



Verificação da Autenticidade de Assinaturas Baseada em Técnicas de Aprendizado de Máquina Profundo

Heitor Boschirolli Comel*, Hélio Pedrini

Resumo

O uso de aprendizado de máquina para verificação automática de assinaturas teve uma grande evolução nos últimos anos. Apesar disso, mesmo os melhores modelos continuam com desempenhos insatisfatórios nas bases de dados mais desafiadoras quando comparados com outras atividades de verificação biométrica. Neste trabalho, uma investigação foi realizada sobre os erros cometidos por modelos de verificação de assinatura automáticos e medido o desempenho de variações feitas em modelos já existentes. Os resultados foram avaliados na base de dados UTSIG e comparados com os resultados de outros modelos já existentes.

Palavras-chave:

Offline signature verification, anomaly detection, image processing.

Introdução

Enquanto algumas áreas de autenticação biométrica já possuem sistemas baseados em aprendizado de máquina adotados na prática, outras ainda não têm um desempenho suficiente para isso. Dentre elas, encontra-se a verificação de assinaturas automatizada, que tem como objetivo classificar uma assinatura como genuína ou falsificada.

O objetivo do trabalho é estudar os erros cometidos por sistemas de verificação de assinaturas existentes e propor métodos para mitigar tais erros.

Resultados e Discussão

Durante a análise das classes de assinaturas em que os modelos testados tinham piores resultados, percebeu-se que as dimensões da imagem eram um fator de impacto: quanto maior as dimensões, pior o desempenho. A provável causa disso é que todas as imagens utilizadas são redimensionadas para um mesmo tamanho antes de serem classificadas.

Outro fator que pareceu influenciar o desempenho do modelo é a quantidade de assinaturas em diferentes ângulos que um autor produz. Se um mesmo autor produz as assinaturas com diferentes inclinações, o desempenho do modelo tende a diminuir.

Por fim, verificou-se que a quantidade de amostras usadas para treinar o modelo causou um grande impacto no resultado. Isso pode ser visto na Tabela 1.

Na tentativa de solucionar o problema da quantidade de dados usada para o treinamento, aplicou-se um aumento dos dados por meio de translações, em que, para cada assinatura, produziu-se outras três compostas por translações da assinatura original na imagem.

Para contornar o problema das diferentes inclinações, empregou-se também um aumento de dados, entretanto, agora com transformações de rotação.

Como não há uma margem nas imagens que contêm as assinaturas, para realizar translações foi preciso adicionar tal margem. Isso resulta em um aumento no tamanho das imagens, o que dificulta a classificação. Nas rotações, um problema similar foi encontrado ao se aplicar uma rotação em uma imagem, em que surge uma margem que não foi removida completamente. O uso de translações piorou o desempenho do modelo, enquanto que o uso de rotações o manteve constante. O uso de normalização nas imagens reduziu a taxa *Equal Error Rate* (EER) para todos os modelos testados.

Tabela 1. Resultados do modelo [1] para diferentes quantidades de assinaturas usadas para treinamento.

Amostras usadas para treinamento	Equal Error Rate (EER)
5	11,80%
10	9,71%
12	9,24%

Conclusão

Dos experimentos realizados, apenas a normalização – que não altera as dimensões das imagens – proporcionou uma melhora no desempenho do modelo.

Para que as técnicas de aumento de dados possam produzir bons resultados, será preciso alterar a rede neural utilizada no modelo de forma que ela utilize imagens maiores. Assim, acredita-se que técnicas de aumento de dados possam ser usadas sem consequências negativas.

Agradecimentos

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa de Iniciação Científica.

[1] Hafemann LG, Sabourin R, Oliveira LS. Learning Features for Offline Handwritten Signature Verification using Deep Convolutional Neural Networks. Pattern Recognition. 2017 Oct 1;70:163-76.