

Article

On the Need for Accurate Brushstroke Segmentation of Tablet-Acquired Kinematic and Pressure Data: The Case of Unconstrained Tracing

Karly S. Franz ^{1,2,*}, Grace Resznetnik ³ and Tom Chau ^{1,2} 

¹ Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, ON M4G 1R8, Canada; tom.chau@utoronto.ca

² Institute of Biomedical Engineering, University of Toronto, Toronto, ON M5S 3G9, Canada

³ Faculty of Medicine, University of Toronto, Toronto, ON M5S 1A8, Canada; grace.resznetnik@mail.utoronto.ca

* Correspondence: karly.franz@mail.utoronto.ca; Tel.: +1-416-425-6220 (ext. 3525)

Abstract: Brushstroke segmentation algorithms are critical in computer-based analysis of fine motor control via handwriting, drawing, or tracing tasks. Current segmentation approaches typically rely only on one type of feature, either spatial, temporal, kinematic, or pressure. We introduce a segmentation algorithm that leverages both spatiotemporal and pressure features to accurately identify brushstrokes during a tracing task. The algorithm was tested on both a clinical and validation dataset. Using validation trials with incorrectly identified brushstrokes, we evaluated the impact of segmentation errors on commonly derived biomechanical features used in the literature to detect graphomotor pathologies. The algorithm exhibited robust performance on validation and clinical datasets, effectively identifying brushstrokes while simultaneously eliminating spurious, noisy data. Spatial and temporal features were most affected by incorrect segmentation, particularly those related to the distance between brushstrokes and in-air time, which experienced propagated errors of 99% and 95%, respectively. In contrast, kinematic features, such as velocity and acceleration, were minimally affected, with propagated errors between 0 to 12%. The proposed algorithm may help improve brushstroke segmentation in future studies of handwriting, drawing, or tracing tasks. Spatial and temporal features derived from tablet-acquired data should be considered with caution, given their sensitivity to segmentation errors and instrumentation characteristics.

Keywords: biomedical signal processing; segmentation algorithm; biomechanical phenomena; error propagation; tracing



Citation: Franz, K.S.; Resznetnik, G.; Chau, T. On the Need for Accurate Brushstroke Segmentation of Tablet-Acquired Kinematic and Pressure Data: The Case of Unconstrained Tracing. *Algorithms* **2024**, *17*, 128. <https://doi.org/10.3390/a17030128>

Academic Editor: Ahmed Shaffie

Received: 18 February 2024

Revised: 15 March 2024

Accepted: 18 March 2024

Published: 20 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The emergence of affordable digital tablets over the past decade has created new opportunities for identifying cognitive and motor control differences among populations through drawing, tracing, and handwriting tasks [1–12]. Compared to conventional pen-and-paper assessments, which limit analysis to the final product, tablet computers allow researchers to study the process of drawing by capturing continuous changes in pen position, pen-on-tablet pressure, and the stylus azimuth or tilt. These data give rise to efficient, computer-based assessment of graphomotor activities, circumventing the need for time-consuming, subjective judgment of conventional pen-and-paper acquired data [13]. For tablet data, various algorithms have been proposed to detect subtle dynamic, kinematic, and static markers of handwriting/drawing difficulties secondary to impairments of perceptual motor control. These brushstroke-derived biomechanical quantities have expanded our understanding of graphomotor processes and set the stage for machine learning-based detection of handwriting or drawing disorders due to developmental differences or disease [2,6–8,12]. For example, Asselborn et al. [2], were able to identify 53 measurable brushstroke features (such as in air-time ratio, drawing velocity, essential

tremor, among others) that together characterized fine motor control difficulties in children with dysgraphia. Most tablet-based handwriting/drawing studies have focused on identifiable features within brushstrokes. However, the actual segmentation, i.e., automatically identifying from the recorded data the end of one stroke and the beginning of the next, are vaguely described, if at all. Brushstroke segmentation is a critical step in the preprocessing of drawing or handwriting data; errors in brushstroke identification propagate to brushstroke-derived measures, which would ultimately yield inaccurate evaluations of task performance.

Few authors have presented approaches to segmenting brushstrokes [1,2,14–16]. Spatial segmentation approaches have categorized strokes based on spatial features such as local minimum/maximal coordinates or have deconstructed characters into predefined components [14,16,17]. Fitjar et al. [16] predefined the strokes needed to trace letters; for example, the letter “A” was decomposed into three lines (two slanted and one horizontal). However, this method did not account for cases where the user writes the letter using a different number of strokes. In alternative schemes, Asselborn et al. [1,2] segmented strokes using predefined bin sizes of sampled points while others separated strokes according to predefined temporal (e.g., 5–10 s segments [14]) rather than spatial boundaries. By adopting a predefined segmentation approach, these algorithms ignored pen lifts that may have occurred within segments. The multiplicity of pen lifts and greater in-air duration has previously yielded insight into cognitive or motor-processing difficulties [10,11,15,18,19]. Other segmentation strategies in the literature include separating written strokes by identifying points in time where the velocity was equal to zero, i.e., directional changes [20,21]. Alternatively, strokes have also been segmented at timepoints where the velocity and acceleration was 0 [14]. However, as participants may pause without lifting the pen, this condition alone is not sufficient to delineate strokes. An alternative is to explicitly instruct participants to start or stop drawing continuously, and observe the time when the pen touches the tablet [4,22,23]. However, explicit start/stop instructions detract from naturalistic drawing and may be difficult for children with cognitive disabilities to comply with.

To avoid issues associated with predefining strokes, time points that indicate specific fluctuations in applied pen pressure have been used to distinguish segments [16,22]. Some authors have defined segments based on pen-up and pen-down positions when the registered pressure equals zero [8,24]. In similar spirit, Rosenblum et al. defined the beginning of a brushstroke as the moment when pen pressure surpassed 4% of the maximum measured pressure and the end as when the pressure dipped below this threshold [15]. These approaches require instrumentation that is sensitive to subtle pressure changes, but they are consequently susceptible to segmentation error whenever unintended palm contact (on capacitive touchscreens) introduces errant pressure measurements.

In this technical note, we illustrate the necessity of accurately segmenting tablet-derived data into brushstrokes. To this end, we developed a simple brushstroke segmentation algorithm that leverages spatial, temporal, and pressure measurements. Using kinematic data from a tablet-based tracing task, we validated the automatically detected strokes against those derived from retrospective video review. We quantified the impact of incorrect brushstroke segmentation on the biomechanical parameters often derived from table-based handwriting and drawing studies.

2. Materials and Methods

2.1. Participants

Fifty-two child–adult dyads were recruited as part of a larger study which included examining brushstroke features while tracing cartoon images. In this technical note, we consider selected tracings from a subset of 11 participants consisting of 5 children (4 diagnosed with autism spectrum disorder, 1 neurotypical; average age = 12.34 ± 0.49 years; 4 males) and 6 adults (average age = 40.64 ± 1.41 years; 5 females).

2.2. Tracing Protocol

A tracing game application for android tablets was developed using the Construct 3 game engine. During the data collection session, which lasted an hour and a half, parent–child dyads were seated at a table across from each other in a quiet, well-illuminated room. Participants were asked to trace images either independently or collaboratively. For each drawing trial, the parent and child took turns selecting one of 20 images they wanted to trace. They then traced with their dominant hand various cartoon characters over a 120 s period using an S Pen stylus (silicone tip stylus 9.7; BoxWave Corp., Kirkland, WA, USA) on a 9.7" electromagnetic resonance touchscreen tablet (Tab A model SM-P550; Samsung, Suwon, South Korea) with 1024 levels of pressure sensitivity. While participants traced the image, time-stamped x- and y-coordinates along with the estimated applied pen pressure (a continuous value between 0 and 1) were recorded at 16.51 ± 0.01 Hz and saved to a server. A sagittal view of the participants was video recorded during the tracing task.

2.3. Tracing Data

A total of 487 tracings were collected and the two most selected images (yoda, selected 44 times; lego, selected 40 times) were used to verify the proposed brushstroke segmentation method presented in this note. From the two selected images, we only used tracings where: (1) the corresponding video recordings provided an unobstructed view of the participant's hand such that pen lifts could be unambiguously confirmed; and (2) there was no evidence of free-hand drawing, i.e., doodling unrelated to the tracing task. For this study, 14 separate tracings met these two criteria and from here on, will be referred to as the clinical dataset.

We created a separate dataset, from here on referred to as the validation dataset, to systematically stress test the segmentation algorithm. The first author traced the same two images as above 12 times, each emulating different tracing behaviors informed by observations of actual participant tracing habits. These conditions comprised variations in speed (slow, fast, and variable), the number of pen lifts (minimal or several), and hand positions (hand resting and not resting on the screen). In total, the validation set consisted of 12 trials for each image, generating a total of 24 tracings. To validate the number of algorithmically identified brushstrokes, actual brushstrokes were manually enumerated by retrospective review of an overhead camera recording that provided a transverse view of the tracing task.

2.4. Segmentation Algorithm

The acquired data were segmented into brushstrokes using a custom semi-automated program developed in MATLAB (v.R2022b) that considered pressure, as well as temporal and spatial measurements. Pen-on-tablet contact was defined as instances where pressure exceeded a threshold, P_T . In this study, $P_T = 0.7$ was empirically determined as the threshold that captured most of the authentic brushstrokes. In complementarity, pen lifts occurred at timepoints where instantaneous pressure, $p(t_i)$, fell below this threshold, namely, $p(t_i) < P_T$. Therefore, the set of times, L_p , associated with pressure-demarcated pen lifts is given by,

$$L_p = \{t_i | p_i(t) < P_T\}, i = 1, \dots, N \quad (1)$$

where N is the number of samples in the entire tracing. This definition identified complete lifts, i.e., where $p(t_i) = 0$, with neither pen nor hand contact, as well as partial lifts where the pen was raised but the hand remained on the screen, yielding non-zero pressure readings, or where the pen tip traversed proximal to the screen without drawing, thereby registering supra-threshold pressure values. However, in these instances, pen lifts could not be detected based on pressure alone.

Concurrent spatial and temporal data were subsequently considered. It was found that the time difference, $\delta(t_i)$, between successive unique x- and y-coordinate values varied as the stylus moved across the screen while tracing (Figure 1a). These variations reflected within-stroke pauses while tracing, natural fluctuations in drawing speed, as well as pen lifts between successive brushstrokes. The latter were an order of magnitude larger than the

time differences associated with pauses and speed fluctuations. Exploiting this observation, we fitted, via maximum likelihood estimation, a Rayleigh probability density to the time differences between successive unique pen positions. The 95th percentile of this density was empirically determined as the temporal threshold, δ_T , that identified authentic pen lifts, i.e., $\delta(t_i) > \delta_T$, from unique x- and y-coordinates with the highest sensitivity and specificity (Figure 1b). Identifying the set of pen lift times, L_S , by spatiotemporal information, that is, the time differences between successive unique coordinates (x_i, y_i) and (x_{i+1}, y_{i+1}) , can, thus, be summarized as follows:

$$L_S = \{t_i | \delta(t_i) > \delta_T, x_{i+1} \neq x_i \text{ and } y_{i+1} \neq y_i\}, i = 1, \dots, N - 1 \text{ where } \delta(t_i) = t_{i+1} - t_i \quad (2)$$

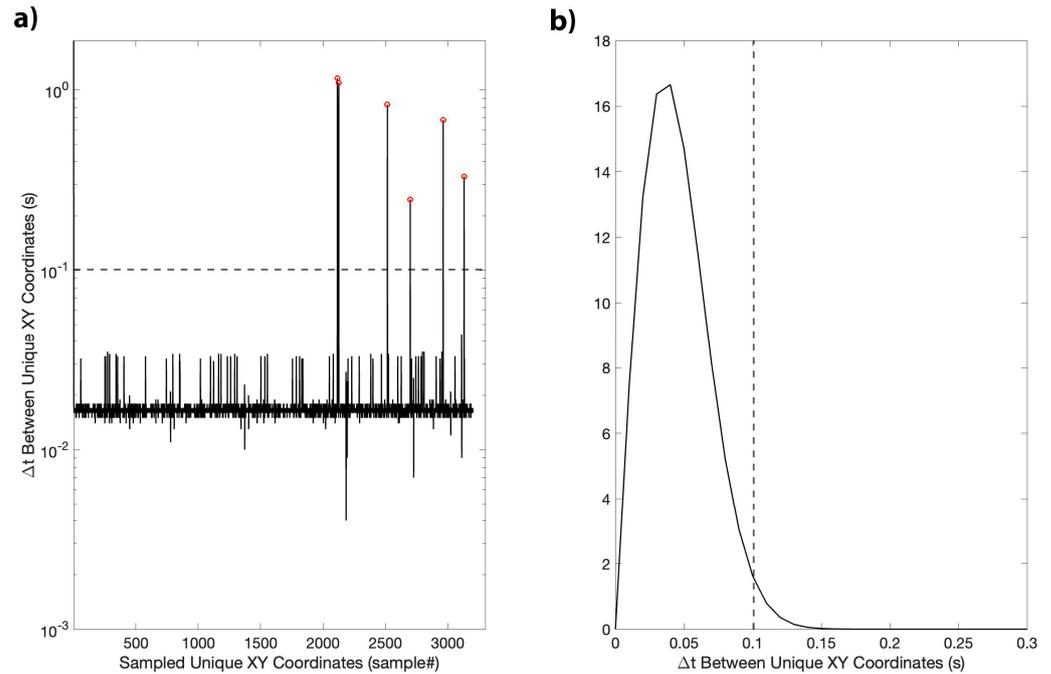


Figure 1. (a) A plot of the time-varying time differences between unique x- and y-coordinates versus the sample number. Time differences above a defined threshold were identified as pen lifts (red peaks) while those below indicated pauses while drawing a brushstroke. (b) A Rayleigh density fit the time differences between unique x- and y-coordinates during the tracing trial. Time differences above the 98th percentile were deemed to be authentic inter-brushstroke pen lifts.

The threshold δ_T is defined as the 98th percentile,

$$\frac{1}{N} \sum_{i=1}^N H(\delta_T - \delta(t_i)) \times 100\% = 98\% \quad (3)$$

where $H(\cdot)$ is the Heaviside function, namely, $H(x) = 1, x > 0$ otherwise $H(x) = 0$.

The times of spatiotemporally determined pen lifts, L_S , were combined with pen lift times detected based on pressure, L_p , to define a final set of pen lift times, $L = t_1, \dots, t_K$ where $t_k \in L_S \cup L_p$ and $K = |L_S \cup L_p|$ where $|\cdot|$ denotes the cardinality of the set. We further define t_0 as the time of the beginning of the first stroke, i.e., the initial time at which $p(t) > P_T$. The k^{th} brushstroke, B_k , is then defined as the set of coordinates between pen lift times,

$$B_k = \{(x_i, y_i) | t_{k-1} + \tau \leq t_i \leq t_k\}, t_k \in L, i = 1, \dots, N, k = 1, \dots, K \quad (4)$$

where, as above, N is the total number of samples in the tracing, K is the number of brushstrokes comprising the tracing, and $\tau > \delta_T$ is a temporal offset guaranteeing a unique point following the last pen lift, wherein $p(t_{k-1} + \tau) > P_T$.

In other words, $t_{k-1} + \tau$ is the first unique point after the last pen lift where pressure exceeds the minimum threshold for legitimate pen-on-tablet contact.

Noisy or spurious points were observed when the stylus traversed the screen during a pen lift or when the hand was resting on the surface while drawing. These spurious points appeared as the first or final six points of some brushstrokes and appeared relatively equidistant but much more widely spaced apart compared to other points within the stroke (Figure 2). To automatically remove these points from an identified brushstroke, we fit a Rayleigh probability density to the distances between unique coordinates via maximum likelihood estimation. The 95th percentile distance demarcated noisy points. In other words,

$$\text{Noise} \triangleq \{(x_i, y_i) | d_i > D_T\}, \forall (x_i, y_i) \in B_k d_i = \sqrt{(x_{i-1} - x_i)^2 + (y_{i-1} - y_i)^2} \quad (5)$$

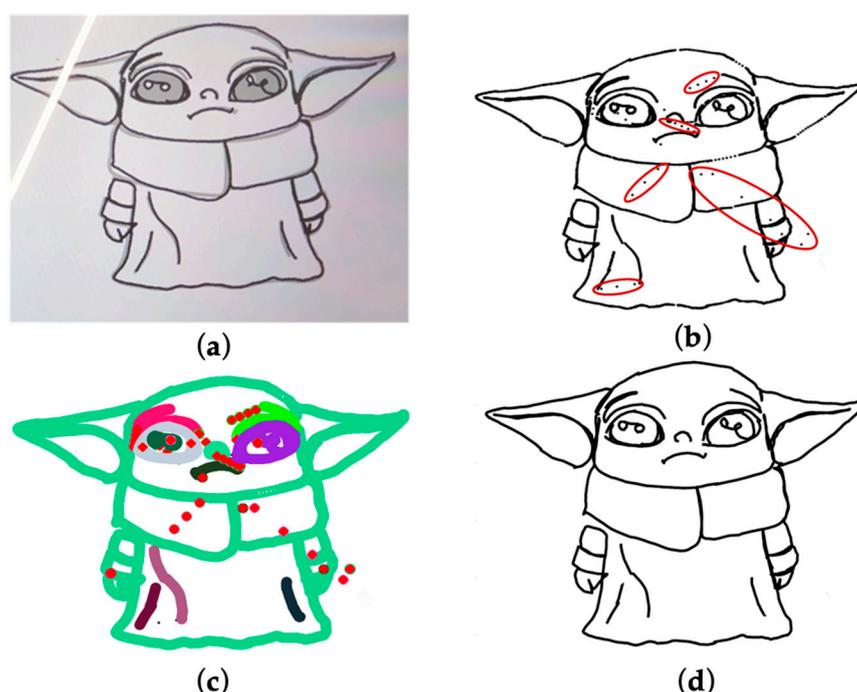


Figure 2. (a) A photo of the original tracing on the tablet. (b) Raw data prior to preprocessing. Points circled in red are examples of noisy points. (c) Detected brushstrokes after applying the segmentation algorithm. Each stroke is represented by a different color and red points indicate identified spurious points. (d) The final reconstructed image after removing noisy points.

2.5. Algorithm Validation

The algorithm was tested with the 12-behavior validation set described above. In the trials where the number of strokes were incorrectly identified by our algorithm, we estimated the propagated error in 37 features typically used in studies to characterize fine motor control [1,2,6,8,25,26]. These features, listed in Table 1, were categorized as temporal, kinematic, or spatial and were quantified using descriptive statistics (average, standard deviation, maximum, and minimum). The algorithm was further tested on the clinical dataset to confirm that the number of strokes identified by the algorithm was equal to the number of pen lifts observed during the recorded trial.

For the temporal features, in-air time was the time that the pen was lifted from the screen between brushstrokes while the on-screen time was the time that elapsed between the first and final coordinates of the stroke, which included both pauses and movement time [8]. Subsequently, the in-air time/on-surface ratio was the total in-air time divided by the calculated total on-surface time [1]. Stop time and movement time were defined,

respectively, as the durations over which the pen remained in contact with the screen while stationary and in motion [6].

Table 1. List of biomechanical features typically derived from tablet-based data to characterize graphomotor performance [1,2,6,8,15,26].

Temporal Features (In Units of Seconds)	Spatial Features (In Units of Pixels)
Total on-screen time	Average distance between brushstrokes
Average on-screen time	Standard deviation distance between brushstrokes
Standard deviation on-screen time	Maximum distance between brushstrokes
Maximum on-screen time	Minimum distance between brushstrokes
Minimum on-screen time	Average brushstroke width
Total in-air time	Standard deviation brushstroke width
Average in-air time	Maximum brushstroke width
Standard deviation in-air time	Minimum brushstroke width
Maximum in-air time	Average brushstroke height
Minimum in-air time	Standard deviation brushstroke height
Total movement time	Maximum brushstroke height
Total stop time	Minimum brushstroke height
In-air/on-paper time ratio (dimensionless)	Average brushstroke length
	Standard deviation brushstroke length
	Maximum brushstroke length
	Minimum brushstroke length
Kinematic features (pixels/s)	
	Average velocity
	Standard deviation stroke velocity
	Maximum stroke velocity
	Average acceleration (pixels/s ²)
	Standard deviation acceleration (pixels/s ²)
	Maximum acceleration (pixels/s ²)
	Minimum acceleration (pixels/s ²)
	Average fluency (number of velocity inversions)

Spatial features such as the brushstroke height and width were defined as the length of the strokes along the x- and y-axes, respectively, while the stroke length was the total Euclidean distance of the brushstroke [6,15]. The distance between brushstrokes was the Euclidean distance between the end point of one stroke and the starting location of the next [2].

The kinematic features such as the velocity and acceleration were computed as instantaneous values between x- and y-coordinates across all brushstrokes [1]. Fluency was defined as the number of inversions in the vector sum of velocity or, equivalently, the number of zero-crossings in the corresponding acceleration signal [26].

When the algorithm incorrectly identified the number of brushstrokes in the validation dataset, we computed the absolute propagated errors in the values of the features in Table 1, namely,

$$Error = \frac{|f_{err} - f_0|}{f_0} \times 100\% \quad (6)$$

where f_{err} is the value of a given feature computed with the erroneously detected number of brushstrokes while f_0 is the same feature computed with the actual number of brushstrokes.

3. Results

On the validation set, the algorithm correctly predicted the number of brushstrokes in 19 of the 24 trials (79%). See Table 2. Of the five instances of incorrect segmentation, all cases yielded one brushstroke more than the original tracing. Discrepancies between the actual and detected number of brushstrokes occurred exclusively in trials where brushstrokes

were deliberately minimized. Hand position and drawing speed seemed to have less of a negative impact on detection accuracy.

Table 2. Segmentation performance of the validation set under different conditions (variations in hand position, drawing speed, and the number of pen lifts). * denotes trials where the number of detected and actual segments differed.

		Condition			Image	Number of Actual Segments	Number of Detected Segments
Trial	Hand Position	Drawing Speed	Number of Pen Lifts				
1	resting	slow	minimum	yoda	11	11	
2				lego	3	3	
3	resting	fast	maximum	yoda	48	48	
4				lego	67	67	
5	resting	variable	minimum	yoda	9 *	10 *	
6				lego	6	6	
7	resting	slow	maximum	yoda	43	43	
8				lego	63	63	
9	resting	fast	minimum	yoda	11	11	
10				lego	5	5	
11	resting	variable	maximum	yoda	41	41	
12				lego	55	55	
13	not resting	slow	minimum	yoda	4 *	5 *	
14				lego	5 *	6 *	
15	not resting	fast	maximum	yoda	45	45	
16				lego	70	70	
17	not resting	variable	minimum	yoda	1	1	
18				lego	5 *	6 *	
19	not resting	slow	maximum	yoda	43	43	
20				lego	52	52	
21	not resting	fast	minimum	yoda	4	4	
22				lego	1 *	2 *	
23	not resting	variable	maximum	yoda	39	39	
24				lego	44	44	

For the clinical dataset, consisting of both adult and pediatric data, the number of pen lifts could only be verified by video review in 14 recordings due to the sagittal view of the participant. See Table 3. Of these 14 video recordings, the algorithm correctly identified the number of brushstrokes in 12 trials (86%).

Table 3. Number of actual and detected brushstrokes for the clinical dataset. The algorithm correctly identified the number of strokes in 12 of 14 trials. * denotes trials where the number of detected and actual segments differed.

Trial	Image	Number of Actual Segments	Number of Detected Segments
1	yoda	19	19
2	yoda	34 *	33 *
3	yoda	26	26
4	lego	22	22
5	lego	24	24
6	yoda	37	37
7	yoda	53	53
8	yoda	12	12
9	yoda	41	41
10	lego	27	27

Table 3. Cont.

Trial	Image	Number of Actual Segments	Number of Detected Segments
11	yoda	26 *	28 *
12	yoda	26	26
13	lego	31	31
14	lego	33	33

Table 4 lists the average errors for each feature, averaged across the 5 trials of the validation set with erroneous brushstroke segmentations.

Table 4. The error for the tested features, averaged across the five trials of the validation set where the incorrect number of strokes was algorithmically identified. Features are sorted according to their corresponding average error from highest to lowest.

Biomechanical Features	Avg Error \pm SD %
Minimum distance between brushstrokes	99.31 \pm 1.08
Minimum in-air time	95.34 \pm 1.96
Standard deviation of in-air time	49.21 \pm 27.33
Standard deviation of on-screen time	30.71 \pm 14.82
Standard deviation of brushstroke length	28.84 \pm 11.60
Maximum on-screen time	24.95 \pm 23.95
Average fluency	22.76 \pm 15.66
Average on-screen time	22.71 \pm 15.70
Average brushstroke length	22.69 \pm 15.70
Average in-air time	22.02 \pm 15.52
Maximum brushstroke length	21.04 \pm 20.22
Average distance between brushstrokes	18.60 \pm 5.76
Average brushstroke width	15.45 \pm 10.81
Standard deviation of brushstroke height	14.21 \pm 4.05
Minimum brushstroke length	11.12 \pm 24.86
Minimum on-screen time	10.77 \pm 22.68
Average acceleration	12.66 \pm 15.19
Standard deviation of brushstroke width	6.26 \pm 5.61
Maximum brushstroke height	6.04 \pm 12.42
Minimum brushstroke width	5.75 \pm 12.85
Average brushstroke height	4.98 \pm 3.97
Maximum brushstroke width	4.07 \pm 9.10
Standard deviation of distance between brushstrokes	2.88 \pm 1.81
Minimum brushstroke height	1.72 \pm 3.84
In-air/on-surface time ratio	0.99 \pm 0.61
Total in-air time	0.92 \pm 0.60
Maximum distance between brushstrokes	0.88 \pm 1.76
Total stop time	0.27 \pm 0.05
Maximum in-air time	0.23 \pm 0.51
Standard deviation of acceleration	0.13 \pm 0.19
Total on-screen time	0.06 \pm 0.02
Average velocity	0.05 \pm 0.02
Standard deviation of velocity	0.04 \pm 0.08
Total movement time	0.04 \pm 0.02
Maximum velocity	0.00 \pm 0.00
Maximum acceleration	0.00 \pm 0.00
Minimum acceleration	0.00 \pm 0.00

The most affected feature was the minimum distance between brushstrokes as well as the minimum in-air time, with errors of 99.31% and 95.34%, respectively. Outside of this, all other spatial features were mildly-to-moderately affected with a cumulative error between 0.88–28.84%. In a similar manner, the temporal features were moderately affected

by erroneously detected strokes, with an error ranging from 0.04% to 49.21%. However, measurements related to total times (such as total stop time, total movement time, total on-screen time, and total in-air time) demonstrated an error of approximately 0.92% or less. Kinematic features were the least affected compared to spatial and temporal features. All kinematic features generally had error rates less than 1%, except the average fluency and acceleration which had discrepancies of 22.76% and 12.66%, respectively.

4. Discussion

We quantified the propagated impact of incorrect brushstroke segmentation on common biomechanical feature values reported in studies utilizing tablet-based drawing, tracing, or handwriting tasks. To do so, we introduced a novel brushstroke segmentation algorithm that leverages both pressure and spatiotemporal data from a digital tablet. The current findings showed that, despite small discrepancies between the algorithmically identified and actual brushstrokes, there can be significant errors in the values of the derived features.

4.1. The Segmentation Algorithm Performance

On the validation set, the algorithm demonstrated robust detection under various drawing conditions (varying the number of strokes, hand position, and drawing speed). Among the five trials where the number of brushstrokes was incorrectly identified, the algorithm was off by only one brushstroke in all five trials. In all instances of erroneous segmentation, the common tracing behavior was that the number of brushstrokes was purposely minimized. The speed and hand position seemed to have a smaller but non-negligible effect on segmentation. In all cases of segmentation error, unique data points were not recorded as the image was traced due to intermittent server connection, creating an inflated mid-stroke pause which led to the incorrect decomposition of a single stroke as multiple brushstrokes. Perhaps the prevalence of pen-contact samples in these trials, which necessitated longer periods of continuous data streaming, increased the risk of the observed hardware fault. The algorithm also demonstrated robust performance with the clinical dataset where it only misidentified two trials and was off by one and two brushstrokes each.

4.2. Impact of Segmentation Errors on Spatial Features

Spatial features were the most affected by incorrect stroke identification, particularly the minimum distance between strokes which had a propagated error of 99.3% as well as the minimum in-air time, a temporal feature, which similarly demonstrated a propagated error above 90%. This is unsurprising given that a single stroke would have closely adjacent x- and y-pixels compared to segmentations that yielded multiple strokes. Given the relationship between distance and time, it is also expected that minimum distance measurements would be positively correlated with minimum in-air time. Incidentally, spacing as a feature is commonly used in handwriting analysis to assess fine motor control in various populations [1,2,15,17,19], although the literature has not consistently associated this feature with neurodivergent writing capabilities. For example, it has been reported that children with dysgraphia [2] and non-proficient writing [19] tend to narrowly space their words. Falk and colleagues found high variability in the inter-letter distance among younger children writers, and concluded that size and space parameters alone could discriminate between proficient and non-proficient handwriting [17]. In contrast, Asselborn et al. found that spatial features, such as the distance between letter strokes, did not provide a clear distinction between children with and without handwriting difficulties [1]. The high possibility of errors in spatial features due to incorrect brushstroke segmentation may explain these inconsistent findings. From our data, segmenting strokes based solely on distance or pressure is vulnerable to error as variable speed and pauses while drawing could falsely emulate a pen lift. It is, thus, recommended that future studies consider the accuracy of their segmentation algorithm or exclude the measures of inter-stroke distance altogether.

All other spatial features were moderately affected by incorrect segmentation. Brushstroke length measures exhibited an error rate between 11.1–28.8%. Again, this was also expected because the erroneous segmentation of one stroke as multiple brushstrokes would positively skew the measured segment length. Alternatively, quantifying brushstroke dimensions in height and width along the y- and x-axis, respectively, proved to be a more robust measurement as the measured errors were between 1.7–15.5%. This is likely due to the dependence of these measurements on the global image rather than individual segments. Some studies have reported correlations between spatial features, particularly those related to height, and writing proficiency [9,17,20]. However, given the sensitivity of these measures to segmentation, spatial features should be considered with caution as quantifiers of fine motor control.

4.3. Impact of Segmentation Errors on Temporal and Kinematic Features

Temporal and kinematic features were less affected by incorrect segmentation. This is a critical finding because these measurements are most commonly used in the literature and have consistently shown correlations to writing or drawing difficulties [1,5,20,22,23,25,27]. Since kinematic and temporal features have been correlated to fine motor control abilities across populations, and have been critical for developing machine learning algorithms for automatically detecting various pathological conditions [2,8,12].

The temporal features explored in this study were moderately but less affected by incorrect brushstroke segmentation as compared to spatial features. The propagated error, outside of the minimum in-air time, ranged from 0.0–49.2%. The temporal features explored in this study are commonly used in the literature and have proven discriminatory between different populations [2,11,15,17–19]. It must be noted that features that measured aggregated time, specifically total in-air time, total stop time, total on-screen time, and total movement time, generally demonstrated the smallest error, between 0.0–0.9%; these features have been associated with graphomotor pathologies in previous studies [2,6,7,11,18,19]. Our analysis supports that total time measures provide reliable insights into writing or drawing capabilities.

Kinematic features proved to be the least sensitive to misidentified brushstrokes. The small, propagated error could be attributed to the fact that these measures rely on instantaneous values. The only kinematic measure that was moderately affected was average acceleration which had a propagated error of 12.7%. This high error rate was due to trials 18 and 22 (validation dataset) with average acceleration errors of 25.9% and 32.2%, respectively, while the other trials (5, 13, and 14) had errors between 0.2–3.0%. The non-robustness in the average explains the sensitivity of these measures to a single outlying observation. The robustness of kinematic features to segmentation errors makes them ideal for assessing fine motor control in different populations.

Several biomechanical features (e.g., minimum brushstroke length, maximum brushstroke width, minimum on-screen time) exhibited large standard deviations in their average propagated errors. This observation can be traced to trial 22 (validation dataset), where the algorithm erroneously identified one brushstroke as two. However, it is important to note that this trial, where the participant traced the entire image with one brushstroke, was the exception. Of the 52 observed clinical trials, only one participant traced in this manner. Thus, features with large variance in propagated error may still be useful.

4.4. Instrumentation Considerations

Our findings suggest that some important technical characteristics of the data capture instrumentation must be considered to minimize propagated errors in biomechanical features in handwriting, drawing, or tracing studies. In this study, erroneous segmentation in the validation dataset could partially be attributed to the tablet's variable sampling rate. While most studies have presumed a fixed sampling rate, others have reported different rates during data collection [21] or have omitted this information altogether [3,5,28]. For the current study, the sampling rate variability could be attributed to the connection between

the server and tablets. While a variable sampling rate may be inherent to the systems used for data acquisition and, thus, unavoidable, this variability could negatively affect common features, such as fluency, entropy, or essential tremor, in the study of handwriting complexities [3,10,29]. For example, if one assumed a constant sampling period between data points when, in fact, the sampling period varied, then all instantaneous measures, such as velocity, that depend on time differences between samples would be inaccurate. Those errors would propagate to derived measures, such as fluency, and would, in turn, be further magnified in any cognate summary statistic, such as average acceleration.

The sampling rate itself is also an important consideration. For example, essential tremor is found to be between 4–12 Hz [29], requiring a minimum sampling rate of 24 Hz to accurately capture this behavior. A low sampling rate may compromise the detection of within-brushstroke pauses, which have been proposed as an important measurement for categorizing divergent drawing or handwriting capabilities [11,30]. Most studies reporting high (100–200 Hz) sampling rates have used specialized WACOM drawing tablets [4,13,20,22,25]. Studies deploying consumer-based tablets, such as the iPad with an Apple Pencil, report a much lower sampling rate of 60 Hz [1], or a Samsung tablet with an unspecified sampling rate [28]. Thus, when using cost-effective, consumer-grade tablets, special attention must be paid to the sampling rate and its consistency over time, and, consequently, the kinematic features that can be reasonably estimated.

Touchscreen technology is another important consideration. Specifically, our findings indicate an effect of hand position on brushstroke segmentation when using consumer-grade tablets outfitted with capacitive screens. Functionally, capacitive screens do not directly measure pen pressure. Rather, they register pen position based on contact with an electrical conductor such as a stylus pen tip or hand. As the current study demonstrated, pen proximity as well as the hand resting on the screen contributed noisy data points that distorted brushstrokes, which, in turn, could propagate to incorrect feature measurements. Since capacitive screens are inherently sensitive to pen proximity and hand position, it is plausible that the same effects were present in other studies using similar hardware. Therefore, it is critical that future studies consider denoising methods, such as the one proposed here, for removing spurious data points.

5. Conclusions

This paper presented a novel brushstroke segmentation algorithm based on spatiotemporal as well as pen pressure measurements. The algorithm robustly identified brushstrokes in both validation and clinical datasets. We showed that incorrect brushstroke segmentation can propagate grave errors to many spatial and selected temporal features commonly cited when characterizing graphomotor activities. Instrumental specifications such as consistency of sampling rate and nature of the touchscreen technology can also contribute to erroneous estimation of biomechanical features. The proposed algorithm can assist in segmenting tablet-acquired data for studies exploring fine motor control through handwriting, drawing, or tracing. Our findings suggest that future studies must provide greater transparency on segmentation approaches, consider spatial features, in particular, with prudence and attend to technical limitations of the acquisition tablet.

Author Contributions: Conceptualization, K.S.F. and T.C.; Methodology, K.S.F. and G.R.; Software, G.R. and K.S.F.; Validation, K.S.F. and G.R.; Formal Analysis, K.S.F.; Investigation, K.S.F. and G.R.; Resources, T.C.; Data Curation, K.S.F.; Writing—Original Draft Preparation, K.S.F.; Writing—Review and Editing, T.C. and K.S.F.; Visualization, K.S.F.; Supervision, T.C.; Funding Acquisition, T.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Sciences and Engineering Research Council of Canada Discovery Grant, grant number RGPIN-2019-06033.

Data Availability Statement: The datasets presented in this article are not readily available due to ethical restrictions. Requests to access the datasets should be directed to tom.chau@utoronto.ca.

Acknowledgments: The work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was granted by the Holland Bloorview Kids Rehabilitation Ethics Board (eREB#0117).

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Asselborn, T.; Chapatte, M.; Dillenbourg, P. Extending the Spectrum of Dysgraphia: A Data Driven Strategy to Estimate Handwriting Quality. *Sci. Rep.* **2020**, *10*, 3140. [[CrossRef](#)]
2. Asselborn, T.; Gargot, T.; Kidziński, Ł.; Johal, W.; Cohen, D.; Jolly, C.; Dillenbourg, P. Automated human-level diagnosis of dysgraphia using a consumer tablet. *npj Digit. Med.* **2018**, *1*, 42. [[CrossRef](#)]
3. Drotar, P.; Mekyska, J.; Rektorova, I.; Masarova, L.; Smekal, Z.; Faundez-Zanuy, M. Decision Support Framework for Parkinson's Disease Based on Novel Handwriting Markers. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2015**, *23*, 508–516. [[CrossRef](#)]
4. Fleury, A.; Kushki, A.; Tanel, N.; Anagnostou, E.; Chau, T. Statistical persistence and timing characteristics of repetitive circle drawing in children with ASD. *Dev. Neurorehabil.* **2013**, *16*, 245–254. [[CrossRef](#)]
5. Cohen, E.J.; Bravi, R.; Minciocchi, D. Assessing the Development of Fine Motor Control in Elementary School Children Using Drawing and Tracing Tasks. *Percept. Mot. Ski.* **2021**, *128*, 605–624. [[CrossRef](#)] [[PubMed](#)]
6. Dui, L.G.; Lomurno, E.; Lunardini, F.; Termine, C.; Campi, A.; Matteucci, M.; Ferrante, S. Identification and characterization of learning weakness from drawing analysis at the pre-literacy stage. *Sci. Rep.* **2022**, *12*, 21624. [[CrossRef](#)] [[PubMed](#)]
7. Rosenblum, S.; Dror, G. Identifying Developmental Dysgraphia Characteristics Utilizing Handwriting Classification Methods. *IEEE Trans. Hum.-Mach. Syst.* **2016**, *47*, 293–298. [[CrossRef](#)]
8. Lopez-De-Ipina, K.; Solé-Casals, J.; Faundez-Zanuy, M.; Calvo, P.M.; Sesa, E.; Roure, J.; Martinez-De-Lizarduy, U.; Beitia, B.; Fernández, E.; Iradi, J.; et al. Automatic Analysis of Archimedes' Spiral for Characterization of Genetic Essential Tremor Based on Shannon's Entropy and Fractal Dimension. *Entropy* **2018**, *20*, 531. [[CrossRef](#)]
9. Johnson, B.P.; Phillips, J.G.; Papadopoulos, N.; Fielding, J.; Tonge, B.; Rinehart, N.J. Understanding macrographia in children with autism spectrum disorders. *Res. Dev. Disabil.* **2013**, *34*, 2917–2926. [[CrossRef](#)]
10. Mekyska, J.; Faundez-Zanuy, M.; Mzourek, Z.; Galaz, Z.; Smekal, Z.; Rosenblum, S. Identification and Rating of Developmental Dysgraphia by Handwriting Analysis. *IEEE Trans. Hum.-Mach. Syst.* **2016**, *47*, 235–248. [[CrossRef](#)]
11. Paz-Villagrán, V.; Danna, J.; Velay, J.-L. Lifts and stops in proficient and dysgraphic handwriting. *Hum. Mov. Sci.* **2014**, *33*, 381–394. [[CrossRef](#)]
12. Parziale, A.; Senatore, R.; Della Cioppa, A.; Marcelli, A. Cartesian genetic programming for diagnosis of Parkinson disease through handwriting analysis: Performance vs. interpretability issues. *Artif. Intell. Med.* **2020**, *111*, 101984. [[CrossRef](#)] [[PubMed](#)]
13. Di Brina, C.; Niels, R.; Overvelde, A.; Levi, G.; Hulstijn, W. Dynamic time warping: A new method in the study of poor handwriting. *Hum. Mov. Sci.* **2008**, *27*, 242–255. [[CrossRef](#)] [[PubMed](#)]
14. Mavrogiorgou, P.; Mergl, R.; Tigges, P.; El Husseini, J.; Schröter, A.; Juckel, G.; Zaudig, M.; Hegerl, U. Kinematic analysis of handwriting movements in patients with obsessive-compulsive disorder. *J. Neurol. Neurosurg. Psychiatry* **2001**, *70*, 605–612. [[CrossRef](#)]
15. Rosenblum, S.; Dvorkin, A.Y.; Weiss, P.L. Automatic segmentation as a tool for examining the handwriting process of children with dysgraphic and proficient handwriting. *Hum. Mov. Sci.* **2006**, *25*, 608–621. [[CrossRef](#)]
16. Fitjar, C.L.; Rønneberg, V.; Nottbusch, G.; Torrance, M. Learning Handwriting: Factors Affecting Pen-Movement Fluency in Beginning Writers. *Front. Psychol.* **2021**, *12*, 663829. [[CrossRef](#)]
17. Falk, T.H.; Tam, C.; Schellnus, H.; Chau, T. On the development of a computer-based handwriting assessment tool to objectively quantify handwriting proficiency in children. *Comput. Methods Programs Biomed.* **2011**, *104*, e102–e111. [[CrossRef](#)]
18. Rosenblum, S.; Parush, S.; Weiss, P.L. Computerized Temporal Handwriting Characteristics of Proficient and Non-Proficient Handwriters. *Am. J. Occup. Ther.* **2003**, *57*, 129–138. [[CrossRef](#)] [[PubMed](#)]
19. Rosenblum, S. Development, Reliability, and Validity of the Handwriting Proficiency Screening Questionnaire (HPSQ). *Am. J. Occup. Ther.* **2008**, *62*, 298–307. [[CrossRef](#)]
20. Van Gemmert, A.W.A.; Adler, C.H.; Stelmach, G.E. Parkinson's disease patients undershoot target size in handwriting and similar tasks. *J. Neurol. Neurosurg. Psychiatry* **2003**, *74*, 1502–1508. [[CrossRef](#)] [[PubMed](#)]
21. Caligiuri, M.P.; Teulings, H.-L.; Dean, C.E.; Niculescu, A.B.; Lohr, J.B. Handwriting movement kinematics for quantifying extrapyramidal side effects in patients treated with atypical antipsychotics. *Psychiatry Res.* **2010**, *177*, 77–83. [[CrossRef](#)]
22. Broderick, M.P.; Van Gemmert, A.W.A.; Shill, H.A.; Stelmach, G.E. Hypometria and bradykinesia during drawing movements in individuals with Parkinson's disease. *Exp. Brain Res.* **2009**, *197*, 223–233. [[CrossRef](#)]
23. Kushki, A.; Schwellnus, H.; Ilyas, F.; Chau, T. Changes in kinetics and kinematics of handwriting during a prolonged writing task in children with and without dysgraphia. *Res. Dev. Disabil.* **2011**, *32*, 1058–1064. [[CrossRef](#)]
24. Fitjar, C.L.; Rønneberg, V.; Torrance, M. Assessing handwriting: A method for detailed analysis of letter-formation accuracy and fluency. *Read. Writ.* **2022**, *37*, 291–327. [[CrossRef](#)]

25. Danna, J.; Paz-Villagrán, V.; Velay, J.-L. Signal-to-Noise velocity peaks difference: A new method for evaluating the handwriting movement fluency in children with dysgraphia. *Res. Dev. Disabil.* **2013**, *34*, 4375–4384. [[CrossRef](#)] [[PubMed](#)]
26. Schenk, T.; Bauer, B.; Steidle, B.; Marquardt, C. Does training improve writer's cramp?: An evaluation of a behavioral treatment approach using kinematic analysis. *J. Hand Ther.* **2004**, *17*, 349–363. [[CrossRef](#)]
27. Kushki, A.; Chau, T.; Anagnostou, E. Handwriting Difficulties in Children with Autism Spectrum Disorders: A Scoping Review. *J. Autism Dev. Disord.* **2011**, *41*, 1706–1716. [[CrossRef](#)] [[PubMed](#)]
28. Degtyarenko, I.; Radyvonenko, O.; Bokhan, K.; Khomenko, V. Text/shape classifier for mobile applications with handwriting input. *Int. J. Doc. Anal. Recognit.* **2016**, *19*, 369–379. [[CrossRef](#)]
29. López-De-Ipiña, K.; Solé-Casals, J.; Faundez-Zanuy, M.; Calvo, P.M.; Sesa, E.; de Lizarduy, U.M.; De La Riva, P.; Marti-Masso, J.F.; Beitia, B.; Bergareche, A. Selection of Entropy Based Features for Automatic Analysis of Essential Tremor. *Entropy* **2016**, *18*, 184. [[CrossRef](#)]
30. Reinders-Messelink, H.; Schoemaker, M.; Snijders, T.; Göeken, L.; Bökkerink, J.; Kamps, W. Analysis of handwriting of children during treatment for acute lymphoblastic leukemia. *Med. Pediatr. Oncol.* **2001**, *37*, 393–399. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.