Face detection in identity documents: an efficient security feature under challenging constraints

Lara Younes, Ahmad Montaser Awal

Research departement - ARIADNEXT, Rennes, France

Abstract

Identity documents are usually highly protected with security features. The ID's holder photo is one of the basic security features as it certifies that the ID belongs to the right person. While face detection has been largely studied in the literature, there is still place for improvement specially in the context of mobile captured images of an ID document. This work tackles the face detection task in ID documents captured in the wild as an efficient security feature for identity verification. It presents a comparison of the mainly used face detectors in the literature. It highlights the remaining effort to be done for detecting faces in documents and suggests a framework for training a new face detector of better accuracy. It also gives recommendations of optimal data augmentation strategies to reach high precision for the detection of low effort photo-related frauds in documents. The result is a face detector for which the effectiveness is also reported over the MIDV-2020 dataset.

Face detection Framework

A unified training and evaluation framework, three pre-trained state-of the art object detection architectures:

- EffecientDet (ED-d0)
- SSD architecture with two different backbones Mobilenet v1 and Mobilenet v2.
- Baseline: MTCNN and opencv detector (cvDnn).

Dataset. A private dataset of identity documents captured in the wild has been collected internally. The dataset includes following sets:

- IDs 1 :images with face zone occlusion (OVDs presence or light reflections)
- IDs 2: difficult dark skinned faces
- Mixed: mixed simple faces with no challenges
- Real frauds whose types ranges from replacing the face by a silhouette, oval shapes, or a smiley or strikes over the face zone
- Selfies

Figure 1. Examples of low effort real frauds.

Experiments without data Augmentation

	Baseline		Trained		
	MTCNN	cvDnn	ED-d0	ssdMNv1	ssdMNv2
IDs 1	62.6%	87.98%	96.2%	96.18%	94.77%
IDs 2	66.10%	39.44%	88.66%	88.13%	85.69%
Mixed	98.89%	99.05%	87.77%	99.5%	99.36%
Total	81.31%	91.03%	97.31%	97.48%	96.7%
Threshold	0.8	0.82	0.94	0.85	0.87

Table 1. Initial results using the raw dataset for training: TAR@0.07 FAR

Augmentation strategy

Augmentation of the dataset with negative examples of similar aspect as low-effort types of document frauds where the user tries to cover a part of the face in the identity document.

Figure 2. Augmentations: geometric shapes over face or a sketch of a smiley generated through the google quickdraw library covering the face.

Experiments with Augmentation

	ED-d0	ssdMNv1	ssdMNv2
IDs 1	86.63%	95.56%	97.92%
IDs 2	56.20%	90.05%	93.54%
Mixed	97.77%	99.20%	99.17%
Total	90.8%	97.98%	98.59%
Threshold	0.9	0.8	0.8
Selfies	98.95%	99.05%	99.2%

Table 2. Results obtained after data augmentation.

Experiments over MIDV2022

The best selected model from our experiments (ssdMNv2) is compared with the public detectors over the 61421 video clips images of the MIDV2020. It's effectiveness is proved through the obtained results.

	mtcnn	cvDnn	ssdMNv2
TAR	96.59%	94.53%	97.59%

Table 3. Tests over the 61421 video clips images of the MIDV2020 for the selected target confidence threshold.

Future work

Future work will focus on the use of the detected faces for face verification of people for a secure remote identity verification.