

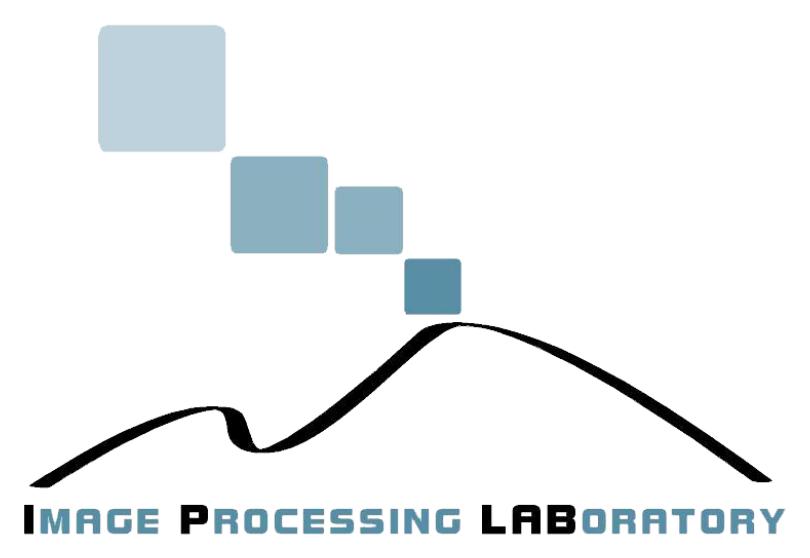
Innovative Methods for Non-Destructive Inspection of Handwritten Documents

Eleonora Breci, Luca Guarnera, Sebastiano Battiato

Department of Mathematics and Computer Science, University of Catania, Italy
 brecieleonora@gmail.com, luca.guarnera@unict.it, sebastiano.battiato@unict.it



Università di Catania

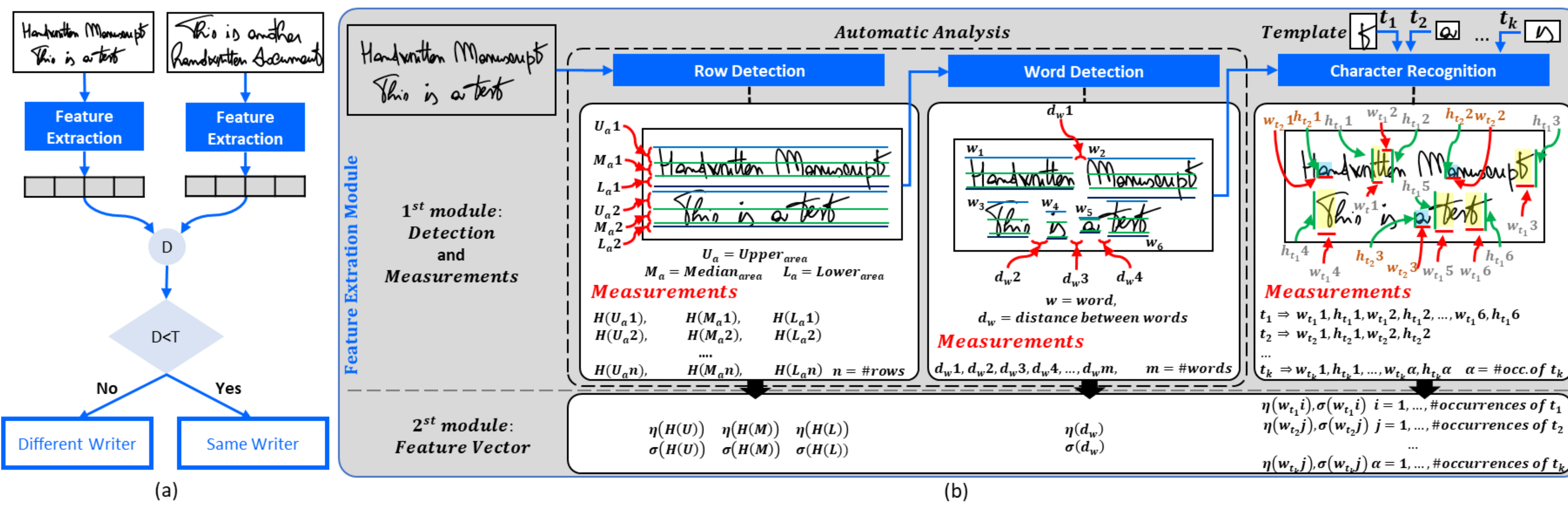


Abstract

Handwritten document analysis is an area of forensic science, with the goal of establishing authorship of documents through examination of inherent characteristics. Law enforcement agencies use standard protocols based on manual processing of handwritten documents. This method is time-consuming, is often subjective in its evaluation, and is not replicable. To overcome these limitations, in this paper we present a framework capable of extracting and analyzing intrinsic measures of manuscript documents related to text line heights, space between words, and character sizes using image processing and deep learning techniques. The final feature vector for each document involved consists of the mean (η) and standard deviation (σ) for every type of measure collected. By quantifying the Euclidean distance between the feature vectors of the documents to be compared, authorship can be discerned. Our study pioneered the comparison between traditionally handwritten documents and those produced with digital tools (e.g., tablets). Experimental results demonstrate the ability of our method to objectively determine authorship in different writing media, outperforming the state of the art

Proposed pipeline

(a) Comparison between documents. (b) Feature extraction module: 1st module): lines of text and words are automatically detected. Then, one or more characters (templates) chosen by the expert can be searched within the document. 2nd module): the feature vector is defined as the means (η) and standard deviations (σ) of all collected measures.



Results

(1) Testing with state-of-the-art datasets

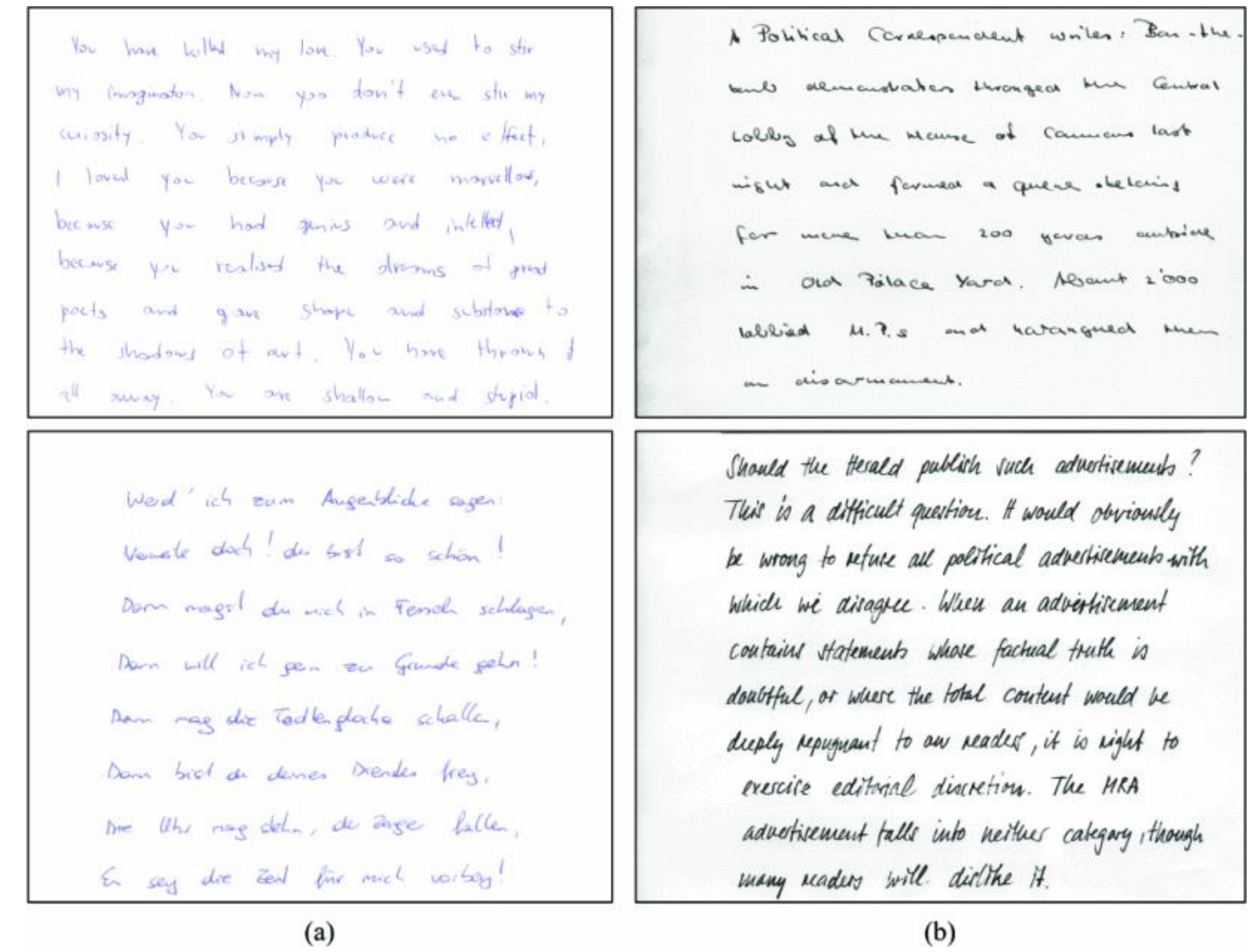


Fig. Examples of (a) CVL and (b) CSAFE digitized texts

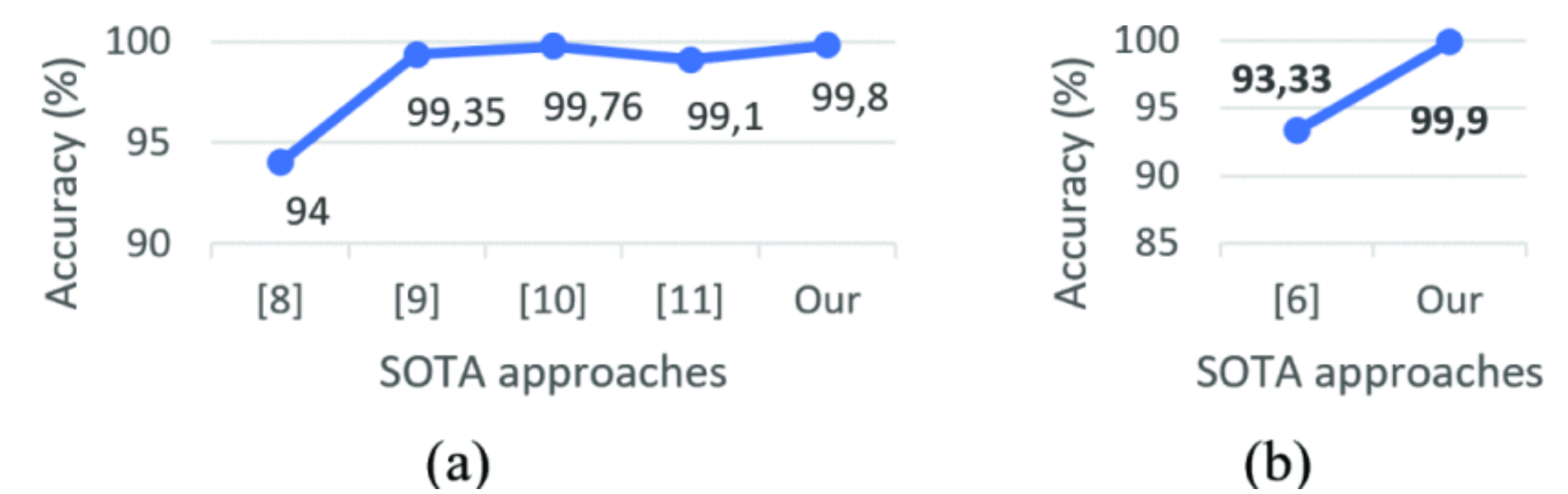


Fig. Comparison with SOTA approaches, by using CVL (a) and CVL + CSAFE datasets (b).

2) Preliminary results with a customized dataset:

- Images acquired from the classic "pen and paper" approach.
- Images acquired using tablets.



All documents were compared with each other, including handwritten and digital documents, resulting in a classification accuracy of 96%.

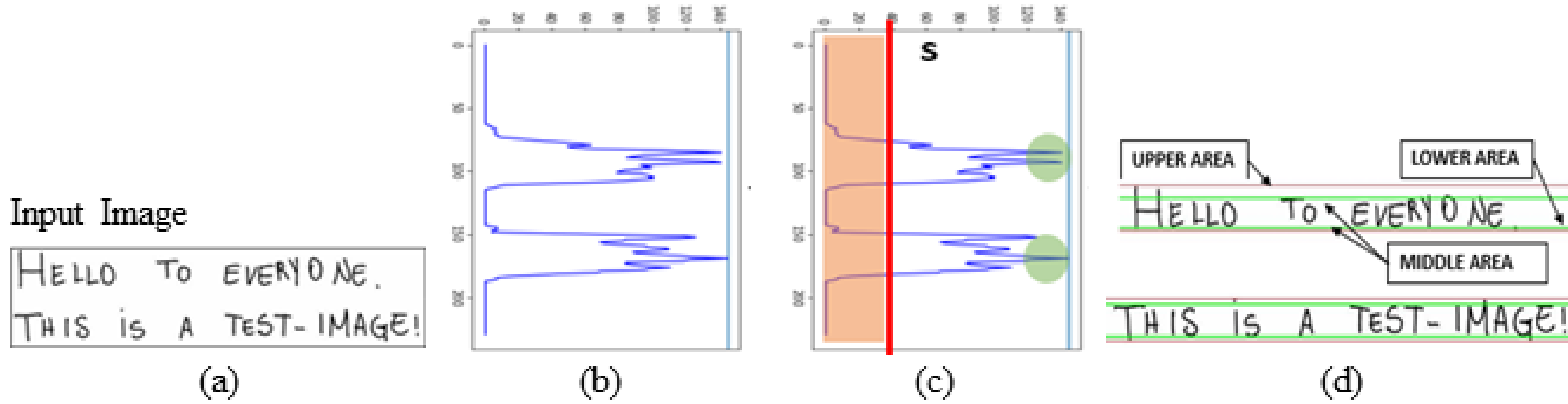
Web Page



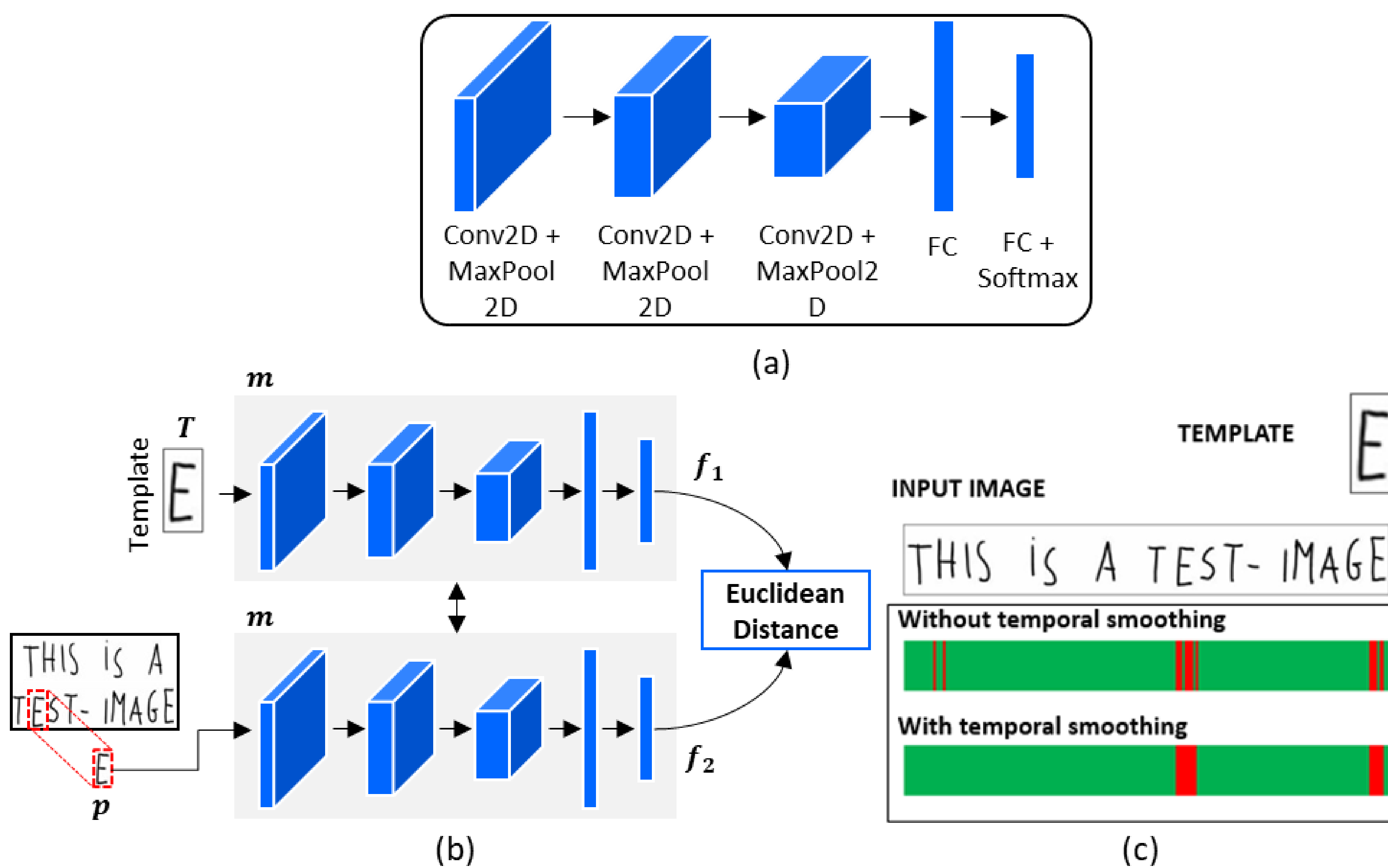
<https://iplab.dmi.unict.it/mfs/forensic-handwriting-analysis/innovative-methods-2023/>

Proposed Approach

Technique 1: Text Line and Word Detection



Technique 2: Character Recognition



References

- Luca Guarnera, et al., "GRAPHJ: A Forensics Tool for Handwriting Analysis," in Image Analysis and Processing-ICIAP 2017: 19th International Conference, Catania, Italy, September 11-15, 2017, Proceedings, Part II 19. Springer, 2017, pp. 591-601
- Eleonora Breci, et al., "A Novel Dataset for Non-Destructive Inspection of Handwritten Documents," in IS&T International Symposium on Electronic Imaging: Media Watermarking, Security, and Forensics, MWSF, 2024
- Luca Guarnera, et al., "Forensic Analysis of Handwritten Documents with GRAPHJ," Journal of Electronic Imaging, vol. 27, no. 5, pp. 051230-051230, 2018.
- "A-z handwritten alphabets in .csv format. kaggle. <https://www.kaggle.com/sachinpatel21/azhandwrittenalphabets-in-csv-format>, 2023.
- Sheng He and Lambert Schomaker, "Fragnet: Writer Identification Using Deep Fragment Networks," IEEE Transactions on Information Forensics and Security, vol. 15, pp. 3013-3022, 2020.
- Amy M Crawford, et al., "A Statistical Approach to Aid Examiners in the Forensic Analysis of Handwriting," Journal of Forensic Sciences, 2023.
- Abdelillah Semma, et al., "Writer Identification Using Deep Learning with Fast Keypoints and Harris Corner Detector," Expert Systems with Applications, vol. 184, pp. 115473, 2021.
- Akram Bennour, et al., "Handwriting Based Writer Recognition Using Implicit Shape Codebook," Forensic Science International, vol. 301, pp. 91-100, 2019.
- Parveen Kumar and Ambalika Sharma, "Segmentation-free Writer Identification Based on Convolutional Neural Network," Computers & Electrical Engineering, vol. 85, pp. 106707, 2020
- Songxuan Lai, et al., "Encoding Pathlet and SIFT Features with Bagged VLAD for Historical Writer Identification," IEEE Transactions on Information Forensics and Security, vol. 15, pp. 3553-3566, 2020.
- Sheng He and Lambert Schomaker, "Fragnet: Writer Identification Using Deep Fragment Networks," IEEE Transactions on Information Forensics and Security, vol. 15, pp. 3013-3022, 2020