

ID documents matching and localization with multi-hypothesis constraints

Guillaume Chiron, Nabil Ghanmi and Ahmad Montaser Awal
AriadNEXT – Rennes, France
{guillaume.chiron, nabil.ghanmi, montaser.awal}@ariadnext.com



ARIADNEXT

Context, Objectives and Challenges

- **Localize** and **classify** ID documents in the wild (webcam, smartphone, scan, digital sources).
- Support a real life ever growing multi-country document coverage.
- Intrinsic variable nature of ID models :



Border detection not suitable as borders/corners are unclear (open booklets with mutli pages)

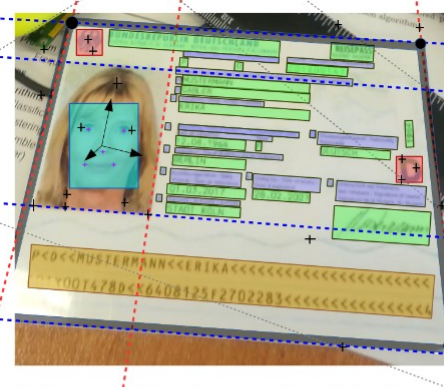
Keypoints matching not suitable due to model variations (barcode, prints multiple neighboring offsets), textureless, few anchorable keypoints

Text Driven not suitable due to few orthogonal text variations (barcode, prints field that are tight and randomly defined).

- Avoiding machine learning approaches: few data for new models, privacy concerns to collect, use or share representative datasets (e.g. limitations of MIDV-500 public dataset)

Exploitable features on ID documents

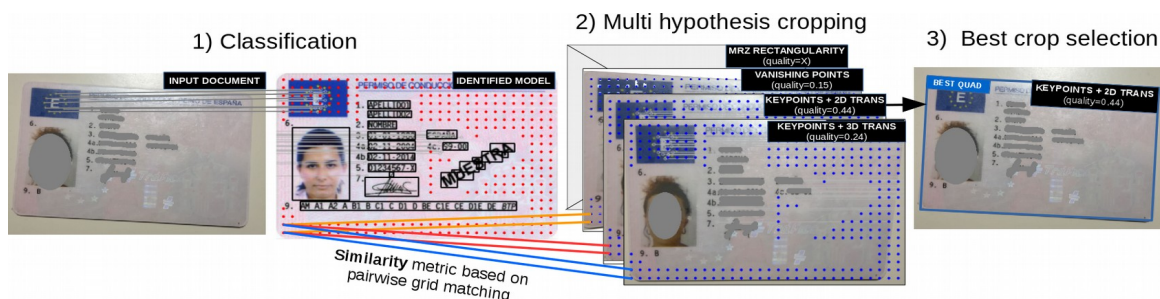
- Overview of potential features that could be used to accurately localize common ID documents :



Usable feature	Extraction method
Keypoints	Surf [5] + Ransac 3d/2d
Vanishing points	HoughNet [17] or H from VPs [6]
MRZ	Contour match. on "<" [15]
Border	(En/De)coder style U-Net[13]
Corner	R-CNN [9]
Face	SSD [25]
Landmarks	Face alignment [10]
Headpose	Keypoint-less network [16]
Logo	One-shot classifier [7]
Variable fields	Text box detector [12]
Fixed labels	Hybrid network det./reco. [27]

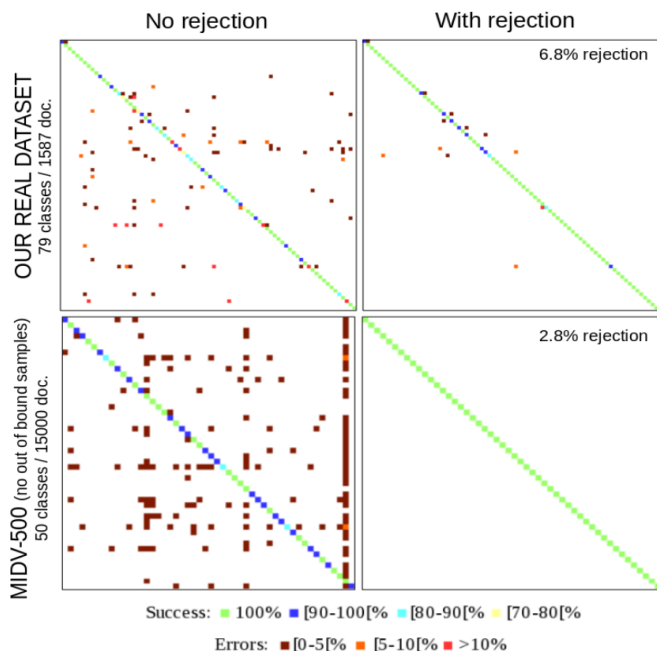
Multi-hypothesis approach proposed

- Assuming the classification is done and a model sample is known [4], our approach tests different crops (built up from complementary features) and then selects the most competing one on the basis of a "model to query" visual similarity metric. The idea is to maximize the coverage of both model diversities and capturing conditions by automatically focusing on the most relevant features/constraints.



Classification results

- Classification (relying on [4]) is an essential early step, so we ensure its robustness independently of the localization.



Localization results

- Ablation study of our multi hypothesis localization approach :

Best selected hypothesis repatriation in docs % (- are ablated hypothesis)					Accepted crops >0.9 Jaccard dist	
Keypoints 3D trans.	Keypoints 2D trans.	MRZ rect.	Vanish. points	Rejected (no crop)	Detected crop	Potential max
45%	23%	5%	27%	0%	92.8%	96.2%
61%	33%	6%	-	0%	92.5%	94.4%
65%	35%	-	-	0%	91.8%	93.5%
48%	24%	-	28%	0%	92.2%	95.6%
-	90%	10%	-	0%	90.2%	90.5%
100%	-	-	-	0%	87.8%	87.8%
-	100%	-	-	0%	89.3%	89.3%
-	-	20%	-	80%	18.7%	18.7%
-	-	-	90%	10%	73.2%	73.2%

Taken individually, no hypothesis is able to overpass an accepted crop rate of 89.3%. However, when features achieve 92.8% as an accepted crop rate.

Conclusion

- **Main contribution:** a novel approach for accurately localizing ID documents, which tests different crop hypotheses and selects the best one. Results have shown the superiority of our solution compared to a more traditional mono-hypothesis one.
- **Side contributions:**
 - 4 crop hypothesis (or combinations of features) performing well together (e.g. vanishing points).
 - Successfully reproduced results of [4] on both an academic dataset of reference (MIDV-500) and supposedly more representative private one.