counts a length of 1, and a walk in

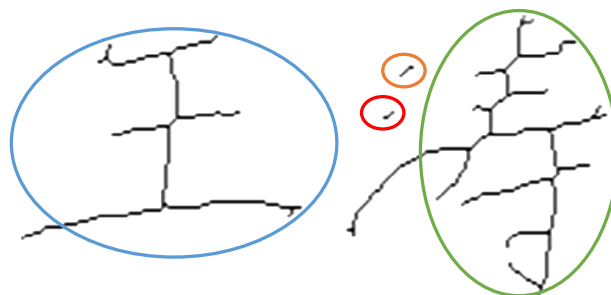


Figure 9. Four connected components.

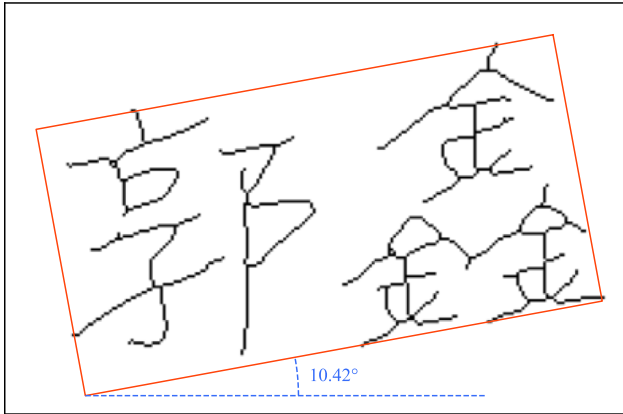


Figure 10. The minimum area rectangle of a signature.

the diagonal direction counts a length of $\sqrt{2}$, as shown in Figure 11. This algorithm can be considered only as an approximation algorithm, taking into account the omitted count of the walks at each closed contour.

7.4. Average Character Spacing

The average character spacing is taken into consideration as a feature. For 2-character signatures, it is rightly the spacing between the two characters, and for 3-character signatures, it is the arithmetic mean value of the two spacings.

We introduce a projection-based method for handwriting character segmentation. The algorithm utilizes vertical projection on the handwriting scripts, where characters are segmented at each location where the pixel frequency of the projection is zero.^[23] Since we have known the exact character number, we

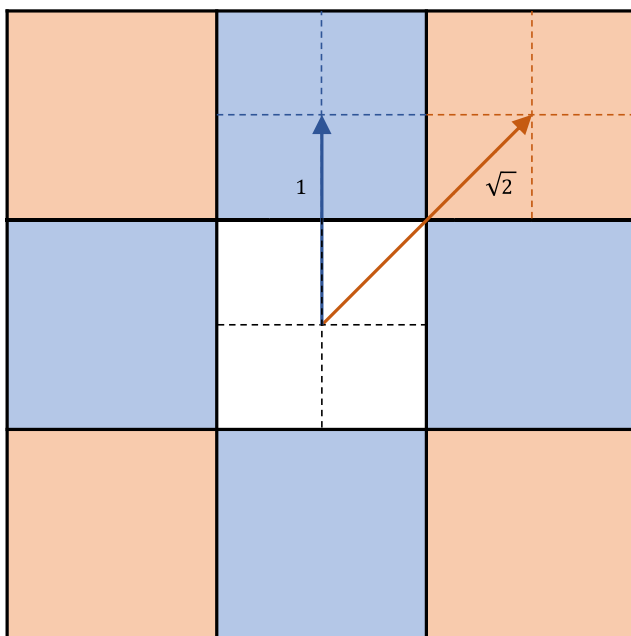


Figure 11. Length calculation by depth-first search.

can do a comparison between the supposed segmentation and algorithm results.

In some cases, volunteers wrote certain parts, such as Chinese radicals and dots, too far away from the main body, resulting in an over-segmentation of the handwriting. Fortunately, these parts usually contain much fewer pixels in simplified Chinese. Therefore, we can sort the segments based on the number of pixels they contain and merge the segments with fewer pixels into the nearest neighboring parts. This process continues until the number of segments matches the number of characters.

Figure 12 illustrates an example of over-segmented handwriting, where green lines indicate the beginning of each segment, red lines indicate the end, and blue bars represent the pixel frequency of the vertical projection. It should be emphasized that, based on the principle of proximity, not all over-segmented Chinese characters can be merged correctly, as shown in Figure 13, where the dashed lines represent an incorrect segmentation that should be corrected to the dotted lines. However, such cases are rare, and on the other hand, regardless of whether the disputed segments are merged to the left or right, the resulting spacing difference is negligible. Therefore, such erroneous segmentation can be ignored.

In other cases, volunteers wrote the Chinese characters so closely together that they cannot be segmented vertically, as shown in Figure 13, where the cyan dotted line is a possible segmentation. In such situations, we consider the spacing between characters to be 0.

7.5. Average Curvature

We implemented curve theory in differential geometry to calculate other features.

To make the calculation independent of parameter selection, we use arc length parameterization for computation, which is rightly defined as the arc length

$$s(t) = \int_0^t |r'(t)| dt \quad (5)$$

Relative curvature k_r and curvature k are used to describe the degree of bending of a curve, defined respectively as^[24]

$$k_r(t) = \frac{1}{|r'(t)|^3} \begin{vmatrix} x'(t) & y'(t) \\ x''(t) & y''(t) \end{vmatrix} \quad (6)$$

$$k(t) = |k_r(t)|$$

Although curvature is not invariant under similarity transformations, it is acceptable to be taken into consideration since the signatures have been preprocessed to almost the same size.

Then, we can define the sum curvature m concerning arc length of a curve C , and the whole arc length l as follows

$$m(C) = \int_{s(0)}^{s(1)} k(\bar{s}) d\bar{s}$$

$$= \int_0^1 |k_r(s(t))| |r'(t)| dt \quad (7)$$

$$l(C) = \int_0^1 |r'(t)| dt$$

where the second equation holds because of the change of variable to transform the integration concerning the arc length parameter s into integration concerning parameter t .

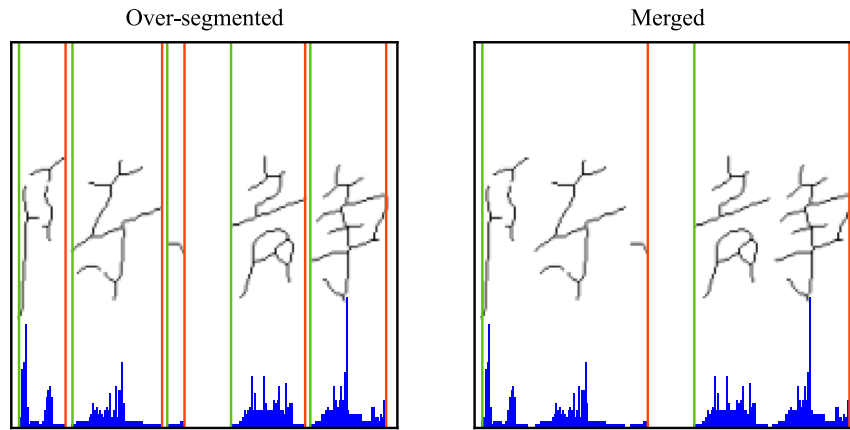


Figure 12. Over-segmented and correctly merged handwriting.

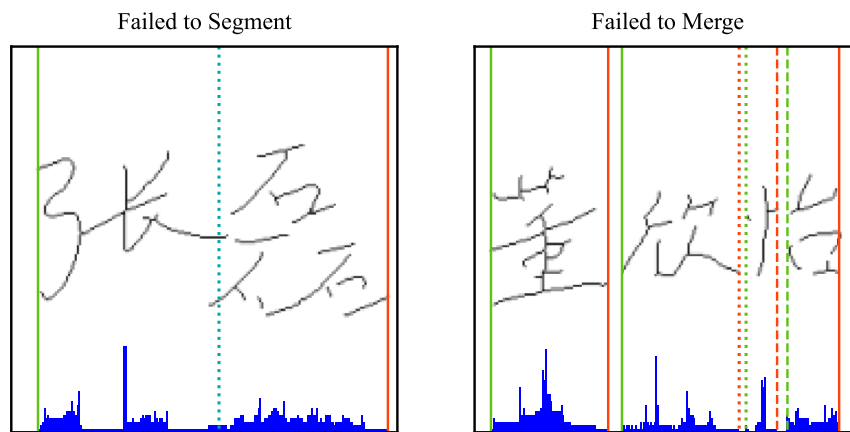


Figure 13. Segmentation and merge failure.

We can calculate the average curvature a of a signature as follows, which represents the degree of bending of the whole curve.

$$a = \frac{\sum_{i=1}^n m(C_i)}{\sum_{i=1}^n l(C_i)} \quad (8)$$

where $m(C_i)$ and $l(C_i)$ are the average curvature and the arc length of each fitted piecewise curve, respectively.

Notice that the average curvature defined here is different from the mean curvature defined on a surface.^[24]

7.6. Winding Degree

If we define θ as the direction angle of curve C at some point s , then we have a relationship between relative curvature and direction angle^[24]

$$k_r(s) = \frac{d\theta}{ds} \quad (9)$$

which implies we can calculate the whole change in direction angle by

$$\theta(s) - \theta(0) = \int_0^s k_r(\bar{s}) d\bar{s} \quad (10)$$

We do not want the changes of direction angle to cancel each other out during accumulation, since we do not know the correct writing direction, usually leading to the incorrect positive/negative sign of changes of the angle where thus we can only calculate the accumulation of their absolute values. Therefore, we define winding degree w of C as follows

$$w(C) = \int_{s(0)}^{s(1)} k(\bar{s}) d\bar{s} = \int_0^1 |k_r(s(t))| |r'(t)| dt \quad (11)$$

Then, we can calculate the accumulation S of the winding degree of a signature as follows, which represents the degree of pen/fingertip rotation during writing the signature.

$$S = \sum_{i=1}^n w(C_i) \quad (12)$$

where $w(C_i)$ is the winding degree of each fitted piecewise curve.

7.7. Statistics of Extracted Features

Paired *t*-test is used to determine if there is a significant difference between two sets of data that are normally distributed and paired.^[25] At first, we performed a normality test on the data by quantile–quantile plot (Q–Q plot)^[26] as normality is required for the paired *t*-test. We then performed the paired *t*-test with a *p*-value of 0.05 for the normally distributed features mentioned earlier, where the first set of features came from paper, and the second one from other media. For samples that do not follow a normal distribution, the Wilcoxon signed-rank test, a non-parametric test method for paired data, is conducted, serving a similar function as the paired *t*-test.^[27]

8. Results

8.1. Neural Networks Performance

We set the learning rate at 0.0005, batch size to be 31 (for the control group to be 11), and used Adam optimizer when training the models. Our calculation platform is Pytorch 2.0 under Python 3.10.16 powered by CUDA under Ubuntu 22.04, with 6 cores of Intel Xeon Skylake 6133 and one Nvidia Tesla V100 GPU. In this environment, it takes approximately 5 min to train a model.

To make graphs clearer, we adopted the semilog plot method for vertical coordinates in loss function plots. The comparison of loss and accuracy of classifiers and comparators are listed in **Table 1** and **2**, respectively. The more detailed corresponding graphs of the training process of the classifier are presented in Figure S1, Supporting Information, for the control group and Figure S3, Supporting Information, for the experimental group. Similarly, the training processes of the comparators are depicted in Figure S2, Supporting Information, for the control group, and Figure S4–S6, Supporting Information, for the experimental group.

8.2. Statistics of Extracted Features

Figure S7, Supporting Information, illustrates the Q–Q plots of the extracted features, with the *x*-axis representing the theoretical

quantiles and the *y*-axis representing the sample quantiles. The *p*-values are computed to generate a heatmap in **Figure 14** to intuitively present the results, using the Wilcoxon signed-rank test for the averaging spacing and the number of connected components, and the paired *t*-test for the remaining features.

9. Discussion

9.1. Neural Network Analysis

For neural network models, our focus should not be paid on the model's performance, but rather on the extent to which they reflect the differences among signatures from different sources.

As shown in the plots in Supporting Information, the loss function and accuracy of the signature classifier on electronic media have a similar trend to those of paper signatures at the beginning, indicating that electronic signatures of different people do have some weak correlation with the graphical features of paper signatures. However, we noticed that before the model overfits the training set, there is a large difference in the network performance between electronic signatures and paper signatures as shown in Table 1 and 2, where Acc means accuracy, and the bold data are for the training whereas the others are for the validation. Moreover, the performance of neural networks in the control group is significantly higher than that in the experimental group. Therefore, it suggests signatures from paper may be different from those from digital devices, which is corresponding to the statistical analysis in Section 9.3.

9.2. Stroke Morphology Analysis

The changes in stroke morphology provide a visual reason for the data from the neural network models, as they can disrupt the identifying structure and features of a signature, contributing to dissimilarities between them.

We must note that such disruptions in stroke morphology to some extent may be caused by a lack of experience in digital writing. In other words, skilled individuals may exhibit less dissimilarity in their signatures on a digital screen.

Table 1. Comparison of classifier performance.

Preprocessing method	Group	Media type	Min loss ↓	Max Acc ↑	Final log–loss difference ↑	Final Acc difference ↓
Binarized	Control	Paper 1	2.91 × 10⁻³	100%	0	0%
		Paper 2	1.65	46.3%	–6.38	54.5%
	Experiment	Paper	6.63 × 10⁻²	100%	0	0%
		iPad	4.46	4.92%	–4.29	95.6%
		Phone (table)	4.54	2.26%	–4.58	99.1%
Skeletonized	Control	Paper 1	1.52 × 10⁻³	100%	0	0%
		Paper 2	1.97	35.5%	–7.17	65.5%
	Experiment	Paper	2.49 × 10⁻²	100%	0	0
		iPad	4.30	6.05%	–5.15	94.5%
		Phone (table)	4.48	2.90%	–5.20	98.3%
		Phone (hand)	4.47	4.35%	–5.19	96.2%

Table 2. Comparison of comparator performance.

Preprocessing method	Group	Media type	Min loss ↓	Max Acc ↑	Final log-loss difference ↑	Final Acc difference ↓
Binarized	Control	Paper 1	4.68×10^{-2}	100%	0	0%
		Paper 2	5.31×10^{-1}	87.9%	-2.71	30.3%
	Experiment with iPad	Paper	4.57×10^{-2}	100%	0	0%
		iPad	1.08	60.2%	-3.34	43.0%
	Experiment with phone (table)	Paper	6.40×10^{-2}	100%	0	0%
		Phone (table)	1.01	62.9%	-2.83	41.9%
	Experiment with phone (hand)	Paper	9.33×10^{-2}	100%	0	0%
		Phone (hand)	1.18	56.2%	-2.57	45.7%
Skeletonized	Control	Paper 1	1.62×10^{-2}	100%	0	0%
		Paper 2	5.90×10^{-1}	84.8%	-3.56	15.2%
	Experiment with iPad	Paper	1.67×10^{-1}	96.7%	0	0%
		iPad	1.06	60.2%	-2.06	36.6%
	Experiment with phone (table)	Paper	8.85×10^{-2}	100%	0	0%
		Phone (table)	1.03	60.8%	-2.59	40.9%
	Experiment with phone (hand)	Paper	4.10×10^{-2}	100%	0	0%
		Phone (hand)	1.11	59.1%	-3.33	43.0%

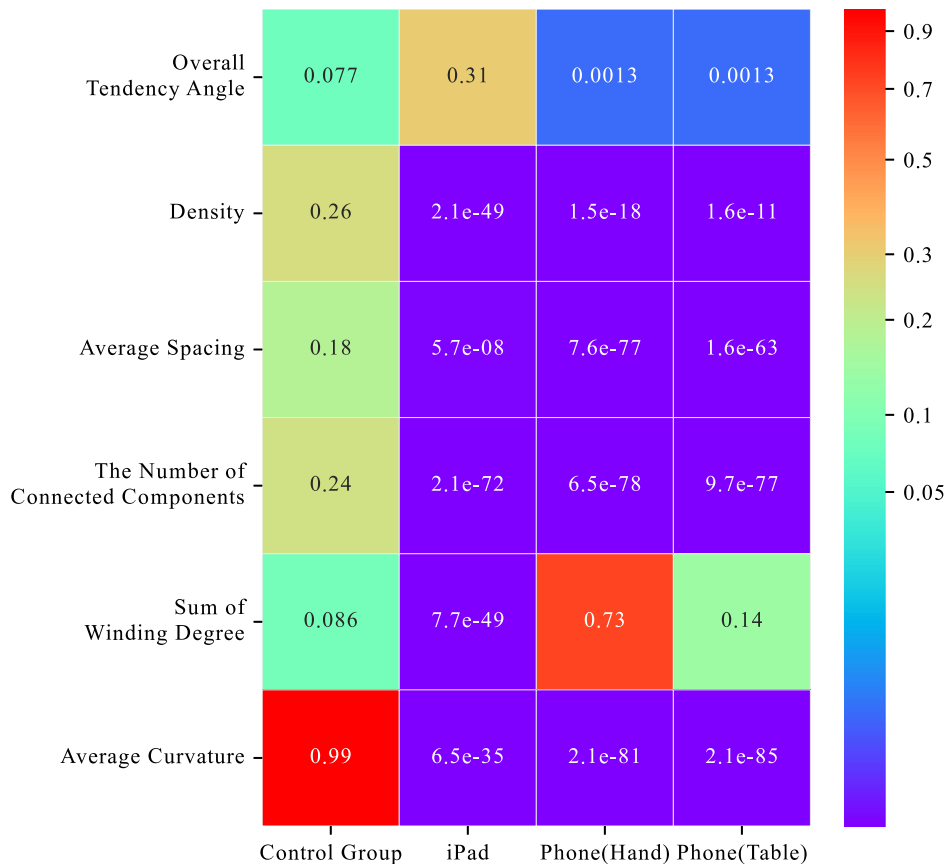


Figure 14. p-Values of paired tests with paper.

9.3. Statistics of Extracted Features

Statistical hypothesis testing of features of fitted curves provides further solid evidence for our former conclusions from neural networks. In Figure S7, Supporting Information, apart from average spacing and the number of connected components, most of the sample points for the remaining features lie approximately on a straight line, indicating that the distribution of these features can be considered normal.^[26] Therefore, the implementation of the paired *t*-test on the remaining features is reasonable. According to Figure 14/Heatmap 14, all the *p*-values of the control group are over 0.05, demonstrating the similarity between a pair of paper signatures, and most of the *p*-values of the experimental group are below 0.05 except 3, strongly supporting the dissimilarity of signatures on paper and from other sources.^[25]

Some of the data presented may indeed reveal interesting facts. Among extracted features, the overall tendency angle and the sum of winding degree exhibit the least similarity across different input media. This may be attributed to the fact that although the angles of individual strokes may vary, the volunteers' extensive practice in writing Chinese characters allows them to maintain a consistent overall angle in their signatures. Additionally, despite changes in the average curvature that may indicate variations in stroke size or bending degree, the volunteers may be still capable of ensuring that straight strokes remain straight and appropriately curving the strokes when necessary.

9.4. Limitations and Future Directions

For the SNN models, although they have shown some ability to capture the relationship and dissimilarity between signatures from different sources and have been previously proven to be canonical and effective for signature verification,^[1] it is important to note that the performance of these networks may vary significantly depending on the construction of paired training data. Therefore, it may be necessary to explore various combinations of paired data to further improve the model's performance.

Regarding the curve fitting model, the piecewise Bézier curves exhibit some lack of smoothness, which contributes to the error in curvature calculation. Further optimization of the fitting algorithm can be implemented to better fit the character,^[28,29] and more zoom-invariant features should be taken into consideration.

When it comes to input touch screen technology, in addition to stylus-touch screens, resistive and capacitive screens are among the most widely used and mature technologies.^[30] However, our study only focused on capacitive and stylus-touch screens and lacks research on signature recognition using resistive ones. Unlike capacitive touch screens, resistive touch screens can record pressure information of handwriting simultaneously.^[31] Moreover, our study only used devices for civilian use and did not investigate specialized devices, which may have differences in touch sampling accuracy and sampling rate compared to these devices, even though Apple has declared,^[32,33] "Apple Pencil delivers pixel-perfect precision". It should be noted that machine-learning techniques have been applied to touch screens^[30] and handwriting applications to enhance the accuracy and aesthetics of handwriting fonts. However, many of these algorithms are proprietary, which raises uncertainty for our research.

9.5. Insight of Signatures

The identification of electronic signatures through physical signatures, or vice versa, is unreliable. However, our investigation indicates that this unreliable method of identity verification is still widely used, particularly among those who work online. To address this issue, we suggest considering the following methods.

Recording different writing styles or fixing writing media is necessary. We have observed that signatures written on mobile devices with different writing styles can be dissimilar. Therefore, when recording digital signatures, samples of different writing styles should be stored and a stationary writing device should be used to reduce the dissimilarity.

Dynamic signature recognition is necessary. Unlike static recognition which only captures the image of the signature, dynamic recognition captures more information, which helps to enhance the accuracy of signature identification and anti-counterfeiting capabilities. In recent years, researchers have used SNNs on a combination of biometric features and signatures, or spatial features from a 3D writing system.^[34,35]

Cryptographic signatures are strongly recommended as surrogates if they can provide the same legal enforcement in the region. Cryptographic digital signature offers technology for identity verification and content integrity protection, and solutions from open source to commercialization are all mature. It overcomes the weakness of traditional signature being easily copied and forged, and has non-repudiation.^[36] Currently, many banks and some companies have adopted tamper-proof chips with non-exportable key storage areas to generate and store signature keys, connected through USB interfaces, providing the highest level of security. Now standards of post-quantum cryptographic signature algorithms and correlated hardware are being developed and selected to resist potential quantum computer attacks against existing algorithms.^[37,38]

10. Conclusion

With the aid of the neural network models and statistical analysis of the curve fitting model, our samples have shown a significant difference between static signatures on paper and those from digital sources, indicating the necessity of recollecting digital signatures, or preferably utilizing dynamic signature recognition. It is highly recommended to implement cryptographic digital signatures as a secure alternative in permissible areas.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This work is sponsored by Medical Research Council Confidence in Concept Award, UK (grant no. MC PC 17171); British Heart Foundation Accelerator Award, UK (grant no. AA/18/3/34220); Biotechnology and Biological Sciences Research Council (BBSRC), UK (grant no. RM32G0178B8); and Hope Foundation for Cancer Research, UK (RM60G0680).

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study, and the source code for the reproduction of all statistical analyses and machine learning methods are openly available at <https://doi.org/10.5281/zenodo.8008661>.

Keywords

Bézier curve, cross-media analysis, neural networks, signature authentication

Received: July 26, 2023

Published online: September 29, 2023

-
- [1] M. M. Hameed, R. Ahmad, M. L. M. Kiah, G. Murtaza, *Signal Process.: Image Commun.* **2021**, *93*, 116139.
- [2] L. Liu, L. Huang, F. Yin, Y. Chen, *Pattern Recognit.* **2021**, *118*, 108009.
- [3] V. Ruiz, I. Linares, A. Sanchez, J. F. Velez, *Neurocomputing* **2020**, *374*, 30.
- [4] M. Liwicki, M. I. Malik, C. E. Van Den Heuvel, X. Chen, C. Berger, R. Stoel, M. Blumenstein, B. Found, in *2011 Int. Conf. Document Analysis and Recognition*, IEEE, Piscataway, NJ **2011**, pp. 1480–1484.
- [5] M. Sharif, M. A. Khan, M. Faisal, M. Yasmin, S. L. Fernandes, *Pattern Recognit. Lett.* **2020**, *139*, 50.
- [6] C. Wencheng, G. Xiaopeng, S. Hong, Z. Limin, in *Advanced Hybrid Information Processing. ADHIP 2017*, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Vol. 219, Springer, Cham, Switzerland **2018**, pp. 33–37.
- [7] P. Wei, H. Li, P. Hu, in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, IEEE, Piscataway, NJ **2019**, pp. 5764–5772.
- [8] T. Longjam, D. R. Kisku, P. Gupta, *Multimedia Tools Appl.* **2023**, *82*, 5839.
- [9] M. Okawa, *Pattern Recognit.* **2018**, *79*, 480.
- [10] F. E. Batool, M. Attique, M. Sharif, K. Javed, M. Nazir, A. A. Abbasi, Z. Iqbal, N. Riaz, *Multimedia Tools Appl.* **2020**, <https://doi.org/10.1007/s11042-020-08851-4>.
- [11] V. Aubin, M. Mora, M. Santos-Peñas, *Pattern Recognit.* **2018**, *79*, 414.
- [12] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, MIT Press, Cambridge, MA **2016**.
- [13] K. He, X. Zhang, S. Ren, J. Sun, in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, IEEE, Piscataway, NJ **2016**, pp. 770–778.
- [14] A. B. Jagtap, D. D. Sawat, R. S. Hegadi, R. S. Hegadi, *Multimedia Tools Appl.* **2020**, *79*, 35109.
- [15] I. Melekhov, J. Kannala, E. Rahtu, in *2016 23rd Int. Conf. Pattern Recognition (ICPR)*, IEEE, Piscataway, NJ **2016**, pp. 378–383.
- [16] J. Björkstén, *Learn to Write Chinese Characters*, Yale University Press, New Haven, CT **1994**.
- [17] W. Li, *Chinese Writing and Calligraphy*, University of Hawaii Press, Honolulu, HI **2010**.
- [18] T. A. Pastva, Master's thesis, Naval Postgraduate School, Monterey, CA, **1998**.
- [19] Q. Li, J.-P. Li, L. Chen, in *2018 IEEE 20th Int. Conf. High Performance Computing and Communications; IEEE 16th Int. Conf. Smart City; IEEE 4th Int. Conf. Data Science and Systems (HPCC/SmartCity/DSS)*, IEEE, Piscataway, NJ **2018**, pp. 1156–1159.
- [20] H.-M. Yang, J.-J. Lu, H.-J. Lee, in *Proc. Sixth Int. Conf. Document Analysis and Recognition*, IEEE, Piscataway, NJ **2001**, pp. 276–280.
- [21] D. B. West, *Introduction to Graph Theory*, Vol. 2, Prentice Hall, Upper Saddle River **2001**.
- [22] P. Schneider, D. H. Eberly, *Geometric Tools for Computer Graphics*, Elsevier, Amsterdam, Netherlands **2002**.
- [23] Y. Liu, Y. Luo, F. Liu, Z. Qiu, in *Proc. 8th Int. Conf. Neural Information Processing (ICONIP-2001)*, Vol. 3, MIT Press, Cambridge, MA **2001**, pp. 1603–1606.
- [24] M. P. Do Carmo, *Differential Geometry of Curves and Surfaces: Revised and Updated Second Edition*, Courier Dover Publications, Mineola, NY **2016**.
- [25] J. Shao, *Mathematical Statistics*, Springer Science & Business Media, Berlin, Germany **2003**.
- [26] H. C. Thode, *Testing for Normality*, Statistics: A Series of Textbooks and Monographs, Vol. 164, CRC Press, Boca Raton, FL **2002**.
- [27] W. J. Conover, *Practical Nonparametric Statistics*, Wiley Series in Probability and Statistics, Vol. 350, John Wiley & Sons, Hoboken, NJ **1999**.
- [28] K. Itoh, Y. Ohno, *Electron. Publishing* **1993**, *6*, 195.
- [29] L. Shao, H. Zhou, *Graph. Models Image Process.* **1996**, *58*, 223.
- [30] H. Nam, K.-H. Seol, J. Lee, H. Cho, S. W. Jung, *Sensors* **2021**, *21*, 4776.
- [31] M. Diaz, M. A. Ferrer, D. Impedovo, M. I. Malik, G. Pirlo, R. Plamondon, *ACM Comput. Surv.* **2019**, *51*, 117.
- [32] Apple Inc., Apple pencil (1st generation) - technical specification **2022**, https://support.apple.com/kb/SP740?locale=en_US (accessed: May 2023).
- [33] Apple Inc., Apple pencil (2nd generation) - technical specifications **2022**, https://support.apple.com/kb/SP786?locale=en_US (accessed: May 2023).
- [34] D. D. Chakladar, P. Kumar, P. P. Roy, D. P. Dogra, E. Scheme, V. Chang, *Inf. Fusion* **2021**, *71*, 17.
- [35] S. Ghosh, S. Ghosh, P. Kumar, E. Scheme, P. P. Roy, *Pattern Recognit. Lett.* **2021**, *144*, 13.
- [36] S. Ullah, J. Zheng, N. Din, M. T. Hussain, F. Ullah, M. Yousaf, *Comput. Sci. Rev.* **2023**, *47*, 100530.
- [37] D. Soni, K. Basu, M. Nabeel, R. Karri, A hardware evaluation study of NIST post-quantum cryptographic signature schemes, in *Second PQC Standardization Conf.*, NIST, Gaithersburg, MD **2019**.
- [38] G. Alagic, D. Apon, D. Cooper, Q. Dang, T. Dang, J. Kelsey, J. Lichtinger, C. Miller, D. Moody, R. Peralta, R. Perlner, *Status Report on the Fourth Round of the NIST Post-Quantum Cryptography Standardization Process*, US Department of Commerce, NIST, to be published **2023**.
- [39] R. Yug, 8 strokes of yong-zh **2010**, https://upload.wikimedia.org/wikipedia/commons/c/c6/8_strokes_of_%E6%B0%B8-zh.svg; <http://creativecommons.org/licenses/by/4.0/>.