How to augment a small learning set for improving the performances of a CNN-based steganalyzer?

Mehdi YEDROUDJ[1,2], Marc CHAUMONT[1,3], Fredéric COMBY[1,2]
LIRMM[1], Univ Montpellier[2], CNRS, Univ Nîmes[3], Montpellier, France

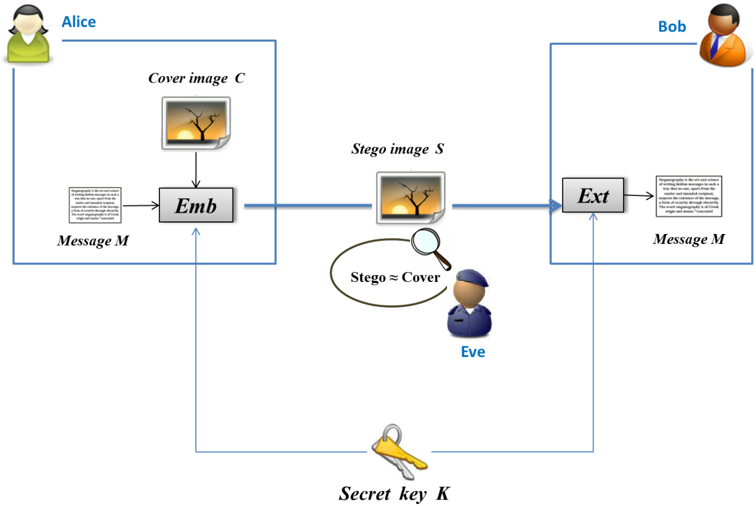
February 3, 2018

Media Watermarking, Security, and Forensics, IS&T Int. Symp. on Electronic Imaging, SF, California, USA, 28 Jan. - 2 Feb. 2018.

Outline

1. Introduction
2. Yedroudj-Net CNN
3. Database augmentation
4. Conclusion

Steganography / Steganalysis



Alice

Cover image C

Message M

Emb

Stego image S

Stego ≈ Cover

Eve

Bob

Ext

Message M

Secret key K

Steganography is the art and science of writing hidden messages in such a way that no-one, apart from the sender and intended recipient, suspects the existence of the message, a form of security through obscurity. The word steganography is of Greek origin and means "concealed
Steganography is the art and science of writing hidden messages in such a way that no-one, apart from the sender and intended recipient, suspects the existence of the message, a form of security through obscurity. The word steganography is of Greek origin and means "concealed

Few observations:

- Current CNNs (spatial or JPEG steganalysis) give similar or even better results than the EC+RM,
- CNNs need a lot of samples when used for steganalysis purposes (at least 5 000 pairs, can reach millions),
- Each time the target distribution is modified, the learning process has to be re-done.

Problem:
In “real world” the learning set can be very small (0, 10, 100, ..).

Question: How to deal with a low regime learning?

Current solutions to improve a CNN efficiency

- Virtual Augmentation [Krizhevsky 2012]
- Transfer Learning [Qian et al. 2016],
- Use of Ensemble [Xu et al. 2016],
- Learn on millions of images? [Zeng et al. 2018],
- Incorporating the knowledge of Selection Channel [Ye et al. 2017],
- Use new networks [Yedroudj et al. ICASSP'2018], ...
- ...
- Augment the data-base ...

"ImageNet Classification with Deep Convolutional Neural Networks", A. Krizhevsky, I. Sutskever, G. E. Hinton, NIPS'2012,
"Learning and transferring representations for image steganalysis using convolutional neural network", Y. Qian, J. Dong, W. Wang, T. Tan, ICIP'2016,
"Ensemble of CNNs for Steganalysis: An Empirical Study", G. Xu, H.-Z. Wu, Y. Q. Shi, IH&MMSec'16,
"Large-scale jpeg image steganalysis using hybrid deep-learning framework", J. Zeng, S. Tan, B. Li, J. Huang, TIFS'2018,
"Deep Learning Hierarchical Representations for Image Steganalysis," J. Ye, J. Ni, and Y. Yi, TIFS'2017,
"Yedroudj-Net: An Efficient CNN for Spatial Steganalysis", M. Yedroudj, F. Comby, M. Chaumont, ICASSP'2018.

Outline

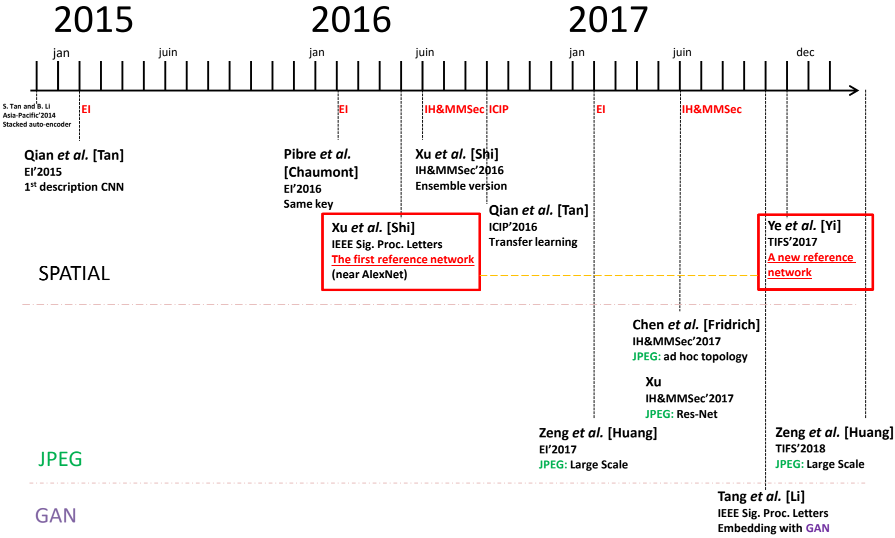
1. Introduction

2. Yedroudj-Net CNN

3. Database augmentation

4. Conclusion

The CNNs for the steganalysis



S. Tan and B. Li
Asia-Pacific'2014
Stacked auto-encoder

Qian *et al.* [Tan]
EI'2015
1st description CNN

Pibre *et al.* [Chaumont]
EI'2016
Same key

Xu *et al.* [Shi]
IEEE Sig. Proc. Letters
The first reference network
(near AlexNet)

Xu *et al.* [Shi]
IH&MMSec'2016
Ensemble version

Qian *et al.* [Tan]
ICIP'2016
Transfer learning

Zeng *et al.* [Huang]
EI'2017
JPEG: Large Scale

Chen *et al.* [Fridrich]
IH&MMSec'2017
JPEG: ad hoc topology

Xu
IH&MMSec'2017
JPEG: Res-Net

Ye *et al.* [Yi]
TIFS'2017
A new reference network

Zeng *et al.* [Huang]
TIFS'2018
JPEG: Large Scale

Tang *et al.* [Li]
IEEE Sig. Proc. Letters
Embedding with GAN

Marc Chaumont, " **The emergence of Deep Learning in steganography and steganalysis (in the period 2015 - 2017)** ", Tutorial given at a research day in France, the 16th of January 2018, Poitiers, France, 2018 **DOI 10.13140/RG.2.2.35080.32005**.
http://www.lirmm.fr/~chaumont/publications/Emergence_of_Deep_Learning-Tutorial_Meeting-France-CHAUMONT_16_01_2018.pdf

A new network: Yedroudj-Net

"Yedroudj-Net: An Efficient CNN for Spatial Steganalysis", M. Yedroudj, F. Comby, M. Chaumont, IEEE ICASSP'2018.

The code can be downloaded here: http://www.lirmm.fr/~chaumont/Yedroudj-Net.html

Yedroudj-Net [Yedroudj et al. ICASSP'2018]

Aggregation of the "most efficient" bricks of newly designed CNNs.
Objective: To have a basic CNN (baseline) at the state-of-the-art.

The essential elements of our network:

- A high-pass filters bank for pre-processing (SRM [1][2]),
- A truncation activation function ("hard tanh") [2],
- The "batch normalization" associated with a "scaling" layer [3][4][5].

[1]: "Ensemble Classifiers for Steganalysis of Digital Media", J. Kodovský, J. Fridrich, V. Holub, TIFS'2012,
[2]: "Deep Learning Hierarchical Representations for Image Steganalysis", J. Ye, J. Ni, Y. Yi, TIFS'2017,
[3]: "BN: Accelerating deep network training by reducing internal covariate shift", S. Ioffe, C. Szegedy, ICML'2015,
[4]: "Deep residual learning for image recognition", K. He, X. Zhang, S. Ren, J. Sun, CVPR'2016,
[5]: "Structural Design of Convolutional Neural Networks for Steganalysis", G. Xu, H. Z. Wu, Y. Q. Shi, IH&MMSec'2016.

Yedroudj-Net

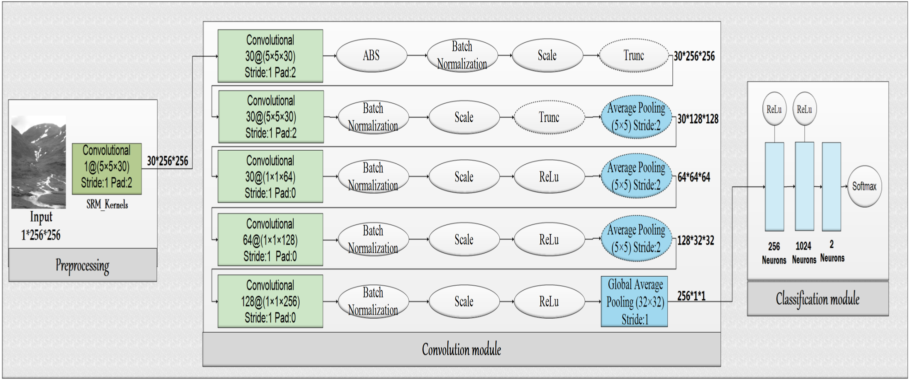


Figure: Yedroudj-Net

Rapid “fair” comparison (1)

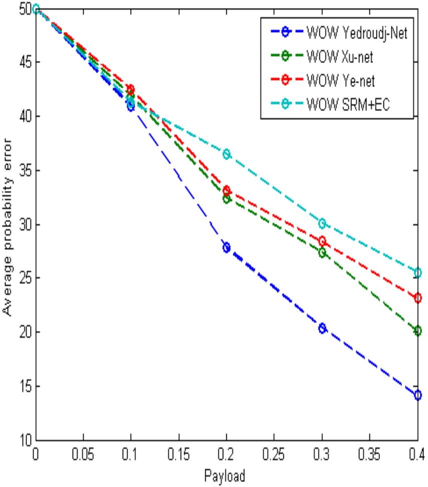
Clairvoyant protocol

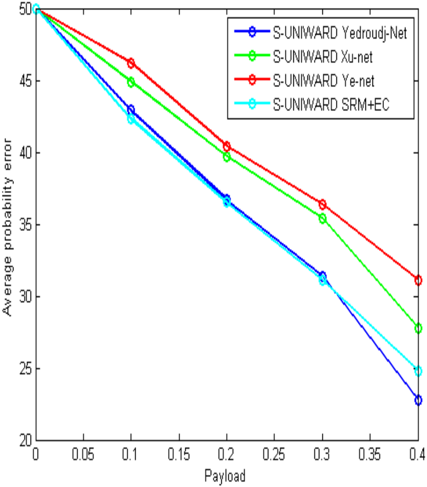
- Resize the 10 000 images of BOSSBase from 512×512 to 256×256,
- Using embedding algorithms WOW [1] and S-UNIWARD [2] to generate the stegos (Matlab Version),
- Selection of 5 000 pairs for learning including 1 000 pairs for validation,
- The other 5 000 pairs are used for the test (evaluation).

[1] "Designing Steganographic Distortion Using Directional Filters", V. Holub, J. Fridrich, WIFS'2012.

[2] "Universal Distortion Function for Steganography in an Arbitrary Domain", V. Holub, J. Fridrich, T. Denemark, JIS'2014.

Rapid "fair" comparison (2)





Outline

1 Introduction

2 Yedroudj-Net CNN

3 Database augmentation

4 Conclusion

Current solutions to improve a CNN efficiency

- Virtual Augmentation [Krizhevsky 2012]
- Transfer Learning [Qian et al. 2016],
- Use of Ensemble [Xu et al. 2016],
- Learn on millions of images? [Zeng et al. 2018],
- Incorporating the knowledge of Selection Channel [Ye et al. 2017],
- Use new networks [Yedroudj et al. ICASSP'2018], ...
- ...
- Augment the data-base ...

"ImageNet Classification with Deep Convolutional Neural Networks", A. Krizhevsky, I. Sutskever, G. E. Hinton, NIPS'2012,
"Learning and transferring representations for image steganalysis using convolutional neural network", Y. Qian, J. Dong, W. Wang, T. Tan, ICIP'2016,
"Ensemble of CNNs for Steganalysis: An Empirical Study", G. Xu, H.-Z. Wu, Y. Q. Shi, IH&MMSec'16,
"Large-scale jpeg image steganalysis using hybrid deep-learning framework", J. Zeng, S. Tan, B. Li, J. Huang, TIFS'2018,
"Deep Learning Hierarchical Representations for Image Steganalysis," J. Ye, J. Ni, and Y. Yi, TIFS'2017,
"Yedroudj-Net: An Efficient CNN for Spatial Steganalysis", M. Yedroudj, F. Comby, M. Chaumont, ICASSP'2018.

Counterproductive enrichment ...

Setup: Enrichment with other cameras (Probability of error):

Training set BOSS+LIRMM = 14 000 pairs.

	WOW 0.2 bpp	S-UNIWARD 0.2 bpp
BOSS	**27.8** %	**36.7** %
BOSS+LIRMM	29.9 %	38.6 %

Setup: Enrichment with strongly dissimilar sources and unbalance proportions (Probability of error):

BOSS + PLACES2 1% = 14 000 pairs; BOSS + PLACES2 10% = 104 000 pairs; BOSS + PLACES2 100% = 1 004 000 pairs.

	WOW 0.2 bpp	S-UNIWARD 0.2 bpp
BOSS	**27.8** %	**36.7** %
BOSS+PLACES2 1%	34.2 %	41.6 %
BOSS+PLACES2 10%	40.0 %	43.9 %
BOSS+PLACES2 100%	44.6 %	45.3 %

Enrichment of the learning base

Setup: Enrichment with the same cameras (Probability of error):

BOSS+BOWS2: 14 000 pairs.

	WOW 0.2 bpp	S-UNIWARD 0.2 bpp
BOSS	**27.8** %	**36.7** %
BOSS+BOWS2	23.7 %	34.4%

Setup: Enrichment with the same RAW images but with different developments (Probability of error): BOSS+all-DEV: 44 000 pairs.

	WOW 0.2 bpp	S-UNIWARD 0.2 bpp
BOSS+all-DEV	23.0 %	33.2 %

.. and the use of VA...: BOSS+BOWS2+VA: 112 000 pairs.

	WOW 0.2 bpp	S-UNIWARD 0.2 bpp
BOSS+VA	24.2 %	34.8 %
BOSS+BOWS2+VA	20.8 %	31.1 %

A conjecture (rule for the increase):

"How to augment a small learning set for improving the performances of a CNN-based steganalyzer?", M. Yedroudj, F. Comby, and M. Chaumont, EI'2018.

Given a target database:

- either Eve (the steganalyst) finds the same camera(s) (used for generating the target database), capture new images, and reproduce the same development than the target database, with a special caution to the resizing,
- either Eve has an access to the original RAW images and reproduce similar developments than the target database with the similar resizing,

The reader should also remember that the Virtual Augmentation is also a good cheap processing measure.

Outline

1 Introduction

2 Yedroudj-Net CNN

3 Database augmentation

4 Conclusion

Conclusion

Enrichment

- Either finds the same camera and re-develop,
- Either use RAW images and re-develop,

And use Virtual Augmentation.

1. We can augment a learning database and it improves the results (around 5% for a classical payload size),
2. CNNs are sensitive to cover-source mismatch (as other machine-learning approaches),
3. Using only 5 000 pairs for the learning ⇒ CNNs have not reached their maximum efficiency,

What about "real world" steganalysis?

Well, our augmentation protocol is not always feasible in real word, but fortunately, sometimes it is!

We are looking for other solutions...

End of talk

CNN is the new state-of-the-art steganalysis tool ...
... there is still things to do...

"Yedroudj-Net: An Efficient CNN for Spatial Steganalysis", M. Yedroudj, F. Comby, M. Chaumont, IEEE ICASSP'2018.

The code can be downloaded here:
http://www.lirmm.fr/~chaumont/Yedroudj-Net.html

The slides can be downloaded here:
http://www.lirmm.fr/~chaumont/Publications.html.