
IDCHECK.IO

Application mobile de contrôle de documents d'identité

Ahmad Montaser Awal — Fabien Chevalier

AriadNEXT
80 Avenue des Buttes de Coësmes
35700 Rennes
{montaser.awal, fabien.chevalier}@ariadnext.com

RÉSUMÉ. Il existe un très grand nombre de techniques de sécurisation des documents d'identité et il n'est pas rare de voir un même document en implémenter plusieurs dizaines pour accroître, voire rendre impossible, le travail du fraudeur. Le contrôle des documents d'identité effectué par notre solution IDCHECK.IO permet de lutter efficacement contre la fraude documentaire qui est devenue la troisième plus grande industrie criminelle au monde et coûte chaque année des milliards de dollars à l'économie mondiale.

ABSTRACT. There is a very large number of identity document security techniques and it is common to implement several techniques to better protect these documents from eventual fraudulent acts. Control of identity documents carried by our IDCHECK.IO solution effectively fights against document fraud that has become the third largest criminal industry in the world and annually costs billions of dollars to the global economy.

MOTS-CLÉS : Application mobile, Contrôle d'identité, Acquisition, Segmentation, OCR, Classification

KEYWORDS: Mobile application, Identity verification, Image capturing, Segmentation, OCR, Classification

1. Introduction

Le fraude documentaire est un enjeu sociétal majeur. IDCHECK.IO est une application mobile qui permet la vérification automatique des documents d'identité d'une manière simple et rapide, accessible au grand public. Nous présentons dans cette démonstration une application mobile capable de :

- 1) En mode hors-ligne (non-connecté) : faire une lecture et des contrôles basiques de la pièce d'identité localement dans l'appareil mobile.
- 2) En mode En-ligne (connecté) : faire une lecture et des contrôles avancés en se connectant à un serveur distant.

2. Description

2.1. Acquisition

L'acquisition de l'image d'un document est guidée par deux modules garantissant une bonne qualité de l'image requise. Le premier module est un détecteur de flou basé sur l'analyse du spectre de fourrier. Le deuxième module vise à détecter le reflet. Ensuite, une acquisition automatique est déclenchée lorsque la qualité de l'image est jugée suffisante. Ce dernier module, nommé 'SmartCrop', se base sur la transformé de Hough pour détecter les bords de document et prendre l'image quand le ratio du document détecté est conforme à un standard ICAO (Organisation de l'aviation civile internationale).

2.2. Analyse locale (*Segmentation et reconnaissance des caractères*) :

La Zone de Lecture Automatique (ZLA) d'un documents d'identité contient les informations essentielles présentes dans la partie visible du document. Le nombre de ligne de cette zone varie en fonction du format de 1 à 3 lignes. Les lignes de cette zone sont détectées automatiquement grâce à un algorithme de segmentation sans aucune information à priori et ensuite segmentées en caractères. Ensuite, une reconnaissance des caractères est effectué en se basant sur un classifieur classique de type SVM. Une fois les caractères décodés, la ZLA est parsée afin d'extraire les informations du document (nom, prénom, numéro, ...). De plus, nous avons développé des parseurs génériques permettant d'avoir une couverture géographique mondiale. En outre, le décodage de la ZLA permet d'avoir un contrôle locale en temps réel sur les chiffres de sécurité présents dans la ZLA conformément à la norme ICAO.

2.3. Analyse distant (*Classification d'images et contrôles avancés*)

Le mode connecté nous permet de faire des traitements avancés sur l'image. Dans un premier temps, l'image est classifiée dans sa totalité afin d'en déduire le type de

documents. Cette classification a deux avantages : 1) contrôler authenticité du document en comparant le type déduit du classifieur avec le type déduit de la ZLA, 2) connaître le type des documents n'ayant pas une ZLA (les versos et les anciennes pièces). Les champs hors-ZLA sont ensuite décodés par un OCR ce qui permet l'extraction de toutes les informations présentes dans le document et de faire un contrôle croisé avec les informations déjà extraites de la ZLA. Finalement, des contrôle très avancés pourront être faits en fonction du modèle de document trouvé (recherche des motifs, des impressions de sécurité, ...).