Forensic Science International: Synergy 4 (2022) 100230

Contents lists available at ScienceDirect

Forensic Science International: Synergy

journal homepage: www.sciencedirect.com/journal/forensic-science-international-synergy

A strawman with machine learning for a brain: A response to Biedermann (2022) the strange persistence of (source) "identification" claims in forensic literature

ARTICLE INFO

Keywords
Forensic inference
Machine learning

ABSTRACT

We agree wholeheartedly with Biedermann (2022) FSI Synergy article 100222 in its criticism of research publications that treat forensic inference in source attribution as an "identification" or "individualization" task. We disagree, however, with its criticism of the use of machine learning for forensic inference. The argument it makes is a strawman argument. There is a growing body of literature on the calculation of well-calibrated likelihood ratios using machine-learning methods and relevant data, and on the validation under casework conditions of such machine-learning-based systems.

Letter to Editor:

Biedermann [1] is critical of research publications that treat forensic inference in source attribution as an "identification" or "individualization" task. Biedermann [1] argues that such publications condone unscientific attitudes and practices, foster unrealistic expectations among consumers of forensic science, and undermine trust in peer-reviewed publications because so-called "original research papers" are not, in fact, well grounded. With respect to these points, we agree wholeheartedly with Biedermann [1].

With respect to criticism of machine learning, however, we feel that Biedermann [1] makes a strawman argument. It defines "standard" machine learning as outputting categorical decisions and then criticizes the use of "standard" machine learning for forensic inference because it outputs categorical decisions. There are indeed research publications that misapply machine learning to forensic-inference problems, including using algorithms that output categorical decisions, e.g. [2]. But we fear that many readers will get the impression from Biedermann [1] that this is the only way (or at least the primary way) that machine learning is applied to forensic inference. There is in fact a growing body of literature on the calculation of well-calibrated likelihood ratios using machine-learning methods and relevant data, and on the validation under casework conditions of such machine-learning-based systems. Recent examples include [3–11].

Disclaimer

All opinions expressed in the present paper are those of the authors, and, unless explicitly stated otherwise, should not be construed as representing the policies or positions of any organizations with which the authors are associated.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contributions

Morrison, Ramos, Ypma: Writing - Original Draft, Writing - Review & Editing. **All other authors:** Writing - Review & Editing.

Acknowledgements

The writing of this response was supported by Research England's Expanding Excellence in England Fund as part of funding for the Aston Institute for Forensic Linguistics 2019–2023.

References

[1] A. Biedermann, The strange persistence of (source) "identification" claims in forensic literature through descriptivism, diagnosticism and machinism, Forensic Sci. Int.: Synergy 4 (2022), https://doi.org/10.1016/j.fsisyn.2022.100222 article 100222.

[2] L. Quijano-Sánchez, F. Liberatore, J. Camacho-Collados, M. Camacho-Collados, Applying automatic text-based detection of deceptive language to police reports: extracting behavioral patterns from a multi-step classification model to understand how we lie to the police, Knowl. Base Syst. 149 (2018) 155–168, https://doi.org/10.1016/j.knosys.2018.03.010.

[3] W. Bosma, S. Dalm, E. van Eijk, R. El Harchaoui, E. Rijgersberg, H.T. Tops, A. Veenstra, R.J.F. Ypma, Establishing phone-pair co-usage by comparing mobility patterns, Sci. Justice 60 (2020) 180–190, https://doi.org/10.1016/j.scijus.2019.10.005.

[4] T. Matzen, C. Kukurin, J. van de Wetering, S. Ariëns, W. Bosma, A. Knijnenberg, A. Stamouli, R.J.F. Ypma, Objectifying evidence evaluation for gunshot residue comparisons using machine learning on criminal case data, Forensic Sci. Int. 335 (2022), https://doi.org/10.1016/j.forsciint.2022.111293 article 111293.

https://doi.org/10.1016/j.fsisyn.2022.100230
Received 5 April 2022; Accepted 25 April 2022
Available online 6 May 2022

[5] G.S. Morrison, E. Enzinger, V. Hughes, M. Jessen, D. Meuwly, C. Neumann, S. Planting, W.C. Thompson, D. van der Vloed, R.J.F. Ypma, C. Zhang, A. Anonymous, B. Anonymous, Consensus on validation of forensic voice comparison, Sci. Justice 61 (2021) 229–309, https://doi.org/10.1016/j.scijus.2021.02.002.
[6] G.S. Morrison, E. Enzinger, D. Ramos, J. González-Rodríguez, A. Lozano-Díez, Statistical models in forensic voice comparison, in: D.L. Banks, K. Kafadar, D. H. Kaye, M. Tackett (Eds.), Handbook of Forensic Statistics, CRC, Boca Raton, FL, 2020, pp. 451–497, https://doi.org/10.1201/9780367527709. Ch. 20.
[7] D. Ramos, D. Meuwly, R. Haraksim, C.E.H. Berger, Validation of forensic automatic likelihood ratio methods, in: D. Banks, K. Kafadar, D.H. Kaye, M. Tackett (Eds.), Handbook of Forensic Statistics, CRC, Boca Raton, FL, 2020, pp. 143–163, https://doi.org/10.1201/9780367527709. Ch. 7.
[8] D. Ramos, J. Maroñas, J. Almirall, Improving calibration of forensic glass comparisons by considering uncertainty in feature-based elemental data, Chemometr. Intell. Lab. Syst. 217 (2021), https://doi.org/10.1016/j.chemolab.2021.104399 article 104399.
[9] P. Vergeer, J.N. Hendrikse, M.M.P. Grutters, L.J.C. Peschier, A method for forensic gasoline comparison in fire debris samples: a numerical likelihood ratio system, Sci. Justice 60 (2020) 438–450, https://doi.org/10.1016/j.scijus.2020.06.002.
[10] P. Weber, E. Enzinger, B. Labrador, A. Lozano-Díez, D. Ramos, J. González-Rodríguez, G.S. Morrison, Validation of the alpha version of the E^3 Forensic Speech Science System (E^3FS^3) core software tools, Forensic Sci. Int.: Synergy 4 (2022), https://doi.org/10.1016/j.fsisyn.2022.100223 article 100223.
[11] R.J.F. Ypma, P.A. Maaskant-van Wijk, R. Gill, M.J. Sjerps, M. van den Berge, Calculating LRs for presence of body fluids from mRNA assay data in mixtures, Forensic Sci. Int.: Genetics 52 (2021), https://doi.org/10.1016/j.fsigen.2020.102455 article 102455.

Geoffrey Stewart Morrison*
Forensic Data Science Laboratory, Aston University, Birmingham, UK
Forensic Evaluation Ltd, Birmingham, UK

Daniel Ramos
AUDIAS – Audio, Data Intelligence and Speech, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Madrid, Spain

Rolf JF Ypma
Netherlands Forensic Institute, The Hague, the Netherlands
Forensic Data Science Laboratory, Aston University, Birmingham, UK

Nabanita Basu
Forensic Data Science Laboratory, Aston University, Birmingham, UK

Kim de Bie
Netherlands Forensic Institute, The Hague, the Netherlands

Ewald Enzinger
Eduworks Corporation, Corvallis, OR, USA
Forensic Data Science Laboratory, Aston University, Birmingham, UK

Zeno Geradts
Netherlands Forensic Institute, The Hague, the Netherlands
University of Amsterdam, Amsterdam, the Netherlands

Didier Meuwly
Netherlands Forensic Institute, The Hague, the Netherlands
University of Twente, Enschede, the Netherlands

David van der Vloed
Netherlands Forensic Institute, The Hague, the Netherlands

Peter Vergeer
Netherlands Forensic Institute, The Hague, the Netherlands

Philip Weber
Forensic Data Science Laboratory, Aston University, Birmingham, UK

* Corresponding author.
E-mail address: geoff-morrison@forensic-evaluation.net (G.S. Morrison).