

Large-Scale Historical Watermark Recognition: dataset and a new consistency-based approach



Xi Shen¹

Ilaria Pastrolin²

Oumayma Bounou¹

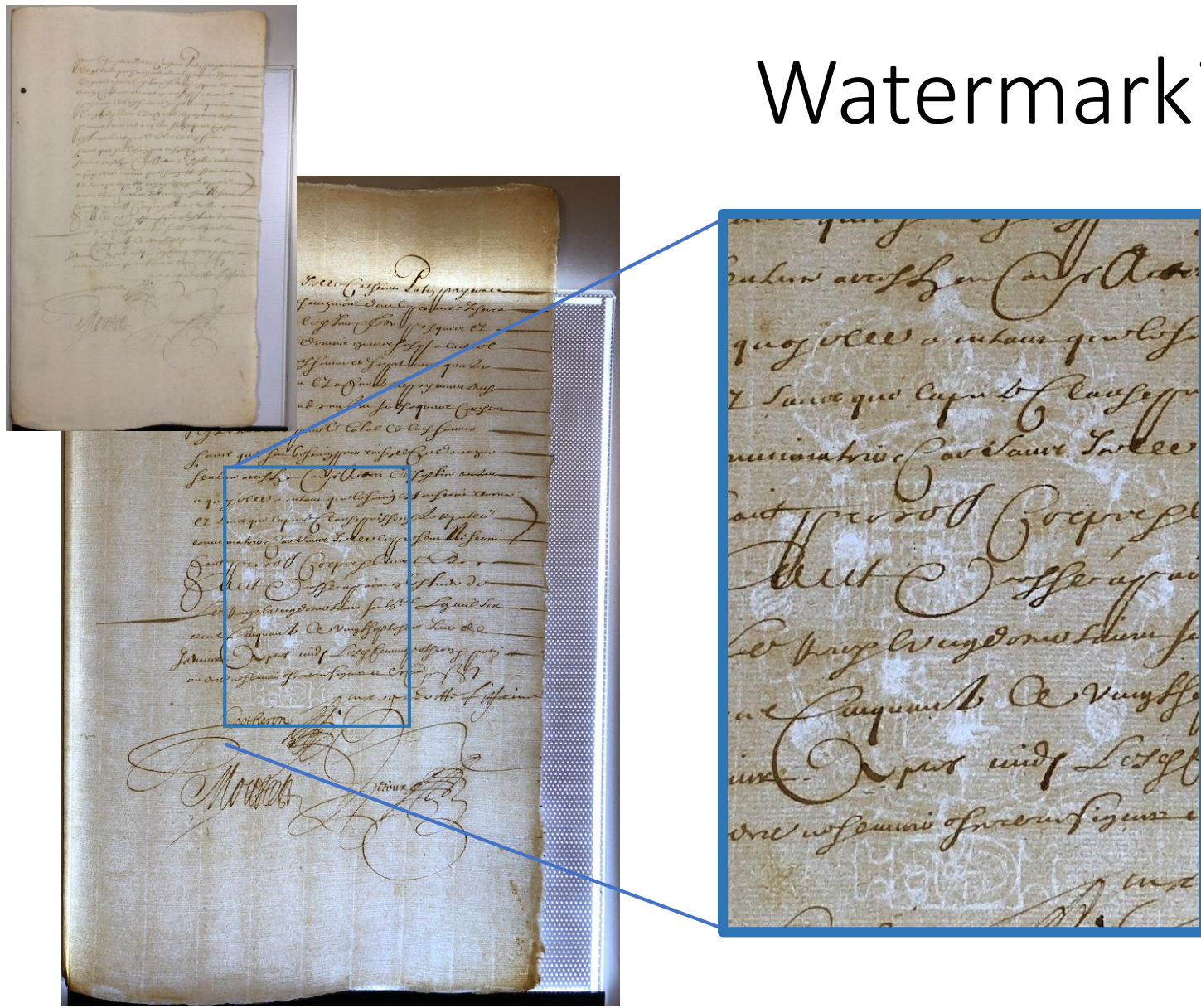
Spyros Gidaris³

Marc Smith²

Olivier Poncet²

Mathieu Aubry¹

Watermark?

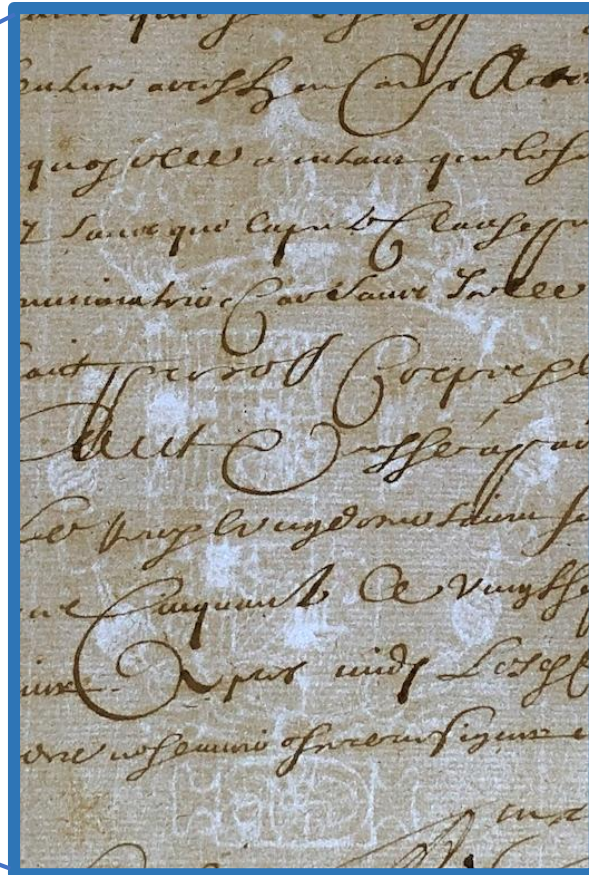


Looking through old paper

Watermark?

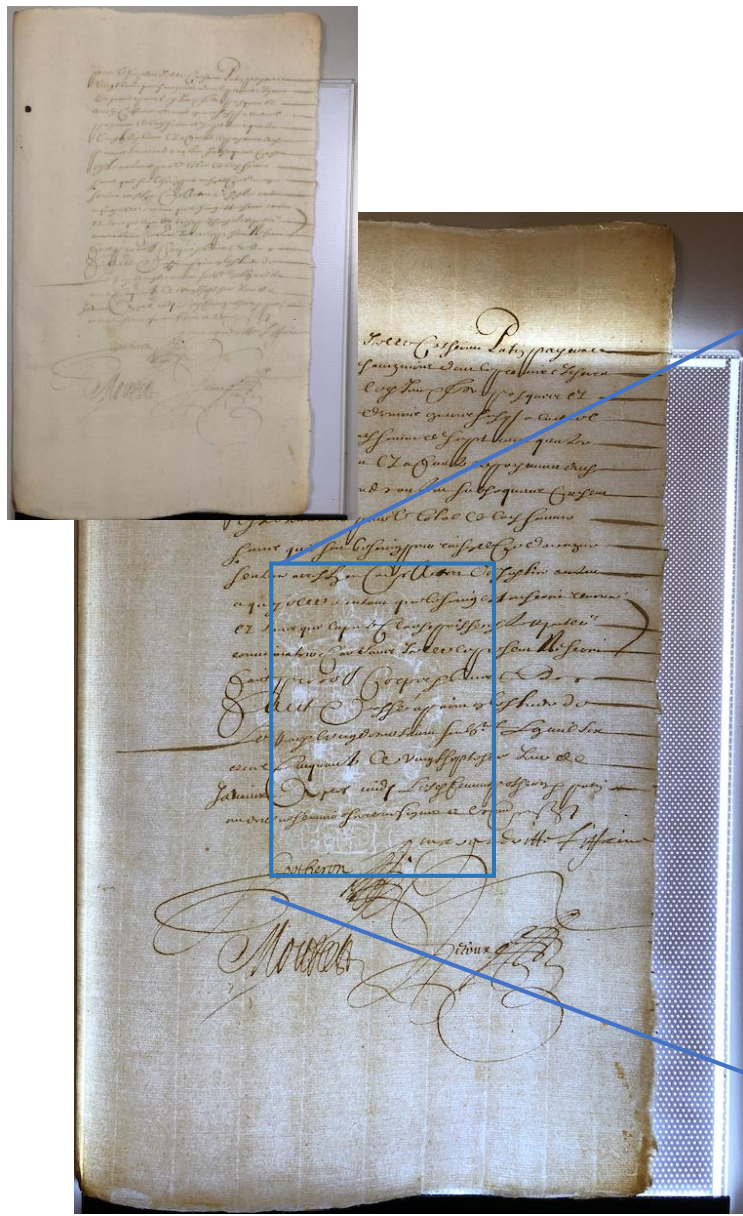


Mold for paper fabrication

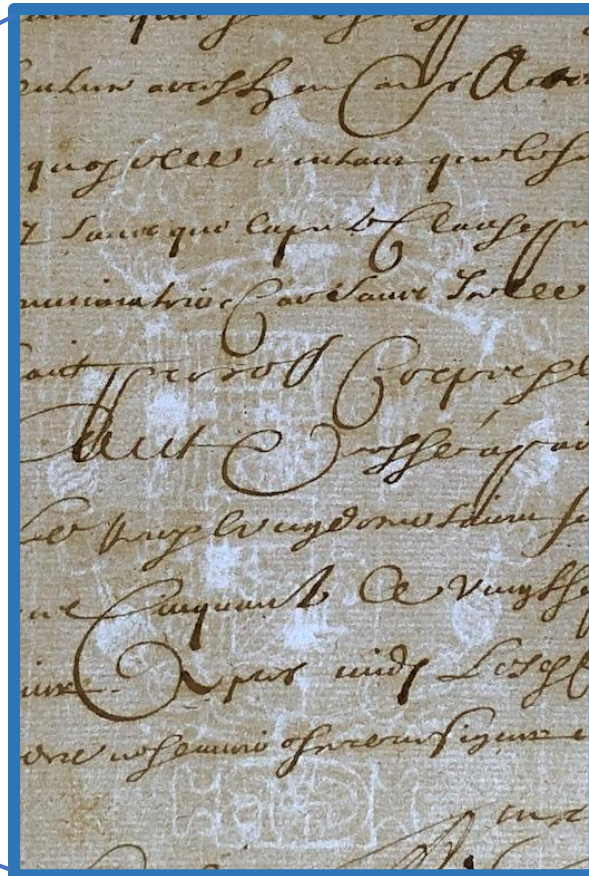


Looking through old paper

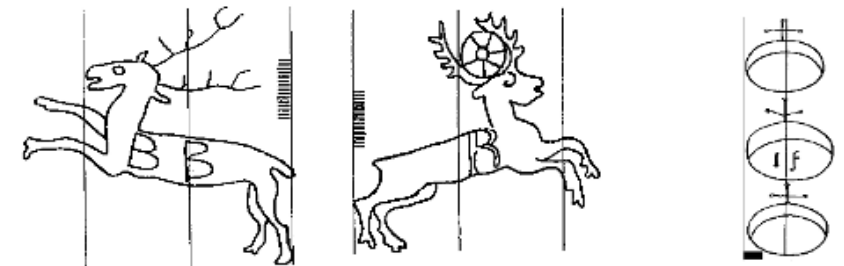
Watermark?



Looking through old paper



Mold for paper fabrication



Watermark Catalog

Contributions

- A large public **dataset**
- Consistency based **local matching** score
- Weakly supervised **feature fine-tuning**

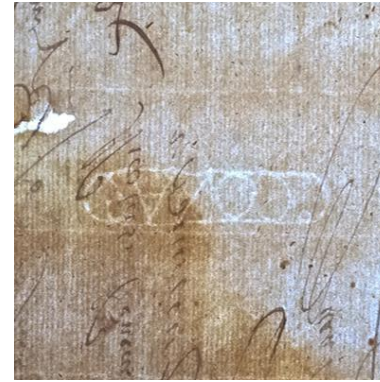
Contributions

- A large public **dataset**
- Consistency based local matching score
- Weakly supervised feature fine-tuning

Classification dataset

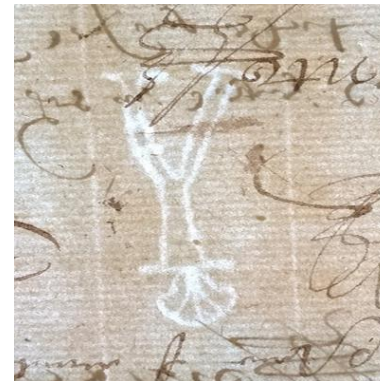
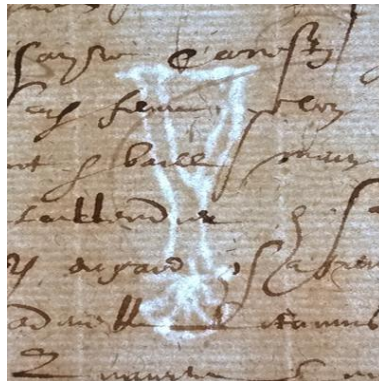
100 classes: **50** images / class for training and **10** images / class for validation.

Class 1



...

Class 2



...

One-shot dataset

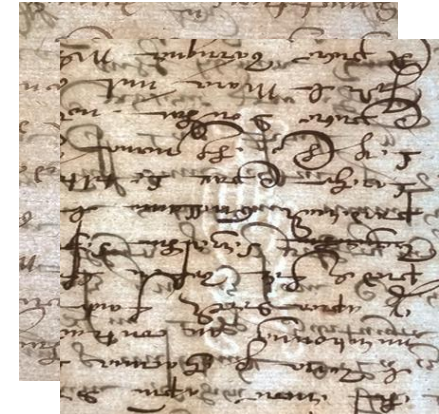
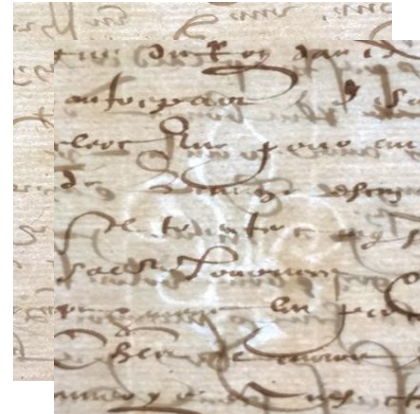
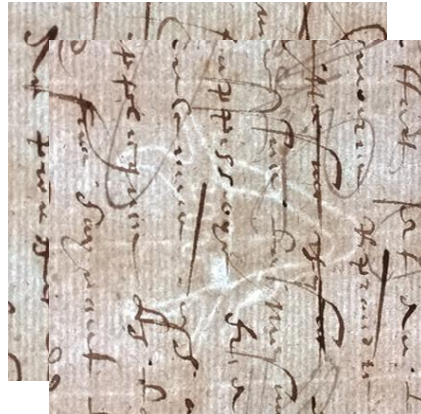
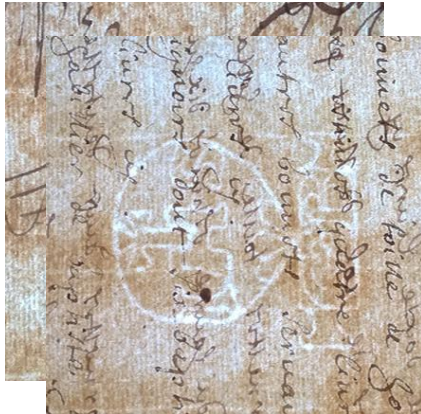
100 other classes with **3** images per class:

1 reference without any text + **2** cluttered query photographs

References



Queries



Cross-domain dataset

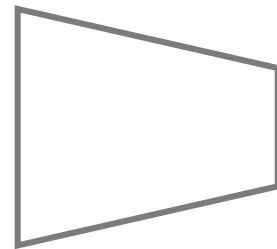
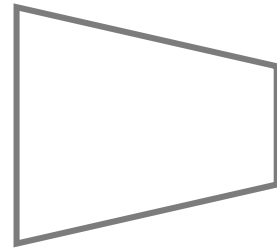
- **140** training classes with **1** drawing as reference + **1-7** query photographs.
- **100** testing classes with **1** drawing as reference + **2** query photographs.



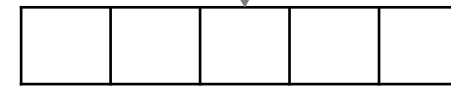
Our contribution

- A large public dataset
- Consistency based **local matching** score
- Weakly supervised feature fine-tuning

Global feature matching

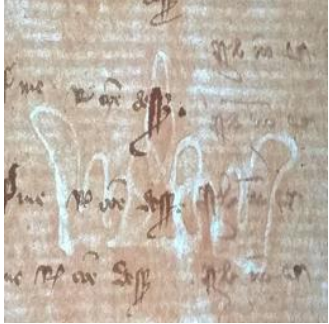
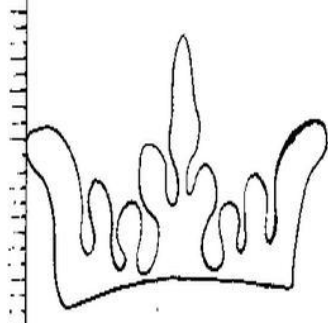
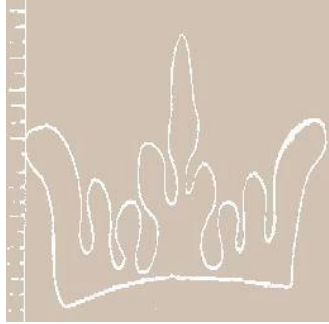


ResNet-18
trained on
classification dataset

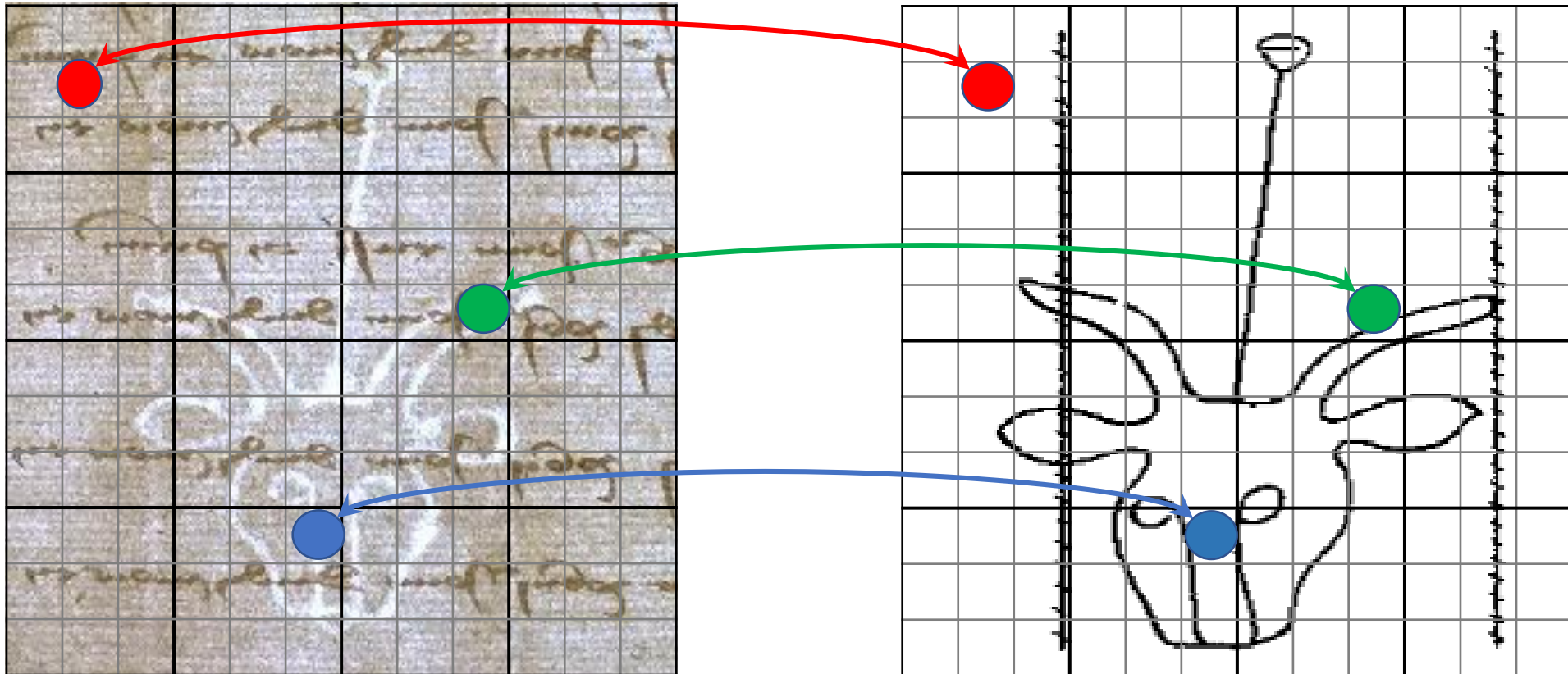


Cosine similarity
between
global pooled features

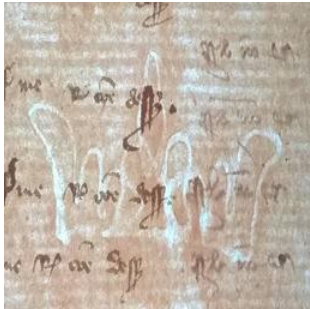
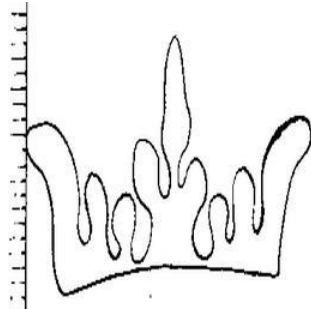
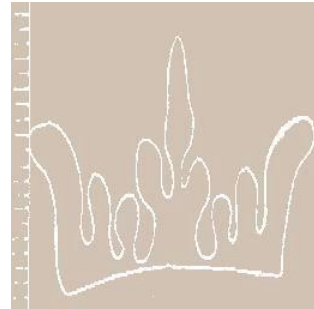
Performances

Dataset, Ref.	One-shot Real	Cross-domain Drawing	Cross-domain Synthetic
Ref.			
Global (Average Pool)	69%	4 %	12%

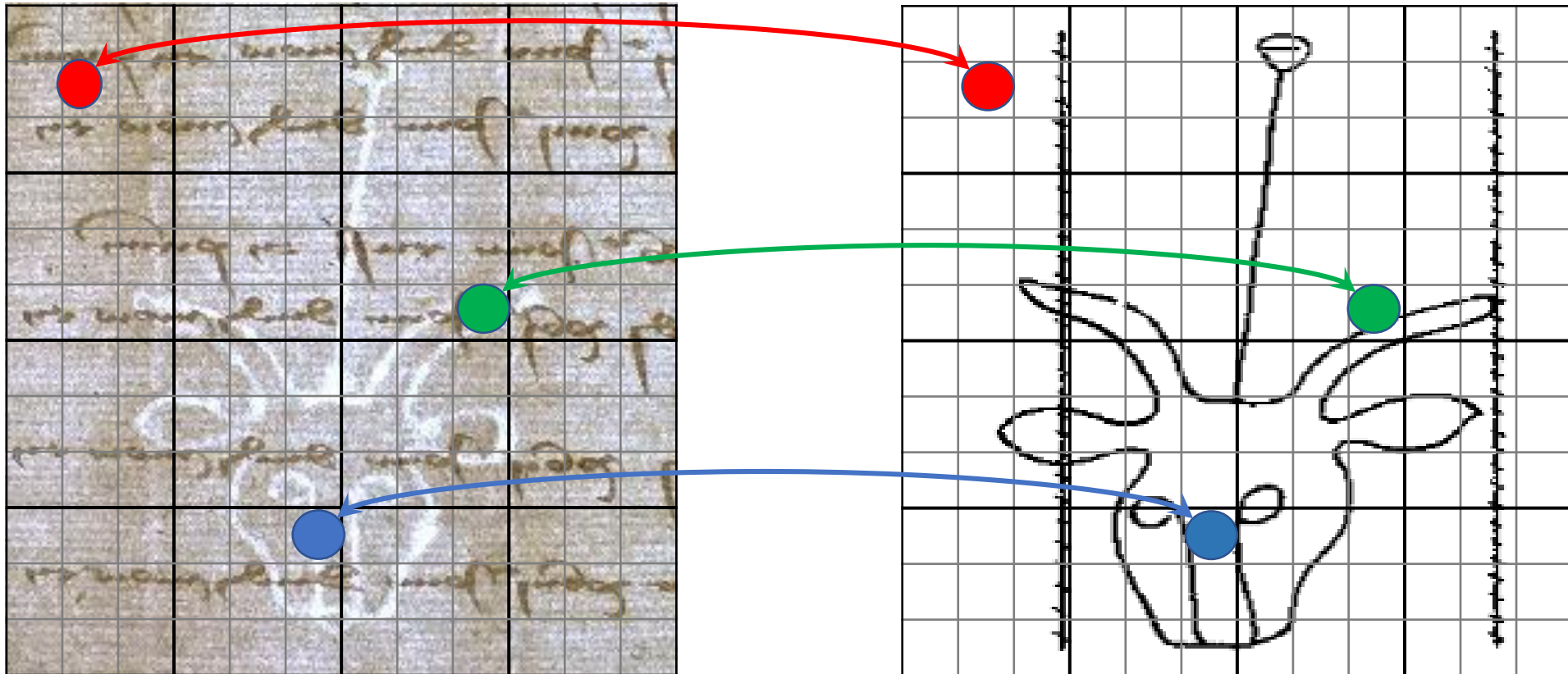
Local feature matching (Conv4)



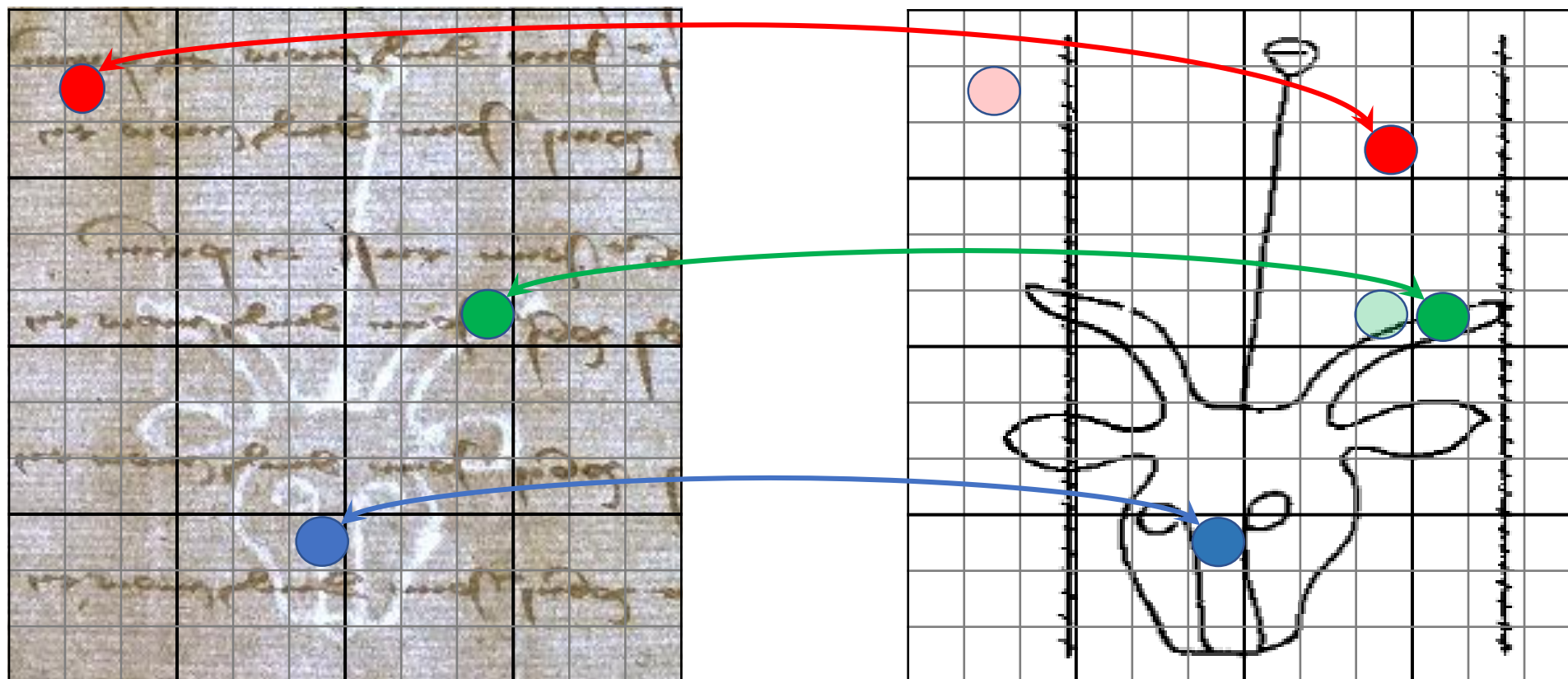
Performances

Dataset, Ref.	One-shot Real	Cross-domain Drawing	Cross-domain Synthetic
Ref.			
Global (Average Pool)	69%	4 %	12%
Local (conv4)	75%	56 %	65 %

Local feature matching (Conv4)



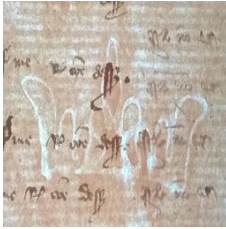
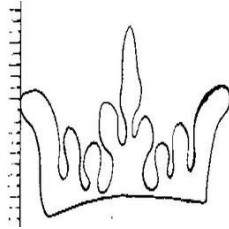
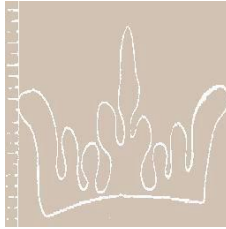
Local matching score



$$S(I_1, I_2) = \sum_{i \in \mathcal{I}} \underbrace{e^{-\frac{\|x_1^i - x_2^i\|^2}{2\sigma^2}}}_{\text{Spatial Consistency}} \underbrace{\cos(f_1^i, f_2^i)}_{\text{Feature Similarity}}$$

Local matching score

See paper for comparisons to few-shot methods

Dataset, Ref.	One-shot Real	Cross-domain Drawing	Cross-domain Synthetic
Ref.			
Global (Average Pool)	69%	4 %	12%
Local (conv4)	75%	56 %	65 %
Local matching (Ours)	90%	66%	72%

Contributions

- A large public dataset
- Consistency based local matching score
- Weakly supervised **feature fine-tuning**

Local matching-based domain adaptation

$$\mathcal{L}(\theta) = \frac{1}{\mathcal{N}} \sum_{n_1, n_2 \in \mathcal{N}} \max(1 - \lambda, \cos(f_\theta(n_1), f_\theta(n_2))) - \frac{1}{\mathcal{P}} \sum_{p_1, p_2 \in \mathcal{P}} \min(\lambda, \cos(f_\theta(p_1), f_\theta(p_2)))$$

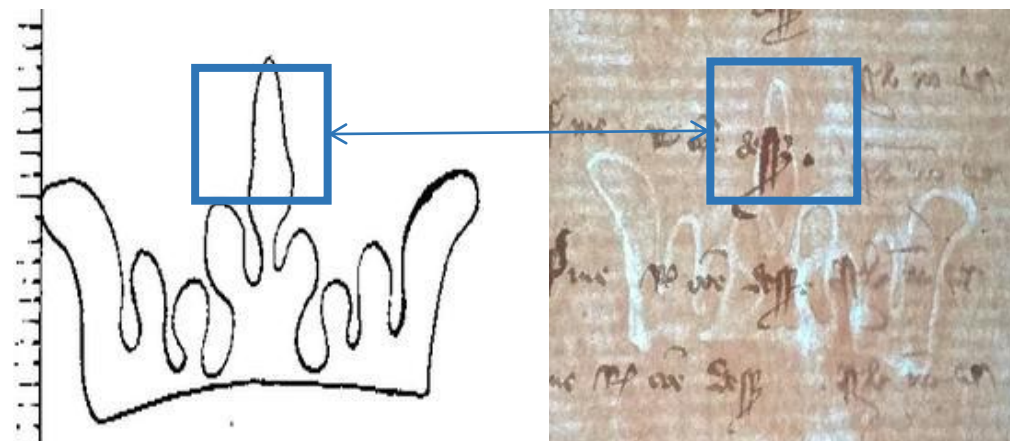
Negative Pairs:

- Different classes

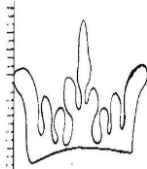


Positive Pairs:

- Same class
- Similar location

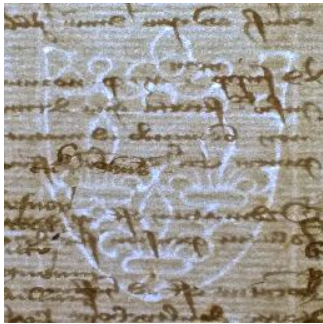


Fine-tuning results

Dataset, Ref.	Cross-domain Drawing	Cross-domain Synthetic
Ref.		
Global (Average Pool)	4 %	12%
Local (conv4)	56 %	65 %
Local matching (Ours)	66%	72%
Fine-tuning + Local matching (Ours)	75 %	83 %

Qualitative results

Query



Local (conv 4)

1st Match



2nd Match



3rd Match



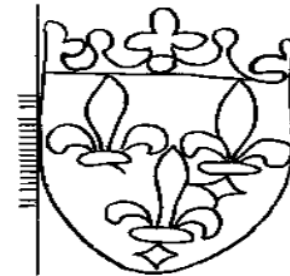
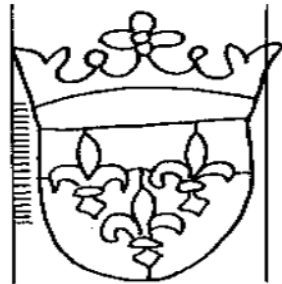
4th Match



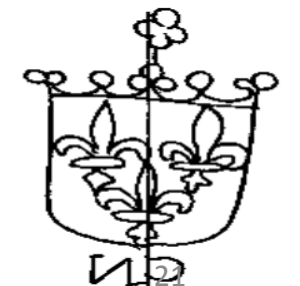
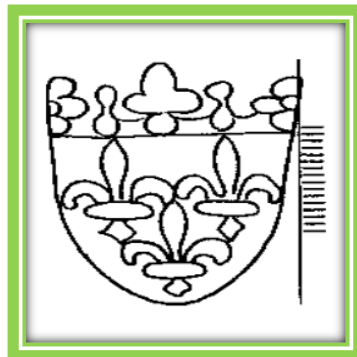
5th Match



Local
Matching



Local
Matching + F.T.



Summary

- Dataset and a consistency based approach for historical watermark recognition
- Code/Visual results: <http://imagine.enpc.fr/~shenx/Watermark/>
- Web application : <https://filigranes.inria.fr/>



A Web Application for Watermark Recognition
2020 Journal of Data Mining & Digital Humanities
Oumayma Bounou, Tom Monnier, Ilaria Pastrolin, Xi Shen, Christine Benevent, Marie-Françoise Limon-Bonnet, François Bougard, Mathieu Aubry, Marc Smith, Olivier Poncet, Pierre-Guillaume Raverdy