

# Aprendizagem de máquina aplicada ao reconhecimento e verificação de assinaturas

Daniel Elias<sup>1</sup>, Paulo Henrique<sup>1</sup> e  
Pedro Afonso<sup>1</sup>

<sup>1</sup>Departamento de Informática – Instituto Federal de Educação,  
Ciência e Tecnologia de Minas Gerais (IFMG)  
Sabará – MG – Brasil

{danielias.santos, paulohenriquercs, pedroafonsouza}@gmail.com

**Resumo.** A verificação da autenticidade de assinaturas manuscritas é tarefa realizada cotidianamente em instituições financeiras e cartórios, entre outros, para comprovar a autoria de intenções e acordos em documentos de interesse. Essa análise de autenticidade quando feita por humanos, está restrita a profissionais especializados que, ainda assim, podem cometer erros. Desta forma, a automatização da atividade pode trazer benefícios, como a redução de custos e a celeridade. Associadas a essa automatização, técnicas de inteligência artificial, tal como a aprendizagem de máquina, podem ser aplicadas para minorar a ocorrência de erros. Portanto, este trabalho busca propor um sistema de verificação da autenticidade de assinaturas utilizando técnicas de aprendizagem de máquina.

## 1. Introdução

Assinaturas manuscritas são o meio mais utilizado, atualmente, para firmar contratos e intenções. Conforme fundamentou Freitas *et al.*[1], sistemas de gerenciamento e controle de informações, inerentes à era digital, são constantemente confrontados com este instituto arcaico, mas estabelecido, de atribuição de autenticidade a documentos, que é a assinatura manuscrita. Neste sentido, diversos são os estudos realizados para harmonizar estes dois aspectos que, em grande medida, se mostram antagônicos. Além disso, as assinaturas representam, juridicamente, a comprovação de intenções e acordos, que geram responsabilidades a seus autores. Vê-se, então, que a assinatura tem valor para o seu autor e, portanto, há o risco de fraudes.

Além disso, diversos fatores subjetivos influenciam a forma das assinaturas manuscritas, tais como nacionalidade, idade, tempo, hábitos, estado psicológico ou mental e condições físicas do autor. A verificação de assinatura tem o objetivo de comprovar a autenticidade por meio de comparação, entre uma assinatura reconhecida como verdadeira e outra que esteja sob questionamento[2]. Ainda segundo Gonçalves[2], os métodos de aquisição de assinaturas podem ser *on-line*, por meio de *hardwares* específicos (tal como mesa digitalizadora) ou *off-line*, método tradicional, no qual a assinatura é disposta em papel, e aí se enquadram os sistemas automáticos de verificação.

Segundo Coetzer *et al.*[3], os seres humanos apresentam altos índices de erros no processo de verificação de assinaturas. Percebe-se, então, que esta é uma tarefa árdua, atribuída a profissionais especialistas, e ainda assim está sujeita a erros. Gonçalves[2]

cita os tipos de erros cometidos, que podem ser de “falsa rejeição”, quando uma assinatura genuína é verificada como falsa, e “falsa aceitação”, quando uma assinatura falsa é classificada erroneamente como genuína.

O intuito da verificação automática de assinaturas, em termos práticos, é prover automatização, velocidade, escalabilidade e segurança à atividade, nos mais diversos contextos aplicáveis. Desta forma, este trabalho pretende desenvolver um sistema que, por meio da aprendizagem de máquina, seja capaz de executar a verificação de assinaturas manuscritas com uma taxa de erro aceitável, em relação ao desempenho humano da tarefa.

## 2. Objetivo

Para cumprir o seu objetivo, este trabalho intenta responder as seguintes perguntas:

1. Qual é a taxa de acertos da técnica utilizada para reconhecimento e validação de assinaturas manuscritas?
2. Qual é a taxa de ocorrência de falsos-positivos?
3. Qual é a taxa de ocorrência de falsos-negativos?
4. A técnica utilizada no aprendizado de máquina tem um resultado melhor, quanto à análise, reconhecimento e validação que a verificação humana?

## 3. Trabalhos Relacionados

A fim de observar metodologias existentes na área de verificação de assinaturas que utilizam técnicas de aprendizado de máquina, foram estudados trabalhos relevantes ao assunto. Para Gonçalves[2] os métodos *on-line* e *off-line* se diferem na concepção da assinatura, porém cada um tem seus prós e contras. A ideia do autor é melhorar a capacidade de análise da abordagem *off-line* uma vez que é a mais utilizada.

Relacionado à maneira de serem obtidas as assinaturas Freitas *et al.*[1] ressalta que para se obter o reconhecimento de uma dada forma escrita é necessário estabelecer um grau de comparação entre as outras formas de treinamento afim de obter um maior grau de semelhança dos itens. Em seu outro estudo Justino[4] explica que existem elementos técnicos genéticos que muitas vezes passam imperceptíveis pela análise computacional, e portanto é necessário aumentar o grau de verificação nesses quesitos. Neste trabalho são usadas essas considerações para a construção de um sistema capaz de verificar as formas escritas.

No estudo de Tebaldi[5], é mostrado um exemplo de comparação de assinaturas. O modelo é feito de forma complexa, utilizando o algoritmo R-prop, que é uma vertente do *back-propagation*, porém o resultado do estudo não é tão satisfatório.

## 4. Metodologia

### 4.1. Coleta dos dados

Foram recolhidas diversos exemplares de assinaturas manuscritas de alunos do IFMG unidade Sabará, com o intuito de alimentar a base de dados em que serão inferidos os processamentos devidos para a utilização de dados viáveis.

## 4.2. Processamento da imagem

Foram realizados recortes nas assinaturas de forma a estabelecer um padrão de dimensões de altura e largura para o consumo da programa ImageJ. A partir da utilização do programa ImageJ, e da execução do *plugin Trainable Weka Segmentation*, se sucederam uma sequência de inferências nas imagens selecionadas para gerar arquivos no formato ARFF (*Attribute-Relation File Format*) um tipo arquivo de texto ASCII que contem uma lista de instâncias, que compartilham um conjunto de atributos de cadeia, data e instâncias esparsas que possuem duas seções, uma de informações de cabeçalho (*Header*) e outra de informações características de dados (*Data*).

O próximo passo foi de criar uma sequencia de rotinas pré-pro gamadas (*Macro*), em que se processa uma automação de ações executadas no diretório onde estão armazenados os objetos referentes as assinaturas coletadas, realizando tratamento das imagens com inferências programadas no macro, em que se seleciona uma classe referente a camada tracejada detectada automaticamente em cada imagem carregada, e assim se processando um arquivo resultante com os dados necessários para a utilização posterior com uma estratégia de aprendizado de máquina selecionada.

## 4.3. Algoritmo de aprendizado de máquina

### Referências

- [1] FREITAS, C. O. D. A.; JUSTINO, E. J. R. A análise e o reconhecimento da escrita manuscrita na língua portuguesa. Pontifícia Universidade Católica do Paraná, 2003.
- [2] GONÇALVES, D. B. Agrupamento de classificadores na verificação de assinaturas *Off-Line*. Pontifícia Universidade Católica do Paraná, 2008.
- [3] COETZER B. M. HERBST, J. A. D. P. J. Off-line signature verification: A comparison between human and machine performance. Tenth International Workshop on Frontiers in Handwriting Recognition, 2006.
- [4] JUSTINO, E. J. R. O grafismo e os modelos escondidos de markov na verificação automática de assinaturas. Pontifícia Universidade Católica do Paraná, 2001.
- [5] TEBALDI, F. B. Estudo de redes neurais artificiais para verificação e classificação de assinaturas através de imagens. Centro Universitário Eurípedes de Marília, 2007.