Using Set Covering to Generate Databases for Holistic Steganalysis

Rony Abecidan, Vincent Itier, Jérémie Boulanger, Patrick Bas, Tomáš Pevný

Using Set Covering to Generate Databases for Holistic Steganalysis

Rony Abecidan
Univ. Lille, CNRS, Centrale Lille,
UMR 9189 CRIStAL,
F-59000 Lille, France
Email: rony.abecidan@univ-lille.fr

Vincent Itier
IMT Nord Europe, Institut Mines-Télécom,
Centre for Digital Systems, Univ. Lille, CNRS, Centrale Lille,
UMR 9189 CRIStAL,F-59000 Lille, France
Email: vincent.itier@imt-nord-europe.fr

Jérémie Boulanger
Univ. Lille, CNRS, Centrale Lille,
UMR 9189 CRIStAL,
F-59000 Lille, France
Email: jeremie.boulanger@univ-lille.fr

Patrick Bas
Univ. Lille, CNRS, Centrale Lille,
UMR 9189 CRIStAL,
F-59000 Lille, France
Email: patrick.bas@cnrs.fr

Tomáš Pevný
Department of Computers and Engineering
Czech Technical University
Prague, Czech Republic
Email: pevnak@protonmail.ch

***Abstract*—Within an operational framework, covers used by a steganographer are likely to come from different sensors and different processing pipelines than the ones used by researchers for training their steganalysis models. Thus, a performance gap is unavoidable when it comes to out-of-distributions covers, an extremely frequent scenario called Cover Source Mismatch (CSM). Here, we explore a grid of processing pipelines to study the origins of CSM, to better understand it, and to better tackle it. A set-covering greedy algorithm is used to select representative pipelines minimizing the maximum regret between the representative and the pipelines within the set. Our main contribution is a methodology for generating relevant bases able to tackle operational CSM. Experimental validation highlights that, for a given number of training samples, our set covering selection is a better strategy than selecting random pipelines or using all the available pipelines. Our analysis also shows that parameters as denoising, sharpening, and downsampling are very important to foster diversity. Finally, different benchmarks for classical and wild databases show the good generalization property of the extracted databases. Additional resources are available at github.com/RonyAbecidan/HolisticSteganalysisWithSetCovering.**

I. Introduction

Cover-source mismatch (a.k.a. CSM) is well-known in modern steganalysis [1] [2] [3]. In the literature, steganalysis models are commonly trained on controlled cover distributions coming from BOSSBASE [4] or ALASKABASE [5] for instance. Meanwhile, in operational steganalysis, it is rarely possible to guess the distributions to which the images belong. Cover distributions, also called cover sources, present a lot of diversity because of several factors such as the image acquisition device (camera, mobile phones, scanner, *etc.*), the quality of the sensor, the settings of this device (ISO, zoom, aperture, shutter time, etc.), the content captured (inside, outside, luminosity, level of details, *etc.*), the post-processing step (sharpening, denoising, white balance, gamma correction, cropping, *etc.*), and also the usual compression step (8-bit conversion, JPEG compression, *etc.*). These development pipelines correspond to a set of transformations associated with parameters impacting the statistics of the developed image before a potential embedding. In this context, a cover source can be seen as a mix of two distributions: the content and the noise. Previous studies have shown that the mismatch between two cover sources is generally much more fostered by the noise distribution than by that of the content [1], [6].

More generally, it is shown in [1] that the processing pipeline is the main perpetrator of CSM. Table I shows how far two sources can mismatch if they differ only by one parameter value in their processing pipelines. This set of transformations is commonly used for aesthetic and compression purposes. In the steganalysis case, it is impacting significantly the noise distribution while keeping the semantics pristine. Even if machine learning schemes are effective for steganalysis tasks, they are very often extremely sensitive to the very nature of the analyzed signal. This is why CSM might occur.

To address this issue, several approaches have been designed. On one hand, atomistic approaches assume that the distribution of the test covers can be reproduced. It requires to have access to covers close to the test ones in terms of distribution. Then, it is possible to create a batch of classifiers trained on specific sources and use them accordingly. For example, in [3], the authors propose to pick up the classifier

Train / Eval	**No Denoising**	**Max Denoising**
No Denoising	5	77
Max Denoising	48	0.18

TABLE I
P_E matrix (Probability of Error in%) between two mismatching sources only differing by a denoising factor in their processing pipelines. Messages are embedded with UERD[7] with a payload of 1.5bpp. Training with a linear classifier on DCTr features [8]. Pipelines 5 and 167 in our directory.

WIFS'2022, December, 12-16, 2022, Shanghai, China. XXX-X-XXXX-XXXX-X/XX/$XX.00

trained on the closest distribution to the test one. Whereas, in [2], the authors use an ensemble classifier to face the CSM problem.

On the other hand, there are holistic approaches for which a mixture of relevant cover distributions should help to cope with the heterogeneity of the test set [3]. The purpose of these methods is to bring a lot of information and diversity to the dataset. For example, to be the most representative possible in terms of content and noise, pictures from ALASKABASE [5] are coming from 479 different cameras with various ISO ranging from 16 to 51200.

Generally, the holistic approach is more suitable because it does not require too much assumption about the cover distribution. In this work, we put ourselves in a realistic scenario where we do not know anything about the distribution of the test covers. Hence, we propose a holistic framework to solve this problem based on an extensive study of processing pipelines. Our goal is two-fold:

- Investigate the role of processing pipelines in CSM.
- Derive from our results a framework to build holistic training Cover databases.

Motivated by these facts, this paper is the first attempt to address the CSM issue by proposing a framework for the generation of databases for holistic steganalysis. Our work has the following contributions:

- We show that a wise selection of development pipelines allows us to generalize the performances of a steganalysis model on several SOTA databases with fewer training samples.
- We also provide a simple yet efficient greedy algorithm for the selection.

The rest of the paper is organized as follows: In Section II, we formalize our objective and formulate the source selection problem as a set covering problem. Then, in Section III, we present some experiments where we want to better understand the origins of CSM meanwhile finding interesting databases for our battle against CSM. Afterwards, in Section IV, we test the potential of these bases in a state-of-the-art framework in steganalysis. Finaly, Section V concludes this work.

II. Source screening with set-covering

A. Formalization

Following [9], we consider that a processing pipeline is entirely defined by a vector $\omega \in \Omega$ which contains all the pipeline parameters (demosaicking algorithm, denoising coefficient, JPEG quality factor, etc.). In the steganalysis context, we also introduce a parameter γ representing steganographer choices, notably embedding strategy and payload. The state of the art for this task essentially lies on machine learning models that can be seen as predictors

$$f(x \mid \theta_{\omega,\gamma}) : \mathcal{X} \to \{cover, stego\}$$
$$x \mapsto y$$

where $\theta_{\omega,\gamma} \in \Theta$ contains all the parameters learnt with covers derived from the pipeline ω and potentially embedded following γ.

Algorithm 1 Greedy covering

Input: $\epsilon > 0$, the regret matrix R
  Let N be the width of R
  Let $Old\_covering$ and $Greedy$ be empty dictionaries
  for $i \in N$ **do**
    Initialize $P_{\epsilon,i}$ as the set of sources with which we get a regret of at most ϵ when we train on the i-th source
    Fill $Old\_covering[i]$ with $P_{\epsilon,i}$
  end for
  #At that stage we have an initial covering and we are going to refine it.
  while $Old\_covering$ is not empty **do**
    Fill $Greedy$ with the source k that is currently covering a maximum of other sources in $Old\_covering$: $Greedy[k] \leftarrow Old\_covering[k]$
    for $i \in N$ **do**
      Delete from $Old\_covering[i]$ the sources already covered by $Greedy[k]$: $Old\_covering[i] \leftarrow Old\_covering[i] \setminus Greedy[k]$
    end for
  end while
Output: $Greedy$

To assess CSM properly, two relevant metrics have been introduced in [1] and [9] :

- The Intrinsic Difficulty of a source that is, the probability of error P_E we obtain after training on images from this source and evaluating on images from this same source.

$$\mathbb{E}_{(x,y)\sim P((x,y)|\omega,\gamma)}(f(x \mid \theta_{\omega,\gamma}) \neq y) \quad (1)$$

- The Regret $R_{s,t}$ between two cover sources s and t defined as the difference between the P_E we obtain by training on s and evaluating on t and the Intrinsic Difficulty of t.

$$\begin{aligned}&\mathbb{E}_{(x,y)\sim P((x,y)|\omega_t,\gamma)}(f(x \mid \theta_{\omega_s,\gamma}) \neq y)\\ -&\mathbb{E}_{(x,y)\sim P((x,y)|\omega_t,\gamma)}(f(x \mid \theta_{\omega_t,\gamma}) \neq y)\end{aligned} \quad (2)$$

Through our study, we are looking for a basis of sources sufficiently rich in order to guarantee a generalization as great as possible on any source. We formalize this covering objective as follows : We want a basis $\Omega_B \subset \Omega^B$ s.t.

$$\forall \omega \in \Omega, \quad \exists \omega_b \in \Omega^B \quad \setminus \quad R_{\omega_b,\omega} \leq \epsilon \quad (3)$$

with ϵ being the maximum level of mismatch we are accepting in terms of Regret.

B. Extracting reference sources using set-covering

Given a steganalysis detector and a finite number of sources N, (3) can be rewritten as the famous set-covering problem [10]. For each $i \in N$, let $P_{\epsilon,i}$ be the set of all the sources with which we get a regret of at most ϵ when we train on the i-th source. Starting from the covering $C = \bigcup_{i\in N} P_{\epsilon,i} = N$, we precisely want to extract a minimal subset $N_\epsilon \subset N$ such that

$$C = \bigcup_{i\in N_\epsilon} P_{\epsilon,i} = N$$

This problem is NP-complete and therefore, finding the optimal covering N_ϵ^* is not an easy task. However, a greedy algorithm with a theoretical upper bound exists [10]. The pseudo-code of this algorithm is presented in Alg. 1. The

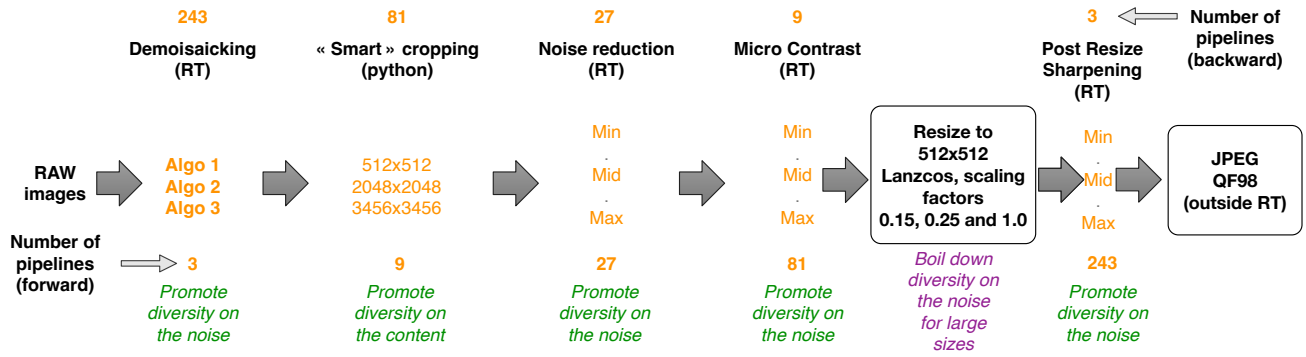

Fig. 1. Generation of 3^5 pipelines. Note that the different scaling factors are only due to crops of different sizes.

algorithm first selects the pipeline with the largest cover-set following the constraint on the regret, then append other pipelines not yet covered using a greedy strategy.

III. Practical Selection of Source Representatives

A. Experimental protocol

For the experiments presented here, we extract 1115 RAW Images of size 5184×3456 from ALASKA coming from the camera CANON-EOS-100D and captured with ISO > 1000. The choice of using one given camera model is not innocuous. We want to precisely study the role of processing pipelines in CSM and hence we are trying to avoid other important factors such as the sensor quality. Moreover, because RAW images associated with high ISO are noisier than low ISO images, this increases the contrast between the different pipelines.

To simulate processing pipelines, we are using RawTherapee[1], an open-source software that handles a large range of processing operations ordering from demosaicking to jpeg compression.

The impact of jpeg compressions on steganalysis is already well-known in the literature [1]. Here, we propose to put at the end of our processing pipelines a JPEG compression with a quality factor of 98 to promote CSM.

Concerning the head of the pipeline, we cherry-picked five operations based on their inherent ability to promote diversity in content or noise and, we study 3^5 combinations of them to better understand the mismatch they could bring together. The details about covers generation is presented in Fig 1. Please note that we didn't use the regular cropping of RawTherapee considering that it leads to low-textured covers in practice. Instead, the "smart" cropping released on ALASKA2 Challenge [5] that is seeking crops with highly textured areas was preferred. It's also important to have in mind that this cropping operation may lead to downsampling according to its size. Furthermore, the final JPEG compression is done using Imagemagick[2] in order to fully control it. The details of the pipelines numbered from 0 to 242 are available in our github repo.

[1]rawtherapee.com
[2]imagemagick.org

From the beginning, the RAW images are randomly split into 50% train and 50% test. Afterward, the covers are generated and their embeddings are done using UERD [7] with a payload of 1.5bpp. To keep things simple and to save computational resources, we then train linear classifiers using DCTR features [8] from our sets of covers and stegos. This payload may seem high but, in practice, it enables obtaining cover sources with rather small Intrinsic Difficulty, ranging from 0% to 14%. Note that using a standard payload of 0.4bpnzac, it resulted in the generation of many cover sources with an important Intrinsic Difficulty, i.e. $\simeq 50\%$. In such a case, the detector is not learning anything and cannot generalize well on other sources, making us blind to the generalization potential of the training source.

Once the cover distributions are generated and the detectors trained on each of them, we study CSM using a regret matrix R where $R[s,t] = R_{s,t}$, $s,t \in N^2 = \{0,...,242\}^2$. Using R, we apply then the greedy algorithm presented in Alg. 1 with $\epsilon = 2\%, 4\%, ..., 10\%$. In Table II, you can find some information about the covering obtained and, in Table III the regret matrix of the 5 sources returned for $\epsilon = 10\%$.

ϵ	$\lvert N_\epsilon \rvert$	$\min \lvert N_\epsilon^* \rvert$
2%	26	6
4%	14	3
6%	11	3
8%	10	2
10%	5	1

TABLE II
Covering size obtained using Alg. 1 for different values of ϵ. The right column is the minimum possible size of the covering-set (see [10]).

As one can expect, the lower ϵ, the more sources we need to guarantee a regret less than ϵ for everyone. Moreover, we also observe an interesting property for the sources returned by the greedy algorithm. By inspecting Tab III, the extracted representatives are "complementary" since they are associated with important regrets regarding all the other selected sources. This feature is expected since the selected sources are representing sources of different types.

Train / Eval	**21**	**22**	**31**	**60**	**229**
21	0	20	12	30	30
22	5	0	6	25	37
31	21	10	0	32	33
60	25	26	20	0	30
229	19	16	29	32	0

TABLE III
REGRET MATRIX (IN %) BETWEEN THE 5 SOURCES IN $N_{10\%}$

B. Analysis: which parameter promotes heterogeneity?

Once the greedy algorithm returns a source covering, we assign labels to each of our 243 sources according to the representatives in N_ϵ that cover them with a radius ϵ. In doing so, we are disclosing clusters of sources that help us to realize which parameters of our processing pipelines are the most discriminating. For instance, for $N_{10\%}$ we end up with 5 clusters and the covering follows a kind of Pareto law: Two clusters are encompassing 94% of the sources. In Figures 2 and 3 we present visually their substance.

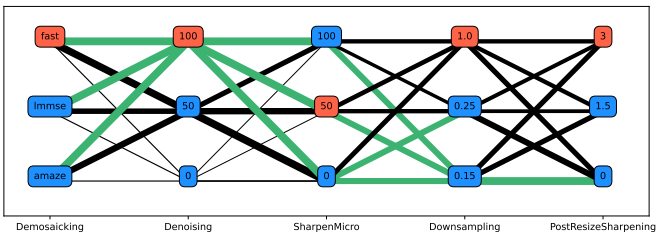

Fig. 2. Content of the sources covered by pipeline #229 in $N_{10\%}$. The red boxes are the parameters of #229 & $|C_{\#229}| = 159$ sources. The green links are the most represented links among the 9 possible at each stage.

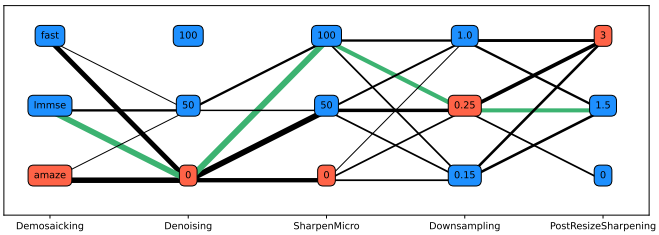

Fig. 3. Content of the sources covered by pipeline #60 in $N_{10\%}$.The red boxes are the parameters of #60 & $|C_{\#60}| = 69$ sources. The green links are the most represented links among the 9 possible at each stage.

From Fig. 2 we realize that pipeline #229 is issued from important denoising followed by an important adding of noise. Concretely, the sensor noise has been well cleaned and an artificial noise is then added based on what is remaining. Surprisingly this intriguing combination enables to cover around 66% of our arsenal of 243 pipelines with a 10% regret radius. The pipelines covered are mostly the ones issued from high denoising. Pipeline #60 is issued from a high cropping factor, hence followed by important downsampling and, a significant adding of noise. This time the sensor noise is not cleaned, we are extracting more content and we end the processing by adding artificial noise through sharpening. This mix enables us to cover around 28% of our sources, corresponding roughly to most noisy sources.

Without surprise, decreasing the maximal level of regret wanted allows one to blatantly reveal the most difficult sources to cover very precisely. The more we decrease ϵ, the more we are building sparse clusters, and the more we can observe what parameters make very singular and specific sources.

Using $\epsilon = 1\%$, we have 30 clusters, most of them with very few members. Instead of analyzing each cluster, we decide this time to train a random forest algorithm with an entropy criterion, trained to guess the clusters of each source according to their pipeline parameters. In Fig. 4, we present the importance of each parameter in the decisions of the model thanks to the Mean Decrease Impurity (MDI), which expresses the classification gain associated with splits of the forest for a given variable [11].

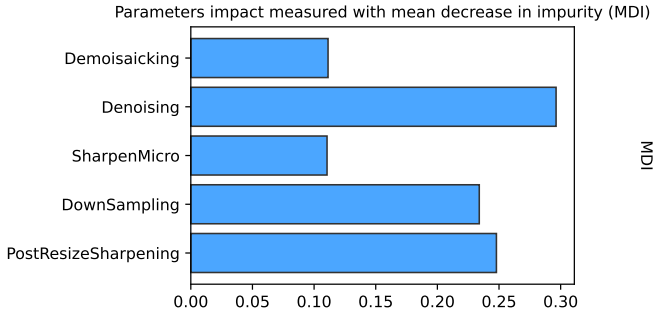

Fig. 4. Study of parameters importance in the creation of singular sources (using $N_{1\%}$)

Fig 4 shows that Denoising, PostResizeSharpening, and Downsampling are the parameters that play the biggest roles in the generation of singular sources. The Demosaicking and the SharpenMicro operations seem much less significant. These observations are consistent with the work of [1] and Figs 2 & 3.

IV. BENCHMARK ON DIFFERENT DATABASES

A. Reference framework

At this point, we extracted some sets of sources enabling us to generalize on the grid, up to predefined levels of regret. However, this ability to generalize may also depend on the embedding strategy and the detector used. Hence, we propose now to test the potential of our set of sources in a reference framework representing the current state-of-the-art.

The pre-trained J-UNIWARD ImageNet (JIN) [12] achieves currently state-of-the-art results on classical steganalysis databases for the J-UNIWARD [13] embedding with a payload between 0.4 and 0.6 bpnzac. Hence, we propose to embed our covers using J-UNIWARD with a payload of 0.5bpnzac and then fine-tune JIN for our experiments. To save computational resources while achieving a satisfying convergence, we fix a maximum cover training budget of 80K 256x256 images. The training, validation, and testing sets are split as follows 80/10/10. All the hyperparameters regarding the optimizer and the scheduler are fixed following [12]. This setup is the one adopted for all the upcoming experiments presented in this paper.

As mentioned previously, we systematically observed a kind of Pareto law for our different coverings. For $\epsilon \leq 10\%$,

Eval |Train	IMAGENET (860K)	RTBASE (64K)	ALASKA (64K)	BOSSBOWS (64K)	FLICKR (64K)	$RT_{2\%}$ (18K)	$RT_{4\%}$ (18K)	$RT_{6\%}$ (14K)	$RT_{8\%}$ (11K)	$RT_{10\%}$ (7K)
RTBASE (16K)	11	0	1	5	8	1	**0**	1	3	3
ALASKA (16K)	9	3	0	4	7	**2**	**2**	**2**	5	5
BOSSBOWS (16K)	10	6	**5**	0	13	8	9	8	13	10
FLICKR (16K)	20	18	19	16	0	18	15	18	**10**	13

TABLE IV
REGRET MATRIX : RESULTS ON OUR BASES OF INTEREST (IN %).

Eval |Train	$RT_{2\%}$	$RT_{4\%}$	$Random_{2\%\&4\%}$	$RT_{6\%}$	$Random_{6\%}$	$RT_{8\%}$	$Random_{8\%}$	$RT_{10\%}$	$Random_{10\%}$
RTBASE	1	**0**	2	**1**	2	**3**	**3**	**3**	10
ALASKA	**2**	**2**	3	**2**	4	**5**	**5**	**5**	16
BOSSBOWS	8	9	**6**	8	**7**	13	**7**	10	**9**
FLICKR	18	**15**	18	**18**	**18**	**10**	17	**13**	19

TABLE V
REGRET MATRIX : COMPARISON BETWEEN COVERING AND RANDOM SELECTION OF PIPELINES (IN %).

there are singular sources, often presenting a high Intrinsic Difficulty, that are covering only a few others. During our experiments, we noticed that these specific sources are impacting negatively training with JIN. Discarding representatives that are not covering at least 10 sources revealed to be an effective strategy enabling to disclose the potential of the most interesting ones. We call RT_ϵ the bases resulting from this filtering strategy.

We also create bases made of pipelines picked randomly so that we can compare the results obtained with our greedy coverings and randomness. For each ϵ selected, we pick randomly as many pipelines like the ones in RT_ϵ and, we call $RANDOM_\epsilon$ the source resulting from the mix of these pipelines. For each ϵ we create 3 variants of $RANDOM_\epsilon$ to be able to compute a mean performance of the random strategy.

B. Bases of interest

Fine-tuning JIN using the main sources in N_ϵ for some ϵ, we would like to see how much we can generalize on different bases of interest. Since we didn't include the JPEG compression in our previous analysis, we decide to only work with pictures compressed at the same quality factor as ours, namely, 98. We propose for instance, the following bases:

- A base of 80K covers generated with all our 3^5 pipelines (RTBASE).
- The 80K covers resulting from the concatenation of BOSS [4] & BOWS-2 [14] compressed with a QF of 98 (BOSSBOWS).

RTBASE	ALASKA20K	BOSSBOWS	FLICKR
27%	28%	13%	24%

TABLE VI
INTRINSIC DIFFICULTIES (P_E) OF OUR EVALUATION BASES.

- 80K of covers chosen randomly from ALASKABASE [5] compressed with a QF of 98 (ALASKA).
- 80K of "wild covers" compressed with a QF of 98 derived from unknown realistic pipelines.

For our wild base, we propose to build it using the well-known website flickr[3] which gathers images shared by millions of users. Considering that we don't have any idea about the processing pipelines used by the users, we are in the worst-case scenario. We are notably completely blind about the quantization tables used for the compressions and our JPEG-based filtering is hence, approximative. Moreover, to focus on CSM mostly caused by unknown processing pipelines, we prefer to select images coming from camera models in the same range as the one we used for the grid (CANONEOS).

We share in VI the Intrinsic Difficulties of all the bases described above, finetuning JIN over 15 epochs. All our evaluation sources are reasonably difficult. This is possible thanks to the quality of the pre-trained weights released by [12]. We tried at first to perform trainings from scratch without these weights. Unfortunately, it wasn't possible to make JIN converge properly with a classic training, that's to say, a training that does not involve any steganalysis tricks like the pair constraint [12].

C. Results

In Table IV, we present the regret matrix on the bases of interest obtained after fine-tuning JIN over 15 epochs. The first column is a special case where we present the regrets without fine-tuning, simply using the pre-trained weights computed using ImageNet [15], [12]. As a first result, we can already observe that using JIN "on the shelf" is not the best strategy if we want to generalize on images not coming from ImageNet. That being said, as explained before, we cannot neglect the positive effect of these pre-trained weights on our trainings.

[3] flickr.com

The BOSSBOWS test database is more challenging for the detectors trained on our extracted bases RT_ϵ. This is partly because we started our study with a set of images captured with High ISO, a pattern not very represented in BOSSBOWS which gathers images taken with rather low ISOs.

It is also interesting to notice that, despite the drastic change in framework, the bases RT_ϵ still enable ensure a regret lower than ϵ on the grid represented by RTBASE. Moreover, for $\epsilon = 2, 4, 6\%$, we obtain a pretty low regret with ALASKA. This is rather surprising considering the great diversity in terms of sensor and ISO of the images contained in ALASKA that is far more superior than our bases. This indicates again that learning with a wide noise diversity is more important than learning with a wide content diversity if we want to be as holistic as possible. Moreover, we obtain comparable performances on ALASKA with respect to RTBASE, even if, fewer training samples and fewer pipelines are used to train the RT_ϵ-bases compared to RTBASE.

Looking at the performances on FLICKR, three of our five extracted bases are outperforming all the other training bases, even in the case where they are made of more samples. The case of $RT_{8\%}$ is particularly interesting since it is a result we cannot reproduce using other source combinations. This shows on its own the importance of cherry-picking the pipelines used for the training base instead of using as many images as possible, or making a random mixture with a maximum of pipelines.

At last, we present in Tab. V a performance comparison after training on RT_ϵ and $Random_\epsilon$. Except for BOSSBOWS, it seems that using our bases is a better strategy than generating random combinations of pipelines, especially if we want to generalize on FLICKR.

V. Conclusions and Perspectives

In this paper, we present a methodology for generating relevant bases enabling to fight CSM in an operational steganalysis framework. We show that our strategy leads to bases more informative for a detector than exploiting a high quantity of images, random augmentations, or as many pipelines as possible. Furthermore, from our different studies, it appears that Denoising, Sharpening, and Downsampling are playing a significant role in the cover source mismatch issue. Broadly speaking, other studies should be conducted to fully harness these parameters to cope with CSM. From our study, it is also easy to derive a batch of "complementary" sources. This may help the community to test strategies that are trying to reduce the mismatch between a set of sources. In the near future, we plan to perform such experiments leveraging domain adaptation methods just like we already did in [16] for digital forensics.

Acknowledgments

Our experiments were possible thanks to computing means of IDRIS through the resource allocation 2021- AD011013285 assigned by GENCI. This work received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021687 (project "UNCOVER") and the French Defense & Innovation Agency. The work of Tomas Pevny was supported by Czech Ministry of Education 19-29680L.

References

[1] Q. Giboulot, R. Cogranne, D. Borghys, and P. Bas, "Effects and Solutions of Cover-Source Mismatch in Image Steganalysis," *Signal Processing: Image Communication*, Aug. 2020. [Online]. Available: https://hal-utt.archives-ouvertes.fr/hal-02631559

[2] J. Pasquet, S. Bringay, and M. Chaumont, "Steganalysis with cover-source mismatch and a small learning database," in *EUSIPCO: European Signal Processing Conference*, Lisbon, Portugal, Sep. 2014, pp. 2425–2429. [Online]. Available: https://hal-lirmm.ccsd.cnrs.fr/lirmm-01234249

[3] J. Kodovský, V. Sedighi, and J. Fridrich, "Study of cover source mismatch in steganalysis and ways to mitigate its impact," in *Media Watermarking, Security, and Forensics 2014*, A. M. Alattar, N. D. Memon, and C. D. Heitzenrater, Eds., vol. 9028, International Society for Optics and Photonics. SPIE, 2014, pp. 204 – 215. [Online]. Available: https://doi.org/10.1117/12.2039693

[4] P. Bas, T. Filler, and T. Pevny, ""Break Our Steganographic System": The Ins and Outs of Organizing BOSS," in *INFORMATION HIDING*, ser. Lecture Notes in Computer Science, vol. 6958/2011, Czech Republic, May 2011, pp. 59–70. [Online]. Available: https://hal.archives-ouvertes.fr/hal-00648057

[5] R. Cogranne, Q. Giboulot, and P. Bas, "The ALASKA Steganalysis Challenge: A First Step Towards Steganalysis "Into The Wild"," in *ACM IH&MMSec (Information Hiding & Multimedia Security)*, ser. ACM IH&MMSec (Information Hiding & Multimedia Security), Paris, France, Jul. 2019. [Online]. Available: https://hal.archives-ouvertes.fr/hal-02147763

[6] G. Quentin, P. Bas, C. Rémi, and D. Borghys, "The Cover Source Mismatch Problem in Deep-Learning Steganalysis," in *European Signal Processing Conference*, Belgrade, Serbia, Aug. 2022. [Online]. Available: https://hal-utt.archives-ouvertes.fr/hal-03694662

[7] L. Guo, J. Ni, W. Su, C. Tang, and Y.-Q. Shi, "Using statistical image model for jpeg steganography: Uniform embedding revisited," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 12, pp. 2669–2680, 2015.

[8] V. Holub and J. Fridrich, "Low-complexity features for jpeg steganalysis using undecimated dct," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 2, pp. 219–228, 2015.

[9] D. Šepák, L. Adam, and T. Pevný, "Formalizing cover-source mismatch as a robust optimization," in *EUSIPCO: European Signal Processing Conference*, Belgrade, Serbia, Sep. 2022.

[10] V. Chvatal, "A greedy heuristic for the set-covering problem," *Mathematics of Operations Research*, vol. 4, no. 3, pp. 233–235, 1979. [Online]. Available: https://doi.org/10.1287/moor.4.3.233

[11] L. Breiman, "Manual on setting up, using, and understanding random forests v3. 1," *Statistics Department University of California Berkeley, CA, USA*, vol. 1, no. 58, pp. 3–42, 2002.

[12] J. Butora, Y. Yousfi, and J. Fridrich, "How to pretrain for steganalysis," in *Proceedings of the 2021 ACM Workshop on Information Hiding and Multimedia Security*, ser. IHamp;MMSec '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 143–148. [Online]. Available: https://doi.org/10.1145/3437880.3460395

[13] V. Holub, J. J. Fridrich, and T. Denemark, "Universal distortion function for steganography in an arbitrary domain," *EURASIP Journal on Information Security*, vol. 2014, pp. 1–13, 2014.

[14] T. Furon and P. Bas, "Broken Arrows," *EURASIP Journal on Information Security*, vol. 2008, p. ID 597040, Oct. 2008. [Online]. Available: https://hal.archives-ouvertes.fr/hal-00335311

[15] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[16] R. Abecidan, V. Itier, J. Boulanger, and P. Bas, "Unsupervised JPEG Domain Adaptation for Practical Digital Image Forensics," in *WIFS 2021 : IEEE International Workshop on Information Forensics and Security*. Montpellier, France: IEEE, Dec. 2021. [Online]. Available: https://hal.archives-ouvertes.fr/hal-03374780