AGE and GENDER PREDICTON from HAND WRITTING TEXT using ATTENTION-BASED CNN

MEGHA MERIN ROY[1]

[1]*Computer Science and Engineering, APJ Abdul Kalam Technological University, India*

---***---

Abstract - The goal is to predict the age range and gender of a person based on their handwriting. Two feature extraction methods, Histogram of Oriented Gradients and Linear Discriminant Analysis, are employed, along with two machine learning techniques, Support Vector Machine and Attention Based Convolutional Neural Network. For each feature extraction method, the system predicts the age range and gender using SVM and ABCNN. The system provides accuracy metrics for each prediction to decide which technique is best, indicating the reliability of the age and gender estimations. This is designed to work in real-time, enabling quick predictions for incoming handwriting samples.

Key Words **:** Histogram of Oriented Gradients (HOG), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), Attention Based Convolutional Neural Network (ABCNN) , Convolutional Neural Network (CNN)

1. INTRODUCTION

Handwriting analysis has proven to be a valuable tool in many organizational processes, including recruitment, interviewing, team-building, counseling, and career planning. It is a reliable indicator of personality and behavior, making it an important subject in modern times. Each individual has a unique writing style that reflects their personality, and their handwriting may vary even when writing the same group of words multiple times.

The handwriting analysis system is responsible for handling formatting, character segmentation, and word recognition. The use of handwriting analysis has grown in real-time applications, particularly in the forensic and information security fields, where it can provide insights into a writer's psychological condition and characteristics.

The two gradient features to predict the writer's age range, gender, and handedness. The first feature is the Histogram of Oriented Gradients, which detects the distribution of gradient orientations in images, while the second feature is the gradient local binary pattern, an enhanced gradient feature that computes the gradient by considering the local binary pattern neighborhood. The Attention-Based CNN, which incorporates attention mechanisms, is used to accomplish the prediction task. The ABCNN model is widely utilized in natural language processing tasks, including modeling a pair of sentences.

2. PROPOSED SYSTEM

The proposed system being developed aims to utilize an Attention-Based Convolutional Neural Network for the purpose of detecting age and gender from handwriting analysis. The application of handwriting analysis has increased in fields such as forensics and information security, and it also enables experts to gain insight into a writer's psychological state and characteristics.

The aim of the system is to automatically classify writers based on their gender and age range, using two gradient features: Histogram of Oriented Gradients and gradient local binary patterns. Histogram of Oriented Gradients (HOG) is a gradient feature descriptor that represents the distribution of gradient orientations within an image. Local Binary Patterns (LBP) is a texture descriptor that considers the relationship between a central pixel and its neighbors in an image. To generate features, images of variable sizes are segmented into a fixed number of cells using a grid, and the gradient features are computed and concatenated for each cell to create the image feature vector. The System Architecture is explained in Fig 1. The prediction task is accomplished using an Attention Based Convolution Neural Network. The Attention-Based Convolutional Neural Network (ABCNN) is used in the proposed system to predict the gender and age range of writers based on extracted gradient features from handwritten text.

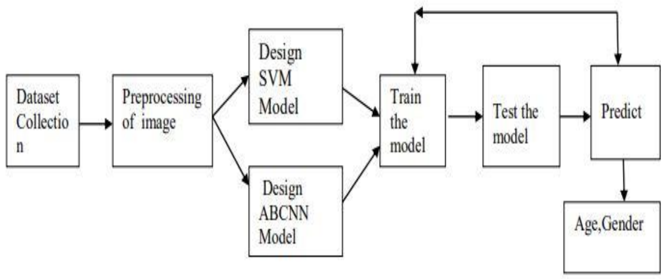

Fig 1 : System Architecture

The IAM Handwriting Database is a publicly available on-line handwriting dataset consisting of forms of handwritten English text. It can be utilized for training and testing handwritten text recognition models as well as conducting writer identification and verification experiments.

The objective of pre-processing is to enhance important image features or suppress unwanted distortions in the image data. Although techniques such as rotation, scaling, and translation fall under geometric transformations of images, they are also classified as pre-processing methods. Pre-processing is a set of operations carried out at the lowest level of abstraction, where both input and output are intensity images represented by a matrix of image function values.

2.1 Histogram of Oriented Gradients :

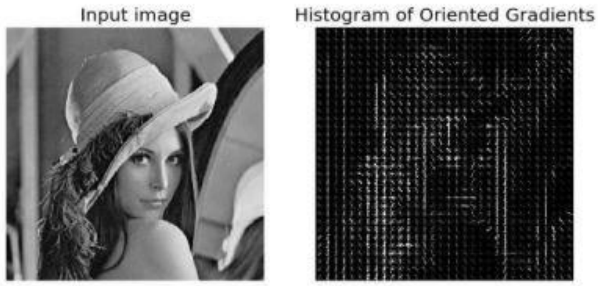

Fig 2 : Histogram of Oriented Gradients

Histogram of Oriented Gradients (HOG) is a widely used feature descriptor in computer vision and image processing is shown in Fig 2. It is typically used for object detection and recognition tasks, where it counts the occurrences of different gradient orientations to describe the shape or structure of objects in an image. Specifically, the HOG descriptor characterizes the shape of objects based on the distribution of local intensity gradients or edge directions. In other words, it calculates the distribution of gradient orientations within an image patch, typically divided into a grid of cells. Each cell contributes a histogram of gradient directions, and these histograms are concatenated into a single feature vector that summarizes the image patch. The resulting HOG feature vector is invariant to small changes in position and illumination, making it an effective descriptor for object detection and recognition tasks. HOG has shown high performance in various applications, including human detection, pedestrian detection, and face detection.

Volume: 07 Issue: 08 | August - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Algorithm :

1)Gradient information is calculated for both the horizontal and vertical directions of each pixel I(x, y), as follows:

$g_x (x, y) = I (x + 1, y) – I (x – 1, y)$

$g_y (x, y) = I (x, y + 1) – I (x, y - 1)$

2)Obtaining the gradient magnitude and phase is achieved by utilizing the following equations:

$M (x, y) = \sqrt{g_x}(x, y)^2 + \sqrt{g_y}(x, y)^2$

$\Theta (x, y) = \arctan (g_x (x,y) / g_y(x,y))$

3)To construct the histogram of gradients, the magnitude is accumulated according to orientations.

2.2 Gradient Local Binary Patterns :

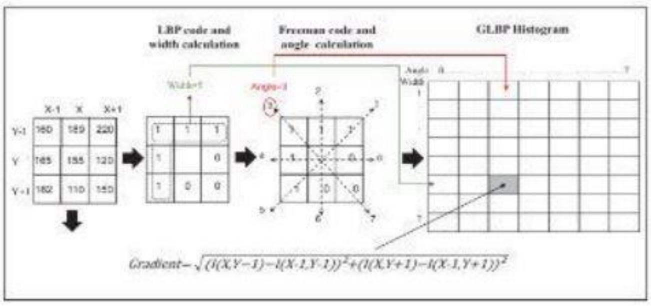

Fig 3 : GLBP

Local Binary Patterns (LBP) is a widely used texture descriptor in computer vision and image analysis is shown in Fig 3. It characterizes the local structure of an image by computing binary patterns on a neighborhood around each pixel. Specifically, it compares the gray level of the center pixel to the gray levels of its neighboring pixels and encodes the result as a binary pattern. LBP patterns are rotation invariant, which means that they do not change when the image is rotated. LBP histograms are commonly used as features for various applications, such as face recognition, texture classification, and human detection. Gradient LBP (GLBP) is an extension of LBP that uses gradients to compute LBP patterns, making it particularly suitable for detecting objects with well-defined edges. GLBP constructs a histogram of directed gradients by using uniform LBP patterns, which correspond to LBP codes that contain only a few transitions from 0 to 1 or from 1 to 0.

Algorithm :

1) Determine the LBP code.

2) Assuming the LBP code corresponds to a uniform pattern, calculate the width and angle values as follows:

• The width values line up with the LBP code's "1" digit.

• The angle value matches the Freeman direction of the centre pixel in the LBP code's "1" region.

3) In the LBP code, calculate the gradient for the 1 to 0 (or 0 to 1) transitions.

4) The position within the GLBP matrix, which is filled by adding gradient values, is determined by the width and angle values.

2.3 Support Vector Machine :

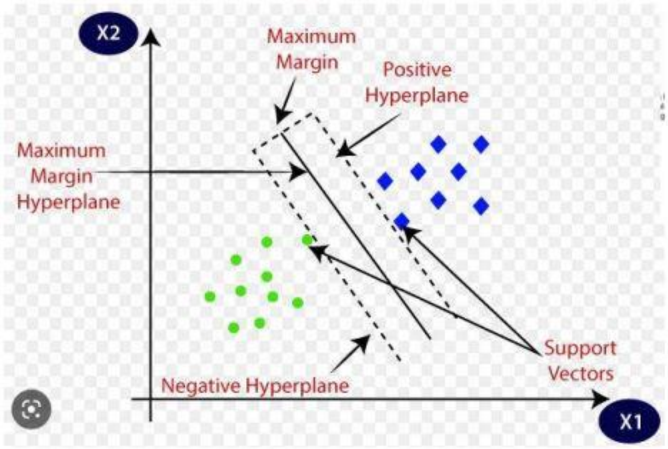

Fig 4 : SVM Model

SVM stands for Support Vector Machine, which is a popular machine learning algorithm used for classification and regression analysis is shown in Fig 4. It is a supervised learning algorithm that is primarily used for classification tasks. The goal of SVM is to find the best possible decision boundary that separates data into different classes.

SVM works by mapping input data into a high-dimensional feature space, where a decision boundary can be more easily identified. It then uses a kernel function to compute the similarity between pairs of data points in this feature space, and finds the hyperplane that maximizes the margin, which is the distance between the hyperplane and the nearest data points from each class. The data points closest to the hyperplane are called support vectors, and they play a crucial role in determining the location of the decision boundary.

SVM has several advantages, such as its ability to handle high-dimensional data and its effectiveness in dealing with small to medium-sized datasets. It is also a versatile algorithm that can be used for both linear and non-linear classification tasks, thanks to its kernel trick. SVM has been successfully applied in various domains, such as image classification, text classification, and bioinformatics. However, SVM can be computationally expensive and sensitive to the choice of kernel function and hyperparameters, which can affect its performance.

Algorithm :

1) Gather training data: This includes labeled data with input features and corresponding output labels.

2) Preprocess the data: Scale and normalize the data to make it more efficient for training and to avoid biases.

3) Choose an appropriate kernel: Select a kernel function, such as linear, polynomial, or radial basis function (RBF), that maps input data into a higher-dimensional feature space. 4) Define the SVM model: Use the kernel and a regularization parameter to define the SVM model.

5) Train the SVM model: Optimize the model parameters to maximize the margin between the decision boundary and the support vectors.

6) Test the SVM model: Evaluate the performance of the model on a separate test set of labeled data.

7) Tune hyperparameters: Adjust the regularization parameter and kernel parameters through cross-validation to achieve optimal performance.

Use the SVM model: Use the trained SVM model to classify new, unlabeled data based on the decision boundary learned during training.

2.4 Attention-Based Convolutional Neural Network

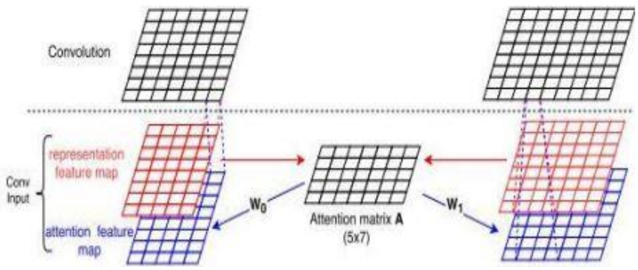

Fig 5 : ABCNN Network

The ABCNN (Attention-Based Convolutional Neural Network) is a neural network architecture designed for modeling pairs of sentences is shown in Fig 5. It is a general architecture that can be applied to a wide range of tasks, such as natural language inference, paraphrase identification, and sentiment analysis.

The ABCNN consists of several layers, including a word embedding layer, a convolutional layer, and an attention layer. The word embedding layer converts each word in the input sentences into a fixed-length vector representation. The convolutional layer applies a set of filters to the sentence embeddings to generate a higher-level representation feature map for each sentence. The attention layer then calculates an attention matrix that captures the relationship between the two input sentences and uses this matrix to guide the convolutional layer to generate "counterpart-biased" sentence representations.

The attention matrix is generated by matching units of the left representation feature map with units of the right representation feature map. The attention values in the matrix indicate the importance of each unit in one sentence with respect to each unit in the other sentence. The attention feature map is then generated by transforming the attention matrix into a new feature map for each sentence. This new feature map is combined with the representation feature map of each sentence and fed into the convolutional layer.

The ABCNN can be extended in several ways to handle different tasks. For example, the ABCNN-1 variant employs an attention feature matrix to influence convolution. The ABCNN-2 variant uses a pooling layer after the convolutional layer to capture global information about the sentence pair. The ABCNN-3 variant uses a recursive neural network to capture hierarchical structures within each sentence. Overall, the ABCNN is a powerful and flexible architecture for modeling pairs of sentences.

The attention feature matrix A is generated by matching units of the left and right representation feature maps, with the attention values of row i in A denoting the attention distribution of the i-th unit of s0 with respect to s1, and the attention values of column j in A denoting the attention distribution of the j-th unit of s1 with respect to s0. A is then transformed into two blue matrices that have the same format as the representation feature maps and combined with them as input to the convolution operation. The attention feature map guides the convolution to learn "counterpart-biased" sentence representations. The higher-level representation feature map generated by convolution is a tensor stacked with the representation feature map and the attention feature map.

Algorithm :

1) Input pre-processing: Convert raw text into numerical features using techniques such as word embedding or bag of words.

2) Convolutional layer: Apply 1D convolution operation to the input features to capture local patterns and correlations among neighboring words.

3) Attention layer: Calculate the attention scores for each pair of sentence embeddings using cosine similarity measure. The attention mechanism enables the model to focus on the most informative parts of the input.

4) Match layer: Compute the similarity score between each pair of sentence embeddings using various matching functions such as cosine similarity, Euclidean distance, or Hadamard product.

5) Pooling layer: Aggregate the matching scores across all pairs of sentence embeddings to obtain a fixed-length representation of the input.

6) Fully connected layer: Apply one or more fully connected layers to the pooled features to perform the final classification.

7) Output: Predict the class label based on the softmax probability distribution over the output classes.

8) Training: Use backpropagation algorithm to adjust the weights of the model to minimize the loss function.

9)Hyperparameter tuning: Fine-tune the hyperparameters of the model such as learning rate, batch size, and regularization strength to achieve optimal performance.

10)Evaluation: Test the model on a held-out dataset and report the accuracy, precision, recall, F1 score, and other performance metrics.

3. EXPERIMENTAL RESULTS

The process involves predicting age and gender groups utilizing the HOG (Histogram of Oriented Gradients) feature extraction technique combined with an SVM (Support Vector Machine) model. The initial step of HOG feature extraction has been successfully completed. Following this, the model is trained and its training accuracy is assessed using the SVM approach.

Age prediction focuses on two distinct groups: individuals under the age of 20 to 35 and those falling within the range of 36 to 50. In the context of age representation, the category of 20 to 35 years is denoted as 0, while the age group of 36 to 50 is represented by the label 1 is shown in Fig 7. The gender classification process classifies individuals into either "male" or "female" categories. Specifically, the gender category "male" is assigned the label 0, while the gender category "female" is denoted by the label 1 is shown in Fig 9.

```
Accuracy: 0.63
```

Fig 6 : Training Accuracy

```
Predicted Age Group : [1]
```

Fig 7 : Predicted Age Group

```
Accuracy: 0.65
```

Fig 8 : Training Accuracy

```
Predicted Gender Group : [1]
```

Fig 9 : Predicted Gender Group

The process involves predicting age and gender groups utilizing the LBP (Local Binary Patterns) feature extraction technique combined with an SVM (Support Vector Machine) model. The initial step of LBP feature extraction has been successfully completed. Following this, the model is trained and its training accuracy is assessed using the SVM approach.

Age prediction focuses on two distinct groups: individuals under the age of 20 to 35 and those falling within the range of 36 to 50. In the context of age representation, the category of 20 to 35 years is denoted as 0, while the age group of 36 to 50 is represented by the label 1 is shown in Fig 11. The gender classification process classifies individuals into either "male" or "female" categories. Specifically, the gender category "male" is assigned the label 0, while the gender category "female" is denoted by the label 1 is shown in Fig 13.

```
Accuracy: 0.66
```

Fig 10 : Training Accuracy

```
Predicted Age Group: [1]
```

Fig 11 : Predicted Age Group

```
Accuracy: 0.5588536335721597
```

Fig 12 : Training Accuracy

```
Predicted Gender : 1
```

Fig 13 : Predicted Gender Group

The process involves predicting age and gender groups utilizing the HOG (Histogram of Oriented Gradients) feature extraction technique combined with an ABCNN (Attention-Based Convolutional Neural Network) model. The initial step of HOG feature extraction has been successfully completed. Following this, the model is trained and its training accuracy is assessed using the ABCNN approach.

Age prediction focuses on two distinct groups: individuals under the age of 20 to 35 and those falling within the range of 36 to 50. In the context of age representation, the category of 20 to 35 years is denoted as 0, while the age group of 36 to 50 is represented by the label 1 is shown in Fig 15. The gender classification process classifies individuals into either "male" or "female" categories. Specifically, the gender category "male" is assigned the label 0, while the gender category "female" is denoted by the label 1 is shown in Fig17.

```
Accuracy: 92.19%
```

Fig 14 : Training Accuracy

```
Predicted age : 1
```

Fig 15 : Predicted Age Group

```
Accuracy: [0.13841998657711792, 0.9375]
```

Fig 16 : Training Accuracy

Predicted gender : 1

Fig 17 : Predicted Gender Group

The process involves predicting age and gender groups utilizing the LBP (Local Binary Patterns) feature extraction technique combined with an ABCNN (Attention-Based Convolutional Neural Network) model. The initial step of LBP feature extraction has been successfully completed. Following this, the model is trained and its training accuracy is assessed using the ABCNN approach.

Age prediction focuses on two distinct groups: individuals under the age of 20 to 35 and those falling within the range of 36 to 50. In the context of age representation, the category of 20 to 35 years is denoted as 0, while the age group of 36 to 50 is represented by the label 1 is shown in Fig 19. The gender classification process classifies individuals into either "male" or "female" categories. Specifically, the gender category "male" is assigned the label 0, while the gender category "female" is denoted by the label 1 is shown in Fig 21.

Test accuracy: 0.9624999761581421

Fig 18 : Training Accuracy

Predicted Age: ['1']

Fig 19 : Predicted Age Group

Test accuracy: 1.0

Fig 20 : Training Accuracy

Predicted Genders: ['1']

Fig 21 : Predicted Gender Group

- Age Prediction Accuracy (%) :

	HOG	LBP
SVM	63	66
ABCNN	92.14	96.24

Table 1 : Age Prediction Accuracy

- Gender Prediction Accuracy (%) :

	HOG	LBP
SVM	65	55.88
ABCNN	93.75	99

Table 2 : Gender Prediction Accuracy

4. CONCLUSION

The study looks into the use of gradient features to extract soft biometric characteristics like age and gender from handwritten text. The experiments were conducted on four English datasets, and Histogram of Gradient Features and Gradient Local Binary Pattern were used for feature extraction. The ABCNN model was used for prediction, which demonstrated high performance for text-based tasks. Independent gender and age tests were conducted on each dataset, confirming the effectiveness of gradient features for these predictions.

As for the future scope of ABCNN in handwriting analysis, the model can be further explored for various applications, such as signature verification, handwriting recognition, and writer identification. These tasks require the extraction of specific features from the text, and ABCNN's attention-based convolutional neural network architecture can be tailored to achieve high performance on these tasks. Additionally, the model's ability to consider the interdependence between sentences makes it a promising approach for more complex tasks, such as document analysis and understanding. Further research can also investigate the use of additional features and data preprocessing techniques to improve the performance of ABCNN in handwriting analysis.

REFERENCES

[1] Nesrine Bouadjenek, Hassiba Nemmour and Youcef Chibani Age, Gender and Handedness Prediction from Handwriting using Gradient Features. 2015 13th International Conference on Document Analysis and Recognition (ICDAR).

[2] V. Basavaraja, Palaiahnakote Shivakumara, D. S. Guru, Umapada Pal, Tong Lu Age Estimation using Disconnectedness Features in Handwriting. 2019 International Conference on Document Analysis and Recognition (ICDAR).

[3] Najla AL-Qawasmeh, Ching Y. Suen Transfer Learning to Detect Age From Handwriting. Advances in Artificial Intelligence and Machine Learning; Research 2 2022.

[4] Francesco Asci, Simone Scardapane Alessandro Zampogna, Valentina D'Onofrio, Lucia Testa, Martina Patera, Marco Falletti, Luca Marsili and Antonio Suppa Handwriting Declines With Human Aging: A Machine Learning Study. 2022 The journal Frontiers in Aging Neuroscience.

[5] Medhavi Bhardwaj, Ayushi Pandey. Handwriting Recognition Using CNN . Advances and Applications in Mathematical Sciences Volume 21, Issue 5, March 2022, Pages 2817-2828 © 2022 Mili Publications, India.

[6] C. C. Ng, M. H. Yap, Y. T. Cheng and G. S. Hsu, "Hybrid ageing patterns for face age estimation", Image and Vision Computing, 69, 2018, pp 92-102.

[7] Guo, H., Wan, J., Wang, H., Wu, H., Xu, C., Miao, L., et al. (2021). Selfpowered intelligent human-machine interaction for handwriting recognition. Research(WashDC)2021:4689869.doi:10.34133/2021/4689869.

[8] Lester, D., McLaughlin, S., Cohen, R., & Dunn, L., Sex-deviant handwriting, femininity, and homosexuality. Perceptual and Motor Skills. 2019, vol.45, p.1156.

[9] Kanno, S., Shinohara, M., Kanno, K., Gomi, Y., Uchiyama, M., Nishio, Y., et al. (2020). Neural substrates underlying progressive micrographia in Parkinson's disease. Brain Behav. 10:e01669. doi: 10.1002/brb3.1.669.

[10] J.-D. Lee, c.-Y. Lin, and C.-H. Huang, "Novel features selection for gender classification," in Mechatronics and Automation (ICMA), 2013 IEEE International Conference on, pp. 785-790, Aug 2013.

[11] M. Liwicki, A. Schlapbach, P. Loretan, and H. Bunke, "Automatic detection of gender and handedness from on-line handwriting," in Conference of' the International Graphonomics Society, pp. 179-183,2007.

[12] N. Bouadjenek, H. Nemmour, and Y. Chibani, "Local descriptors to improve off-line handwriting-based gender prediction," in Soft Computingand Pal/ern Recognition (SoCPaR), 2014 6th International Conference of, pp. 43- 47, Aug 2014.

[13] S. A. Mahmoud, I. Ahmad, W. G. AI-Khatib, M. Alshayeb, M. T. Parvez,V. Margner, and G. A. Fink, "KHATT: an open arabic offline handwritten text database," Pattern Recognition, vol. 47, no. 3, pp. 1096-11 12, 2014.

[14] Ramdan J, Omar K, Faidzul M, Mady A. Arabic Handwriting Data Base for Text Recognition. Procedia Technology, 11:580–584, 2013. 4th International Conference on Electrical Engineering and Informatics, ICEEI 2013.

[15] Djeddi C, Gattal A, Souici-Meslati L, Siddiqi I, Chibani Y, et al. LamisMshd: A Multi-Script Offline Handwriting Database. In The 14th International Conference on Frontiers in Handwriting Recognition, ICFHR, Crete, Greece, September 1-4, 2014, IEEE. 2014: 93-97.

[16] Garoot AH, Safar M, Suen CY. A Comprehensive Survey on Handwriting and Computerized Graphology. In 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), pages 621–626, Los Alamitos, CA, USA, 2017.

[17] J. Zunic, P. L. Rosin and V. Ilic, “Disconnectedness: A new moment invariant for multi-component shapes”, Pattern Recognition, 78, 2018,pp 91- 102.

[18] K. S. Raghunandan, P. Shivakumara, B. J. Navya, G. Pooja, Navya, N.Prakash, G. H. Kumar. U. Pal and T. Lu, “Fourier coefficients for fraud handwritten document classification through age analysis”, In Proc. ICFHR, pp 25-30, 2016.

[19] Delazer, M., Zamarian, L., and Djamshidian, A. (2021). Handwriting in Alzheimer’sdisease. JAD 82, 727–735. doi: 10.3233/JAD-210279.

[20] Caligiuri, M. P., Kim, C., and Landy, K. M. (2014). Kinematics of signature writing in healthy aging. J. Forensic Sci. 59, 1020–1024. doi: 10.1111/1556-4029.12437.