Alzheimer's Disease Screening with Limited and Noisy On-line Pentagon Drawings

Nina Hosseini-Kivanani
University of Luxembourg
Luxembourg
nina.hosseinikivanani@uni.lu

Elena Salobrar-García
Universidad Complutense de Madrid
Spain
elenasalobrar@med.ucm.es

Moises Diaz
Universidad de Las Palmas de Gran Canaria
Spain
moises.diaz@ulpgc.es

Miguel A. Ferrer
Universidad de Las Palmas de Gran Canaria
Spain
miguelangel.ferrer@ulpgc.es

Luis A. Leiva
University of Luxembourg
Luxembourg
luis.leiva@uni.lu

***Abstract*—** Alzheimer's Disease (AD) is a progressive neurodegenerative disorder that leads to cognitive and motor decline. Handwriting analysis has gained attention as a non-invasive approach to detecting early neurodegenerative changes, leveraging cognitive and motor impairments associated with AD. This study explores on-line handwriting for AD screening using a standard neuropsychological assessment task (Pentagon drawings) where the collected data are both scarce and extremely noisy. We propose a novel data augmentation framework grounded in the Kinematic Theory to reconstruct noisy velocity profiles from scratch to generate realistic handwriting variations. We consider several classification tasks with Healthy Controls (HC), Mild Cognitive Impairment (MCI), and AD patients. Our results show that our proposed method achieves the best performance, offering thus a scalable, cost-effective, and non-invasive tool for early neurodegenerative disease screening, with practical applications in digital medicine and healthcare.

***Keywords*—** *Alzheimer's disease, handwriting analysis, data augmentation, velocity profile reconstruction*

I. INTRODUCTION

Neurodegenerative diseases progressively impair neural functioning, affecting cognition and motor control [1], [2]. Among them, Alzheimer's disease (AD) is the most prevalent and severe form of dementia, impacting millions worldwide [3]. While pharmacological treatments provide symptom relief, they do not stop disease progression. Therefore, early detection is critical for managing symptoms and slowing functional decline [4].

Existing diagnostic methods, including e.g. neuroimaging (MRI, PET) and cerebrospinal fluid (CSF) analysis, provide valuable clinical insights but remain costly, invasive, and inaccessible for routine screening [5], [6]. On the other hand, cognitive assessments offer a non-invasive alternative but are subject to variability due to anxiety, reduced patient cooperation, and subjective evaluation [7]. These limitations highlight the need for more objective and more accessible diagnostic tools.

Handwriting analysis has gained attention as a method for detecting early motor and cognitive impairments in AD [8]–[10], making it an interesting source of digital biomarkers [11], [12]. In this paper, we build upon this line of research and investigate Pentagon drawings, a standard neuropsychological assessment task. A particular challenge of our data is that they are both scarce and extremely noisy. We thus propose a novel data augmentation framework grounded in the Kinematic Theory to reconstruct the velocity profiles from scratch and generate realistic handwriting variations that we use for classification.

II. RELATED WORK

Handwriting analysis has gained attention as a method for detecting early motor and cognitive impairments in AD [8]–[10], making it an interesting source of digital biomarkers [11], [12]. Several review and experimental studies have confirmed the relevance of this modality in AD screening, including a comprehensive review by De Stefano et al. [13], and more recent experimental work on spiral dynamics and transfer learning by Carfora et al. [14].

Some studies highlight that parameters such as writing speed, pen pressure, and movement fluency enable accurate classification of AD and mild cognitive impairment (MCI) patients [15], [16]. Computational models have used various features to differentiate AD patients from healthy individuals [17], [18]. Further, researchers have proposed Generative Adversarial Networks [19] and Diffusion models [20] for AD classification, however they require large datasets to be trained on, and most real-world datasets of AD patients are very limited in size. Recent work has investigated the Kinematic Theory with classic Machine Learning models such as Support Vector Machines [21]. To the best of our knowledge, we are the first to explore Deep Learning models in this context, since researchers have worked mostly with off-line data only [22]–[25].

The Kinematic Theory [26] and its latest instantiation, the Sigma-Lognormal model [27], describes handwriting movements by modeling velocity profiles using lognormal functions, and is among the most accurate approaches so far. Following this, we introduce velocity profile reconstruction and controlled distortion levels to create realistic samples, suitable AD screening. Our results show that our proposed method achieves the best performance, offering thus a scalable, cost-effective, and non-invasive tool for early neurodegenerative disease screening, with practical applications in digital medicine and healthcare.

III. MATERIALS AND METHODOLOGY

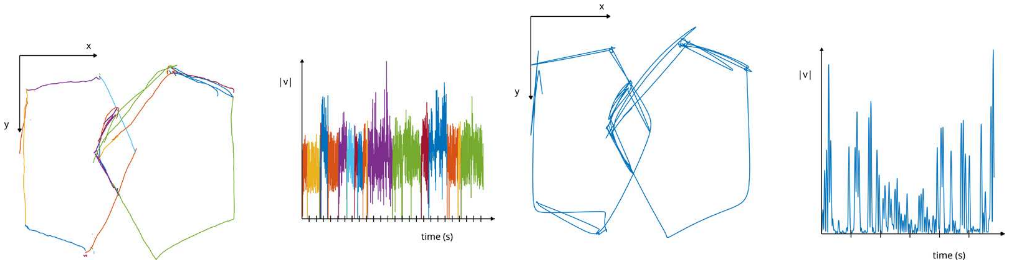
Previous work has used high-resolution digitizing tablets to capture data from AD patients, particularly through tasks like spiral drawing and handwriting analysis, which can reveal bradykinesia, hypometria, and sequencing impairments characteristic of the disease [16], [28]. However, these methods are considered unnatural to many users, who prefer the classic pen-and-pencil method [18]. To enhance ecological validity and user comfort, we used a 10.9-inch Repaper tablet[1], which enables the use of a real ballpoint pen on a blank sheet of paper while digitizing strokes in real time.

A standard pencil was instrumented with a magnetic ring that was sensed by the tablet, which sent the data via Bluetooth for later storage. This setup preserved a natural drawing experience while simultaneously recording on-line trajectories.

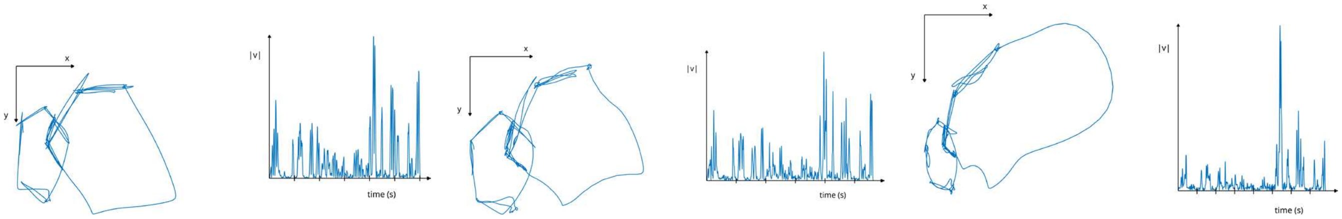
A. Data Collection

We collected 84 Pentagon drawings, commonly used in neuropsychological assessments [29], from 33 AD patients, 38 MCI patients, and 13 Healthy Controls (HCs). This corresponds to one drawing collected per participant.

The age of participants ranged between 76 and 83 years. The Repaper tablet provided the `x,y` coordinates without sampling rate information, so it was assumed to be 100 Hz. However, the velocity profiles were extremely noisy so they had to be entirely synthesized, as described next.



(a) Original trajectory (b) Reconstructed trajectory



(c) Synthetic trajectory, small distortion (d) Synthetic trajectory, medium distortion (e) Synthetic trajectory, large distortion

Fig. 1: Trajectories and velocity profiles of a Pentagon, originally drawn by an MCI patient.

B. Velocity Profile Reconstruction and Synthesis

Although the original trajectories contain multiple pen-down sequences (or *components*, see Fig. 1a), we used the entire trajectory without segmentation for reconstruction, as a single sequence, and ignoring the aforementioned sampling rate of 100 Hz. We applied iDeLog [30] for reconstruction, which is optimized for long and complex handwriting movements. First, iDeLog identifies virtual target points (high-curvature points) and designates them as stroke boundaries. Given that the timing of these points is known, iDeLog inserts a lognormal distribution between the start and end times of each stroke and resamples the trajectory. The lognormal distribution is centered at the midpoint of each stroke, ensuring that 98% of its area falls within the duration of the strokes, allowing for slight overlap between consecutive strokes. As a result, the reconstruction simulates a Pentagon drawing without lifting the pen from the paper, as shown in Fig. 1b.

Then, we synthesized new samples by applying a series of affine transformations [31] to the reconstructed data, in the following order:

1) Gaussian noise distortion, centered on each point with variance in $(1, 5, 10) \times 10^{-3}$, according to three chosen distortion levels; i.e., small, medium, and large.
2) Sinusoidal function [32] to modify stroke lengths non-uniformly. The sinusoid periods are selected from the intervals $[(0.6, 1.2), (0.5, 1.3), (0.6, 1.4)]$, while the corresponding amplitude ranges are $[(11, 14), (8, 17), (5, 20)]$, each representing one of the three distortion levels.

---

[1] https://www.iskn.co/eu

3) Perspective distortion, with a shearing effect with values from $[(4,5),(2,7),(0,9)]\times 10^{-2}$ and depth-related size changes adjusted within the range $[(5,7),(2,9.9),(-1,13)]\times 10^{-4}$.
4) Rotational distortion, applied randomly in either a positive or negative direction. Positive rotation angles are sampled from $[(5,20),(10,30),(15,40)]$ deg, while negative rotations are drawn from $[(-20,-5),(-30,-10),(-40,-15)]$ deg.
5) Finally, the trajectory is either enlarged or shrunk based on randomly chosen scaling factors. The scaling intervals are $[(7,9),(5,1),(3,1.1)]\times 10^{-1}$ or $[(1.1,1.3),(0.9,1.5),(0.7,1.7)]$, depending on the transformation type.

A normally distributed random value was sampled for each parameter within the specified intervals. Then, the resulting trajectory was then parameterized using iDeLog. Subsequently, the virtual target point positions were modified. The minimum distance m between adjacent virtual target points (vtp) is calculated. A circular region with radius $m\cdot\delta_r$ is generated, where $\delta_r \in (0.1, 0.3, 0.5)$ varies according to a distortion level (small, medium, large). The new virtual target point positions (vtp') are then computed as follows:

$$vtp'_x = vtp_x + m\cdot\delta_r\cdot r\cdot\cos(2\pi r) \qquad vtp'_y = vtp_y + m\cdot\delta_r\cdot r\cdot\sin(2\pi r) \tag{1}$$

where r is a randomly generated factor within $(0,1)$, introducing thus a controlled distortion.

The percentage of stroke duration was then randomly modified within the intervals $\delta_t \in \pm(10\%, 30\%, 50\%)$, affecting the total duration. The same percentage modification is applied to the σ_i parameter of each lognormal component in the Pentagon decomposition. Due to the relationship between μ and σ in neuromuscular systems, the μ_i values were updated as:

$$\mu'_i = \ln(\exp(\mu_i - \sigma_i^2)) + \sigma'_i \tag{2}$$

where $\sigma'_i = \sigma_i\cdot\delta_t$. Using these updated parameters, the final trajectory was synthesized. Figures 1c to 1e provide visual examples of synthesized data with small, medium, and large distortions for a randomly selected Pentagon drawing.

IV. EXPERIMENTS

We used the same classifier in all experiments, informed by previous work [33]: a GRU model with a single layer of 200 units (embedding size) using tanh activation and 0.15 dropout. The model takes as input the offsets between consecutive points $(\Delta x, \Delta y)$ to make the trajectories scale-invariant.

The model was trained on half of the data (stratified sampling) for 100 epochs with early stopping of 10 epochs, using Adam optimizer (learning rate $\eta = 0.001$ and decay rates $\beta_1 = 0.9$, $\beta_2 = 0.999$). Only the training data was augmented, leaving the test data as an unseen partition for final model evaluation.

We compared our method against no data augmentation and a jitter-based method using Gaussian noise centered at each coordinate $\mathcal{N}(\mu = 0, \sigma = 5)$, as described in previous work [18], [34]. The value of $\sigma = 5$ was set according to the mean value of the offsets between consecutive points observed in our dataset. We generated 50 variations per human sample with each method. The results of our experiments are summarized in Table I.

TABLE I: Performance metrics on the test data. Best results highlighted in bold.

Method	**Adj. Precision (%)**	**Adj. Recall (%)**	**Adj. F-measure (%)**	**Accuracy (%)**	**AUC ROC (%)**
Classification Task: Healthy Controls vs. AD patients					
No augmentation	48.39	69.57	57.08	69.57	50.00
Jitter baseline	55.21	60.87	57.21	60.87	47.77
Ours, small distortions	88.32	88.46	88.23	88.46	84.72
Ours, medium distortions	**96.58**	**96.15**	**96.21**	**96.15**	**97.22**
Ours, large distortions	74.11	73.08	73.48	73.08	70.14
Classification Task: Healthy Controls vs. MCI patients					
No augmentation	82.46	76.92	69.84	76.92	57.14
Jitter baseline	82.46	76.92	69.84	76.92	57.14
Ours, small distortions	86.44	84.62	85.02	84.62	85.42
Ours, medium distortions	**97.37**	**97.22**	**97.22**	**97.22**	**97.22**
Ours, large distortions	65.73	69.23	65.64	69.23	56.94
Classification Task: MCI vs. AD patients					
No augmentation	64.47	63.89	62.88	63.89	63.00
Jitter baseline	63.82	63.89	63.80	63.89	63.62
Ours, small distortions	83.75	83.33	83.28	83.33	83.33
Ours, medium distortions	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**
Ours, large distortions	59.03	58.33	57.51	58.33	58.33

V. DISCUSSION AND FUTURE WORK

Our experiments show that it is possible to identify AD and MCI patients from HCs with outstanding performance using scarce and noisy on-line data, outperforming previous work on similar tasks [18], [34], [35] as well as previous work that used off-line data [22]–[25]. We have found that medium-level distortions work best.

Although we did not perform an exhaustive grid search over the augmentation parameter space, we systematically evaluated three representative distortion levels—small, medium, and large—which were sufficient to identify the optimal performance range across all classification tasks. These levels were selected to ensure a balance between meaningful variability and physiological plausibility, and they yielded consistent trends throughout our experiments.

With respect to the sinusoidal function used to modulate stroke lengths, we acknowledge that its period values were fixed in this study. Nevertheless, previous work [36], [37] has shown that variations in the sinusoidal period exert minimal influence on the resulting handwriting synthesis and its downstream classification utility. Therefore, our selection of period intervals is consistent with these prior findings and justified within the scope of this study.

The crux of our method lies in working on the sigma-lognormal space. That is, working with reconstructions of human samples and synthetically generated samples from those reconstructions. Future work will explore other drawing symbols used in neuropsychological assessment tasks, such as Clocks and Houses, as well as multi-class classification tasks. Further, our GRU model is very lightweight, so we plan to deploy it in clinical settings as part of a mobile app.

VI. CONCLUSION

We have explored on-line handwriting for AD screening using a standard neuropsychological assessment task (Pentagon drawings) where the collected data are both scarce and extremely noisy. We have proposed a novel data augmentation framework grounded in the Kinematic Theory to reconstruct noisy velocity profiles from scratch to generate realistic handwriting variations. Through several classification tasks, our results show that our proposed method achieves the best performance.

While promising, our approach has certain limitations. Notably, the augmentation parameters were fixed and not optimized for generalizability. Thus, further sensitivity analysis and validation are needed to ensure robustness across different settings and symbols. In this light, our method can be viewed as a potentially scalable, cost-effective, and non-invasive aid for early neurodegenerative disease screening, with future applications in digital medicine and healthcare pending further validation.

ACKNOWLEDGMENTS

Research supported by PID2023-146620OB-I00 (funded by MICIU/AEI 10.13039/501100011033), the European Union's FEDER program, CajaCanaria and LaCaixa (grant 2023DIG05), and the European Innovation Council under grant 101071147.

REFERENCES

[1] C. P. Fernandes, G. Montalvo, M. Caligiuri, M. Pertsinakis, and J. Guimaraes, "Handwriting changes in Alzheimer's disease: A systematic review," *J. Alzheimer's Dis.*, vol. 96, no. 1, 2023.

[2] N. D. Cilia, T. D'Alessandro, C. De Stefano, F. Fontanella, and M. Molinara, "From online handwriting to synthetic images for Alzheimer's disease detection using a deep transfer learning approach," *IEEE J. Biomed. Health Inform.*, vol. 25, no. 12, 2021.

[3] H. Qi, R. Zhang, Z. Wei, C. Zhang, L. Wang, Q. Lang, K. Zhang, and X. Tian, "A study of auxiliary screening for Alzheimer's disease based on handwriting characteristics," *Front. Aging Neurosci.*, vol. 15, 2023.

[4] C. A. Anacur, A. G. Yilmaz, and M. Aykut, "Handwriting analysis for Alzheimer's diagnosis: Feature enrichment and machine learning classification," in *Proc. IDAP*, 2024.

[5] S. Shimizu, D. Hirose, H. Hatanaka, N. Takenoshita, Y. Kaneko, Y. Ogawa, H. Sakurai, and H. Hanyu, "Role of neuroimaging as a biomarker for neurodegenerative diseases," *Front. Neurol.*, vol. 9, 2018.

[6] S. M. McGinnis, "Neuroimaging in neurodegenerative dementias," *Seminars in neurology*, vol. 32, no. 4, 2012.

[7] H. Hampel, L. M. Shaw, P. Aisen, C. Chen, A. Lleó, T. Iwatsubo, A. Iwata, M. Yamada, T. Ikeuchi, J. Jia *et al.*, "State-of-the-art of lumbar puncture and its place in the journey of patients with Alzheimer's disease," *Alzheimer's & Dementia*, vol. 18, no. 1, 2022.

[8] S. Toffoli, F. Lunardini, C. G. de Isla, and S. Ferrante, "AI-based ecological monitoring of handwriting to early detect cognitive decline," in *Proc. BHI*, 2023.

[9] U. Mitra and S. U. Rehman, "ML-powered handwriting analysis for early detection of Alzheimer's disease," *IEEE Access*, 2024.

[10] Y. Luo and Z. Yang, "DynGAN: Solving mode collapse in GANs with dynamic clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2024.

[11] M. Vimaladevi, R. Thangamani, S. Priyadharshini, B. Varshini, and A. Tarun, "Prediction of Alzheimer's disease by analyzing handwriting dynamics using machine learning algorithms," in *Proc. ICESC*, 2024.

[12] N. D. Cilia, C. De Stefano, F. Fontanella, M. Molinara, and A. Scotto Di Freca, "Handwriting analysis to support Alzheimer's disease diagnosis: a preliminary study," in *Proc. CAIP*, 2019.

[13] C. De Stefano, F. Fontanella, D. Impedovo, G. Pirlo, and A. S. di Freca, "Handwriting analysis to support neurodegenerative diseases diagnosis: A review," *Pattern Recognit. Lett.*, vol. 121, 2019.

[14] D. Carfora, S. Kim, N. Houmani, S. Garcia-Salicetti, and A.-S. Rigaud, "On extracting digitized spiral dynamics' representations: A study on transfer learning for early alzheimer's detection," *Bioengineering*, vol. 9, no. 8, p. 375, 2022.

[15] P. Ghaderyan, A. Abbasi, and S. Saber, "A new algorithm for kinematic analysis of handwriting data; towards a reliable handwriting-based tool for early detection of Alzheimer's disease," *Expert Syst. Appl.*, vol. 114, 2018.

[16] C. Kahindo, M. El-Yacoubi, S. Garcia-Salicetti, V. Cristancho-Lacroix, H. Kerhervé, and A.-S. Rigaud, "Semi-global parameterization of online handwriting features for characterizing early-stage Alzheimer and mild cognitive impairment," *IRBM*, vol. 39, no. 6, 2018.

[17] J. A. Meltzer, "Towards early prediction of Alzheimer's disease through language samples," *EClinicalMedicine*, vol. 29, 2020.

[18] N. Hosseini-Kivanani, E. Salobrar-García, L. Elvira-Hurtado, M. Salas, C. Schommer, and L. A. Leiva, "Blueprint of tomorrow: Contrasting off-line and on-line drawing tasks for Alzheimer's disease screening," in *Proc. IDEAL*, 2024.

[19] B. Wang, J. Lan, and J. Gao, “Multi-class GAN for generating multi-class images in object recognition,” *J. Opt. Soc. Am. A*, vol. 39, no. 5, 2022.

[20] X. Yang, T. Ye, X. Yuan, W. Zhu, X. Mei, and F. Zhou, “A novel data augmentation method based on denoising diffusion probabilistic model for fault diagnosis under imbalanced data,” *IEEE Trans. Ind. Inform.*, 2024.

[21] T. D’Alessandro, C. Carmona-Duarte, C. De Stefano, M. Diaz, M. A. Ferrer, and F. Fontanella, “A machine learning approach to analyze the effects of Alzheimer’s disease on handwriting through lognormal features,” in *Proc. IGC*, 2023.

[22] S. Chen, D. Stromer, H. A. Alabdalrahim, S. Schwab, M. Weih, and A. Maier, “Automatic dementia screening and scoring by applying deep learning on clock-drawing tests,” *Sci. Rep.*, vol. 10, no. 1, 2020.

[23] N. Ruengchaijatuporn, I. Chatnuntawech, S. Teerapittayanon, S. Sriswasdi, S. Itthipuripat, S. Hemrungrojn, P. Bunyabukkana, A. Petchlorlian, S. Chunamchai, T. Chotibut, and C. Chunharas, “An explainable self-attention deep neural network for detecting mild cognitive impairment using multi-input digital drawing tasks,” *Alzheimers Res. Ther.*, vol. 78, no. 14, 2022.

[24] R. Raksasat, S. Teerapittayanon, S. Itthipuripat, K. Praditpornsilpa, A. Petchlorlian, T. Chotibut, C. Chunharas, and I. Chatnuntawech, “Attentive pairwise interaction network for ai-assisted clock drawing test assessment of early visuospatial deficits,” *Sci. Rep.*, vol. 13, no. 1, 2023.

[25] N. Hosseini-Kivanani, I. C. Oliveira, S. Kilinç, and L. A. Leiva, “Efficient automatic data augmentation of CDT images to support cognitive screening,” in *Proc. ICAART*, 2025.

[26] R. Plamondon, “A kinematic theory of rapid human movements: Part ii. movement time and control,” *Biol. Cybern.*, vol. 72, 1995.

[27] R. Plamondon and M. Djioua, “A multi-level representation paradigm for handwriting stroke generation,” *Hum. Mov. Sci.*, vol. 25, no. 4, 2006.

[28] M. Kobayashi, Y. Yamada, K. Shinkawa, M. Nemoto, K. Nemoto, and T. Arai, “Automated early detection of alzheimer’s disease by capturing impairments in multiple cognitive domains with multiple drawing tasks,” *Journal of Alzheimer’s Disease*, vol. 88, no. 3, pp. 1075–1089, 2022.

[29] E. Helmes, “Cognitive screening of older adults: the utility of pentagon drawing,” *International psychogeriatrics*, vol. 25, no. 3, 2013.

[30] M. A. Ferrer, M. Diaz, C. Carmona-Duarte, and R. Plamondon, “iDeLog: iterative dual spatial and kinematic extraction of sigma-lognormal parameters,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 1, 2018.

[31] M. Maslych, E. M. Taranta, M. Aldilati, and J. J. Laviola, “Effective 2d stroke-based gesture augmentation for RNNs,” in *Proc. CHI*, 2023.

[32] M. Diaz, A. Fischer, R. Plamondon, and M. A. Ferrer, “Towards an automatic on-line signature verifier using only one reference per signer,” in *Proc. ICDAR*, 2015.

[33] L. A. Leiva, M. Diaz, M. A. Ferrer, and R. Plamondon, “Human or machine? it is not what you write, but how you write it,” in *Proc. ICPR*, 2020.

[34] N. Hosseini-Kivanani, E. Salobrar-García, L. Elvira-Hurtado, M. Salas, C. Schommer, and L. A. Leiva, “Predicting Alzheimer’s disease and mild cognitive impairment with off-line and on-line house drawing tests,” in *Proc. eScience*, 2024.

[35] N. Hosseini-Kivanani, E. Salobrar-García, L. Elvira-Hurtado, I. López-Cuenca, R. de Hoz, J. M. Ramírez, P. Gil, M. Salas, C. Schommer, and L. A. Leiva, “Better together: Combining different handwriting input sources improves dementia screening,” in *Proc. eScience*, 2023.

[36] M. A. Ferrer, M. Diaz, C. Carmona-Duarte, and A. Morales, “A behavioral handwriting model for static and dynamic signature synthesis,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 6, pp. 1041–1053, 2016.

[37] M. Diaz, M. A. Ferrer, G. S. Eskander, and R. Sabourin, “Generation of duplicated off-line signature images for verification systems,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 5, pp. 951–964, 2016.