gipsa-lab

GRAFT: Unsupervised Adaptation to Resizing for Detection of Image Manipulation

Advisors: François Cayre et Kai Wang
PhD: Ludovic Darmet

April 11, 2019

GIPSA-lab, Grenoble

Table of contents

Image Forensics

The field

- Related to security of digital images and passive image authentication

1. Detecting a targeted and specific modification
2. Universal detector (same model for every modification)

The field

- Related to security of digital images and passive image authentication

  1. Detecting a targeted and specific modification
  2. Universal detector (same model for every modification)

- This work is focused on 2.

  a. Explicit modeling and spot inconsistencies with the model
  b. Implicit modeling and extract statistics of this model as features
  c. Deep-learning approaches

The field

⪢ Related to security of digital images and passive image authentication

1. Detecting a targeted and specific modification
2. Universal detector (same model for every modification)

⪢ This work is focused on 2.

a. Explicit modeling and spot inconsistencies with the model
b. Implicit modeling and extract statistics of this model as features
c. Deep-learning approaches

⪢ This work is focused on a. and b. as they are the only methods able to deal with very small patches

Objectives

- Subject definition: "Image Forensics: an image-model based approach"
- Local and spatially accurate detection
- Universal detector in an unsupervised or weakly supervised setting
- Participation in DEFALS challenge with REVEAL team supervised by Patrick BAS

Problem statement

Introduction to the problem

- Image Forensics in the Wild: usually images are pre-processed (storage, enhancement, etc) then retrieved and finally altered

Introduction to the problem

- Image Forensics in the Wild: usually images are pre-processed (storage, enhancement, etc) then retrieved and finally altered
- Most training databases in the literature are not pre-processed
- Pre-processing could introduce statistical differences between training database and testing sets in the wild (or between dataset)

Introduction to the problem

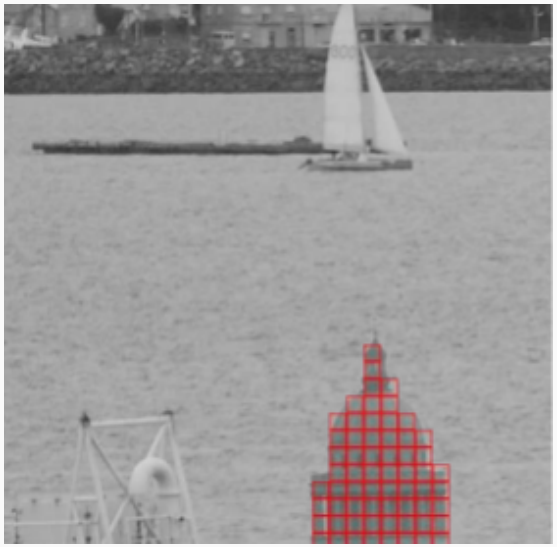
- Image Forensics in the Wild: usually images are pre-processed (storage, enhancement, etc) then retrieved and finally altered
- Most training databases in the literature are not pre-processed
- Pre-processing could introduce statistical differences between training database and testing sets in the wild (or between dataset)
- We are focused here on re-sizing as pre-processing (up and down-scaling) operation, with bi-cubic interpolation (harder scenario)
- Source = training data (without pre-processing) with labels and Target = test data (with pre-processing) without labels

Introduction to the problem

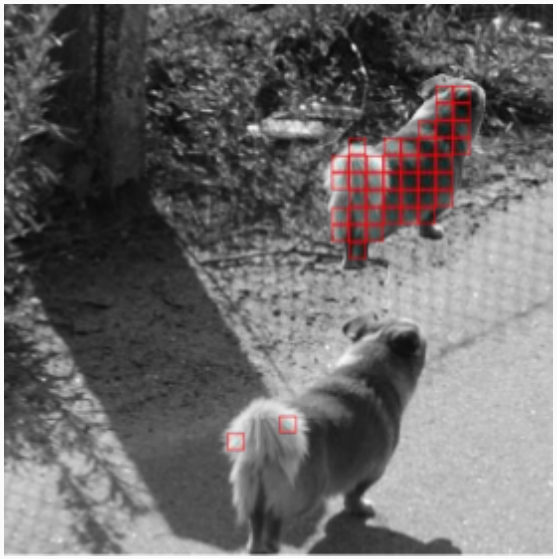
- Working on patches (8x8) in order to obtain good spatial accuracy (localization of the modification)
- Focused on elementary image processing operations as we assume they will be present in more complex tampering (to cover fingerprints)

Table 1: List of modifications applied on full-sized images.

ORI	No image modification
GF	Gaussian filtering with 3×3 kernel and $\sigma = 0.5$
MF	Median filtering with 3×3 kernel
USM	Unsharp masking with window size 3×3, and parameter 0.5 for the Laplacian filter to generate the sharpening filter kernel
WGN	White Gaussian noise addition with $\sigma = 2$
JPEG	JPEG compression with $Q = 90$



(a) JPEG compression



(b) Re-sampling

Figure 1: Copy-move and splicing detected because of elementary operations

Detector: Gaussian Mixture Models

- Based on the method of Wei FAN [1]
- Model patches as Mixture of Gaussian:

$$\sum_{k}^{N} \pi_k \mathcal{N}(x_i | \mu_k, \Sigma_k)$$

- ONE GMM is trained on ONLY original patches and ANOTHER ONE on ONLY modified patches

Detector: Gaussian Mixture Models

- Based on the method of Wei FAN [1]
- Model patches as Mixture of Gaussian:

$$\sum_{k}^{N} \pi_k \mathcal{N}(x_i | \mu_k, \Sigma_k)$$

- ONE GMM is trained on ONLY original patches and ANOTHER ONE on ONLY modified patches
- During test phase, a ratio of likelihood is then computed to predict the class of a patch x_i:

$$r(x_i) = \frac{\mathcal{L}_{GMM_{manip}}(x_i)}{\mathcal{L}_{GMM_{ori}}(x_i)}.$$

If $r(x_i) > 1$ then patch is predicted as manipulated and original otherwise.

Drops in performances: GMMs

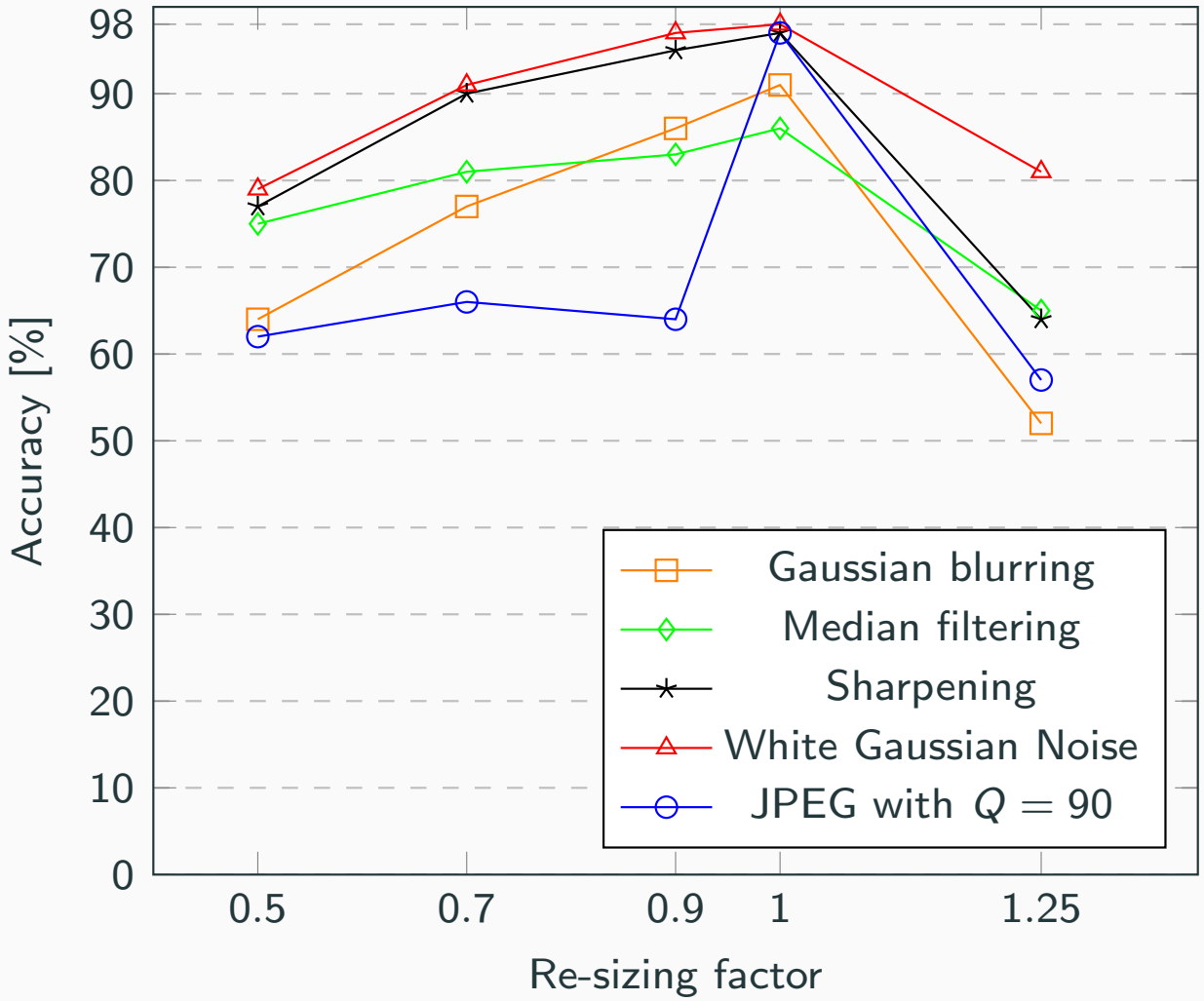


Figure 2: Accuracy of GMM based method under different resizing factors (bi-cubic interpolation) for several manipulations.

Statistical Test: Maximum Mean Discrepancy (MMD) [2]

- MMD formulation:

$$MMD[\mathcal{F}, s, t] = \sup_{f \in \mathcal{F}} \left(\mathop{\mathbb{E}}_{x \sim s}[f(x)] - \mathop{\mathbb{E}}_{y \sim t}[f(y)] \right)$$

Statistical Test: Maximum Mean Discrepancy (MMD) [2]

- MMD formulation:

$$MMD[\mathcal{F}, s, t] = \sup_{f \in \mathcal{F}} \left(\mathop{\mathbb{E}}_{x \sim s}[f(x)] - \mathop{\mathbb{E}}_{y \sim t}[f(y)] \right)$$

- Permutation (randomization test):

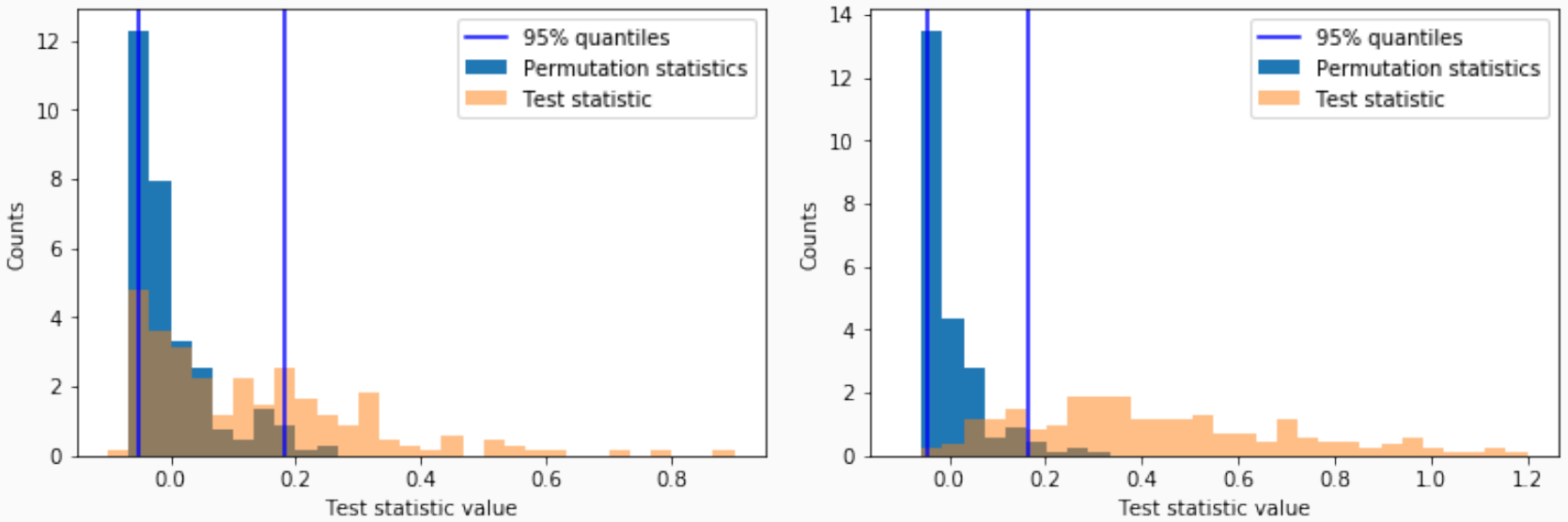


Figure 3: Histograms of MMD distances between training and testing data: (a) on original-sized testing images, and (b) on testing images subject to resizing pre-processing (bi-cubic interpolation with factor of 0.53).

GMM Resizing Adaptation by Fine-Tunning: GRAFT

Model Adaptation

- $\mathcal{C}_1$ (covariance matrices of the two GMMs) and $\mathcal{C}_2$: empirical covariance matrix on testing set $\rightarrow$ find a transformation of $\mathcal{C}_1$ to bring it closer to $\mathcal{C}_2$

Model Adaptation

- $\mathcal{C}_1$ (covariance matrices of the two GMMs) and $\mathcal{C}_2$: empirical covariance matrix on testing set $\rightarrow$ find a transformation of $\mathcal{C}_1$ to bring it closer to $\mathcal{C}_2$
- Similar problems studied in BCI: they are using covariance matrices as features and independent experiments introduce changes in domains
- RPA method from Pedro et al. [3] $\rightarrow$ perform translation, stretching and rotation to bring closer two sets of covariance matrices, near identity to train a classifier on these adapted sets

Model Adaptation

- $\mathcal{C}_1$ (covariance matrices of the two GMMs) and $\mathcal{C}_2$: empirical covariance matrix on testing set $\rightarrow$ find a transformation of $\mathcal{C}_1$ to bring it closer to $\mathcal{C}_2$
- Similar problems studied in BCI: they are using covariance matrices as features and independent experiments introduce changes in domains
- RPA method from Pedro et al. [3] $\rightarrow$ perform translation, stretching and rotation to bring closer two sets of covariance matrices, near identity to train a classifier on these adapted sets
- Differences with our framework: covariance matrices are parameters and not features and should not be centered near identity
- Adding of an additional translation at the end of the procedure so $\mathcal{C}_1^{adp}$ have approximately $\mathcal{C}_2$ as geometric mean

Model Adaptation

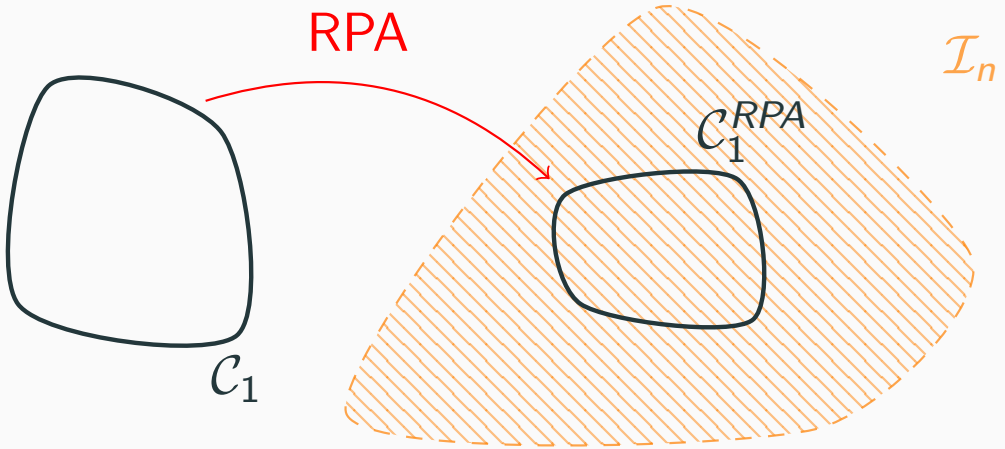


Figure 4: Adding a translation

Model Adaptation

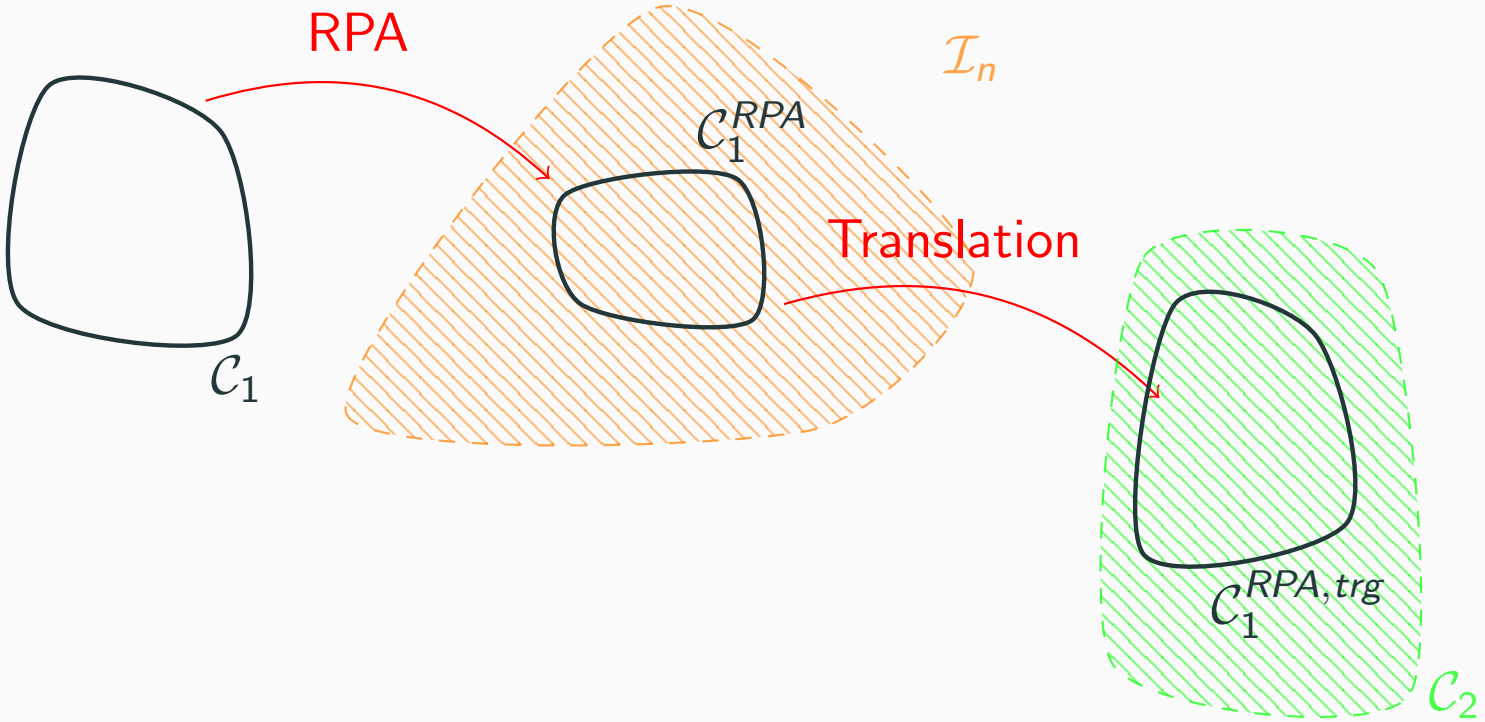


Figure 4: Adding a translation

Model Adaptation

- Problem: GMMs are now too much descriptive and not discriminative enough $\rightarrow$ need to bring back some discriminative power and log-likelihood maximization of the original model
- Additional step to improve discriminative and descriptive power $\rightarrow$ interpolation

$$\begin{aligned} \mathcal{C}_{adp}^{ori} &= \mathcal{C}_1^{trg,ori} * (1-\alpha_1) + \mathcal{C}_1^{RPA,ori} * \alpha_1, \\ \mathcal{C}_{adp}^{mnp} &= \mathcal{C}_1^{trg,mnp} * (1-\alpha_2) + \mathcal{C}_1^{RPA,mnp} * \alpha_2. \end{aligned} \tag{1}$$

- Interpolations are performed separately
- Log-likelihood maximization for GMMs and not only geometric considerations on covariance matrices

Model Adaptation

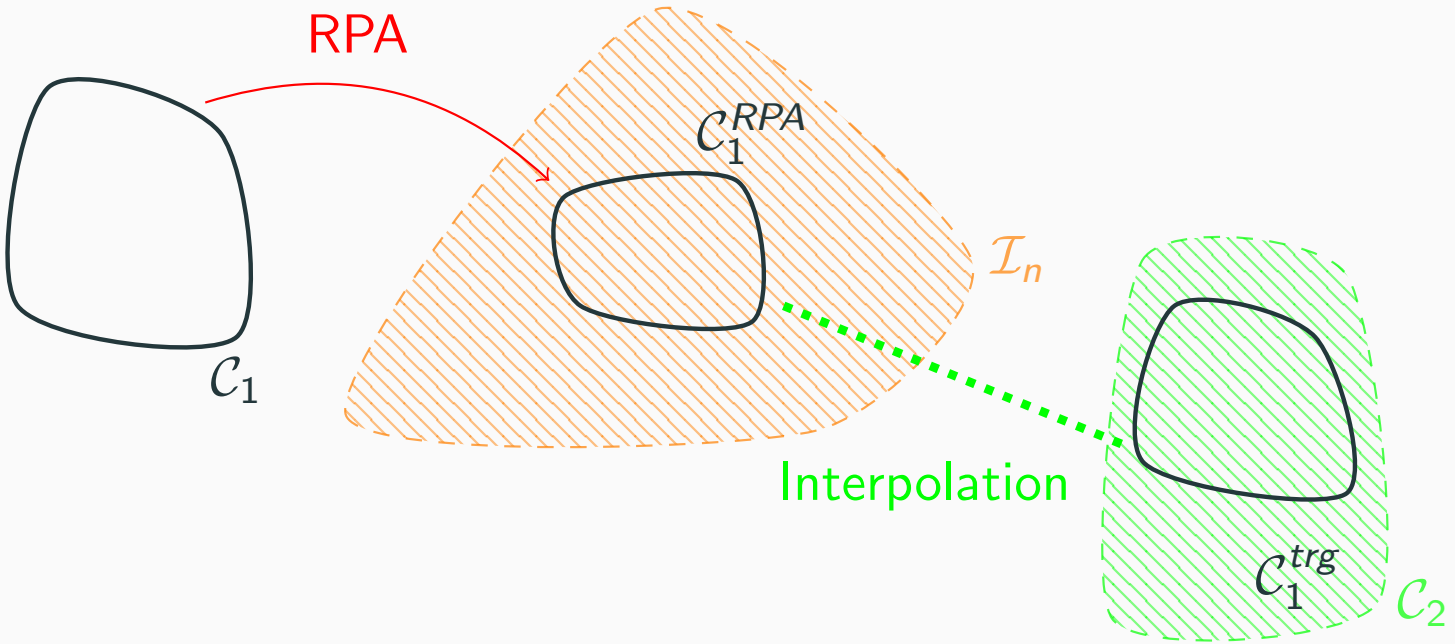


Figure 5: Transformation and interpolation in GRAFT.

Coefficients of interpolation

- GMMs without adaptation still have discriminative power (accuracy is not dropping to 50%) $\rightarrow$ it is still possible to identify almost-surely some modified patches

Coefficients of interpolation

- GMMs without adaptation still have discriminative power (accuracy is not dropping to 50%) $\rightarrow$ it is still possible to identify almost-surely some modified patches
- "pseudo-labeled" patches from test set (resized one): 15% + 15% of the samples with ratio of log-likelihood farther from 1
- Excluding the 5% + 5% extreme values as these examples are probably too close from training set (without re-sizing)

Coefficients of interpolation

- GMMs without adaptation still have discriminative power (accuracy is not dropping to 50%) $\rightarrow$ it is still possible to identify almost-surely some modified patches
- "pseudo-labeled" patches from test set (resized one): 15% + 15% of the samples with ratio of log-likelihood farther from 1
- Excluding the 5% + 5% extreme values as these examples are probably too close from training set (without re-sizing)
- Algorithm is robust against these percentages values
- Sanity check for hypothesis testing and indeed 95% of accuracy on these samples

Coefficients of interpolation

- GMMs without adaptation still have discriminative power (accuracy is not dropping to 50%) $\rightarrow$ it is still possible to identify almost-surely some modified patches
- "pseudo-labeled" patches from test set (resized one): 15% + 15% of the samples with ratio of log-likelihood farther from 1
- Excluding the 5% + 5% extreme values as these examples are probably too close from training set (without re-sizing)
- Algorithm is robust against these percentages values
- Sanity check for hypothesis testing and indeed 95% of accuracy on these samples
- Interpolation coefficients maximize log-likelihood on these samples

Results

	GF	MF	USM	WGN	JPEG
Resizing $\times 0.5$ (without adaptation)	**64**	75	77	79	**62**
Resizing $\times 0.5$ (retraining from scratch with 10%)	**77** (+13)	79 (+4)	82 (+5)	84 (+5)	**81** (+19)
Resizing $\times 0.5$ (semi-supervised with 10%)	**79** (+15)	80 (+5)	89 (+12)	89 (+10)	**76** (+14)
Resizing $\times 0.5$ (unsupervised, GRAFT)	**79** (+15)	79 (+4)	88 (+11)	89 (+10)	**73** (+11)

Conclusion and perspectives

Conclusion

- We introduce new concerns and a new problem and provide a solution
- We provide a real shortcut (complexity and time)
- Code on: https://forge.uvolante.org/darmet/GRAFT

Perspectives

- Semi-supervised extension of GRAFT method
- Feature Adaptation is pretty complex ! Need to spend more time on this
- Submitted to IEEE Transactions on Information Forensics and Security, waiting for the reviews

Questions?

References i

W. Fan, K. Wang, and F. Cayre.
General-purpose image forensics using patch likelihood under image statistical models.
In *2015 IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 1–6, Nov 2015.

A. Gretton, K. M. Borgwardt, M. Rasch, B. Scholköpf, and A. J. Smola.
A kernel approach to comparing distributions.
In *Proceedings of the National Conference on Artificial Intelligence*, pages 1637–1641, 2007.

References ii


P. L. C. Rodrigues, C. Jutten, and M. Congedo.
Riemannian procrustes analysis: transfer learning for brain-computer interfaces.
IEEE Transactions on Biomedical Engineering, pages 1–12, 2019 (doi: 10.1109/TBME.2018.2889705).