Cognitive Computation (2025) 17:13
https://doi.org/10.1007/s12559-024-10360-7

RESEARCH

Computer-Aided Diagnosis of Graphomotor Difficulties Utilizing Direction-Based Fractional Order Derivatives

Michal Gavenciak[1] · Jan Mucha[1,3] · Jiri Mekyska[1] · Zoltan Galaz[1] · Katarina Zvoncakova[2] · Marcos Faundez-Zanuy[3]

Received: 3 February 2024 / Accepted: 3 November 2024

Abstract
Children who do not sufficiently develop graphomotor skills essential for handwriting often develop graphomotor disabilities (GD), impacting the self-esteem and academic performance of the individual. Current examination methods of GD consist of scales and questionaries, which lack objectivity, rely on the perceptual abilities of the examiner, and may lead to inadequately targeted remediation. Nowadays, one way to address the factor of subjectivity is to incorporate supportive machine learning (ML) based assessment. However, even with the increasing popularity of decision-support systems facilitating the diagnosis and assessment of GD, this field still lacks an understanding of deficient kinematics concerning the direction of pen movement. This study aims to explore the impact of movement direction on the manifestations of graphomotor difficulties in school-aged. We introduced a new fractional-order derivative-based approach enabling quantification of kinematic aspects of handwriting concerning the direction of movement using polar plot representation. We validated the novel features in a barrage of machine learning scenarios, testing various training methods based on extreme gradient boosting trees (XGBboost), Bayesian, and random search hyperparameter tuning methods. Results show that our novel features outperformed the baseline and provided a balanced accuracy of 87 % (sensitivity = 82 %, specificity = 92 %), performing binary classification (children with/without graphomotor difficulties). The final model peaked when using only 43 out of 250 novel features, showing that XGBoost can benefit from feature selection methods. Proposed features provide additional information to an automated classifier with the potential of human interpretability thanks to the possibility of easy visualization using polar plots.

Keywords Feature extraction · Graphomotor difficulties · Fractional order derivatives · Polar plot · Computer-aided diagnosis · Machine learning

Introduction

Prior to embarking on the task of handwriting, a child must first acquire the ability to draw [1]. Typically, between infancy and the age of six, a child undergoes a developmental process that involves cultivating a range of motor and non-motor skills. This includes developing motor planning and execution, visual-perceptual abilities, orthographic coding, kinesthetic feedback, and visual-motor coordination [2–4]. These skills, collectively known as graphomotor skills (GS) [2, 5], become automated around the age of 8 to 9 [6]. They serve as the building blocks for drawing and, subsequently, handwriting, accompanying individuals throughout their lives.

Children who do not sufficiently develop graphomotor skills and fail to master drawing and handwriting despite having adequate cognitive capacity, access to learning opportunities, and no neurological issues, often develop graphomotor disabilities (GD). GD are closely associated with the term developmental dysgraphia (DD); however, DD typically encompasses additional higher-level cognitive functions. In fact, the presence of GD can lead to a diagnosis of DD, particularly in grades 3 to 4 when handwriting should become automatic and proficient [6]. The prevalence of GD among school-aged children ranges between 7 and 34 % [6, 7], with boys being diagnosed with GD 2 to 3 times more frequently than girls [8].

Children spend up to 60 % of their time at school performing handwriting and other fine motor tasks [9]. Despite the significant negative impact of GD/DD on children's self-esteem, academic performance, and overall quality of life [2,

Extended author information available on the last page of the article

Published online: 27 November 2024

3], the diagnosis/screening of GD/DD is currently primarily subjective, e.g., based on the assessment of scales such as the Handwriting Proficiency Screening Questionnaire (HPSQ) [10], the shortened version of the Concise Assessment Methods of Children Handwriting (SOS: BHK) [11], or the Handwriting Legibility Scale (HLS) [12]. Besides the subjectivity, assessment based on these scales has some other limitations, e.g., they lack sufficient inter-rater consensus [13], or/and they rely on the perceptual abilities of a rater. As a result, this subjective nature of diagnosis can lead to poorly targeted remediation, thereby resulting in negative consequences for the affected children themselves.

It has been experimentally shown that one way to address the factor of subjectivity is to incorporate supportive machine learning (ML) based assessment into the diagnostic chain [14–17]. This approach was first pioneered in 2016. For example, Mekyska et al. automatically stratified a cohort of 54 children attending the third grade of a primary school (27 intact and 27 diagnosed with DD; all participants performed a simple graphomotor task). The authors utilized a random forest (RF) classifier, achieving 96 % sensitivity and specificity in the leave-one-out cross-validation policy [15]. Similarly, in the same year, Rosenblum et al. analyzed the handwriting of 99 third graders (50 intact and 49 diagnosed with DD). They were able to discriminate between both groups with 90 % sensitivity and specificity using an SVM (support vector machine) classifier in a 10-fold cross-validation approach [18]. Asselborn et al. employed quantitative analysis of online handwriting in a group of 298 children attending first to the fifth grade of a primary school (242 intact and 56 diagnosed with DD). Using the RF classifier in a 25-fold cross-validation approach, they achieved 96.6 % sensitivity and 99.2 % specificity [19].

Since 2018, the field of computer-aided diagnosis of GD/DD has experienced increasing popularity. Researchers have explored different pre-processing steps (e.g., intra-writer normalization [20]), parameterization algorithms (e.g., those based on the tunable Q-factor wavelet transform [21] or modulation spectra [22]), and handwriting/graphomotor tasks [23]. Attention has also been given to the advancement of the ML part of decision support systems. Recent solutions are based on various techniques, such as XGBoost (extreme gradient boosting) [24], adaptive boosting (AdaBoost) [25], convolutional neural network (CNN) [26, 27], vision transformer (ViT) [28], non-discrimination regularization in rotational region convolutional neural network (NDR-R2CNN) [29]. Some of the algorithms are applied directly to handwriting time-series, while others are applied to the image representation (also known as offline handwriting). For further information, we refer to the following reviews [14, 17].

Drawing/handwriting is a complex activity that requires coordinated movement of 43 muscles and joints of the arm, hand, and fingers [30]. Kushki et al. observed that, compared to the wrist system, the finger system is more susceptible to the effects of psychological and muscular fatigue [3]. From an anatomical perspective, the finger system is more involved in the vertical movement of a pen, which is more complex due to the involvement of a higher number of muscles and joints, including the interphalangeal and metacarpophalangeal joints [31]. Since vertical movement requires finer flexions and extensions, it could potentially accentuate poor kinematics resulting from GD [32]. This has been proven in several studies, where features such as velocity, acceleration, and jerk extracted from the vertical projection of pen movement have provided higher accuracies when discriminating between intact children and children diagnosed with GD/DD, or have shown better correlations with scores of, e.g., the Handwriting Proficiency Screening Questionnaire for Children [22, 23]. Nevertheless, although previous studies suggest that specific movement directions could accentuate some manifestations of GD/DD, the body of research has been limited to vertical and horizontal projections. To the best of our knowledge, no one has explored the kinematic abilities of children experiencing GD/DD with respect to the full range of angles. For example, we have no knowledge about the kinematic abilities of these children when performing diagonal movements, such as in the upper loops (spring) task.

Since handwriting is a result of several interacting physiological mechanisms, children with deficient fine motor skills, poor dexterity, poor muscle tone, or unspecified motor clumsiness could introduce irregularity and increased complexity into online handwriting signals [15]. Even though conventional kinematic measures such as velocity or acceleration could potentially quantify these alterations [14, 19, 33], they have limitations. These measures operate in the integer domain, which limits their ability to capture more nuanced variations in the handwriting signal. Additionally, the derivation is a noisy operator [34], amplifying fluctuations in the data, which can lead to less robust quantification of complex movements. Furthermore, differential measures are local operators, meaning they only focus on immediate changes in the signal, making them less suited for capturing broader patterns or dependencies in handwriting. This is why researchers introduced a novel framework for advanced kinematic analysis of drawing/handwriting based on fractional order derivatives (FD) [24, 35–37], offering non-integer and non-local derivation operators.

Generally, FD falls into the field of fractional calculus (FC), which is the theory of integrals and derivatives of an arbitrary order [38]. Its concept was introduced almost at the same age as the well-known differential, integral, or other standard calculus [39]. It attracted the interest of many famous mathematicians, such as Euler, Liouville, Laplace, Abel, Leibniz, Riemann, Grünwald, and Letnikov. The principles of FC have been used in the modeling of many physical

and chemical processes, as well as in modern engineering and science in general. Recently, the potential of FC has been analyzed in computer vision, such as image restoration, super-resolution, image segmentation, and motion estimation [40–42]. Moreover, fractional artificial neural networks (FANNs) have been examined for advanced modeling of various systems [43], including glucose levels from blood samples [44], noise removal from EEG signals [45], and disease modeling (human immunodeficiency virus — HIV, or malaria [46, 47]). More possibilities of FC employment in machine learning have been explored by [48], where authors proposed a fractional generative adversarial network (FGAN), and by [49], where the authors explored the potential of Caputo's operator in backpropagation for neural networks.

To sum up, despite the increasing popularity of decision-support systems facilitating the diagnosis and assessment of GD/DD, this field still has several knowledge gaps, such as a lack of understanding about deficient kinematics in relation to the direction of pen movement. To address the above-mentioned limitations and bridge the knowledge gap, the goal of this study is to explore the impact of movement direction on the manifestations of GD. More specifically, we aim to:

1. Introduce a new FD-based approach enabling quantification of kinematic aspects of handwriting with respect to the direction of movement,
2. Explore how the GD manifest in the newly proposed features (digital endpoints),
3. Evaluate discrimination power of the features in computer-aided diagnosis.

The rest of the paper is structured as follows: in the "Methodology" section, the dataset, feature extraction, and selected statistical methods are explained. In the "Results" section, the results are summarized. The discussion can be found in the "Discussion" section and the conclusions are drawn in the "Conclusion" section.

Methodology

Dataset

Altogether, our dataset consists of online handwriting data from 106 Czech-speaking children (66 boys and 40 girls) attending the third and fourth grades of several elementary schools throughout the Czech Republic. Thirteen of these children were diagnosed with DD. All the children were assigned a Handwriting Deficiency Criterion (HDC) score [6]. HDC is a scoring method that combines scores of the Handwriting Proficiency Screening Questionnaire for Children (HPSQ–C) [50, 51] with an evaluative score provided by an expert remedial teacher (OEE) and can take integer values in the range of [0; 4], where 0 denotes intact handwriting and 4 denotes severely deficient handwriting [6]. For the purposes of our study, we used a threshold to categorize the subjects into two classes based on the HDC score, resulting in a thresholded HDC score (HDC–T). We classified children with HDC scores of 0 or 1 as intact, and those with HDC scores above 1 as having moderate to severe disability. The breakdown of the HDC–T score can be seen in Table 1 and a more thorough description of the HDC evaluation process, including the OEE and HPSQ-C rating process can be seen in a paper by Mekyska et al. [6]. Overall, the dataset consists of 32 samples classified with an HDC–T of 1. The demographic data of the participants can be seen in Table 2.

Table 1 Handwriting disabilities criterion

OEE	OEE(t)	HPSQ–C(t)	HDC	HDC–T
0	0	0	0	0
1	0	0	0	
2	0	0	0	
2	0	1	0	
3	1	0	1	0
3	1	1	2	1
4	1	0	2	
4	1	1	3	1

[1] *OEE*, overall expert evaluation; *OEE(t)*, overall expert evaluation (thresholded); *HPSQ–C(t)*, child's self-evaluation (thresholded); *HDC*, new handwriting disabilities criterion; *HTC-T*, handwriting disabilities criterion (thresholded)

All children were asked to copy the upper loops (spring) task, which can be seen in Fig. 1. During the acquisition phase, the children were shown the figure printed on an A4 sheet of paper. Then the children were asked to copy the figure on another sheet of paper that was securely fixed to a digitizing tablet (Wacom Intuos Pro L PHT-80) with a sampling frequency of 150 Hz. Using the Wacom Inking Pen, the children were instructed to copy the figure at their own comfortable pace. This setup provides the following benefits:

Table 2 Demographic data

	3rd grade				
Gender	N	DD	Age [y]	HDC	HDC>1
Boys	17	3	8.75±0.56	1.06±1.02	5
Girls	16	2	8.91±0.59	0.75±0.86	4
Total	33	5	8.83±0.57	0.91±0.95	9
	4th grade				
Gender	N	DD	Age [y]	HDC	HDC>1
Boys	49	8	9.90±0.53	1.22±1.06	20
Girls	24	0	9.90±0.51	0.33±0.70	3
Total	73	8	9.90±0.52	0.93±1.04	23

[1] *N*, number of samples; *DD*, developmental dysgraphia; *y*, years; *HDC*, handwriting deficiency criterion

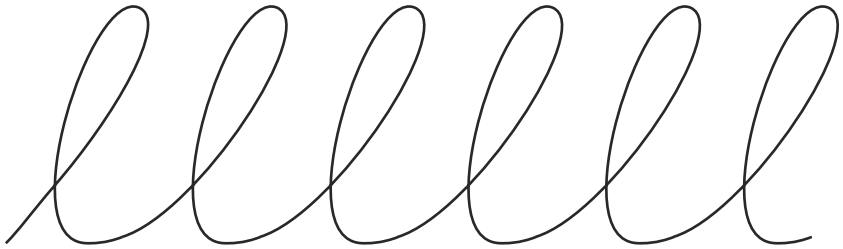


Fig. 1 Template for the upper loops task (real width = 220 mm, height = 65 mm)

- Using the inking pen in combination with the paper provides a child with the usual tactile feedback during handwriting, as well as immediate visual feedback for both the child and an examiner (who administered the test).
- The digitizing tablet is capable of sampling the handwriting process, providing a wealth of data as a discrete function of time, including position (x and y coordinates), pressure, tilt, and azimuth. Moreover, it records the movement of the tip of the pen up to 1.5 cm above the surface of the tablet, i.e., the in-air movement.

Examples of a couple of drawn tasks can be seen in Fig. 2. The project this study is a part of was approved by the Ethics Committee of the Masaryk University and parents of all children enrolled into this study signed an informed consent form. Throughout the whole study, the Ethical Principles of Psychologists and Code of Conduct released by the American Psychological Association were followed [52]. The methodology overview can also be seen in Fig. 3.

Fractional Order Derivatives

This study's fundamentals lie in using fractional order derivatives in the feature extraction process. The employment of FD in handwriting and drawing parameterization was developed to substitute the conventional differential derivatives in order to improve the quantitative analysis of GD [22, 24, 36, 53–55].

While numerous fractional derivative (FD) definitions exist, the most significant contributions to the field relevant to our study come from M. Caputo's work [56]. An advantage of Caputo's FD approach is based on the unnecessity to define the initial FD conditions [38, 57]. Furthermore, Caputo's FD exhibits the property of temporal memory, which can be influenced by the nature of the handwriting signal and related GD. Generally, most of the operators consist of two mathematical operations: convolution and differentiation. The Caputo derivative differentiates input data before the convolution operation is applied. More specifically, it transforms the input handwritten signal into its velocity before the effect of temporal memory is applied to the hand's movement velocity. The right-sided Caputo's fractional operator definition from 1967 is:

$$ {}^{C}\mathrm{D}^{\alpha} z(t) = \frac{1}{\Gamma(n-\alpha)} \int_{0}^{t} (t-\tau)^{n-\alpha-1} z^{n}(\tau)\, \mathrm{d}t, \tag{1} $$

where ${}^{C}\mathrm{D}^{\alpha} z(t)$ denotes the Caputo derivatives of order α of the function $z(t)$, Γ is the gamma function, and $n-1 < \alpha \leq n, n \in \mathbf{N}, t > 0$.

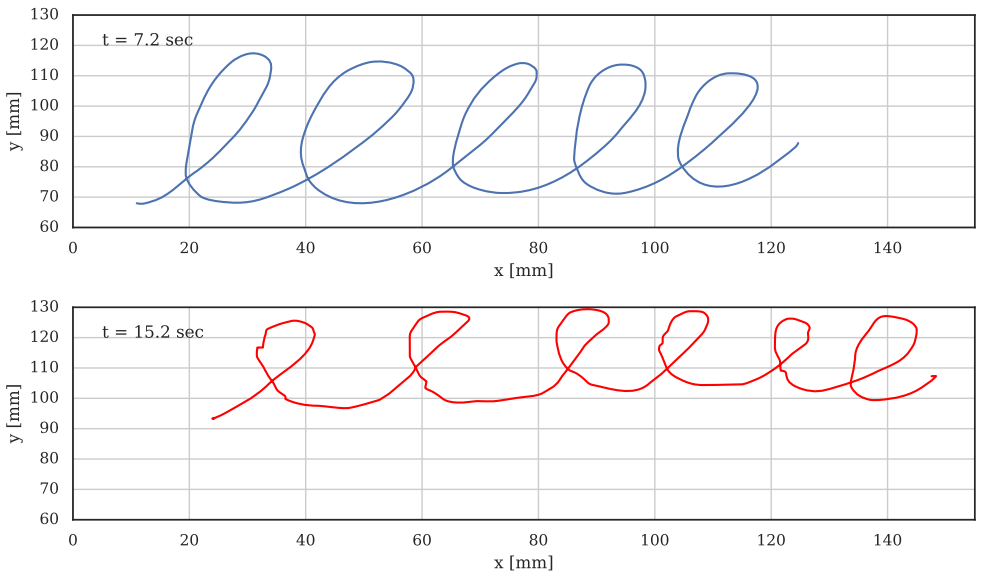


Fig. 2 Upper loops performed by an intact child (top in blue) and a child diagnosed with GD (bottom in red)

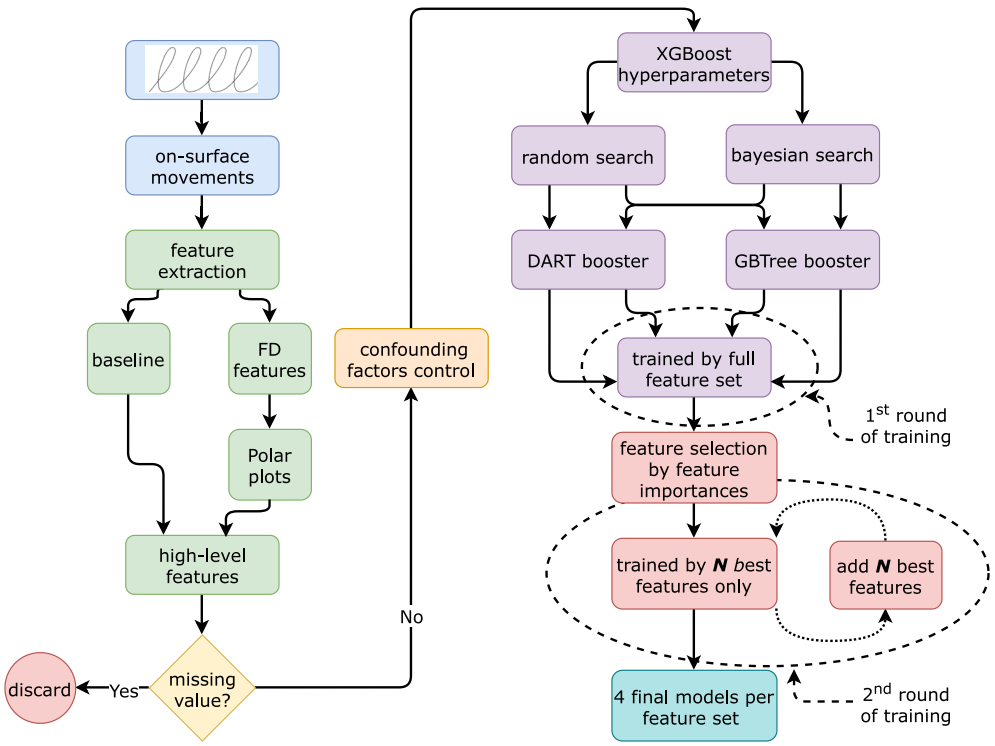


Fig. 3 An overview of the methodology applied in the study

Feature Extraction

In order to enhance the accuracy of automated GD diagnosis, we introduce novel features based on FD and polar plots. We extracted the positional data from the sampled handwriting of each child, yielding vectors of x and y coordinates. These coordinates served as inputs to calculate the FD for $\alpha = [0.1; 1]$ with a step size of 0.1, resulting in ten different vectors for each processed exercise. It is important to note that with $\alpha = 1$, we obtain a full derivative, representing the velocity of the pen tip. Due to the property of temporal memory, this computation produces high peak values at points of discontinuity in the underlying handwriting, such as the beginning of the exercise and typically when the child lifted the pen off the digitizing surface. These high-peak values inside each computed vector were removed as outliers

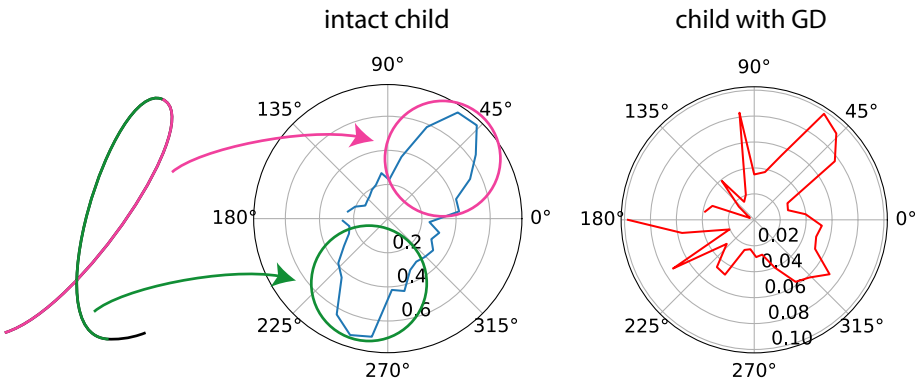


Fig. 4 An example of FD data extracted from a child in the intact group (displayed in blue on the left) and from the group associated with GD (shown in red on the right) can be seen in polar plots with ten-degree binning and the α value of 0.1. It is evident that in the case of the intact child, the majority of movements were carried out in approximately the directions of 50° and 260°, corresponding to the long and relatively straight sections of the upper loop task. Conversely, the movement of the child diagnosed with GD appears to be distributed more chaotically across the plot

during feature extraction. In summary, we extracted 20 feature vectors (10 for the x-axis and 10 for the y-axis), which were subsequently processed into the proposed novel handwriting features.

Next, we calculated the direction of movement for every two consecutive samples by determining the angle between points $p_n[x_n; y_n]$ and $p_{n+1}[x_{n+1}; y_{n+1}]$. The fixed line, relative to which the angle was calculated, is a right-facing, horizontal line, denoted as the 0°angle in Fig. 4. Combining this vector of angles with the derivatives provides a representation of the movement dependent on direction. For example, in the case of $\alpha = 1$, we obtain vectors representing both the velocity and the direction of the corresponding movement. The purpose of this combination is to assess the children's ability to execute movements in various directions.

To evaluate the discrimination power of these novel features, we also extracted a set of conventionally used baseline kinematic parameters [6, 14, 16, 25], specifically based on velocity and angular velocity [58]. These measures were extracted from the global trajectory as well as the horizontal and vertical projections of the movement. To transform these time series into scalar values, we used the median, 95th percentile, and non-parametric coefficient of variation defined as iqr/median, where iqr stands for the interquartile range. Finally, we included the number of changes in the velocity profile and its relative form. In summary, we extracted 25 baseline velocity-based features. A complete list of the computed features is provided in the supplementary material (see Supplementary File S1).

Polar Plots

One of the goals of handwriting featurization is undoubtedly the extraction of features that are easily interpretable by non-technicians, such as psychologists or remedial teachers, who conduct clinical evaluations. Such features are beneficial to the overall evaluation process as they provide more insight into the handwriting, even without the use of an automated

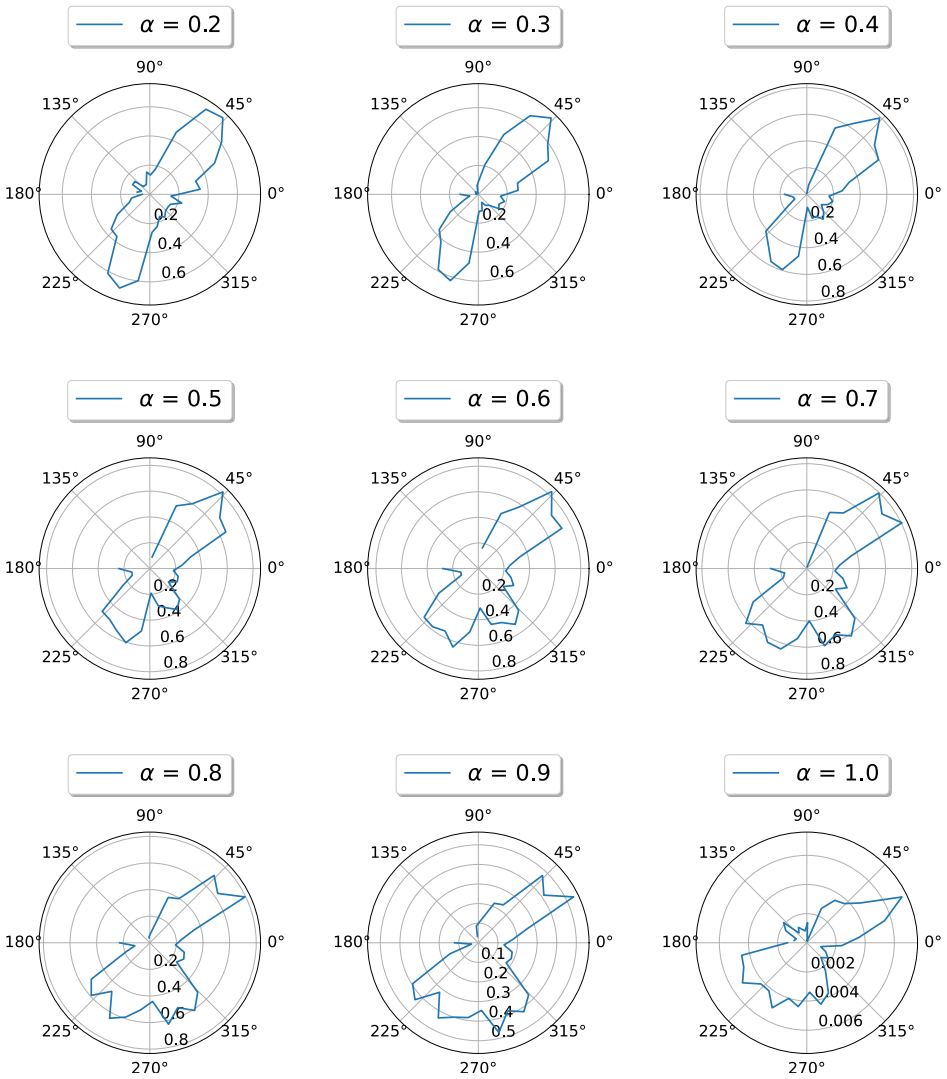


Fig. 5 The figure displays the progression of the FDE with α values ranging from 0.2 to 1.0 in multiple polar plots for the intact child

evaluator, and allow experts to link these features to specific manifestations of graphomotor difficulties.

Since the FD-based and direction-based features are represented by vectors, we have chosen to use polar plots to interpret the data. Polar plots are typically used to visualize data that is a function of angles. However, in our dataset, there can be more than one FD data point for a specific movement direction. To address this, we employed a pooling method in which we grouped the data into angle bins. Specifically, we chose a step size of 0.1°, resulting in 3600 bins. These bins were then processed using the following methods:

1. Calculating the median of data points in each bin (Bin median),
2. Calculating the 95th percentile of data points in each bin (Bin percentile),
3. Calculating the interquartile range (IQR) of data points in each bin (Bin IQR).

Additionally, we also calculated the number of elements of each bin and included this metric in the features, however, as the number of datapoints is equal across different α they were calculated only once.

Examples of polar plots (one for an intact child and another for a child diagnosed with GD) created using the second binning method with FD data can be seen in Fig. 4. Furthermore, corresponding polar plots with α values ranging from 0.2 to 1.0 can be found in Figs. 5 and 6.

To easily process these polar plots computationally, we need to represent them by scalar values. Still in the idea of easy interpretability, we took the approach of describing the plots using simple quantifiers, i.e., the total area of the plot and its centroid in different coordinate systems (Cartesian $[x, y]$ and polar $[r, \theta]$, where r is the distance from the center of the plot, and θ is the angle in relation to a line going from the center of the graph to the right). Considering that we can create at least two different plots for each α value, based on the two different bin processing methods we used, we can

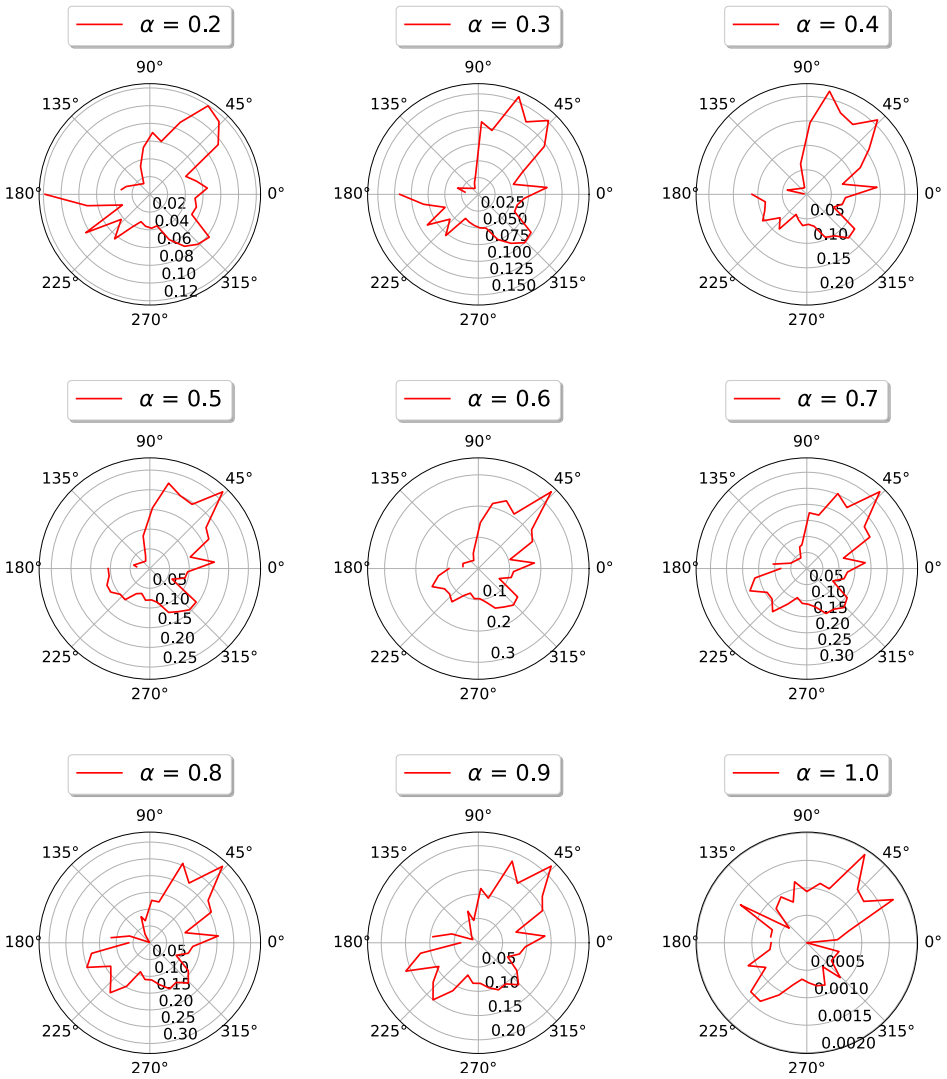


Fig. 6 The figure displays the progression of the FDE with α values ranging from 0.2 to 1.0 in multiple polar plots for the child diagnosed with GD

extract these descriptors for both of them. Additionally, even though the raw vector data cannot be plotted, we can still calculate the same descriptive values. The extracted features, along with their labels, can be seen in Table 3.

Even though the centroid-based features roughly describe the dominant movement direction, they might not fully capture the nature of the plot's distribution. For instance, a plot with a strong and narrow peak in one direction might result in the same centroid position as a plot with smaller values distributed more widely around the same direction. Furthermore, if a plot is more or less balanced around the zero-point, the centroid will be located near the center of the plot.

While it might seem redundant to include the coordinates of the centroid in two systems, the information each coordinate carries is fundamentally different. The Cartesian system captures the centroid's horizontal and vertical position, describing the balance between left-right and up-down movements. The polar system clearly shows the distance from the center (indicating the strength of dominance in a specific movement direction) and its direction.

Finally, we also extended the featureset by processing the vectors in bins with a larger size of 45°, creating eight sections (octants).

To sum up, each FD order allowed us to construct three polar plots by binning the vectors with 0.1° step while utilizing the three bin processing methods (Taking the median, 95th percentile, and IQR of each bin). From each polar plot, we then extract five features (Area, CentroidX, CentroidY, CentroidTheta, and CentroidR). These five features can be extracted from the raw vectors. Furthermore, this process extraction process is then repeated for a different bin size of 45°. After the removal of features containing NaN values and/or outliers, we reached the number of 250 features. For a detailed breakdown of the features, please refer to Supplementary File S1.

Machine Learning

To discriminate between intact children and those diagnosed with GD, we employed the XGBoost algorithm [59]. XGBoost (eXtreme Gradient Boosting) is a machine learning technique that builds an ensemble of decision trees through a boosting process, where each subsequent tree attempts to correct the errors of the previous ones. Its advantages include handling imbalanced datasets, high computational efficiency, and the ability to naturally perform feature selection by assigning importance scores to features [60]. These characteristics make XGBoost particularly well-suited for our study, as it allows us to efficiently handle the complex and potentially imbalanced dataset of handwriting features while identifying the most relevant kinematic parameters for distinguishing between the groups of children.

Prior to training, we controlled for confounding factors (i.e., sex, age, and class) by training a linear regression model on each feature with respect to each confounding factor. Only the residuals from these regressions were used for further analysis [61]. To be able to evaluate the performance of the novel features compared to the baseline feature set, we trained our models in three scenarios:

- **Baseline** — using velocity-based baseline features (25 features),
- **Novel** — using the novel features proposed in the previous section (250 features),
- **Combined** — combining the novel and baseline feature sets into one, measuring their combined performance (275 features).

Additionally, we trained multiple models with the same dataset, varying their settings to compare their suitability for our use case. The key differences we compared in the XGBoost variants were:

- Hyperparameter tuning strategies, where we compared Bayesian search with random search,
- Gradient boosting strategies, where we compared the GBtree (gradient boosted tree) algorithm with the DART (dropout additive regression trees) booster. GBTree is one of the core algorithms used in XGBoost, where gradient boosting is applied to decision trees. DART is an advanced version of boosting that uses dropout to reduce overfitting by randomly dropping trees during training.

Table 3 Features extracted from the polar plots

Feature basis	Raw vector	Bin percentile
Area	alpha-[α]_Area	Q[p]_alpha-[α]_Area
Centroid X	alpha-[α]_CentroidX	Q[p]_alpha-[α]_CentroidX
Centroid Y	alpha-[α]_CentroidY	Q[p]_alpha-[α]_CentroidY
Centroid Theta	alpha-[α]_CentroidTheta	Q[p]_alpha-[α]_CentroidTheta
Centroid R	alpha-[α]_CentroidR	Q[p]_alpha-[α]_CentroidR

[1] alpha-[α], the order of the fractional derivative used; Q[p], the percentile based value extracted from the bin (*Q0.5* denotes median, *Q0* denotes IQR, *Q0.95* denotes 95th percentile). For example, Q0_alpha−0.1_Area is the code of a feature calculated as the area of a plot, where each data point is the IQR of a bin from fractional order derivative of order 0.1

Table 4 Hyperparameter search space

Hyperparameter	DART		GBTree	
	Min	Max	Min	Max
Learning rate	0.01	0.5	0.01	0.5
Max depth	4	8	4	8
Subsample	0.6	1.0	0.6	1.0
ColSample (level)	0.1	1	0.1	1
ColSample (tree)	0.1	1	0.1	1
ColSample (node)	0.1	1	0.1	1
Min child weight	0.1	3	0.1	3
Negative class ratio	1	7	1	7
Drop rate (DART)	0	0.15	0	1
Drop skip (DART)	0	0.2	0	1
Number of trees	50	250	50	1000

Random search is an uninformed hyperparameter tuning algorithm that treats every set of hyperparameter values separately. The values were randomly selected from a predefined grid, with the number of iterations set to 1000. While random search is generally faster than typical grid search, which systematically iterates over the entire search space, it carries the potential risk of missing the hyperparameter combination that would result in the most performant model. This also brings the apparent disadvantage stemming from the fact that random search will not explore areas around well-performing hyperparameter combinations to potentially find even better sets.

On the other hand, Bayesian search is an informed tuning algorithm, as it takes into account the performance of its previous iterations. While there are multiple possible implementations of Bayesian search, a surrogate model is created and optimized in principle rather than the objective function itself. After each iteration, selected hyperparameters are tested on the objective function, and the resulting performance is used to update the surrogate model. While the initial iterations might resemble the random search algorithm, Bayesian search will gradually start optimizing the model and converge to a solution. Overall, a Bayesian search should require fewer iterations to reach a comparable or better result than a random search.

The DART algorithm is based on providing a dropout technique to the more general tree-boosting machine learning method to address overfitting. After every boosting iteration, a proportion of the trees is dropped [62]. As such, the DART booster introduces additional hyperparameters: drop rate, which represents the proportion of dropped trees, and drop skip, which is the probability that dropout will be skipped. Both hyperparameters can be set in the range of [0, 1], with a zero drop rate indicating no dropouts and thus effectively functioning as a general GBTree booster.

Before hyperparameter tuning, we upsampled the minority class. The models were all trained for binary classification, modeling the HDC–T value, using repeated stratified k-fold cross-validation with 5 repeats and 5 folds. After hyperparameter tuning, we extracted the ordered list of the most important features from the trained model. We then trained the model again on a subset of n most important features, in the range from 10 features up to the whole featureset. Effectively, we trained each model twice, once on the whole featureset and then again, using the already best-performing hyperparameters, on a subset of features that proved beneficial to the model's performance. Even though XGBoost should be able to select only the most relevant features, we still wanted to reduce the complexity of the final model. Hyperparameters [59] were searched within the ranges shown in Table 4. Model performance was evaluated in terms of balanced accuracy (BACC), sensitivity (SEN), and specificity (SPE). The final model performance was evaluated using leave-one-out cross-validation. We also conducted a post-hoc ROC analysis and further tuned the decision threshold to achieve a better trade-off between SEN and SPE. After this step, the evaluation measures are denoted as BACC-pos, SEN-pos, and SPE-pos. Finally, in order to interpret the models and gain insights, we explored the feature importance. Hyperparameters for the DART and GBTree boosters were mostly set to the same range, except for the *number of trees* parameter, which was set to a smaller range

Table 5 Classification results for the GBTree booster-based models

Scenario	Tuner	BACC	SEN	SPE	BACC-pos	SEN-pos	SPE-pos	N
Baseline	R	0.82	0.86	0.77	0.83	0.77	0.89	13
	B	0.76	0.82	0.69	0.81	0.80	0.82	17
Novel	R	0.82	0.88	0.77	0.83	0.88	0.78	46
	B	**0.79**	0.86	0.72	**0.87**	0.82	0.92	43
Combined	R	0.78	0.81	0.74	0.85	0.77	0.93	106
	B	**0.78**	0.84	0.72	**0.86**	0.80	0.92	61

[1] *BACC*, balanced accuracy; *SEN*, sensitivity; *SPE*, specificity; *BACC-pos*, balanced accuracy after decision threshold adjustment; *SEN-pos*, sensitivity after decision threshold adjustment; *SPE-pos*, specificity after decision threshold adjustment; *N*, number of features used to train the model

Table 6 Classification results for the DART booster-based models

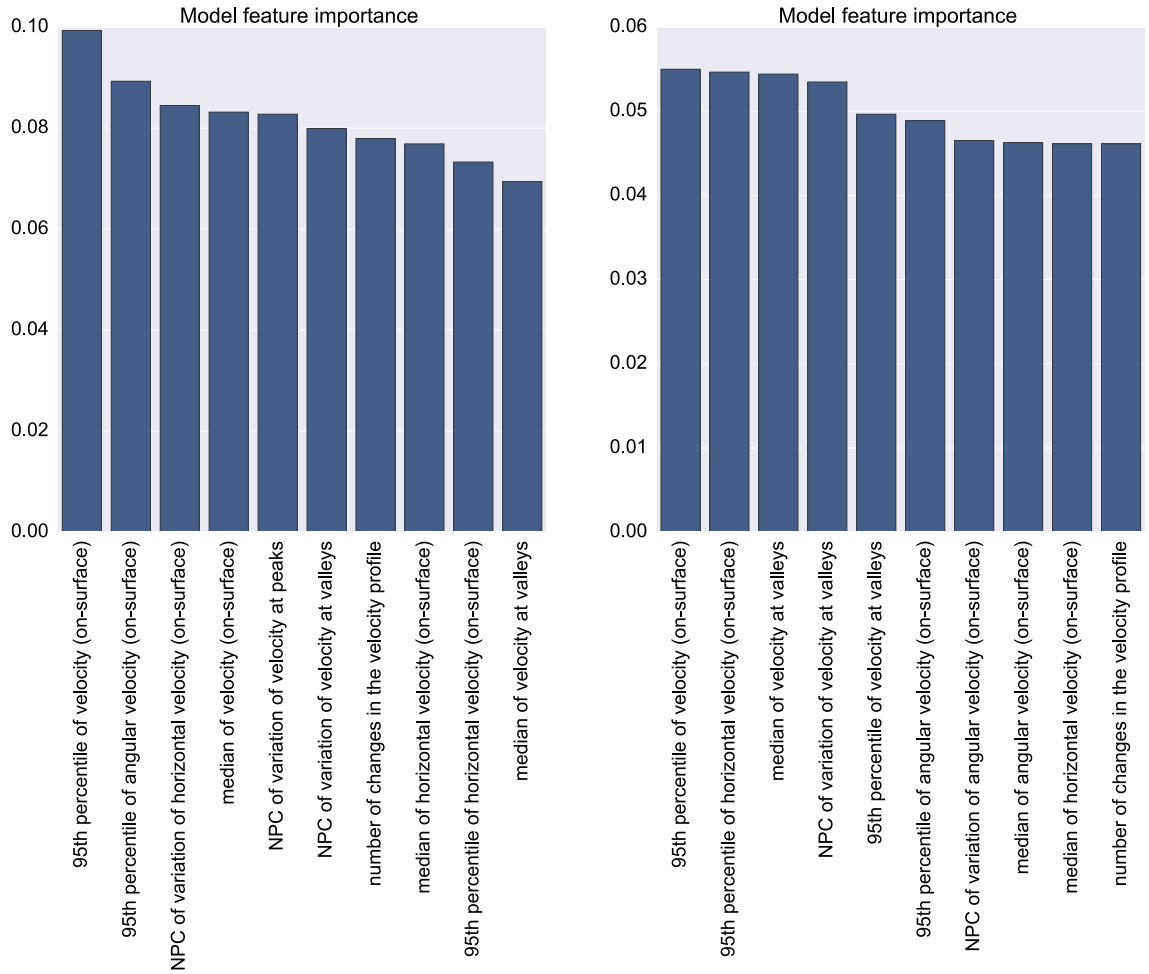
Scenario	Tuner	BACC	SEN	SPE	BACC-pos	SEN-pos	SPE-pos	N
Baseline	R	0.80	0.86	0.73	0.81	0.81	0.81	14
	B	**0.81**	0.88	0.74	**0.84**	0.88	0.80	22
Novel	R	0.78	0.84	0.72	0.83	0.80	0.86	110
	B	0.8	0.84	0.76	0.83	0.8	0.86	80
Combined	R	0.81	0.82	0.80	0.83	0.74	0.92	95
	B	0.79	0.82	0.76	0.82	0.77	0.86	60

[1] *BACC*, balanced accuracy; *SEN*, sensitivity; *SPE*, specificity; *BACC-pos*, balanced accuracy after decision threshold adjustment; *SEN-pos*, sensitivity after decision threshold adjustment; *SPE-pos*, specificity after decision threshold adjustment; *N*, number of features used to train the model

for DART booster as training time increases dramatically with the number of estimators. The same is true for the second round of training, where we didn't test performance for every amount n of the most important feature, but tested the range with a step of 5.

Results

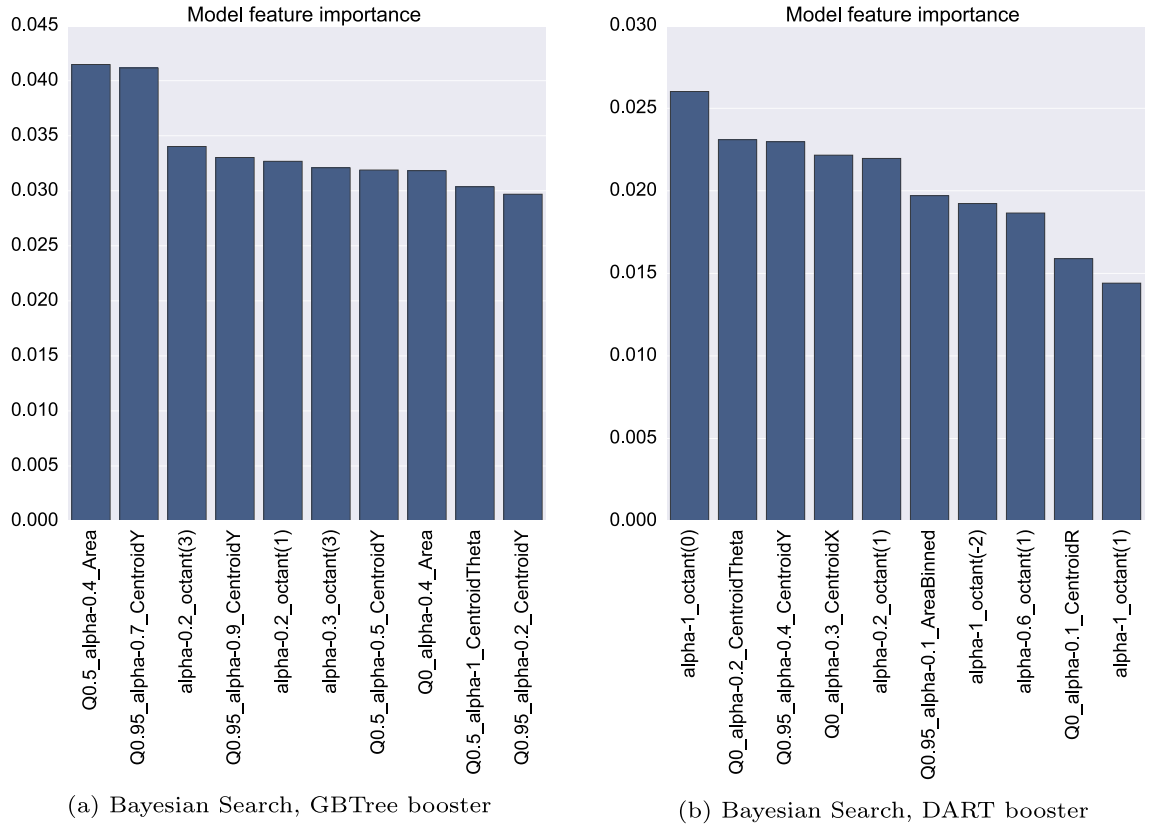
Results based on the GBTree booster are reported in Table 5. Results for the random search tuning method can be found in Table 6. Regarding GBTree-trained models, the Bayesian



(a) Random Search, GBTree booster

(b) Bayesian Search, DART booster

Fig. 7 Feature importances for models trained in the baseline scenario



(a) Bayesian Search, GBTree booster

(b) Bayesian Search, DART booster

Fig. 8 Feature importances for models trained in the novel scenario

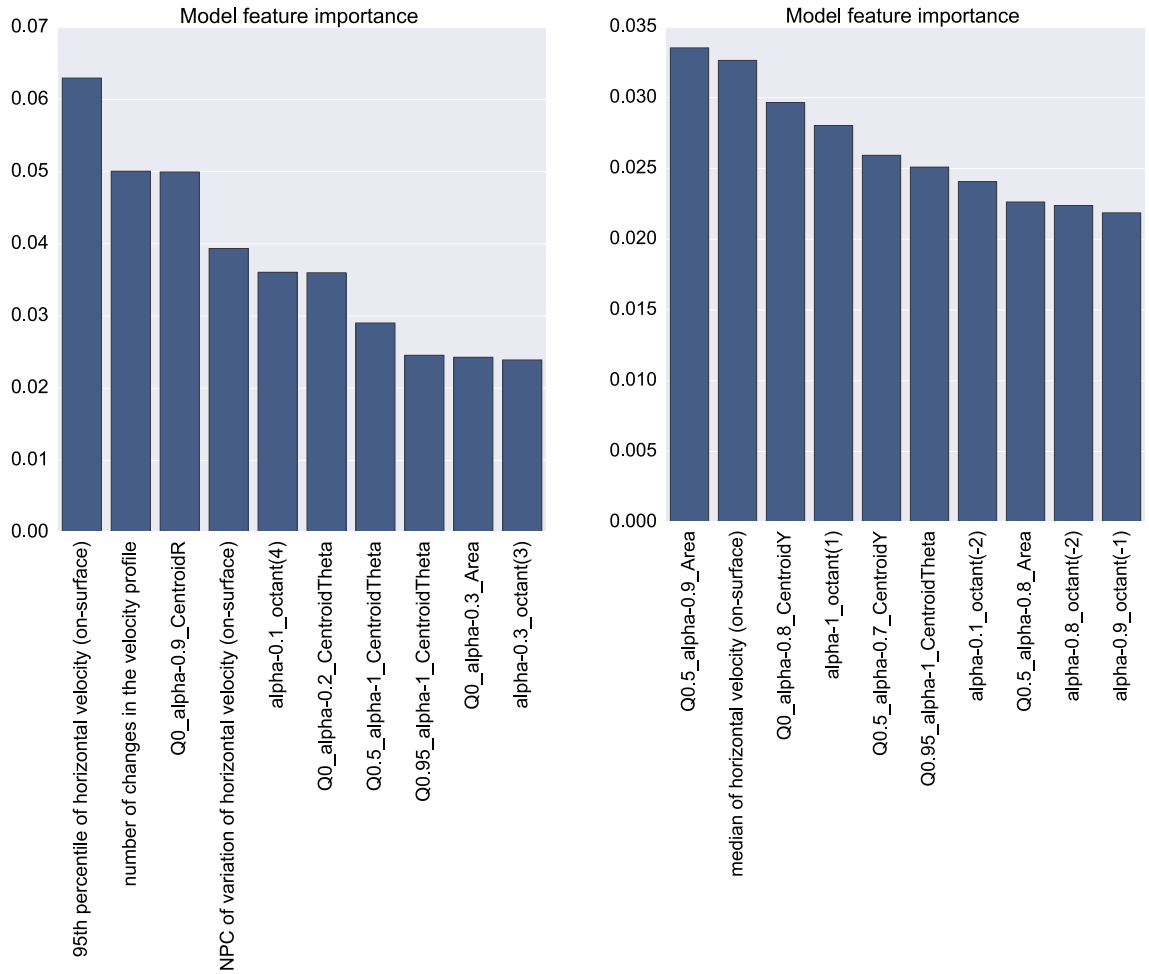
search tuning provided a better performant model in two of the three scenarios, which is a result that was repeated by the DART-trained models. The best-performing model resulted from the combination of the novel feature set and Bayesian search, achieving a balanced accuracy of 87 % (SEN = 82 %, SPE = 92 %). This provides a 3 percentage point increase over the best performant model in the baseline scenario (trained using Bayesian search and DART booster), which achieved BACC of 84 % (SEN = 88 %, SPE = 80 %). Another notable model was trained in the combined scenario, using the Bayesian tuner and GBTree booster, reaching BACC of 86 % (SEN = 80 %, SPE = 92 %). These three models are highlighted in the corresponding tables. All three models were trained using the Bayesian tuning method. Regarding the boosters, we see that the models which were trained on feature sets containing the novel features (novel and combined scenarios) reached better results when trained with the GBTree booster with a significant lead over their DART-trained counterparts (performance lead of 3 and 4 percentage points respectively). While the best performant baseline model was achieved using the DART booster, its advantage is only 1 percentage point.

The best-performing model reached its highest BACC accuracy score while using 43 features, which is roughly one-sixth of the whole feature set. With the whole featureset, the same model would reach BACC of only 79 %. The mean and median BACC over the whole range of n both reached a value of 82 %. The BACC vs. number of features plot can be seen in Fig. 11.

To provide a closer look on the most performant models from each scenario and booster, their feature importances are plotted in Figs. 7, 8, and 9. Corresponding ROC curves are plotted in Fig. 10.

Discussion

Generally, vertical and horizontal movements during handwriting are facilitated by distinct anatomical and biomechanical processes, allowing for effective task execution while accommodating variations in speed, pressure, and fluency. Horizontal movement is primarily produced through wrist actions; specifically, horizontal strokes are generated by ulnar abductions (outward movement) and adduction (inward movement) of the wrist [3, 32]. This type of movement allows for smooth progression across the page and facilitates the formation of letters along the line of writing. On the other hand, vertical movement involves more complex, finer movements controlled largely by the finger flexors and



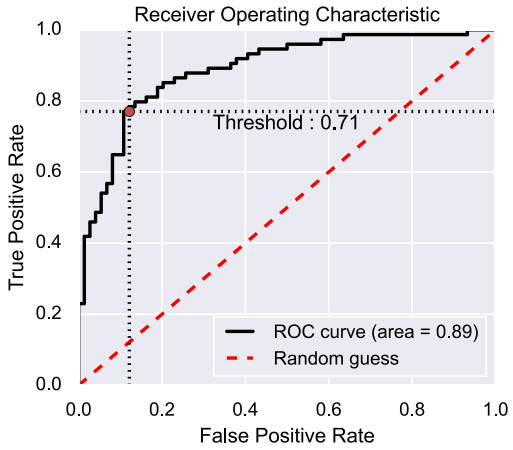
(a) Bayesian Search, GBTree booster (b) Random Search, DART booster

Fig. 9 Feature importances for models trained in the combined scenario

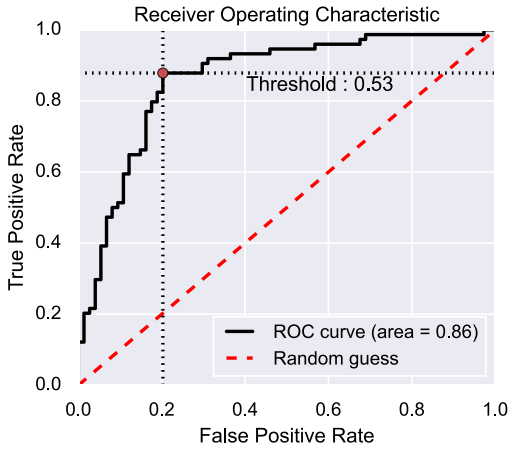
extensors [3, 32]. Due to the finger system's multiple joints, it demands greater coordination and neuromotor control, which can reduce movement speed and fluency compared to horizontal strokes. This complexity is associated with increased neuromotor noise and lower efficiency, potentially contributing to fatigue over prolonged writing tasks [31, 32]. This may also explain why several studies report that GD is more pronounced in vertical movement [22, 23]. However, during writing, vertical and horizontal movements are performed simultaneously, involving a trade-off between fluidity and precision and balancing motor efficiency with biomechanical control. When producing both vertical and horizontal movements simultaneously, the motor system often sacrifices speed for accuracy [3, 32]. This balance is mediated by a process called motor coordination, which is critical for maintaining letter structure and legibility. Nevertheless, to date, the simultaneous execution of vertical and horizontal movement in children with GD has not been fully explored.

The ability to visualize the trade-off between vertical and horizontal movements brings new benefits, as it can aid in interpreting certain manifestations of GD and build intuition in the assessment process. For example, in Fig. 2, we can see a comparison of polar plots of an intact child and a child diagnosed with GD. It seems that the intact child was more comfortable with the nature of the movements required to properly finish the exercise. The two petals on the plot roughly correspond to the relatively straight and long parts of the loop exercise, where it is expected the strokes should reach a higher velocity. On the other hand, looking at the plot of the child with GD, we can see the distribution is more chaotic. While there is a pronounced node around the 45° direction, we can also see similar increases in other directions, including many local maxima. It is notable that the GD-example plot reaches much lesser values, further suggesting the child was not able to consistently reach higher velocities and produced similar velocities in the long and

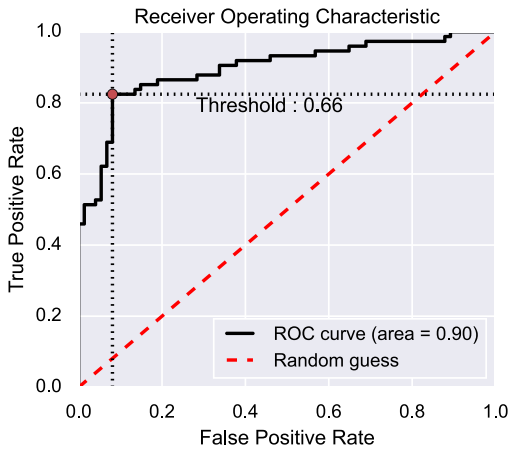
GBTree booster DART booster



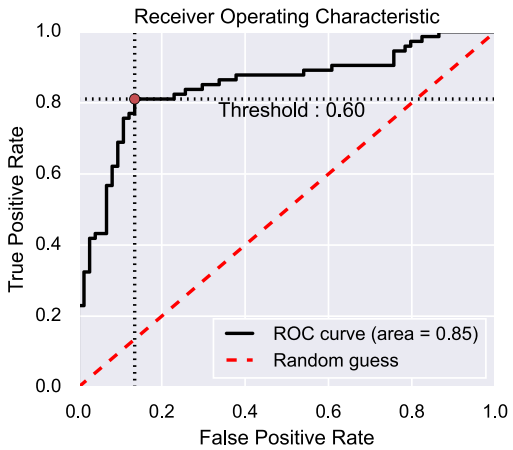
(a) Baseline scenario, Random Search



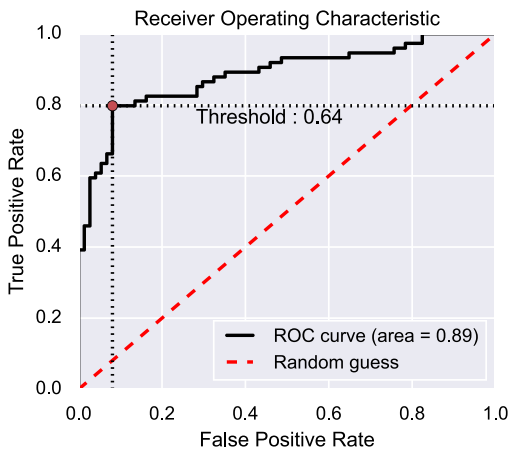
(b) Baseline scenario, Bayesian Search



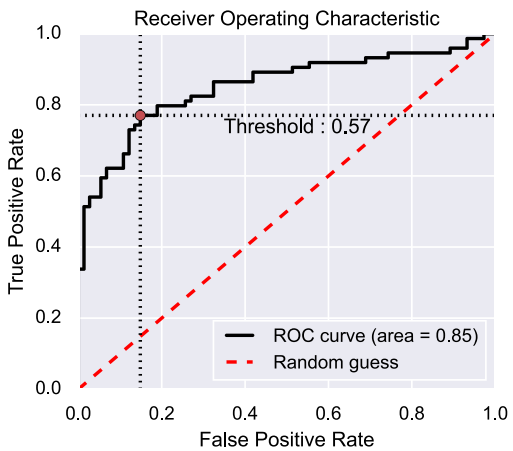
(c) Novel scenario, Bayesian Search



(d) Novel scenario, Bayesian Search



(e) Combined scenario, Bayesian Search



(f) Combined scenario, Random Search

Fig. 10 ROC curves for the best-performing models for each scenario in each hyperparameter tuning method. Rows contain baseline, novel, and combined scenarios respectively, while columns show Bayesian search and random search tuning methods, respectively

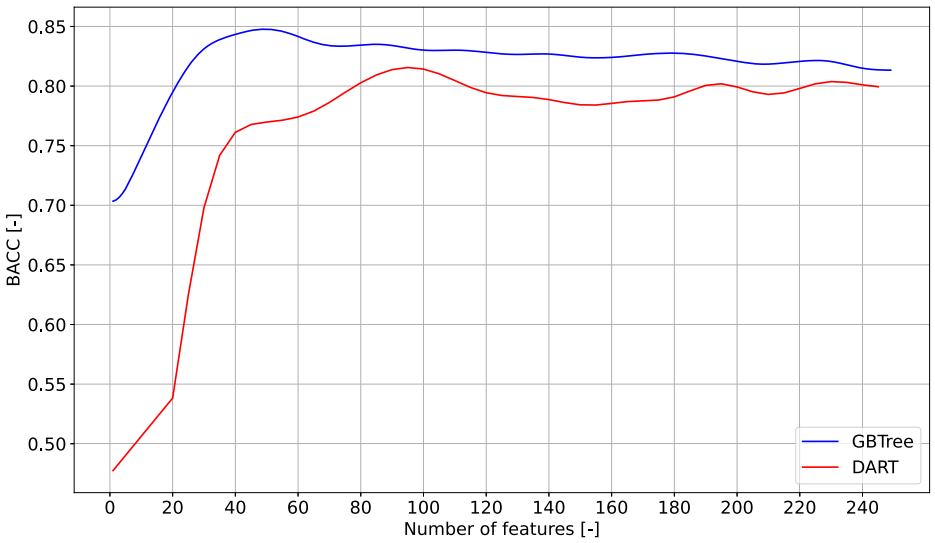


Fig. 11 BACC score dependence on the number of features included in two of the final XGBoost models. Both models were trained with novel features using the Bayesian search hyperparameter tuning method. However, one was trained using the GBTree booster (in blue), while the other was using the DART booster (in red). Both lines were smoothed using Gaussian smoothing for better visual clarity

straight parts of the exercise as well as in the tighter curves. These differences correspond with the principles of handwriting isochrony, which suggests that strokes with longer trajectories should be carried out to reach higher velocity [63]. It has been shown that children with developmental dyslexia and DD violate these principles [64], and features based on the principles of handwriting isochrony can be used to differentiate intact children and children with GD [65].

Using the same polar plots, we can observe the progression of the fractional order derivative as the order increases from $\alpha = 0.2$ to $\alpha = 1.0$ in Figs. 5 and 6 in an intact child and a child with GD, respectively ($\alpha = 0.1$ was omitted for spacing reasons, but the respective plots can be seen in Fig. 4). It seems the plots for the intact child for lower values of α are relatively smooth, with two dominant nodes roughly in the directions of the largest movements in the exercise. With increasing α values, the plot gets more jagged as local maxima start to get more pronounced. With α from 0.5 to 0.7 we can start noticing a new node starting to form roughly at the 305° direction.

Perhaps, as the order of the fractional derivation increases from values 0 close to 1, the output also gradually shifts to a function resembling the velocity curve. This interpretation could offer some insight into the gradual shape-shifting of the plot. With higher α values, the most prominent parts of the plot will be in the directions, where the child reached the highest velocities. This would suggest, that the node appearing for $\alpha = 0.5$ is in part an effect of velocity. Given the direction of the node, it could correspond to the movement at the lower part of the loop before connecting to the next one. It is interesting, that the node in the first quadrant remains isolated throughout all the α values, while the nodes in the lower half of the plot eventually blend together. The effect however seems to be more pronounced in the plot for the intact child, where the lower nodes are separated for $\alpha = 0.7$ and well blended for $\alpha = 0.9$. In contrast, in the plot for the child diagnosed with GD, we can see that while the nodes are not as pronounced for higher α values, there are still soft, but obvious maxima. This might be an indication of a problem with fluent change of direction of the pen tip movement in the later part of the loop, where the child has to slowly go from a left-down facing direction to a mostly right-facing direction with fluency and control. The lack of fluency of change of direction of the pen movement for the child diagnosed with GD can be also seen in Fig. 2, where it seems the changes in direction are more obvious and are executed in pronounced steps, followed by a straighter section.

When looking at the α progression for the child with GD (see Fig. 6), it is evident right away that the values are overall smaller when compared to the intact child, which might be explained by the overall smaller handwriting sample (as can be seen in Fig. 2). However, if we evaluate the shape alone, we also see a dominant node in the first quadrant across the orders, yet, it is much wider and in some α values, tends to have many local maxima. The width of the nodes, or lack, thereof, might be an indication of consistency in the exercise and similarity between the individual loops. As children with proficient handwriting will produce loops that are all more similar to each other, the portions of the exercise reaching the highest velocities should point in the same direction. The difference in consistency of the loop shape can be observed in the rendered samples in Fig. 2. Another notable difference in the GD-child plots, when looking at $\alpha > 0.7$, is that the values reached in the bottom half are much lower than the maximum in the upper half, whereas, in the intact child's plots, they seem to be closer. The bottom halves of the two samples are distinct as well, the GD-child's bottom half of the polar plot seems to be covered more uniformly, rather than with a pronounced node, suggesting inconsistency with the movement direction. This might be interpreted, as while even

the GD-child is able to execute motion in the 45° direction, they have problems with the southward motion of the pentip.

As can be noticed in Fig. 4, the loop task consists of two segments. The first segment is the movement from the loop origin to the top part of the loop (the turning point), and we will refer to it as the upward segment (pink line in Fig. 4). The second segment is the movement from the turning point to the end of the loop, and we will refer to it as the downward segment (green line in Fig. 4). Observing polar plots for intact children indicates that the directional values of trajectory data visualized can be associated with these two segments. The upward segment is represented by directions in the quadrant around 50°, and the downward segment represents the quadrant around 260°. Imagining the ideal movement during the task performance for an intact child (except the top part of the loop), the directions on each loop segment should be around the same degree level. Therefore, we can claim this representation as an accurate measurement of the writing directionality. The downward segment should be more automatic and effortless, resulting in higher velocity. With higher velocity, the variability between the movement directions will increase, and the derivation progress should stress out this phenomenon. Considering the temporal memory effect of the FDEs (Caputo's) operator, the variability principle should be more visible with the increasing alpha order of the derivation, confirmed by the observations in Fig. 5. Therefore, we should analyze this effect separately in future studies to describe it precisely since the potential of FDEs can be utilized even more.

The novel features seem to be beneficial to the computer-aided GD diagnosis when using XGBoost classification models, as seen in Tables 5 and 6. Comparing the two tables and considering the BACC and BACC-pos values, it is evident, that Bayesian search has provided better results in more scenarios than random search. With exceptions being Baseline GBTree and Combined DART scenarios, where performance leans toward random search, and the Novel DART scenario, where the highest BACC (adjusted for positive class) score is tied, but non-adjusted BACC is higher with Bayesian search.

Except for the baseline scenario using the Bayesian tuner, DART does not seem to offer obvious improvement in model performance. Despite the fact, that the hyperparameter tuning algorithms could set all the DART-specific parameters to zero, effectively transforming it to the GBTree booster, none of the best DART models chose this tactic. It might be important to point out, that as the computation time for DART booster training is significantly longer, we set the number of estimators lower than in the GBTree models to roughly match the computational resources. It is possible, that given unlimited resources, the DART-boosted models would reach comparable or better results, but such an experiment is beyond the scope of this study.

The model trained using the Novel featureset accomplished a 3 percentage point improvement over the baseline-trained model when considering the BACC score. The fact that the baseline features carry only the information about velocity, while the novel features also provide information in the fractional order derivatives, suggests that featurising a broader dimensional range could be an important factor in feature engineering. This is also supported by the Feature importance plots for both of the most performant novel-scenario models, where some features appear multiple times in the top 10 feature list, but with a different derivative order and quartile, e.g., the features CentroidY (α = {0.7, 0.9, 0.5, 0.2}, in that order) for the bayesian trained GBTree model (Fig. 8). We can also see the feature repeating throughout the feature importance plots of other trained models.

While the combined scenario did not provide the best performant model (2 percentage point improvement in BACC over baseline), observing the feature importance rating of the models trained on the combined datasets, could provide some insight into how the features impact the overall model. From Fig. 9 we can see that the GBTree model only selected 3 features from the baseline dataset and only two novel features that were calculated from an integer-order derivation, which suggests that a combination of fractional order and full order derivatives is beneficial to the overall performance of the model. This is further supported by the feature importance list of the DART-trained model, which selected only novel features in the 10 most important features, but we can still see, that full-order derivative features are present (2 out of 10). For both models, we can see that the novel features are represented in greater numbers in the plots.

The method of training the models in two rounds was beneficial across the board. Examples for two models from the GBTree and DART-based models can be seen in Fig. 11. In the GBTree example is a clearly visible trend, where the BACC score grows until peaking between 40 and 60 features, after which it slowly decreases. As the BACC plot against the number of features has a noisy character, the performance of the model trained with the full featureset lies in a local minimum (BACC = 79 %) and thus comparing it with the best result might be misleading. However, the mean and median BACC score of the model is 82 %, so only supplying a subset of the features provided a 5 percentage point increase in accuracy. A similar effect can be seen in the DART example, however, it is notable, that this model reaches its peak performance later and the decrease in BACC with a growing number of features is not as notable. Given the increase in performance when only using a subset of the featureset, we conclude that this technique might prove beneficial and should be tested with more datasets of various sizes.

While the resulting balanced accuracy of our best-performing model is 87 %, which might appear lagging

behind the current state-of-the-art approaches, it is important to realize, that we only trained our models with access to limited information regarding the subjects' handwriting ability. Firstly, we only used data from one single graphomotor exercise, rather than from a barrage of exercises or handwriting samples. Secondly, we limited the dimensionality of features to only include kinematic features based on handwriting velocity. Furthermore, we show that features based on fractional-order calculus improve the performance of classifier-based approaches to automated handwriting assessment and their employment in more complex systems should be considered.

Conclusion

This study had several goals related to the field of featurisation and objective assessment of handwriting deficiencies in children. First and foremost, we proposed novel features based on fractional calculus and polar plots, extracted from the position data of online handwriting. These features allow for additional information availability to an automated classifier with the potential of human interpretability thanks to the possibility of easy visualization using polar plots.

We validated the novel features in a barrage of machine learning scenarios, which allowed us to test the performance of various XGBboost training methods, namely comparing the GBTree and DART boosters, and Bayesian and random search hyperparameter tuning methods. Lastly, after hyperparameter tuning, we trained the models with a subset of the featureset, using only N most important features, iterating N in the range from 1 to the total amount of features. In the end, novel features provided a balanced accuracy of 87 % (SEN = 82 %, SPE = 92 %), while using the GBTree booster and Bayesian search tuning method, accomplishing 3 percentage point increase of BACC score compared to a model trained using conventional features. The model reached its peak performance when using only 43 out of 250 newly proposed features, where that peak BACC was 5 percentage points above the mean and median accuracy and 8 percentage points above the accuracy reached with the full featureset, showing that XGBoost can benefit from feature selection methods. However, this improvement comes at the cost of increased model complexity, requiring 43 features compared to just 13 in the best baseline configuration. This trade-off between performance and complexity underscores the need for further exploration into optimizing both feature selection and model simplicity.

The study has several identified limitations. To our best knowledge, this is the first study exploring the usage of FDE in combination with polar plots, thus more analysis is required to enable a more thorough understanding of their viability for handwriting assessment. More specifically, our case is limited to only one handwriting exercise we only used velocity-based features, and the α step of 0.1 is not sensitive enough. Also, only Caputo's approach to FDEs was tested, and others should be explored (e.g., Baleanu, Riemann-Liouville). The size of the dataset the study was conducted on can also be considered a limitation. Moreover, the target variable in training was obtained as a combination of HPSQ-C and expert assessment provided by only one rater. To overcome this, further studies should be conducted on different datasets, not limited to children's handwriting (e.g., Parkinson's disease handwriting assessment). Finally, while we tested more variants of the XGBoost ML approach, testing more models is also desirable in future studies.

Supplementary Information

List of supplementary files:
S1 list of extracted features

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12559-024-10360-7.

Author Contribution Michal Gavenciak: investigation, data curation, analysis, methodology, software, writing—original draft preparation. Jan Mucha (corresponding author): conceptualization, methodology, investigation, data curation, writing-review and editing. Jiri Mekyska: supervision, d ata curation, project administration, formal analysis, writing—review and editing, funding acquisition. Zoltan Galaz: resources, data curation, methodology. Katarina Zvoncakova: validation, formal analysis, investigation, writing—review and editing. Marcos Faundez-Zanuy: methodology, supervision, funding acquisition.

Funding This work was supported by the Czech Science Foundation, project no. GN23-06074O (Research of advanced methods of graphomotor disabilities analysis based on fractional calculus), by the Technology Agency of the Czech Republic, project no. TL03000287 (Software for advanced diagnosis of graphomotor disabilities), by EU – Next Generation EU (project no. LX22NPO5107 (MEYS)), and by Spanish grant of the Ministerio de Ciencia e Innovacion no. PID2020-113242RB-I00.

Data Availability The whole database is available upon a reasonable request sent to the corresponding author.

Materials Availability Not applicable

Code Availability The drawing/handwriting data were recorded using the HandAQUS acquisition software [66]. Most of the drawing/handwriting features could be extracted using the Python library handwriting-features (v 1.0.8) [67].

Declarations

Ethics Approval and Consent to Participate This study was reviewed and approved by the Ethical Board of the Department of Psychology of the Masaryk University.

Consent for Publication Not applicable

Conflict of Interest The authors declare no competing interests.

References

1. Matijević-Mikelić V, Košiček T, Crnković M, Trifunović-Maček Z, Grazio S. Development of early graphomotor skills in children with neurodevelopmental risks. Acta Clinica Croatica. 2011;50(3):317–21.
2. Feder KP, Majnemer A. Handwriting development, competency, and intervention. Developmental Medicine & Child Neurology. 2007;49(4):312–7.
3. 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. Research in developmental disabilities. 2011;32(3):1058–64.
4. Rosenblum S. Inter-relationships between objective handwriting features and executive control among children with developmental dysgraphia. PloS one. 2018;13(4):0196098.
5. Alhusaini AA, Melam GR, Buragadda S. Short-term sensorimotor-based intervention for handwriting performance in elementary school children. Pediatrics International. 2016;58(11):1118–23.
6. Mekyska J, Safarova K, Urbanek T, Bednarova J, Zvoncak V, Havigerova JM, Cunek L, Galaz Z, Mucha J, Klauszova C, et al. Graphomotor and handwriting disabilities rating scale (GHDRS): towards complex and objective assessment. Australian J Learn Difficulties. 2024;29(1):1–34. https://doi.org/10.1080/19404158.2024.2326686.
7. Karlsdottir R, Stefansson T. Problems in developing functional handwriting. Perceptual and motor skills. 2002;94(2):623–62.
8. Katusic SK, Colligan RC, Weaver AL, Barbaresi WJ. The forgotten learning disability: epidemiology of written-language disorder in a population-based birth cohort (1976–1982), Rochester. Minn Pediatr. 2009;123(5):1306–13.
9. McHale K, Cermak SA. Fine motor activities in elementary school: preliminary findings and provisional implications for children with fine motor problems. Am J Occup Ther. 1992;46(10):898–903.
10. Rosenblum S. Development, reliability, and validity of the handwriting proficiency screening questionnaire (HPSQ). Am J Occup Ther. 2008;62(3):298–307.
11. Van Waelvelde H, Hellinckx T, Peersman W, Smits-Engelsman BC. SOS: a screening instrument to identify children with handwriting impairments. Phys Occup Ther Pediatr. 2012;32(3):306–19.
12. Barnett AL, Prunty M, Rosenblum S. Development of the handwriting legibility scale (HLS): a preliminary examination of reliability and validity. Res Dev Disabil. 2018;72:240–7.
13. Deschamps L, Devillaine L, Gaffet C, Lambert R, Aloui S, Boutet J, Brault V, Labyt E, Jolly C. Development of a pre-diagnosis tool based on machine learning algorithms on the BHK test to improve the diagnosis of dysgraphia. Adv Artif Intell Mach Learn. 2021;1(2):114–35.
14. Kunhoth J, Al-Maadeed S, Kunhoth S, Akbari Y. Automated systems for diagnosis of dysgraphia in children: a survey and novel framework 2022.
15. Mekyska J, Faundez-Zanuy M, Mzourek Z, Galaz Z, Smekal Z, Rosenblum S. Identification and rating of developmental dysgraphia by handwriting analysis. IEEE Trans Human-Mach Syst. 2016;47(2):235–48.
16. 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(1):42.
17. Mittal D, Yadav V, Sangwan A. Identification of dysgraphia: a comparative review. In: International Conference on Emerging Technologies in Computer Engineering. 2022. pp. 52–62. Springer.
18. Rosenblum S, Dror G. Identifying developmental dysgraphia characteristics utilizing handwriting classification methods. IEEE Trans Human-Mach Syst. 2017;47(2):293–8.
19. 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 Digital Medicine. 2018;1(1):42.
20. Zvoncak V, Mekyska J, Safarova K, Galaz Z, Mucha J, Kiska T, Smekal Z, Losenicka B, Cechova B, Francova P, et al. Effect of stroke-level intra-writer normalization on computerized assessment of developmental dysgraphia. In: 2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). 2018. pp. 1–50.
21. Zvoncak V, Mekyska J, Safarova K, Smekal Z, Brezany P. New approach of dysgraphic handwriting analysis based on the tunable Q-Factor wavelet transform. In: 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). 2019. pp. 289–294.
22. Galaz Z, Mucha J, Zvoncak V, Mekyska J, Smekal Z, Safarova K, Ondrackova A, Urbanek T, Havigerova JM, Bednarova J, Faundez-Zanuy M. Advanced parametrization of graphomotor difficulties in school-aged children. IEEE Access. 2020;8:112883–97.
23. Mekyska J, Galaz Z, Safarova K, Zvoncak V, Mucha J, Smekal Z, Ondrackova A, Urbanek T, Havigerova JM, Bednarova J, Faundez-Zanuy M. Computerised assessment of graphomotor difficulties in a cohort of school-aged children. In: 2019 11th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). 2019. pp. 1–6.
24. Mucha J, Mekyska J, Galaz Z, Faundez-Zanuy M, Zvoncak V, Safarova K, Urbanek T, Havigerova JM, Bednarova J, Smekal Z. Analysis of various fractional order derivatives approaches in assessment of graphomotor difficulties. IEEE Access. 2020;8:218234–44.
25. Drotár P, Dobeš M. Dysgraphia detection through machine learning. Scientific reports. 2020;10(1):21541.
26. Kunhoth J, Al Maadeed S, Saleh M, Akbari Y. CNN feature and classifier fusion on novel transformed image dataset for dysgraphia diagnosis in children. Expert Syst Appl. 2023;120740.
27. Dui LG, 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. Scientific Reports. 2022;12(1):21624.
28. Vilasini V, Rekha BB, Sandeep V, Venkatesh VC. Deep learning techniques to detect learning disabilities among children using handwriting. In: 2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT). 2022. pp. 1710–1717. IEEE.

29. Ghouse, F., Paranjothi, K., Vaithiyanathan, R.: Dysgraphia classification based on the non-discrimination regularization in rotational region convolutional neural network. Int J Intell Engineer Syst. 2022;15(1).
30. Newell K, Van Emmerik R. The acquisition of coordination: preliminary analysis of learning to write. Hum Mov Sci. 1989;8(1):17–32.
31. Van Den Heuvel CE, Galen GP, Teulings H-L, Gemmert AW. Axial pen force increases with processing demands in handwriting. Acta Psychol. 1998;100(1–2):145–59.
32. Van Galen GP. Handwriting: issues for a psychomotor theory. Hum Mov Sci. 1991;10(2–3):165–91.
33. Gargot T, Asselborn T, Pellerin H, Zammouri I, Anzalone MS, Casteran L, Johal W, Dillenbourg P, Cohen D, Jolly C. Acquisition of handwriting in children with and without dysgraphia: a computational approach. Plos one. 2020;15(9):0237575.
34. Blom J, Haar Romeny BM, Bel A, Koenderink JJ. Spatial derivatives and the propagation of noise in gaussian scale space. J Vis Commun Image Represent. 1993;4(1):1–13.
35. Mucha J, Faundez-Zanuy M, Mekyska J, Zvoncak V, Galaz Z, Kiska T, Smekal Z, Brabenec L, Rektorova I, Lopez-de-Ipina K. Analysis of Parkinson's disease dysgraphia based on optimized fractional order derivative features. In: 2019 27th European Signal Processing Conference (EUSIPCO). 2019. pp. 1–5.
36. Zvoncak V, Mucha J, Galaz Z, Mekyska J, Safarova K, Faundez-Zanuy M, Smekal Z. Fractional order derivatives evaluation in computerized assessment of handwriting difficulties in school-aged children. In: 2019 11th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). 2019. pp. 1–6.
37. Mucha J, Galaz Z, Mekyska J, Faundez-Zanuy M, Zvoncak V, Smekal Z, Brabenec L, Rektorova I. Exploration of various fractional order derivatives in Parkinson's disease dysgraphia analysis. In: International Graphonomics Conference. 2022. pp. 308–321. Springer.
38. Podlubny I. Fractional differential equations an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications. San Diego: Academic Press; 1999.
39. Lazarević M. Further results on fractional order control of a mechatronic system. Scientific Technical Review, ISSN. 2013;206.
40. Uchaikin VV. Fractional derivatives for physicists and engineers vol. 2. Springer, ???. 2013.
41. Sun H, Zhang Y, Baleanu D, Chen W, Chen Y. A new collection of real world applications of fractional calculus in science and engineering. Communications in Nonlinear Science and Numerical Simulation. 2018;64:213–31. https://doi.org/10.1016/j.cnsns.2018.04.019.
42. Persechino A. An introduction to fractional calculus. Advanced Electromagnetics. 2020;9(1):19–30.
43. Joshi M, Bhosale S, Vyawahare VA. A survey of fractional calculus applications in artificial neural networks. Artif Intell Rev. 2023;1–54.
44. Zuñiga-Aguilar CJ, Gomez-Aguilar J, Franc S, Charpentier G, Doron M, Benhamou P, Romero-ugalde H. Blood glucose prediction with a fractional order neural network. Diabetes Technol Ther. 2020;22:82–82.
45. Nagar S, Kumar A. Orthogonal features based EEG signals denoising using fractional and compressed one-dimensional CNN autoencoder. IEEE Trans Neural Syst Rehabil Eng. 2022;30:2474–85.
46. Arshad S, Baleanu D, Bu W, Tang Y. Effects of HIV infection on cd4+ t-cell population based on a fractional-order model. Adv Difference Equ. 2017;2017(1):92. https://doi.org/10.1186/s13662-017-1143-0.
47. Pinto CMA, Machado JAT. Fractional model for malaria transmission under control strategies. Computers & Mathematics with Applications. 2013;66(5):908–16. https://doi.org/10.1016/j.camwa.2012.11.017. Fractional Differentiation and its Applications.
48. Herrera-Alcántara O, Castelán-Aguilar JR. Fractional gradient optimizers for PyTorch: enhancing GAN and BERT. Fractal and Fractional. 2023;7(7). https://doi.org/10.3390/fractalfract7070500.
49. Altan G, Alkan S, Baleanu D. A novel fractional operator application for neural networks using proportional Caputo derivative. Neural Comput Appl. 2023;35(4):3101–14.
50. Rosenblum S, Gafni-Lachter L. Handwriting proficiency screening questionnaire for children (HPSQ-C): development, reliability, and validity. The American Journal of Occupational Therapy. 2015;69(3):6903220030–169032200309.
51. Šafárová K, Mekyska J, Zvončák V, Galáž Z, Francová P, Čechová B, Losenická B, Smékal Z, Urbánek T, Havigerová JM, et al. Psychometric properties of screening questionnaires for children with handwriting issues. Front Psychol. 2020;10:2937.
52. Association AP, et al. Ethical principles of psychologists and code of conduct. American Psychol. 2002;57(12):1060–73.
53. Mucha J, Zvoncak V, Galaz Z, Faundez-Zanuy M, Mekyska J, Kiska T, Smekal Z, Brabenec L, Rektorova I, Lopez-de-Ipina K. Fractional derivatives of online handwriting: a new approach of Parkinsonic dysgraphia analysis. In: 2018 41st International Conference on Telecommunications and Signal Processing (TSP). 2018. pp. 214–217. IEEE
54. Mucha J, Mekyska J, Faundez-Zanuy M, Lopez-de-Ipina K, Zvoncak V, Galaz Z, Kiska T, Smekal Z, Brabenec L, Rektorova I. Advanced Parkinson's disease dysgraphia analysis based on fractional derivatives of online handwriting. In: 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). 2018. pp. 158–165
55. Mucha J, Mekyska J, Galaz Z, Faundez-Zanuy M, Lopez-de-Ipina K, Zvoncak V, Kiska T, Smekal Z, Brabenec L, Rektorova I. Identification and monitoring of Parkinson's disease dysgraphia based on fractional-order derivatives of online handwriting. Appl Sci. 2018;8(12):2566.
56. Caputo M. Linear models of dissipation whose Q is almost frequency independent—II. Geophys J Int. 1967;13(5):529–39. https://doi.org/10.1111/j.1365-246X.1967.tb02303.x. https://academic.oup.com/gji/article-pdf/13/5/529/1600098/13-5-529.pdf.
57. Luchko Y, Gorenflo R. An operational method for solving fractional differential equations with the Caputo derivatives. Acta Mathematica Vietnamica. 1999;24(2):207–33.
58. Luria G, Rosenblum S. A computerized multidimensional measurement of mental workload via handwriting analysis. Behav Res Methods. 2012;44:575–86.
59. Chen T, Guestrin C. XGboost: a scalable tree boosting system. In: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining. 2016. pp. 785–794. ACM
60. Zhang P, Jia Y, Shang Y. Research and application of XGboost in imbalanced data. Int J Distrib Sensor Netw. 2022;18(6):15501329221106936.
61. Todd MT, Nystrom LE, Cohen JD. Confounds in multivariate pattern analysis: theory and rule representation case study. Neuroimage. 2013;77:157–65.
62. Rashmi KV, Gilad-Bachrach R. DART: dropouts meet multiple additive regression trees. arXiv 2015. https://doi.org/10.48550/ARXIV.1505.01866. arxiv:1505.01866
63. Pagliarini E, Scocchia L, Vernice M, Zoppello M, Balottin U, Bouamama S, Guasti MT, Stucchi N. Children's first handwriting productions show a rhythmic structure. Scientific reports. 2017;7(1):5516.

64. Pagliarini E, Guasti MT, Toneatto C, Granocchio E, Riva F, Sarti D, Molteni B, Stucchi N. Dyslexic children fail to comply with the rhythmic constraints of handwriting. Hum Mov Sci. 2015;42. https://doi.org/10.1016/j.humov.2015.04.012.
65. Gavenciak M, Zvoncak V, Mekyska J, Safarova K, Cunek L, Urbanek T, Havigerova JM, Bednarova J, Galaz Z, Mucha J. Exploring the contribution of isochrony-based features to computerized assessment of handwriting disabilities. In: 2022 45th International Conference on Telecommunications and Signal Processing (TSP). 2022. pp. 355–359. https://doi.org/10.1109/TSP55681.2022.9851254.
66. Mucha J, Mekyska J, Zvoncak V, Galaz Z, Smekal Z. HandAQUS – handwriting acquisition software. GitHub 2022.
67. Galaz Z, Mucha J, Zvoncak V, Mekyska J. Handwriting features. GitHub 2022.

Authors and Affiliations

Michal Gavenciak[1] · Jan Mucha[1,3] · Jiri Mekyska[1] · Zoltan Galaz[1] · Katarina Zvoncakova[2] · Marcos Faundez-Zanuy[3]

✉ Jan Mucha
mucha@vutbr.cz

Michal Gavenciak
xgaven07@vut.cz

Jiri Mekyska
mekyska@vutbr.cz

Zoltan Galaz
galaz@vutbr.cz

Katarina Zvoncakova
zvoncakova@psu.cas.cz

Marcos Faundez-Zanuy
faundez@tecnocampus.cat

[1] Department of Telecommunications, Faculty of Electrical Engineering and Communication, Brno University of Technology, Technicka 12, Brno 61200, Czech Republic

[2] Department of Research Methodology, Institute of Psychology, The Czech Academy of Sciences, Veveri 97, Brno 60200, Czech Republic

[3] TecnoCampus Mataró, Universitat Pompeu Fabra, Ernest Lluch 32, Mataro 08302, Spain