Journal of the American Society of Questioned Document Examiners, Inc.
Volume 25, Number 1, June 2022

Technical Note: Signature Verification Algorithms: What They are, How They are Used, and How to Evaluate Them

Kevin P. Kulbacki, D-ABFDE, M.S.[1], Nikolaos Kalantzis[2,3] MSc, FSSocDip, Batya Miller Fuchs MSc[4]

[1]KDX Forensic Consulting LLC, 3712 N. Broadway, Suite 332, Chicago, IL 60613
[2]Chartoularios Institute, Kolokotroni 92-94, 18535, Piraeus, Greece
[3]Staffordshire University, Stoke-on-Trent, UK
[4]Questioned Documents Lab. DIFS, Israel Police

Over the last decade there has been a significant shift towards conducting business in a paperless and digital manner. As part of this transition, the collection of Digitally Captured Signatures (DCS—a.k.a. biometric signatures, online signatures or dynamic signatures) has increasingly replaced traditional static signatures written on paper. Traditionally, digitization of signatures has been accomplished with flat (2 dimensional) scanning of a signature executed on paper, that leads to the creation of a bitmap[1] file (i.e. static signatures) [1]. While static signatures represent visual information in the X and Y coordinates only (and hence lack information regarding the dynamic aspects of signing), dynamic signatures capture the X and Y coordinates as well as Time and Force. In conjunction with the faster pace of the digital world, businesses have had to adapt their fraud prevention efforts to handle not only more transactions, but also more sophisticated attempts at fraud. The numerical nature of the DCS data allows the contemplation of statistical methods to characterize a group of "known" DCS and then compare a "questioned" DCS versus that group, effectively performing a statistical evaluation of authenticity; various companies have invested in such algorithms that evaluate signatures in order to potentially determine whether or not a given signature was actually prepared by the purported author. This technical note discusses what these algorithms are, how they work and how they can be evaluated as per their forensic capabilities.

1. Introduction: What is a Signature Verification Algorithm?

The use of algorithms, or other automated procedures, to analyze handwriting is not a new concept [2, 3, 4]. As applied to the analysis of signatures, the first attempts at automated analysis predate DCS technology as we know it [2]. The introduction of the DCS as a medium for signing allowed this technology to utilize new quantifiable and measurable aspects of the signing behavior, mainly due to the numerical nature of the captured data (such as time of execution or quantifiable force values), thereby allowing for a more in-depth investigation of the individualistic characteristics of signature execution (through pen movement) in a computational manner. Still, the purpose of development of Signature Verification Algorithms (SVAs) largely affects the criteria and the design of the process, and inevitably the results [2, 5, 6].

The most obvious uses—which are already in effect [7] of an SVA would be to increase security either live during the transaction (teller deployment) or retroactively. These scenarios though do impose a specific set of conditions that might limit the scope of application of the SVA but at the same time ameliorate the possible results of the process. For example, Herbst and Liu [2] created a procedure with a 2.9 percent rejection of valid signatures and a 2.1 perfect acceptance of forgeries, but only after rejecting the outliers of the known material and not taking into account disguise as a scenario. These limiting conditions might make sense for specific deployment, such

[1] A bitmap file will represent an image in the form of an array, where each pixel is represented by a cell and its value corresponds to a color (depending on the resolution of the file). This is used in contrast to a vector file.

as the WonderNet Solution used by the Israeli Air force [8], but broadening the spectrum of the parameters analyzed may allow such SVAs to also act as an assistant tool to the Forensic Document Examiner (FDE).

Use of similar tools can be found within fingerprint examination—when the AFIS or other similar software databases [9] are utilized, the fingerprint examiner is given a list of possible matches by similarity score order. The expert then performs the analysis on the possible matches and reaches a conclusion. Hence, the software does not replace the examiner; rather, the software assists them. A forensic implementation of an algorithmic solution is found within DNA Analysis, where a profile is searched against CODIS and if there is a candidate match, the examiner will then go through procedures to confirm the match [10]. Another example for a supporting comparative algorithm is the Next Step system for forensic footwear impressions developed in the DIFS which floats possible matches according to shoeprint element similarity and geographic proximity [11]. The use of algorithmic solutions for examination and analysis of trace evidence does not necessarily stop there as the use of such algorithms can exceed the obvious comparison for determination of authenticity tasks and be used in the evaluation or validation phase of the process [12, 13].

If the SVA is tasked with assisting the expert during the analysis phase, or the validation phase (even as a safeguard aiming at highlighting all the aspects or characteristics a FHE should assess), it may suggest a list of parameters in the presented questioned signature that do not agree with values established in the known comparison set for the specific writer.

This could help focus the examiner on specific features during their examination; however, such an implementation should only be done after a comprehensive study of the potential for introducing cognitive bias.

2. Overview of Signature Verification Algorithms: How they work and how to use them

2.a. What do SVAs do?

SVAs are designed to compare a questioned entry/signature to a control group and to present a result; however, specifically how they are implemented, the conditions they take into account, and how the results are utilized are subjects of discussion that are beyond the scope of this paper due to the significant differences that are present on an SVA-by-SVA and/or implementation-by-implementation basis.

SVAs can be developed for both static and dynamic signatures. Depending on the comparison metrics utilized by an SVA to analyze a signature, it is also possible within some SVAs to mix the two genres[2] of signatures in the same manner that an FDE can use both dynamic and static signatures during casework. [14].

Whilst the programming behind most SVAs are trade secrets, there are several common approaches that have been noted across the industry. For SVAs that examine static signatures, the SVAs can often be described as pictorial similarity calculators; they compare an image of a presented questioned signature to a series of known signatures, in order to determine whether or not there is agreement among any of the utilized geometric comparison metrics. For SVAs that examine dynamic signatures, SVAs calculate a conformity threshold—i.e. a level of conformity that characterizes the individual members of the control group (the known material) and keep that threshold as a pass/fail criterium to which the questioned signature is compared. As depicted in Figure 1, the general architecture of SVAs is the same between static and dynamic comparison; what changes is the data (depending on how the static image is digitized) and some of the actual metrics used.

Metrics that can be common between the two genres are geometric oriented parameters such as size and angles, area covered, placement etc., either of the entire configuration of writing or of specific segments of it [15]. Dynamic parameters that cannot be derived by static images can involve time derivatives of the captured parameters; the mainstream captured parameters (data channels) for DCS are the spatial coordinates X and Y (i.e. position on the vertical and horizontal axis), Force (usually expressed in pressure levels [16] and time. For example, the time derivative of X and Y can give velocity, acceleration and jerk[3] for both axis but also for distance travelled (combining X and Y parameters via the Pythagorean theorem).

The selected metrics are calculated and compared, and according to the specifics of each SVA,

[2] "genre" as utilized herein refers to the classes of signatures being utilized by an SVA, i.e. static vs dynamic.
[3] The third time derivative of distance.

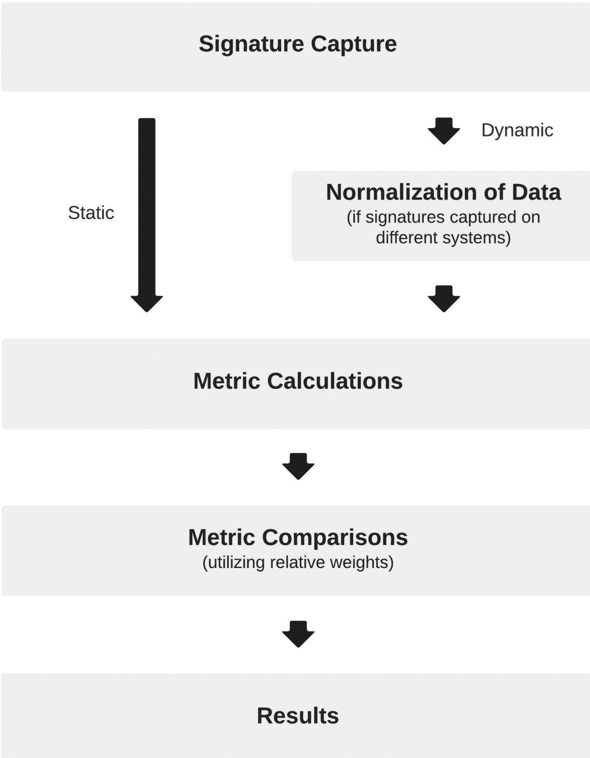

Figure 1.

each metric will be evaluated with a relative weight parameter (i.e. not all metrics are considered equally important). The metrics that are taken into account, as well as the relative weight given to each of them during analysis, is the core of an SVA and, as such, they are seldomly presented in the User Interface (UI) or even made public. Notwithstanding, the metrics and their relative weights need to be considered in the development of a validation study of an SVA [6, 17]. It has been observed that some analysis software will include metrics that are not trivial or obvious to the lay person or even the trained expert, e.g. area of bounding rectangle (D2_BOUNDING) or root mean square pen speed (V_RMS_SPEED) from SignatureScope [15]. It is to be expected that such metrics may be used in SVAs as well.

Furthermore, it has been observed [6, 17] that different alphabets, writing systems and/or scripts may require specific "tweaking" of those relative weights [18]. Therefore, class characteristics of the writing, such as writing system or alphabet, are a variable that should be taken into account in the development of a validation study depending on the proposed use case of an SVA.

2.b. How can SVAs be used?

SVAs have already been deployed as a means of identifying potential fraud in teller and retroactive deployment scenarios. Under these scenarios, the SVA produces a simple result stating whether a questioned signature has been executed by the author of the known signatures, based on the parameters of the deployment. In most circumstances, these results do not include detailed information of the metrics utilized to inform the result.

On the contrary, the theoretical forensic assistance use of SVAs would require exactly that—a reporting of the metrics used and their compatibility scores, either in the form of data sheets or as visual information (i.e. highlighting areas of interest on the signature XY visualization). This type of information can be used to assist the FHE by pointing out the features that they should focus their attention on (analysis phase) or to point out the features that they should have assessed during their analysis phase (validation phase). The potential of such use of SVAs is such that a possible integration of SVAs in the peer review process appears feasible as in other trace evidence fields [19]. Notwithstanding, the potential for introducing bias in a forensic application of an SVA must be studied in detail prior to implementation in any manner.

3. SVA Evaluation methodology

3.a. Methodology reasoning

With the various SVA uses in mind, an evaluation methodology is required to ascertain the validity of the results. Traditional logic for the collection of signatures revolves around obtaining writing samples that are representative of the range of variation of a particular writer, while minimizing the effects of any unwanted variables. Some variables that are generally considered include the collection timing as well as physical conditions [20].

When traditional logic for the collection of signatures has been utilized to generate samples for research, there are further considerations to ensure that the samples are representative of the various outcomes that may be possible. Standard outcomes may include genuine signatures, simulations, disguised signatures, and signatures that may be influenced by a miscellaneous factor [21]. These outcomes may be further elaborated to include simulations in which there is an attempt to copy the form, simulations in which there is not an attempt to copy the form, writers with poor signature repeatability/larger ranges of variation, alternate signature styles, medical issues, or unusual writing conditions.

Designing a signature sample set for a validation study of an SVA must take into account not only all of the above factors and outcomes, but many others as well.

For the evaluation of SVAs that compare static signatures, particular emphasis must be placed on the design of the collection sheets utilized for collecting the signatures. The collection sheets must be widely varied in design as the presence of a baseline, text, an overlapping signature, or any extraneous detail may lead to a misinterpretation of the algorithm [22].

For the evaluation of SVAs that compare dynamic signatures, the sample set should consider various properties of the deployed hardware. While the ideal scenario for the deployment of an SVA would include the deployment of one consistent hardware solution (i.e. digitizer and stylus) across the entire system, it is likely that there may be other hardware solutions deployed or even the intermingling of various hardware solutions. Furthermore, digitizers and stylus of different makes and models have different specifications, particularly with respect to their ability to record Force and the rate of capture of data points [16]. Simultaneously, hardware may fail or degrade over time due to wear and tear or aging of the pressure sensor [23]. As such, the sample set should be designed so that an SVA can be validated against numerous combinations of hardware configurations that may influence the data they evaluate. Finally, the sample set should also account for the physical conditions associated with the digitizer and stylus' physical properties, such as whether they are tethered, slanted, or in some other configuration.

3.b. Signature sample creation

Signature sample creation can be divided in two categories: the samples intended for use as the group of "known" signatures (i.e. the reference point) and the samples intended for use as "questioned" ones to test the SVA. Even though there are different strategies in how to produce and maintain the "known" signature group (to account for changes in range of variation), this is beyond the scope of the present study and is not explored.

Sample signature creation should aim at accommodating all possible scenarios. As such, samples (for both categories) should include signatures executed over a period of time and not at one go. Different physical execution conditions should be taken into account if they are a possibility for the given intended use; e.g. if an SVA is to be evaluated for deployment at the teller, then the physical conditions (such as the height of the teller's desk) should be replicated in the sample collection. An indicative flowchart of considerations is illustrated in Figure 2.

In addition to physical condition parameters, when examining DCS it may also be necessary to account for variables introduced by the hardware and software utilized as part of the deployed SVA. While an ideal scenario for an SVA is a completely consistent hardware/software solution, it is far more likely that some components of the solution may be replaced, such as individual digitizers or styli, each of which may have different technical specifications. It may be possible to overcome such differences (depending on how the SVA works when it comes to the unnormalized data), or to manufacture an SVA for specific deployment scenarios where only one type of software and hardware is deployed. Regardless, wear and tear (especially in EMR solutions [23]) can also affect the faithfulness of the dynamic data and hence possibly lead to false interpretations and rejections of samples by an SVA.

3.c. Signature sample considerations

Due to the variety of deployment scenarios and implementations for which SVAs are being developed, the testing procedures of a validation study for any SVA should mirror the procedures utilized by the particular SVA's intended deployment scenario. It has been observed that in some studies [24] "problematic" samples or conditions are being left out in order to produce good results—this practice should not be followed when conducting validation testing of an SVA due to the potential for encountering these samples under real world scenarios, especially if aiming at a forensic application of such a tool.

The scenarios that need to be investigated for SVA performance (arguably after the "normal" samples have been dealt with) should certainly include simulations (of various degrees of sophis-

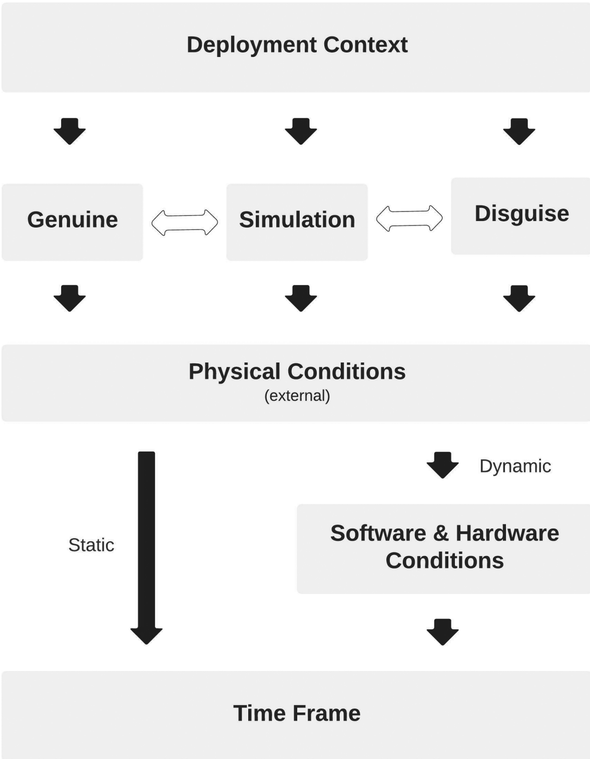

Figure 2.

tication), disguise, writers with poor signature repeatability/larger ranges of variation, alternate signature styles, medical issues, or unusual writing conditions, all of which may be encountered under real world scenarios.

4. Discussion

From instant screening of point-of-sale transactions and secondary evaluation of signatures for fraud to potential forensic applications, SVAs have the possibility to be powerful tools in the fight to identify and combat fraud; however, a rigorous evaluation of SVAs is required to understand their limits and must be structured to ensure that all the relevant parameters are considered. In SVA deployment scenarios where there is variety in DCS software/hardware solutions used, normalization issues might also need to be considered. If parameters for creation and maintenance of the known signature group are available, these should be explored as well. Furthermore, extensive research of the potential introduction of cognitive bias is necessary prior to implementation of an SVA in a forensic application. A rigorous testing scenario, materialized by the creation and handling of sample signature groups incorporating all the aforementioned parameters, will allow the evaluation of the strength of the SVA but also will allow for the exploration of different algorithm tweaking strategies, either per style or script used, or even on a writer-based approach.

5. References

[1] Plamondon, R. and Srihari, S.N., "Online and off-line handwriting recognition: a comprehensive survey," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 63-84, Jan. 2000, doi: 10.1109/34.824821.

[2] Herbst, N.M., and Liu, C.N., "Automatic signature verification based on accelerometry", *IBM J. Re,. Develop.*, 21 (1977)245-253. https://doi.org/10.1147/rd.213.0245

[3] Guyon, I., Albrecht, P., LeCun Y., Denker, J.S. and Hubbard, W., "A Time Delay Neural Network Character Recognizer for a Touch Terminal", *Pattern Recognition, (1990).*

[4] Lorette, G. and Plamondon, R., "Dynamic approaches to handwritten signature verifica- tion", in *Computer processing of handwriting,* Eds. R. Plamondon and C. G. Leedham, World Scientific, 1990. https://doi.org/10.1142/9789814439329_0002

[5] Gideon, S.J., Kandulna, A., Kujur, A.A., Diana, A., Raimond, K.. Handwritten Signature Forgery Detection using Convolutional Neural Networks, *Procedia Computer Science*, Volume 143, 2018, Pages 978-987, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2018.10.336.

[6] Malik, M.I., Liwicki, M., Dengel, A., and Found, B.. (2013). *Man vs. Machine: A Comparative Analysis for Forensic Signature Verification*. 9-13. https://doi.org/10.31974/jfde24-21-35

[7] Wacom. (2021, November 17). *Preventing signature fraud in banking*. YouTube. Retrieved July 4, 2022, from https://www.youtube.com/watch?v=uVLruwy0rAs&t=652s

[8] *WonderNet*. Electronic Signature Solutions. (n.d.). Retrieved May 12, 2022, from https://www.wondernet.co.il/customers/

[9] Komarinski, P. D., Considerations for Improving Latent Print Processing (2009). NIST Latent Fingerprint Testing Workshop. NIST.

[10]https://www.fbi.gov/services/laboratory/biometric-analysis/codis/codis-and-ndis-fact-sheet - CODIS-DNA%20Databases FBI. (2016, June 8). *CODIS and NDIS Fact Sheet*. FBI. Retrieved May 9, 2022, https://www.fbi.gov/services/laboratory/biometric-analysis/codis/codis-and-ndis-fact-sheet#CODIS-DNA%20Databases

[11] Daniel, O., Levi, A., Pertsev, R., Issan, Y., Pasternak, Z., & Cohen, A. (2022). The Next Step–a semi-automatic coding and comparison system for forensic footwear impressions. *Forensic Science International*, 111378. https://doi.org/10.1016/j.forsciint.2022.111378.

[12] Daniel Ramos, D., Haraksim, R., Meuwly, D. , Likelihood ratio data to report the validation of a forensic fingerprint evaluation method, Data in Brief, Volume 10, 2017, Pages 75-92, ISSN 2352-3409, https://doi.org/10.1016/j.dib.2016.11.008.

[13] Henry Swofford:Toward Computational Algorithms in Forensic Fingermark Examination: Navigating a Path Forward, EAFS 2022, Stockholm, 30-05, 3-06-2022

[14] European Network of Forensic Science Institutes. Appendix 3—Overview procedure for forensic handwriting examinations and comparisons. In: European Network of Forensic Science Institutes, editor. Best practice manual for the forensic examination of handwriting. Version 03. Wiesbaden: ENFHEX; 2020. p. 26–38.

[15] Strach, S., Novotny, M., Devlin, A., Graydon, C., and Westwood, P. (2016) Development of a Draft Methodology for Electronic Signature Examination using Wacom SignatureScope, *Presented at the 74th Annual General Meeting of the American Society of Questioned Document Examiners*, Pensacola Beach, Florida, August 20-25 2016

[16] Kalantzis, N., Platt, A.W.G.. Digitally captured signatures: A method for the normalization of force through calibration and the use of the zeta function. *J Forensic Sci*. 2022; 67: 651– 668. https://doi.org/10.1111/1556-4029.14927

[17] Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P., "Gradient-based learning applied to document recognition," in *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791.

[18] Impedovo, D., & Pirlo, G. (2008). Automatic signature verification: The state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *38*(5), 609–635. https://doi.org/10.1109/tsmcc.2008.923866

[19] Christophe Champod, Didier Meuwly, Forensic biometrics: Let's face it, EAFS 2022, Stockholm, 30-05, 3-06-2022

[20] Kelly, J. S., & Lindblom, B. S. (2006). *Scientific Examination of questioned documents* (Second Edition). CRC/Taylor & Francis.

[21] Kam, M., Abichandani, P. and Hewett, T. (2015), Simulation Detection in Handwritten Documents by Forensic Document Examiners. J Forensic Sci, 60: 936-941. https://doi.org/10.1111/1556-4029.12801

[22] S. Lam, L. Javanbakth and S. N. Srihari, "Anatomy of a Form Reader," *Proceedings Second International Conference on Document Analysis and Recognition*, Tsukuba, Japan, pp. 506-509. https://doi.org/10.1109/ICDAR.1993.395685

[23] Kalantzis N. Normalization and comparability of digitally captured signatures (DCS). In: Proceedings of the 73rd Annual Scientific Meeting of the American Academy of Forensic Sciences; 2021 Feb 15–19; held virtually. Colorado Springs, CO: American Academy of Forensic Sciences; 2021. p. 717.

[24] Bromley, J., Bentz, J., Bottou, L., Guyon, I., Lecun, Y., Moore, C., Säckinger, E., & Shah, R. (1993). Signature verification using a "Siamese" time delay neural network. *International Journal of Pattern Recognition and Artificial Intelligence*, *07*(04), 669–688. https://doi.org/10.1142/s0218001493000339