



Universidade de Pernambuco  
Escola Politécnica de Pernambuco  
Programa de Pós-Graduação Acadêmica em Engenharia de Computação

Victor Kléber Santos Leite Melo

# Deep Generative Models to Synthesize Handwritten Signatures

Dissertação de Mestrado

Recife, setembro de 2016



Universidade de Pernambuco  
Escola Politécnica de Pernambuco  
Programa de Pós-Graduação Acadêmica em Engenharia de Computação

Victor Kléber Santos Leite Melo

# Deep Generative Models to Synthesize Handwritten Signatures

Dissertação de Mestrado

Dissertação apresentada ao Programa de Pós-Graduação acadêmico em ENGENHARIA DE COMPUTAÇÃO da Universidade de Pernambuco como requisito parcial para obtenção do título de Mestre em Engenharia de Computação.

Prof. Dr. Byron Leite Dantas Bezerra  
Orientador

Recife, setembro de 2016

*Este trabalho é dedicado às crianças adultas que,  
quando pequenas, sonharam em se tornar cientistas.*

O epígrafe é uma pagina onde o autor pode colocar uma citação ou pensamento que, de alguma forma, influenciou em seu trabalho acadêmico.

# Abstract

This is the abstract

# Resumo

Segundo a [??, 3.1-3.2], o resumo deve ressaltar o objetivo, o método, os resultados e as conclusões do documento. A ordem e a extensão destes itens dependem do tipo de resumo (informativo ou indicativo) e do tratamento que cada item recebe no documento original. O resumo deve ser precedido da referência do documento, com exceção do resumo inserido no próprio documento. (...) As palavras-chave devem figurar logo abaixo do resumo, antecidas da expressão Palavras-chave:, separadas entre si por ponto e finalizadas também por ponto. This is the english abstract.

# Contents

<b>List of Figures</b> . . . . .	<b>vi</b>
<b>List of Tables</b> . . . . .	<b>vii</b>
<b>List of symbols and abbreviations</b> . . . . .	<b>viii</b>
<b>1 Introduction</b> . . . . .	<b>1</b>
1.1 Problem statement . . . . .	2
1.2 Proposal outline . . . . .	2
1.3 Statements of the contribution . . . . .	2
1.4 Dissertation structure . . . . .	2
<b>2 Synthetic handwritten signatures</b> . . . . .	<b>5</b>
2.1 Automatic Handwritten signature verification systems . . . . .	5
2.2 Generation of duplicated samples . . . . .	6
2.3 Generation of new identities . . . . .	6
<b>3 Deep Generative Models</b> . . . . .	<b>7</b>
3.1 Neural Networks . . . . .	7
3.2 Generative Adversarial Networks (GAN) . . . . .	7
3.3 Autoencoder . . . . .	7
3.4 Variational Autoencoder . . . . .	7
3.4.1 DRAW . . . . .	7
<b>4 Offline Signature Generation From Dynamic Data</b> . . . . .	<b>8</b>
<b>5 Offline Forgery Generation</b> . . . . .	<b>9</b>
<b>6 Experiments and results</b> . . . . .	<b>10</b>
6.1 Datasets . . . . .	10
6.2 Experiments protocol . . . . .	10
<b>7 Conclusion</b> . . . . .	<b>11</b>
7.1 Contributions . . . . .	11
7.2 Future work . . . . .	11
<b>Bibliography</b> . . . . .	<b>12</b>

# List of Figures



# List of Tables

# List of symbols and abbreviations

ABNT	Associação Brasileira de Normas Técnicas
abnTeX	ABsurdas Normas para TeX
$\Gamma$	Letra grega Gama
$\Lambda$	Lambda
$\zeta$	Letra grega minúscula zeta
$\in$	Pertence

# Acknowledgements

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

# Introduction

In the modern society, biometric technology is used in several security applications for personal authentication. The aim of such systems is to confirm the identity of a given subject based on physiological or behavioral traits. In the first case, recognition is based on biological characteristics such as fingerprint, palmprint, iris, face, etc. The latter relies on behavioral traits such as voice pattern and handwritten signature [1].

Handwritten signature stands as one of the main approaches for identity authentication. One of the reasons for its widespread is the fact that signature acquisition is easy and non-invasive, and most individuals are familiar with its use in their daily life [2]. Due to its convenient nature, signatures can be employed as a sign of confirmation in a wide set of important documents, specifically, bank checks, credit card transactions, identification documents and a variety of business certificates and contracts.

As a behavioral trait, signatures are susceptible to spoof attacks, which is the attempt to forge the signature of a legitimate subject [1]. Two types of impostors are considered, namely: casual impostors (producing random forgeries) when no information about authentic writer signature is known, and real impostors (producing skilled forgeries) when some information of the signature is used [3].

If a signature on a document is forged, this document is also considered invalid. Thus, preventing frauds in the signature verification process has been a challenge for researchers around the world. However, manual signature-based authentication of a large set of documents is difficult and also a time-consuming and labor-intensive task. Hence, several Automatic Handwritten Signature Verification Systems (AHSVS) have been proposed to support this task. These systems aim to automatically decide if a query signature is in fact of a particular person or not.

AHSVS are essentially a pattern recognition application that works by receiving a signature as input, extracting a feature set from the data and classifying the sample using a template database as reference. As any pattern recognition system, AHSVS are learning-based, which requires a dataset that can be used to assess their performances to create accurate signature verification methods. These datasets contain signatures digitized by either using an optical scanner to obtain the signature directly from the paper or by using an acquisition device such as digitizing tablets or electronic pens with digital ink.

The two approaches are identified as offline (static) and online (dynamic), respectively. In the online modality, data is stored during the writing process and consists of a temporal sequence of the two-dimensional coordinates (x, y) of consecutive points, whereas in the offline case, only a static representation of the completed writing is available as an image. Moreover, each representation has specific attributes not present in the other [4]. Specifically, online data do not include information about the width of the strokes and the texture of the ink on the paper [5]. While the offline representation has lost all dynamic information of the writing process. As a result, features such as pen trajectory, which can be easily computed in the online domain, can only be inferred from a static image [6].

## **1.1 Problem statement**

Encouraged by the motivations depicted previously, the goal of this dissertation can be stated as follows: This work defines two approaches for testing software product line architectures defining activities, steps, inputs, outputs and roles in order to be confident that modifications (correction or evolution) are conform with the architecture specification, do not introduce unexpected errors and that the new versions work as expected.

## **1.2 Proposal outline**

## **1.3 Statements of the contribution**

## **1.4 Dissertation structure**

In the last few years, several handwritten signature datasets have been created and some made publicly available. The general corpus consists of a set of genuine and forgery signatures for each writer and can be categorized on different dimensions including the modality (online or offline), script and size.

Although many datasets containing samples for offline, online or both modalities combined haven been proposed, those datasets normally do not convey some important real world

challenges, not assessing the robustness of the systems on real world scenarios. Consequently, said systems often fail to deliver the expected results when employed in practice [7].

In practical scenarios, signatures are acquired on a wide set of conditions and in both modalities. Different conditions for online acquisition includes signatures acquired in several types of devices, e.g., using smartphones or different models of digitizing tablets. Moreover, when dealing with offline signatures, most of the samples are present in documents with complex backgrounds and with different signing area constraints. Examples of such documents include bank checks, contracts, identification documents, forms, etc. [7, 8]. Those distinct types of signatures often need to be integrated into the same system in an interoperable manner.

In regards to signature verification interoperability, many research problems are open to investigation, such as (i) development of complete document authentication systems involving both signature segmentation and verification process taking into account different signing area constraints (ii) analysis of the implications on AHSVS of the combination of signatures acquired on smartphones and conventional digitizing tablets (iii) development of systems capable of integrating both online and offline samples interchangeably, towards a unified signature biometry. With the currently available datasets, investigation on the direction of the listed research problems is limited to samples acquired in controlled environments or can not be made at all.

Works have been done on topics directed towards signature verification interoperability. Qiao et al. [9] proposed an offline signature verification system that uses online handwriting signatures instead of images in the registration phase, however, in the experiments the authors used synthetic offline images generated by the interpolation of online signature samples. Upalapati [10] proposed a system to integrate both modalities of handwritten signatures, not only providing a method to match offline signatures against an online and vice-versa, but also using both static and dynamic features, when available, to improve the system performance. Ahmed et al. [8] proposed a method for signature extraction from documents, it is noteworthy that the segmentation accuracy was evaluated only on the patch level. According to the authors, it is due to the lack of publicly available datasets containing ground truth of signatures on the stroke level. Ahmed et al. in [7] discuss the currently non-applicability of most signature verification systems and the lack of complete document authentication systems involving signature segmentation and verification. According to the authors, it is due to the absence of datasets suitable for the development of such systems, containing both patch and stroke level ground truth. Pirlo et al. [11] investigated the effects of signing area constraints on geometric features of online signatures. Diaz et al. [5] proposed several approaches to synthetically generate offline signatures simulating the pen ink deposition on the paper based on dynamic information from online signatures. Zareen and Jabin [12] presented a comprehensive survey of mobile-biometric systems and proposed a method for online signature verification. The approach was evaluated on the SVC [13] dataset and a database acquired using a mobile device [14].

Aiming to overcome the limitations of the current state of handwritten signature datasets,

we present the RPPDI-SigData, an evaluation dataset for AHSVS that includes signatures captured for both online and offline modalities and from different signing conditions. Samples for the online modality were acquired on smartphones and digitizing tablets and for the offline domain acquired in documents with complex backgrounds (including a stroke level ground-truth) and different signing area constraints.

Alongside with the description of the RPPDI-SigData, this chapter also summarizes 17 publicly available handwritten signature datasets. Our goal is to provide the reader an overview of the existing evaluation datasets and its main characteristics such as number of samples, protocols, type of forgeries, script.

The rest of the chapter is structured as follows: Section 2 presents an overview of the existing evaluation datasets for handwritten signature verification. Section 3 describes RPPDI-SigData, our proposed dataset and we conclude the chapter in Section 4.

# Synthetic handwritten signatures

## 2.1 Automatic Handwritten signature verification systems

In this context, the scientific literature has proposed two ways to synthesize the signatures [7], [8]:

1) Generation of duplicated samples. In this case algorithms create artificial intra-personal variability from reference specimens. These reference specimens are the seed of these algorithms. Therefore, they do not create new individuals but possible signatures executed by the signers. Both in on-line [9], [10], [11], [12] and off-line [13], [14], [15], [16], [17], [18], [19], [20], the algorithms are able to generate new samples of a real user. Duplicated signatures present several advantages such as training improvement in ASV, statistically meaningful evaluations, enlarging the number of samples in a database, etc.

2) Generation of new identities. Algorithms are designed in this modality by taking into account different rules and logics in order to simulate realistic effects. In [21] a model is described to generate dynamic signatures based on visual characteristics extracted from the time domain. The method was evaluated from a visual validation. In Galbally et al. [7], [8] the full generation of on-line flourish-based signatures was carried out through two algorithms based on spectral analysis and the kinematic theory of rapid human movements. The first attempt to generate static flourishes in context of static signature image is due to [22]. Two follow-up works were presented in [1], [2] where authors coped with the generation of Western static and dynamic signatures (text and flourish) as well as skilled forgeries under a neuromotor approach.



## **2.2 Generation of duplicated samples**

## **2.3 Generation of new identities**

# Deep Generative Models

## 3.1 Neural Networks

## 3.2 Generative Adversarial Networks (GAN)

## 3.3 Autoencoder

## 3.4 Variational Autoencoder

### 3.4.1 DRAW

## **Offline Signature Generation From Dynamic Data**

CHAPTER

# 5

## Offline Forgery Generation

# Experiments and results

## 6.1 Datasets

## 6.2 Experiments protocol



# **Conclusion**

## **7.1 Contributions**

## **7.2 Future work**

# Bibliography

- [1] JAIN, A. K.; ROSS, A.; PRABHAKAR, S. An introduction to biometric recognition. *IEEE Transactions on circuits and systems for video technology*, IEEE, v. 14, n. 1, p. 4–20, 2004.
- [2] IMPEDOVO, D.; PIRLO, G. Automatic signature verification: the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, IEEE, v. 38, n. 5, p. 609–635, 2008.
- [3] FIERREZ, J.; ORTEGA-GARCIA, J. On-line signature verification. In: *Handbook of biometrics*. [S.l.]: Springer, 2008. p. 189–209.
- [4] VIARD-GAUDIN, C. et al. The ireste on/off (ironoff) dual handwriting database. In: IEEE. *Document Analysis and Recognition, 1999. ICDAR'99. Proceedings of the Fifth International Conference on*. [S.l.], 1999. p. 455–458.
- [5] DIAZ-CABRERA, M. et al. Generation of enhanced synthetic off-line signatures based on real on-line data. In: IEEE. *Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on*. [S.l.], 2014. p. 482–487.
- [6] NEL, E.-M.; PREEZ, J. A. D.; HERBST, B. M. Estimating the pen trajectories of static signatures using Hidden Markov models. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 27, n. 11, p. 1733–1746, 2005.
- [7] AHMED, S. et al. Towards Signature Segmentation & Verification in Real World Applications. In: *Proceedings of the 16th Biennial Conference of the International Graphonomics Society*. [S.l.: s.n.], 2013. p. 139–142.
- [8] AHMED, S. et al. Signature segmentation from document images. In: IEEE. *Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on*. [S.l.], 2012. p. 425–429.
- [9] QIAO, Y.; LIU, J.; TANG, X. Offline signature verification using online handwriting registration. In: IEEE. *2007 IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.], 2007. p. 1–8.
- [10] UPPALAPATI, D. Integration of Offline and Online Signature Verification systems. *Department of Computer Science and Engineering, IIT, Kanpur*, 2007.
- [11] PIRLO, G. et al. Interoperability of biometric systems: Analysis of geometric characteristics of handwritten signatures. In: SPRINGER. *International Conference on Image Analysis and Processing*. [S.l.], 2015. p. 242–249.
- [12] ZAREEN, F. J.; JABIN, S. Authentic mobile-biometric signature verification system. *IET Biometrics*, IET, v. 5, n. 1, p. 13–19, 2016.

- 
- [13] YEUNG, D.-Y. et al. SVC2004: First international signature verification competition. In: *Biometric Authentication*. [S.l.]: Springer, 2004. p. 16–22.
- [14] MARTINEZ-DIAZ, M. et al. Mobile signature verification: feature robustness and performance comparison. *IET Biometrics*, IET, v. 3, n. 4, p. 267–277, 2014.