

Resumo

A presente dissertação, com o título “*Sistema automático de reconhecimento do montante de um cheque*”, foi realizada no *Instituto de Engenharia e Sistemas e Computadores do Porto* (INESC Porto), de Outubro de 2007 a Março de 2008.

Apesar do aparecimento de métodos electrónicos mais seguros, os cheques bancários continuam a ser utilizados frequentemente como forma de pagamento, e a sua verificação e validação envolvem actualmente o reconhecimento manual de cheques manuscritos com grande variedade e complexidade de caligrafias. Existe assim a necessidade de garantir um serviço mais rápido, eficaz, e automático, capaz de responder ao elevado número de cheques bancários utilizados.

Pretende-se com este trabalho elaborar a especificação de um sistema que, após submissão de imagens digitalizadas de cheques manuscritos, possa reconhecer os montantes especificados nos mesmos. Para atingir os objectivos propostos, foi necessário determinar uma arquitectura adequada, bem como investigar e comparar algoritmos de processamento de texto manuscrito.

Foi efectuada uma análise do problema proposto, das soluções existentes e das plataformas que melhor se adequam ao desenvolvimento do sistema pretendido, tendo sido escolhida a plataforma .NET 2.0 e tecnologias Microsoft associadas, bem como as bibliotecas AForge.NET 1.51 (*Processamento Digital de Imagem*) e Weka 3.4 (*Machine Learning*).

Foi criada uma base de dados de cheques portugueses manuscritos, de forma a permitir testar os métodos e algoritmos implementados no protótipo desenvolvido, o qual serviu de base para validar as conclusões e resultados obtidos.

Os cheques digitalizados podem ser afectados por ruído visual e/ou um grau de rotação significativo (resultantes do processo de digitalização), implicando assim uma fase de pré-processamento. Esta fase recorre a algoritmos de *Processamento Digital de Imagem* para a obtenção dos resultados pretendidos.

Para o reconhecimento do montante manuscrito (*Optical Character Recognition* – OCR e *Intelligent Character Recognition* - ICR) é necessária ainda a segmentação do cheque para extracção dos campos pretendidos (valor de cortesia e valor legal). São utilizados

algoritmos de *Machine Learning* para reconhecimento e classificação dos valores escritos no cheque. Com base na bibliografia e artigos relacionados com a área, foram analisados os paradigmas das *Redes Neurais* e *Máquinas de Vectores de Suporte* para o reconhecimento do valor de cortesia, e *Modelos Escondidos de Markov* para o reconhecimento do valor legal.