

Article

Towards Parkinson's Disease Detection Through Analysis of Everyday Handwriting

Jeferson David Gallo-Aristizabal ¹, Daniel Escobar-Grisales ¹, Cristian David Ríos-Urrego ¹,
Jesús Francisco Vargas-Bonilla ¹, Adolfo M. García ^{2,3,4,5}  and Juan Rafael Orozco-Arroyave ^{1,6,*} 

¹ GITA Lab., Faculty of Engineering, University of Antioquia, Medellín 510010, Colombia; jeferson.gallo@udea.edu.co (J.D.G.-A.); daniel.esobar@udea.edu.co (D.E.-G.); cdavid.rios@udea.edu.co (C.D.R.-U.); jesus.vargas@udea.edu.co (J.F.V.-B.)

² Cognitive Neuroscience Center, Universidad de San Andrés, Buenos Aires B1644BID, Argentina; adolfomartingarcia@gmail.com

³ Global Brain Health Institute (GBHI), University of California San Francisco, San Francisco, CA 94143, USA

⁴ Trinity College Dublin, D02 R590 Dublin, Ireland

⁵ Departamento de Lingüística y Literatura, Facultad de Humanidades, Universidad de Santiago de Chile, Santiago 9170020, Chile

⁶ LME Lab., University of Erlangen, 91054 Erlangen, Germany

* Correspondence: rafael.orocho@udea.edu.co

Abstract: Background: Parkinson's disease (PD) is the second most prevalent neurodegenerative disorder worldwide. People suffering from PD exhibit motor symptoms that affect the control of upper and lower limb movement. Among daily activities that depend on proper upper limb control is the handwriting process, which has been studied in state-of-the-art research, mainly considering non-semantic drawings like spirals, geometric figures, cursive lines, and others. **Objectives:** This paper analyzes the suitability of modeling the handwriting process of digits from 0 to 9 to automatically discriminate between PD patients and healthy control subjects. The main hypothesis is that modeling these numbers allows a more natural evaluation of upper limb control. **Methods:** Two approaches are considered: modeling of the images resulting from the strokes collected by the digital tablet and modeling of the time series yielded by the digital tablet while performing the strokes, i.e., time-dependent signals. The first approach is implemented by fine-tuning a CNN-based architecture, while the second approach is based on hand-crafted features measured upon the time series, namely pressure and kinematic measurements. Features extracted from time-dependent signals are represented following two strategies, one based on statistical functionals and the other one based on creating Gaussian Mixture Models (GMMs). **Results:** The experiments indicate that pressure-based features modeled with functionals are the ones that yield the highest accuracy, indicating that PD-related symptoms are better modeled with dynamic approaches than those based on images. **Conclusions:** The dynamic approach outperformed the image-based model, indicating that the writing process, modeled with signals collected over time, reveals motor symptoms more clearly than images resulting from handwriting. This finding is in line with previous results in the state-of-the-art research and constitutes a step forward to create more accurate and informative methods to detect and monitor PD symptoms.



Academic Editor: Benedetta Nacmias

Received: 6 January 2025

Revised: 29 January 2025

Accepted: 30 January 2025

Published: 5 February 2025

Citation: Gallo-Aristizabal, J.D.; Escobar-Grisales D.; Ríos-Urrego, C.D.; Vargas-Bonilla, J.F.; García, A.M.; Orozco-Arroyave, J.R. Towards Parkinson's Disease Detection Through Analysis of Everyday Handwriting. *Diagnostics* **2025**, *15*, 381. <https://doi.org/10.3390/diagnostics15030381>

Copyright: © 2025 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/>).

Keywords: Parkinson's disease; handwriting; convolutional neural networks; dynamic analysis; natural handwriting tasks

1. Introduction

Parkinson's disease (PD) is a neurodegenerative disorder that affects the central nervous system, and it is characterized by the progressive loss of dopaminergic neurons in the midbrain [1]. These neurons are mainly responsible for the production of dopamine, which is a neurotransmitter in charge of keeping operational neural pathways associated with mood, motor control, and other functions [2]. Among motor symptoms derived from PD are muscular rigidity, resting tremor, postural instability, and bradykinesia [3]. Motor symptoms negatively impact activities that require highly coordinated movements like handwriting [4]. Anomalies in handwriting include micrographia (abnormal reduction in writing size) and dysgraphia (deficits in graphomotor production) [5]. Other neurodegenerative disorders also affect the handwriting process, for instance, Alzheimer's disease patients are known for producing abnormal strokes in different handwriting tasks [6,7]. Handwriting analyses are carried out via online data or offline data. Online data refer to handwriting signals acquired using digital devices, such as tablets or smart pens; offline data consist of traditional methods like writing with an ink pen on paper. Handwriting using conventional methods could be more natural for elderly patients; however, online data allow dynamic analyses of the handwriting process and enable sensing-relevant bio-signals such as pressure, in-air movements, dynamic horizontal and vertical positions, pen inclinations, and other relevant information [8,9]. For a comprehensive review of methods used to evaluate neurodegenerative disorders considering handwriting biosignals, please see [10].

Online and offline handwriting studies typically consider geometric-based tasks, namely, drawings that involve figures like spirals, meanders, geometric figures, and others [11]. Studies have explored the use of different features including kinematic [12–14], geometric, and pressure [15]. One of the main advantages of online handwriting is that it provides the sensor-based signals and also the resulting image, i.e., the drawing. This characteristic has motivated researchers to use image processing methods, mainly based on Convolutional Neural Networks (CNNs), to model those images. For instance, in [16], the authors used time series-based images from drawing tasks. Raw signals collected from the tablet were transformed into matrices to form images which were further processed by a CNN architecture pre-trained on the ImageNet dataset. In another study presented in [17], the authors created spectral representations from tablet signals to model tremor and other symptoms. These signal representations were stacked to form an image and then feed a CNN model with kernel sizes of 1×5 and 1×3 . In [18], authors analyzed drawings from three datasets (PaHaW [19], HandPD [20], and NewHandPD [11]) to increase sample-size. They used an AlexNet model pre-trained on ImageNet and data augmentation techniques. Finally, in [21], authors implemented a CNN with two blocks containing three convolutional layers and one pooling layer. Similarly, in [22], the authors used pre-trained models based on ImageNet to improve the model's generalization capability. Other datasets like MNIST and UJIPenchars2, which are semantically similar to writing tasks, have also been used to pre-train models [23].

Regarding modeling dynamic information from raw tablet signals, previous studies suggest that dynamic analysis of online handwriting provides suitable information to detect PD motor symptoms such as tremor and bradykinesia [9]. Regarding the offline approach, deep learning architectures mainly based on CNNs have been used to analyze handwritten drawings. On the one hand, the majority of handwriting-related studies consider drawings of geometric figures and spirals. On the other hand, writing tasks that request the patient to write numbers, letters, words, or sentences [19] are more natural and closer to activities of daily living, which is what expert neurologists usually focus on during clinical check-ups. Models focused on these kinds of tasks eventually produce better systems to perform accurate monitoring of patients. Additionally, these semantic writing tasks may have a

higher cognitive load to patients, therefore providing information about motor planning (e.g., between-letter or word transitions), which is affected due to lack of coordination in PD [10]. Dynamic signals extracted from writing tasks have been modeled using different dimensions such as kinematic [24,25], spectro-temporal [26], and non-linear [27].

State-of-the-art research shows that deep learning architectures (mainly based on CNN architectures) have been used to model images resulting either from offline handwriting images or from images created by considering online signals. Most studies with promising results are based on drawing tasks [22,28,29]. However, these approaches are not as natural as those based on regular writing tasks, i.e., writing of letters, numbers, words, or sentences, making those models not suitable for performing non-intrusive monitoring of the neurological state of patients.

This paper introduces and compares two different approaches based on online handwriting signals collected from PD patients and healthy controls (HCs). One approach considers resulting images of the tablet handwriting to feed a CNN architecture pre-trained with adapted samples of the MNIST corpus and fine-tuned with Parkinson's data. The second approach consists of considering online signals and extracting different features, including those related to the pressure and movements of the pen (namely, kinematic). The resulting features are compressed and represented with two statistical approaches: statistical functionals and Gaussian Mixture Models (GMMs). Statistical representations are used as input to a Support Vector Machine (SVM) classifier, the parameters of which are optimized following a k-fold cross-validation strategy. Results show that image-based features are outperformed by dynamic analyses with the pressure feature set using statistical functionals. This likely indicates that vertical control of hand movement is more challenging for PD patients.

The rest of the paper is organized as follows. Section 2 introduces the data considered for this study, Section 3 explains the methodology followed in the paper, Section 4 shows experiments and results, Section 5 elaborates a discussion about findings reported in the study, and finally, Section 6 presents the conclusions derived from this work.

2. Data Acquisition and Participants

Our handwriting database (Hw-DB) consists of 104 participants, including 51 PD patients and 53 HC subjects. Forty-seven of the patients were evaluated by an expert neurologist according to the MDS-UPDRS-III scale. The remaining four patients could not come to the clinic for their neurological evaluation because they live in the countryside and were not available for these clinical screenings. Subjects were asked to write the numbers from 0 to 9 using a Wacom Cintiq 13 HD tablet, with a sampling frequency of 180 Hz. This tablet provides six signals: x and y position, z (distance from the tablet surface to the pen tip), azimuth and altitude angles, and pen pressure. Additional details about the participants are shown in Table 1.

Table 1. Demographic and clinical information of the participants.

	PD Patients		HC	
	Male	Female	Male	Female
Number of subjects	24	27	32	21
Age ($\mu \pm \sigma$) *	69.2 $\pm$ 10.0	62.1 $\pm$ 12.0	67.1 $\pm$ 10.6	58.8 $\pm$ 10.8
Age range	50–90	29–84	49–85	43–83
MDS-UPDRS-III ($\mu \pm \sigma$)	39.2 $\pm$ 17.0	32.3 $\pm$ 15.9		
Range of MDS-UPDRS-III	16–82	14–77		

* Indicates that there is no statistical difference between the two groups according to a t -test with p -value = 0.45. Gender bias was discarded through a Chi-squared test with p -value = 0.2439.

3. Methodology

Figure 1 summarizes the methodology addressed in this paper. The information from the sequence of numbers is analyzed following two approaches: image-based and dynamic. The first approach only uses the data of the image generated by the position information of the tablet, making this task similar to offline handwriting. This approach considers a Convolutional Neural Network (CNN) pre-trained using sequences created with numbers from the MNIST database. Sizes of digits in the MNIST corpus are changed to artificially create the effect of micrographia. Then, the CNN model is fine-tuned using the handwriting digits of PD and HC subjects. The second approach constitutes a dynamic analysis where we use all information available when data are captured using a digital tablet. Two sets of classical handwriting features are computed from the 6 time-series signals available in online handwriting (x and y position, z (distance from the tablet surface to the pen tip), azimuth and altitude angles, and pen pressure). Finally, feature vectors obtained from the image-based and dynamic analyses are used independently to train two Support Vector Machine (SVM) classifiers to discriminate between PD patients and HC subjects.

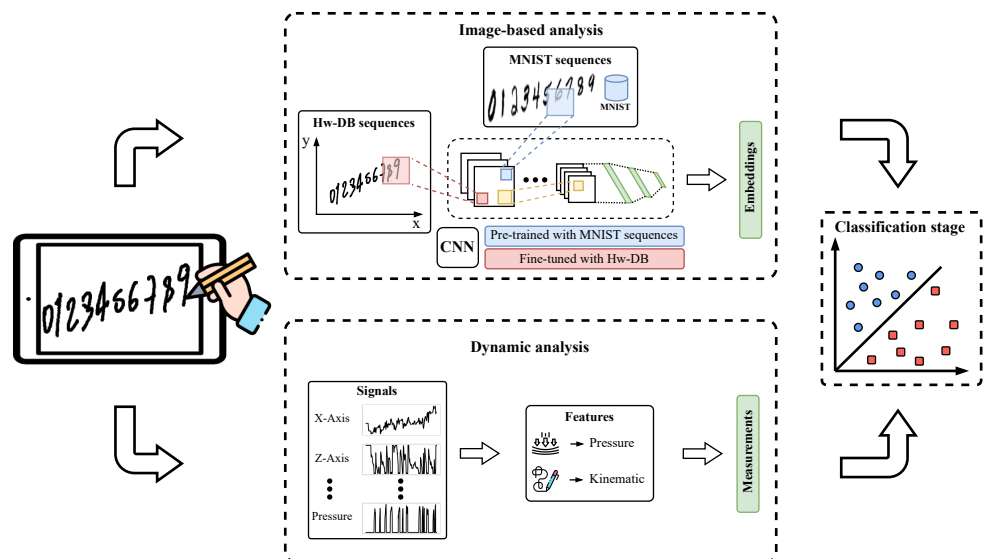


Figure 1. General methodology proposed in this study.

3.1. Feature Extraction

3.1.1. Image-Based Analysis

This model is based on a CNN architecture, a widely used modeling strategy in image processing [30]. In each layer, different filters are trained to detect patterns while the spatial dimensions of the image are reduced to obtain a compressed image representation. The CNN model training requires many samples to generalize patterns and avoid model overfitting; this represents a limitation in the study of medical data, where only a few samples are available. Transfer Learning (TL) and Data Augmentation (DA) techniques have emerged to address the problem of data scarcity [28,31–33].

In this study, we created 7000 synthetic sequences of numbers from 0 to 9 using the digits of the MNIST database [34]. Half of the sequences contain concatenated digits, while the rest are modified to model micrographia, which is a distinct symptom of handwriting in PD patients. The idea behind this approach is to pre-train the architecture to recognize non-modified sequences (synthetic samples of HC subjects, i.e., without micrographia) vs. modified sequences (synthetic samples of PD patients, i.e., with micrographia), which implicitly means having an architecture properly trained to detect micrographia. The starting digit for the modified sequence is randomly chosen from 0 to 5. Once the first

modified digit is chosen, the rest of the sequence, until 9, is progressively reduced until the last digit (9) results in half of its original size. Examples of the modified and non-modified sequences of digits from MNIST are shown in Figure 2. In addition, given the fact that several sequences in the Hw-DB exhibited slight rotations resulting from the writing process of participants, all the synthetic sequences (modified and non-modified) were randomly rotated at angles of -15° , -5° , 0° , 5° , 15° . From the set of generated sequences, a subset of 6000 synthetic sequences was used to train a CNN architecture to classify between modified and non-modified sequences. The remaining synthetic sequences were used as a test set. The CNN model employed in this experiment is based on a LeNet architecture because it constitutes a simple model appropriate for the size of the corpus that we could access to develop this study. Additionally, this architecture showed as good results as in a previous work where we modeled digits [35]. This network consists of 3 convolutional layers with 16, 8, and 4 filters, employing a (5×5) kernel. Max-pooling layers were used to reduce the size of each feature map by half. The output feature map was flattened to feed three fully connected layers with 512, 64, and 2 neurons, respectively. The first two fully connected layers employed a rectified linear unit (ReLU) activation function, and the last one made the final decision using a Softmax activation function. The shape of input images is 144 in height and 216 in width.

The pre-trained architecture and its weights are used as the starting point to fine-tune a specific architecture for the automatic discrimination between PD and HC subjects. The fine-tuning process is performed with data from the Hw-DB corpus. Once the architecture is fine-tuned, we take the embedding of the flattened feature map as a feature vector to feed the SVM classifier.

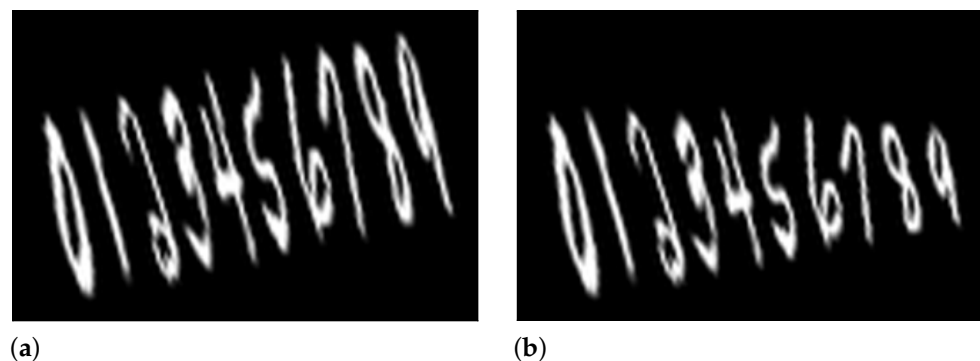


Figure 2. Sequences created by MNIST database. (a) Non-modified sequence. (b) Modified sequence.

3.1.2. Dynamic Analysis

Although the image-based representation method has the advantage of being suitable for online and offline handwriting, it does not allow modeling motor problems such as rigidity, bradykinesia, or freezing of the upper limbs. These motor problems affect the dynamics of handwriting, leading to changes in velocity, acceleration, and fluency, which cannot be captured from the resulting image. In contrast, the dynamic approach leverages all the available data from online handwriting, where different feature sets are computed using the time series signals provided by the digital tablet. Table 2 shows a summary of all computed features, the details of which are described below.

Pressure-based features: Pressure features describe the mechanical force exerted on the pen tip during on-surface movements produced along the writing process. Common pressure features are based on statistical functionals computed over (1) the raw pressure signal ($p[n]$), (2) changes in pressure signal ($\Delta p = p[n+1] - p[n]$), (3) the rate of pressure changes over time ($p'[n] = \frac{\Delta p}{\Delta t}$), (4) the rate of pressure variability ($p''[n] = \frac{\Delta^2 p}{\Delta t^2}$), and (5) the

pressure jerk ($p'''[n] = \frac{\Delta^3 p}{\Delta t^3}$). Additionally, we can measure the number of changes in pressure (NCP) and the relative number of changes in pressure ($RNCP = \frac{NCP}{time}$), using (Δp , $p'[n]$, and $p''[n]$). The resulting feature set includes a total of 26 pressure features.

Kinematic features: According to the status of the z -axis signal, handwriting signals can be grouped into on-surface and in-air movements. On-surface samples correspond to digits' strokes; conversely, in-air samples correspond to hand movements during the transition between digits. To compute kinematic features from on-surface and in-air movements, we employed the set of signals $\{x, y, azimuth, altitud\}$. Notice that features computed from the z -axis signal contain samples of in-air movements only. Furthermore, we employ the x and y axes to compute the pen trajectory $r[n]$ defined in Equation (1), and the pen displacement $Di[n]$ defined in Equation (2). Then, the set of kinematic features contains (1) movement descriptors as $\{\Delta x, \Delta y, \Delta azimuth, \Delta altitud, \Delta z, r[n], Di[n]\}$, (2) velocity descriptors computed as changes of the previous signals over time ($sig_{vel}[n] = \frac{\Delta sig}{\Delta t}$), (3) acceleration descriptors as changes of velocity over time ($sig_{acc}[n] = \frac{\Delta sig_{vel}}{\Delta t}$), (4) jerk as the changes of acceleration over time ($sig_{jerk}[n] = \frac{\Delta sig_{acc}}{\Delta t}$), and (5) number of changes of velocity and acceleration, NCV and NCA, respectively. Finally, we extracted a total of 120 and 140 kinematic features from on-surface and in-air movements, respectively. Table 2 presents a comprehensive summary of the features extracted in this work.

$$r[n] = \sqrt{x[n]^2 + y[n]^2} \quad (1)$$

$$Di[n] = \sqrt{(x[n+1] - x[n])^2 + (y[n+1] - y[n])^2} \quad (2)$$

Table 2. Summary of the computed feature in each set. s = scalar value, and v = vector of elements.

Set	Feature	s/v	Description
Pressure	$p[n], \Delta p, p'[n], p''[n], \text{ and } p'''[n]$	v	Raw pressure, pressure changes, first, second and third derivatives
	NCP, and RNCP	s	Number of local extrema of pressure
Kinematic	$r[n], Di[n], \Delta x, \Delta y, \Delta azimuth, \Delta altitud, \Delta z$	v	Trajectory, displacements, and signal changes.
	Velocity	v	Velocity computed as changes in signals w.r.t. time
	Acceleration	v	Acceleration computed as changes in signal velocity w.r.t. time
	Jerk	v	Jerk computed as changes in signal acceleration w.r.t. time
	NCV and RNCV	s	Number of local extrema for velocity
	NCA and RNCA	s	Number of local extrema for acceleration

3.2. Statistical Modeling

The feature computation process introduced above can result in either a scalar value or a vector (see the third column in Table 2). The resulting vectors generate a feature matrix per subject. To obtain a static vector representation per subject, we consider two statistical modeling strategies, one based on statistical functionals and the other one based on Gaussian Mixture Models (GMMs). Scalar values are concatenated with resulting statistical representations.

Dynamic feature vectors that change over time are commonly represented using statistical functionals to describe their statistical distribution. However, there exist other

robust methods like the GMM that can be applied to represent dynamic phenomena. In this section, we compare classical statistical modeling based on functionals vs. GMM modeling.

3.2.1. Statistical Functionals

Statistical functionals are used to describe properties of data distributions. These functionals capture specific characteristics such as central tendency, variability, or higher-order moments. This strategy is commonly used to obtain fixed representations from time-dependent signals. We consider 4 statistical functionals, namely, mean, standard deviation, skewness, and kurtosis, forming a static representation per feature vector. The dimensionality of the kinematic feature vector is now $4 \times 52 = 208$, while the dimensionality of the pressure feature vector is $4 \times 5 = 20$.

3.2.2. Gaussian Mixture Models (GMMs)

GMM representations enable modeling complex dynamics of time-dependent signals. GMM is a probabilistic model that combines multiple Gaussian distributions to obtain a tighter representation of the data distribution. Equation (3) defines a GMM with M Gaussians, where each Gaussian's contribution is weighted by the parameter c . The parameters μ_m and Σ_m represent the mean vector and covariance matrix of the m -th Gaussian, respectively.

$$f(x) = \sum_{m=1}^M c_m \mathcal{N}(x; \mu_m, \Sigma_m), \quad (3)$$

The idea behind the GMM model is that each Gaussian in the mixture models a sub-population along the temporal dynamics. Once the GMM is calculated, a fixed representation is obtained by combining the mean vectors and covariance matrices of all the Gaussians in the mixture. This representation is known as the GMM supervector (λ) [36]. The dimension of the supervector depends on the number of features in the input matrix and the number of Gaussians, where $\lambda_{dimension} = M \times F \times 2$. M is the number of Gaussian components, and F is the number of vector measurements in the feature matrix. We optimized the number of Gaussians based on the obtained accuracy during training. The dimensionality of the supervector that results to represent a given phenomenon or feature vector is computed as follows: optimal number of Gaussians $\times$ number of feature vectors $\times 2$. This last number appears because the supervector is formed with the entries of the mean vector and the diagonal of the covariance matrix. For instance, when the optimal number of Gaussians is GMM with $M = 4$, the cardinality of the resulting λ is 416-dimensional ($4 \times 52 \times 2$).

4. Experiments and Results

4.1. Experimental Setup

We conducted two experiments in this work: image-based analysis and dynamic analysis. The image-based analysis involved fine-tuning a pre-trained CNN model, described in Section 3.1.1. The dynamic analysis is based on two feature sets: pressure and kinematic. Each feature set was modeled considering two approaches for comparison purposes, statistical functionals and GMMs. All models were trained, fine-tuned, and tested following the same partitions according to a 5-fold cross-validation strategy. The hyperparameters of the SVM classifier, namely C , γ , and the *kernel*, were optimized using a grid-search, where C and $\gamma \in \{1e^{-3}, 1e^{-2}, \dots, 1e^3\}$, and *kernel* $\in \{\text{linear, rbf}\}$. Notice that we evaluated two more classification methods, namely Random Forest and Gradient Boosting (GB); however, we decided to report only the ones obtained with the SVM because all accuracies with that classifier were higher.

4.2. Experiment 1: Image-Based Analysis

The pre-trained CNN architecture used to discriminate between non-modified and modified sequences was further used as a feature extractor for images in the Hw-DB corpus. We took the flattened output of the last convolutional layer as a feature vector to feed an SVM classifier to distinguish between PD patients and HC subjects. To improve the characterization capability of the pre-trained CNN, we considered four fine-tuning schemes over the base model using data from Hw-DB: fully frozen, partially frozen, semi-frozen, and unfrozen. In the fully frozen schema, the weights and biases of all layers in the CNN were fixed during the training process; therefore, the pre-trained CNN only used the weights and biases obtained with the synthetic data. In the partially frozen schema, the first convolutional layer is unfrozen; thus, only the weights and biases of this layer can be fine-tuned with the Hw-DB data. In the semi-frozen schema, the first and second convolutional layers were fine-tuned. Finally, in the unfrozen schema, the weights and bias of all convolutional layers were fine-tuned. Results of the image-based experiments are shown in Table 3.

Table 3. Results from the image-based analysis of the digit sequences using a pre-trained CNN. Values reported in terms of ($\mu \pm \sigma$).

Fine-Tuning	Accuracy (%)	Specificity (%)	Sensitivity (%)	F1-Score (%)
Fully frozen	55.7 $\pm$ 9.3	59.6 $\pm$ 19.9	50.7 $\pm$ 13.3	0.52 $\pm$ 0.10
Partially frozen	57.6 $\pm$ 8.4	61.5 $\pm$ 19.8	52.9 $\pm$ 13.3	0.54 $\pm$ 0.09
Semi-frozen	62.3 $\pm$ 11.8	65.5 $\pm$ 18.2	58.7 $\pm$ 18.5	0.60 $\pm$ 0.13
Unfrozen	61.4 $\pm$ 9.4	63.6 $\pm$ 17.8	58.7 $\pm$ 13.5	0.59 $\pm$ 0.09

Results indicate a progressive improvement in terms of accuracy when more convolutional layers are unfrozen. This suggests that the pre-trained model needs more knowledge from Parkinson's data (i.e., the Hw-DB corpus). However, a slight decrease in the model's accuracy is observed when the three convolutional layers are unfrozen. This could be attributed to the limited amount of data in the Hw-DB, which may not be sufficient to properly fine-tune all filters in the layers. The semi-frozen schema shows the best results, with an accuracy of up to 62.3%.

4.3. Experiment 2: Dynamic Analysis

Pressure and kinematics features, extracted from the time series provided by the digital tablet, are used in this experiment. An early fusion of these two feature sets was also considered. We consider two strategies to obtain a fixed representation from these time-dependent characterization strategies, namely statistical functionals and GMMs. Although the first strategy is simple, it constitutes a well-established approach to statically represent phenomena that originate from time-dependent feature sets. Concerning the GMM-based strategy, the number of Gaussians was optimized according to the accuracy obtained in training such that $M = 2, 4, 6, \dots, 30$. Figure 3 shows the results obtained in the training process with each number of Gaussians per feature set. Notice that they are sorted from the lowest to the highest accuracy. Therefore, the optimal number of Gaussians is the one indicated at the right-hand side of each figure.

Table 4 summarizes the results obtained in this experiment with both strategies (statistical functionals and GMMs). Results indicate that pressure is the feature set with the best performance, achieving an accuracy of 75% when using the statistical functionals. Notice that results based on statistical functionals outperformed those obtained with GMMs when the pressure feature set was considered. In contrast, when only kinematic features were

considered, both strategies showed similar results. Finally, results with the early fusion strategy did not show improvement.

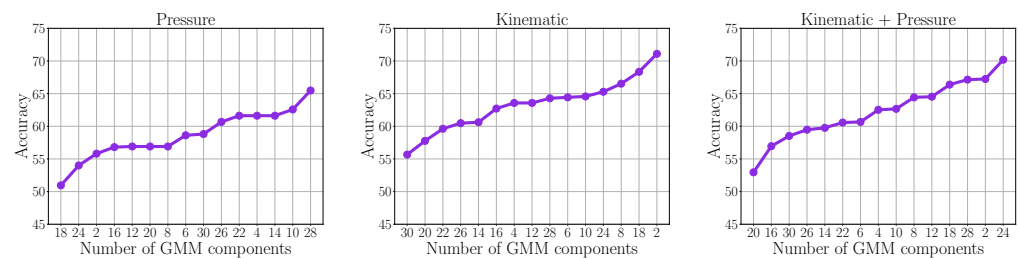


Figure 3. Accuracy mean obtained in training folds with different feature sets when changing the number of Gaussians in the GMM models.

Table 4. Values are reported as $(\mu \pm \sigma)$.

Features	Accuracy (%)	Specificity (%)	Sensitivity (%)	F1-Score (%)
Statistical functional				
Pressure	75.0 $\pm$ 5.2	79.3 $\pm$ 12.1	70.5 $\pm$ 7.2	73.4 $\pm$ 5.1
Kinematic	71.3 $\pm$ 12.4	73.8 $\pm$ 21.4	68.5 $\pm$ 15.0	69.9 $\pm$ 11.7
Kinematic + Pressure	71.3 $\pm$ 14.1	70.1 $\pm$ 19.2	72.5 $\pm$ 13.0	71.4 $\pm$ 12.7
GMMs				
Pressure (with 28 Gaussians)	65.5 $\pm$ 8.2	62.2 $\pm$ 15.7	68.5 $\pm$ 4.9	66.3 $\pm$ 3.3
Kinematic with 28 Gaussians	71.1 $\pm$ 7.1	73.5 $\pm$ 15.5	68.4 $\pm$ 13.5	69.5 $\pm$ 8.3
Kinematic + Pressure (with 28 Gaussians)	70.2 $\pm$ 15.6	69.5 $\pm$ 17.4	71.1 $\pm$ 21.8	69.5 $\pm$ 17.1

5. Discussion

Three different approaches were presented in this paper: one based on images resulting from the handwriting process and two based on time-dependent signals collected from a digital tablet. This discussion includes remarkable aspects of each approach and a comparison among them.

Our experiments show that classification results with the image-based features improve as the architecture adjusts more weights. The best result is obtained in the semi-frozen schema, where we allowed the fine-tuning of weights in two convolutional layers. The unfrozen schema shows slightly less accuracy. This behavior suggests that the CNN needs more real samples from the database to obtain better classification results.

The dynamic analysis yields the best results when considering features extracted from the pressure signal. These results outperform those obtained with the kinematic feature set, suggesting that vertical control of hand movement while writing is more challenging for PD patients than the horizontal ones.

When comparing models based on statistical functionals vs. GMMs, we found that the first approach is more accurate when considering the pressure feature set. Kinematic features yielded similar results with both statistical modeling strategies.

Figure 4 shows the ROC curves and their corresponding distribution of scores for the best results achieved with each modeling approach. The AUC values confirm the aforementioned claims. Image-based features are outperformed by the dynamic analyses with the pressure feature set using statistical functionals. This result is in line with the findings reported in Tables 3 and 4, where the best accuracy obtained with the dynamic approach outperformed by up to 13% the best accuracy obtained with the image-based approach. We believe that the low performance of the image-based approach is because, although micrographia is a key feature in handwriting deficiencies, the writing process

itself (i.e., the dynamic of handwriting over time) is the one that reveals motor symptoms more clearly.

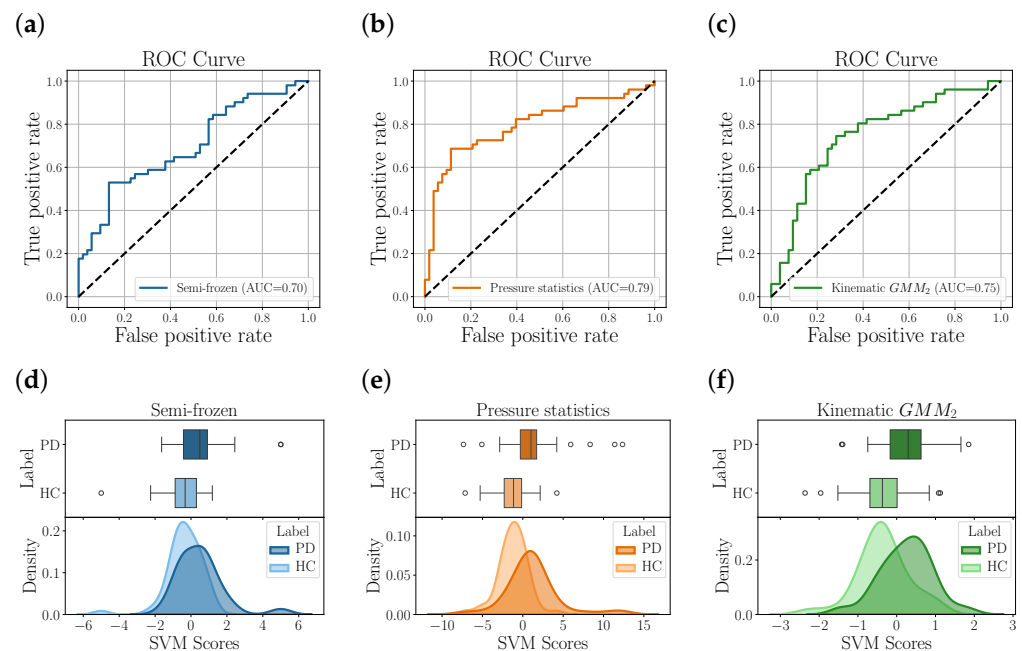


Figure 4. Results of the image-based analysis and the two dynamic analyses (modeled with statistical functionals and GMMs). The first row presents the ROC curves for (a) semi-frozen scheme, (b) pressure features modeled by statistical functionals, and (c) kinematic features modeled by two GMM components. The second row (d–f) depicts the distribution of the corresponding SVM decision scores resulting from the classification with the three feature sets.

We believe that this study constitutes a step forward in the development of automated systems that could be used to diagnose and monitor PD progression. For future work, we will focus on modeling transitions that occur while writing because our previous works show promising results [37] and because other studies have recently reported interesting insights towards this direction in neurodegenerative and immune diseases [38].

6. Conclusions

The present study showed that time-dependent feature sets extracted from digital tablets (namely, pressure and kinematic) are more accurate than those features extracted from images resulting from the writing process, when discriminating between PD patients and HC subjects. Additionally, our results show that the fusion of pressure and kinematic features does not improve the accuracy. Time-dependent features are represented following two strategies, one based on statistical functionals and another one based on GMMs. The comparison between using statistical functionals and using GMMs showed that the first is more suitable and yields higher accuracies. The main advantage of using statistical functionals is their direct interpretability, while the use of GMMs requires more sophisticated computations and their interpretation is not as easy, therefore limiting their use in clinical practice. This paper only considered the writing of numbers from 0 to 9; we believe that the results will stay the same when considering other tasks, like those based on words or sentences, because all these tasks are natural and belong to the set of daily living activities. Therefore, findings reported in this paper constitute a step forward in the process of creating clinically informative methods to model Parkinson's symptoms. Further research is required considering more writing tasks and also a larger number of patients to

make it possible to create more sophisticated models like those based on recurrent neural networks or long short-term memory (LSTM) models.

Author Contributions: Conceptualization, D.E.-G., J.F.V.-B., A.M.G. and J.R.O.-A.; Methodology, D.E.-G., C.D.R.-U., A.M.G. and J.R.O.-A.; Software, J.D.G.-A., D.E.-G. and C.D.R.-U.; Validation, J.D.G.-A. and D.E.-G.; Formal analysis, J.D.G.-A., D.E.-G. and J.R.O.-A.; Investigation, J.D.G.-A. and D.E.-G.; Data curation, J.D.G.-A. and C.D.R.-U.; Writing—original draft, J.D.G.-A. and D.E.-G.; Writing—review & editing, D.E.-G. and J.R.O.-A.; Visualization, J.D.G.-A.; Supervision, D.E.-G., A.M.G. and J.R.O.-A.; Project administration, J.F.V.-B. and J.R.O.-A.; Funding acquisition, J.F.V.-B. and J.R.O.-A. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been funded by the School of Engineering at UdeA and Pratech Group S.A.S. grants # IAPFI23-1-01 and # PI2023-58010. Adolfo M. García is an Atlantic Fellow at the Global Brain Health Institute (GBHI) and is partially supported by the National Institute On Aging of the National Institutes of Health (R01AG075775, 2P01AG019724); ANID (FONDECYT Regular 1210176, 1210195); DICYT-USACH (032351G-DAS); Agencia Nacional de Promoción Científica y Tecnológica (01-PICTE-2022-05-00103); Agencia Nacional de Investigación e Innovación (EI-X-2023-1-176993); and the Multi-partner Consortium to Expand Dementia Research in Latin America (ReDLat), which is supported by the Fogarty International Center and the National Institutes of Health, the National Institute on Aging (R01AG057234, R01AG075775, R01AG21051, and CARDS-NIH), Alzheimer’s Association (SG-20- 725707), Rainwater Charitable Foundation’s Tau Consortium, the Bluefield Project to Cure Frontotemporal Dementia, and the Global Brain Health Institute.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Hornykiewicz, O. Biochemical aspects of Parkinson’s disease. *Neurology* **1998**, *51*, S2–S9. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Meder, D.; Herz, D.M.; Rowe, J.B.; Lehericy, S.; Siebner, H.R. The role of dopamine in the brain—lessons learned from Parkinson’s disease. *Neuroimage* **2019**, *190*, 79–93. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Sveinbjornsdottir, S. The clinical symptoms of Parkinson’s disease. *J. Neurochem.* **2016**, *139*, 318–324. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Thomas, M.; Lenka, A.; Kumar Pal, P. Handwriting analysis in Parkinson’s disease: Current status and future directions. *Mov. Disord. Clin. Pract.* **2017**, *4*, 806–818. [\[CrossRef\]](#)
5. Letanneux, A.; Danna, J.; Velay, J.; Viallet, F.; Pinto, S. From micrographia to Parkinson’s disease dysgraphia. *Mov. Disord.* **2014**, *29*, 1467–1475. [\[CrossRef\]](#)
6. D’Alessandro, T.; Carmona-Duarte, C.; De Stefano, C.; Diaz, M.; Ferrer, M.; Fontanella, F. A Machine Learning Approach to Analyze the Effects of Alzheimer’s Disease on Handwriting Through Lognormal Features. In Proceedings of the Graphonomics in Human Body Movement: Bridging Research and Practice from Motor Control to Handwriting Analysis and Recognition (IGS), Évora, Portugal, 16–19 October 2023; pp. 103–121.
7. D’Alessandro, T.; De Stefano, C.; Fontanella, F.; Nardone, E.; Pace, C. From Handwriting Analysis to Alzheimer’s Disease Prediction: An Experimental Comparison of Classifier Combination Methods. In Proceedings of the Document Analysis and Recognition—ICDAR 2024, Athens, Greece, 30 August–4 September 2024; pp. 334–351.
8. Impedovo, D.; Pirlo, G. Dynamic Handwriting Analysis for the Assessment of Neurodegenerative Diseases: A Pattern Recognition Perspective. *IEEE Rev. Biomed. Eng.* **2019**, *12*, 209–220. [\[CrossRef\]](#)
9. Aouraghe, I.; Khaissidi, G.; Mrabti, M. A literature review of online handwriting analysis to detect Parkinson’s disease at an early stage. *Multimed. Tools Appl.* **2023**, *82*, 11923–11948. [\[CrossRef\]](#)
10. Vessio, G. Dynamic Handwriting Analysis for Neurodegenerative Disease Assessment: A Literary Review. *Appl. Sci.* **2019**, *9*, 4666. [\[CrossRef\]](#)
11. Pereira, C.R.; Weber, S.A.; Hook, C.; Rosa, G.H.; Papa, J.P. Deep learning-aided Parkinson’s disease diagnosis from handwritten dynamics. In Proceedings of the 2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Sao Paulo, Brazil, 4–7 October 2016; pp. 340–346.

12. Kotsavasiloglou, C.; Kostikis, N.; Hristu-Varsakelis, D.; Arnaoutoglou, M. Machine learning-based classification of simple drawing movements in Parkinson's disease. *Biomed. Signal Process. Control* **2017**, *31*, 174–180. [\[CrossRef\]](#)
13. Mucha, J.; Zvoncak, V.; Galaz, Z.; Faundez-Zanuy, M.; Mekyska, J.; Kiska, T.; Smekal, Z.; Brabenec, L.; Rektorová, I.; Lopez-de Ipina, K. Fractional derivatives of online handwriting: A new approach of parkinsonic dysgraphia analysis. In Proceedings of the 2018 41st International Conference on Telecommunications and Signal Processing (TSP), Athens, Greece, 4–6 July 2018; pp. 1–4.
14. Zham, P.; Arjunan, S.P.; Raghav, S.; Kumar, D.K. Efficacy of guided spiral drawing in the classification of Parkinson's disease. *IEEE J. Biomed. Health Inform.* **2017**, *22*, 1648–1652. [\[CrossRef\]](#)
15. Valla, E.; Nomm, S.; Medijainen, K.; Taba, P.; Toomela, A. Tremor-related feature engineering for machine learning based Parkinson's disease diagnostics. *Biomed. Signal Process. Control* **2022**, *75*, 103551. [\[CrossRef\]](#)
16. Pereira, C.R.; Pereira, D.R.; Rosa, G.H.; Albuquerque, V.H.; Weber, S.A.; Hook, C.; Papa, J.P. Handwritten dynamics assessment through convolutional neural networks: An application to Parkinson's disease identification. *Artif. Intell. Med.* **2018**, *87*, 67–77. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Gil-Martín, M.; Montero, J.M.; San-Segundo, R. Parkinson's disease detection from drawing movements using convolutional neural networks. *Electronics* **2019**, *8*, 907. [\[CrossRef\]](#)
18. Kamran, I.; Naz, S.; Razzak, I.; Imran, M. Handwriting dynamics assessment using deep neural network for early identification of Parkinson's disease. *Future Gener. Comput. Syst.* **2021**, *117*, 234–244. [\[CrossRef\]](#)
19. Drotár, P.; Mekyska, J.; Rektorová, I.; Masarová, L.; Smékal, Z.; Faundez-Zanuy, M. Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease. *Artif. Intell. Med.* **2016**, *67*, 39–46. [\[CrossRef\]](#)
20. Pereira, C.R.; Pereira, D.R.; Silva, F.A.; Masieiro, J.P.; Weber, S.A.; Hook, C.; Papa, J.P. A new computer vision-based approach to aid the diagnosis of Parkinson's disease. *Comput. Methods Programs Biomed.* **2016**, *136*, 79–88. [\[CrossRef\]](#)
21. Li, Z.; Yang, J.; Wang, Y.; Cai, M.; Liu, X.; Lu, K. Early diagnosis of Parkinson's disease using Continuous Convolution Network: Handwriting recognition based on off-line hand drawing without template. *J. Biomed. Inform.* **2022**, *130*, 104085. [\[CrossRef\]](#)
22. Galaz, Z.; Drotar, P.; Mekyska, J.; Gazda, M.; Mucha, J.; Zvoncak, V.; Smekal, Z.; Faundez-Zanuy, M.; Castrillon, R.; Orozco-Arroyave, J.R.; et al. Comparison of CNN-Learned vs. Handcrafted Features for Detection of Parkinson's Disease Dysgraphia in a Multilingual Dataset. *Front. Neuroinform.* **2022**, *16*, 877139. [\[CrossRef\]](#)
23. Gazda, M.; Hireš, M.; Drotár, P. Multiple-fine-tuned convolutional neural networks for Parkinson's disease diagnosis from offline handwriting. *IEEE Trans. Syst. Man Cybern. Syst.* **2021**, *52*, 78–89. [\[CrossRef\]](#)
24. Miler Jerkovic, V.; Kojic, V.; Dragasevic Miskovic, N.; Djukic, T.; Kostic, V.S.; Popovic, M.B. Analysis of on-surface and in-air movement in handwriting of subjects with Parkinson's disease and atypical parkinsonism. *Biomed. Eng. Tech.* **2019**, *64*, 187–194. [\[CrossRef\]](#)
25. Angelillo, M.T.; Impedovo, D.; Pirlo, G.; Vessio, G. Performance-driven handwriting task selection for Parkinson's disease classification. In Proceedings of the AI*IA 2019—Advances in Artificial Intelligence: XVIIIth International Conference of the Italian Association for Artificial Intelligence, Rende, Italy, 19–22 November 2019; Proceedings 18. Springer: Cham, Switzerland, 2019; pp. 281–293.
26. Aouraghe, I.; Alae, A.; Ghizlane, K.; Mrabti, M.; Aboulem, G.; Faouzi, B. A novel approach combining temporal and spectral features of Arabic online handwriting for Parkinson's disease prediction. *J. Neurosci. Methods* **2020**, *339*, 108727. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Rios-Urrego, C.D.; Vásquez-Correa, J.C.; Vargas-Bonilla, J.F.; Nöth, E.; Lopera, F.; Orozco-Arroyave, J.R. Analysis and evaluation of handwriting in patients with Parkinson's disease using kinematic, geometrical, and non-linear features. *Comput. Methods Programs Biomed.* **2019**, *173*, 43–52. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Diaz, M.; Ferrer, M.A.; Impedovo, D.; Pirlo, G.; Vessio, G. Dynamically enhanced static handwriting representation for Parkinson's disease detection. *Pattern Recognit. Lett.* **2019**, *128*, 204–210. [\[CrossRef\]](#)
29. Moetesum, M.; Siddiqi, I.; Vincent, N.; Cloppet, F. Assessing visual attributes of handwriting for prediction of neurological disorders—A case study on Parkinson's disease. *Pattern Recognit. Lett.* **2019**, *121*, 19–27. [\[CrossRef\]](#)
30. Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 6999–7019. [\[CrossRef\]](#)
31. Tan, C.; Sun, F.; Kong, T.; Zhang, W.; Yang, C.; Liu, C. A survey on deep transfer learning. In Proceedings of the Artificial Neural Networks and Machine Learning—ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, 4–7 October 2018; Proceedings, Part III 27; Springer: Cham, Switzerland, 2018; pp. 270–279.
32. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. *J. Big Data* **2019**, *6*, 60. [\[CrossRef\]](#)
33. Cilia, N.; D'Alessandro, T.; De Stefano, C.; Fontanella, F. Deep transfer learning algorithms applied to synthetic drawing images as a tool for supporting Alzheimer's disease prediction. *Mach. Vis. Appl.* **2022**, *33*, 49. [\[CrossRef\]](#)
34. Cohen, G.; Afshar, S.; Tapson, J.; Van Schaik, A. EMNIST: Extending MNIST to handwritten letters. In Proceedings of the Proceedings of IJCNN, Anchorage, AK, USA, 14–19 May 2017; pp. 2921–2926.

35. Gallo-Ariscitizabal, J.; Escobar-Grisales, D.; Ríos-Urrego, C.; Pérez-Toro, P.; Nöth, E.; Maier, A.; Orozco-Arroyave, J. Assessment of handwriting in patients with Parkinson's disease using non-intrusive tasks. In Proceedings of the International Symposium on Biomedical Imaging (ISBI), Cartagena, Colombia, 18–21 April 2023; pp. 1–4.
36. Ríos-Urrego, C.; Rusz, J.; Nöth, E.; Orozco-Arroyave, J. Automatic Classification of Hypokinetic and Hyperkinetic Dysarthria based on GMM-Supervectors. In Proceedings of the Annual Conference of the International Speech Communication Association, Dublin, Ireland, 20–24 August 2023; INTERSPEECH 2023.
37. Vásquez-Correa, J.; Arias-Vergara, T.; Orozco-Arroyave, J.; Eskofier, B.; Klucken, J.; Nöth, E. Multimodal assessment of Parkinson's disease: A deep learning approach. *IEEE J. Biomed. Health Inform.* **2019**, *23*, 1618–1630. [[CrossRef](#)]
38. Romijnders, R.; Atrsaie, A.; Rehman, R.Z.U.; Strehlow, L.; Massoud, J.; Hinchliffe, C.; Macrae, V.; Emmert, K.; Reilmann, R.; van der Woude, C.J.; et al. Association of real life postural transitions kinematics with fatigue in neurodegenerative and immune diseases. *npj Digit. Med.* **2025**, *8*, 12. [[CrossRef](#)]

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.