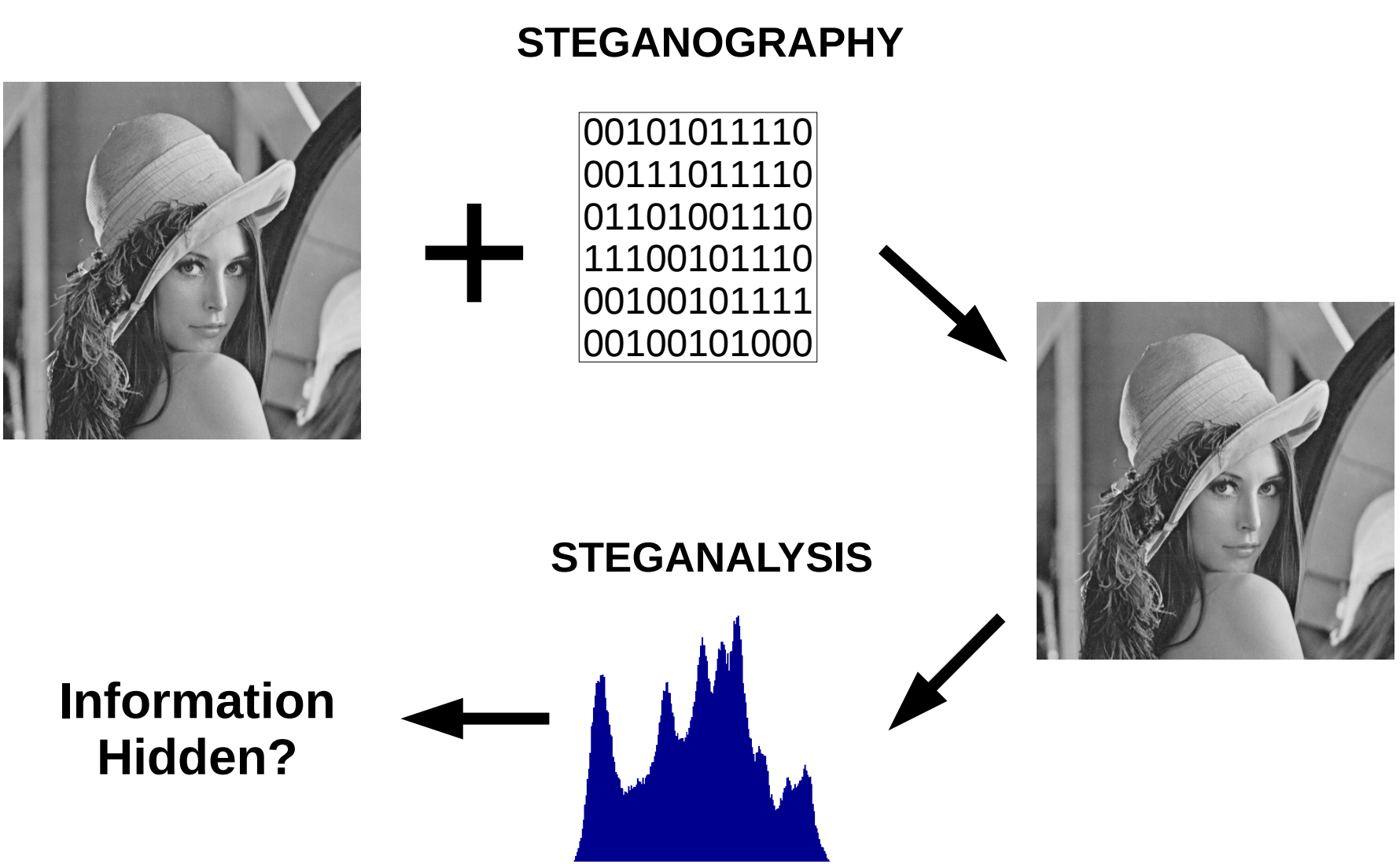
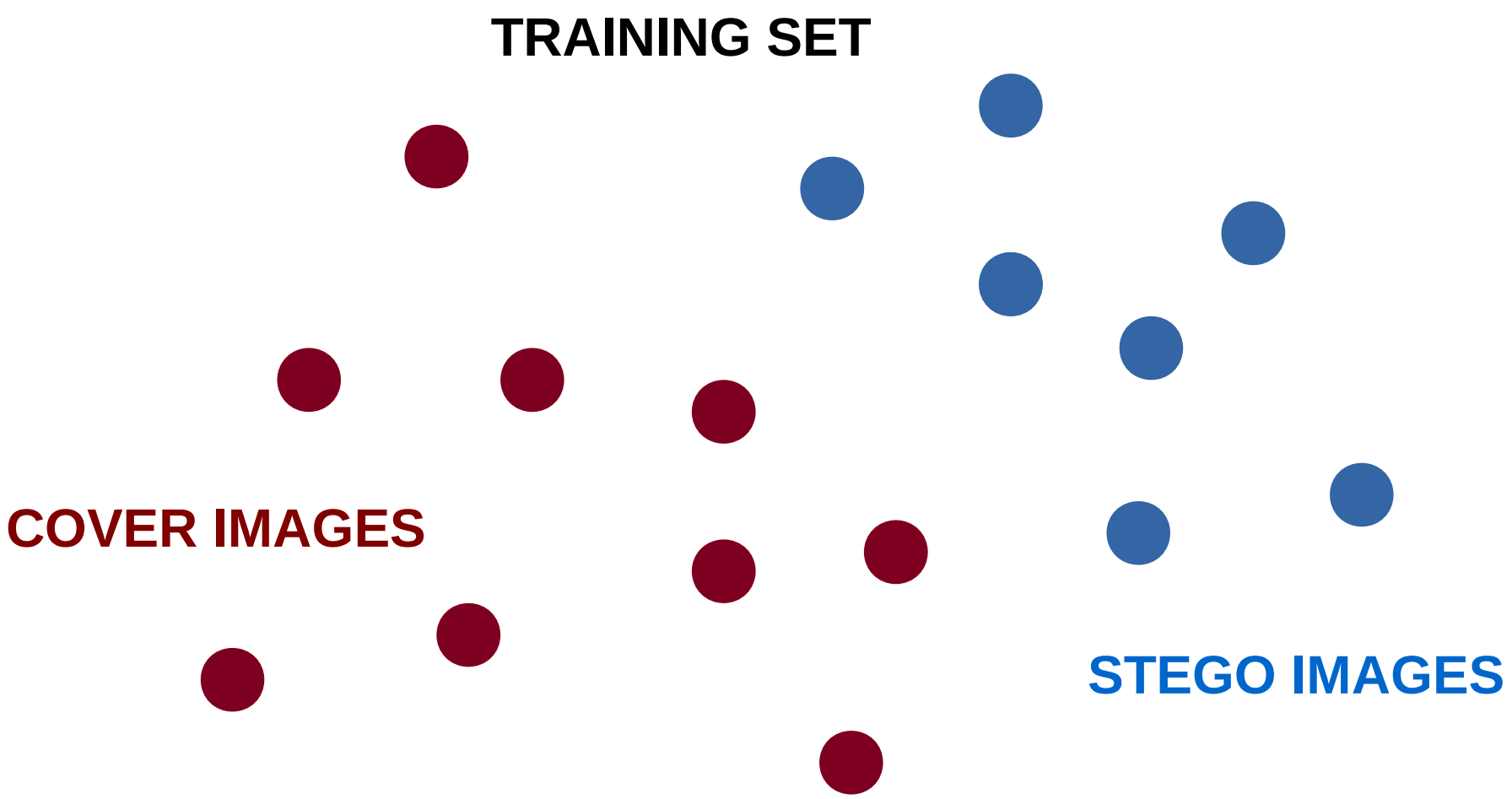
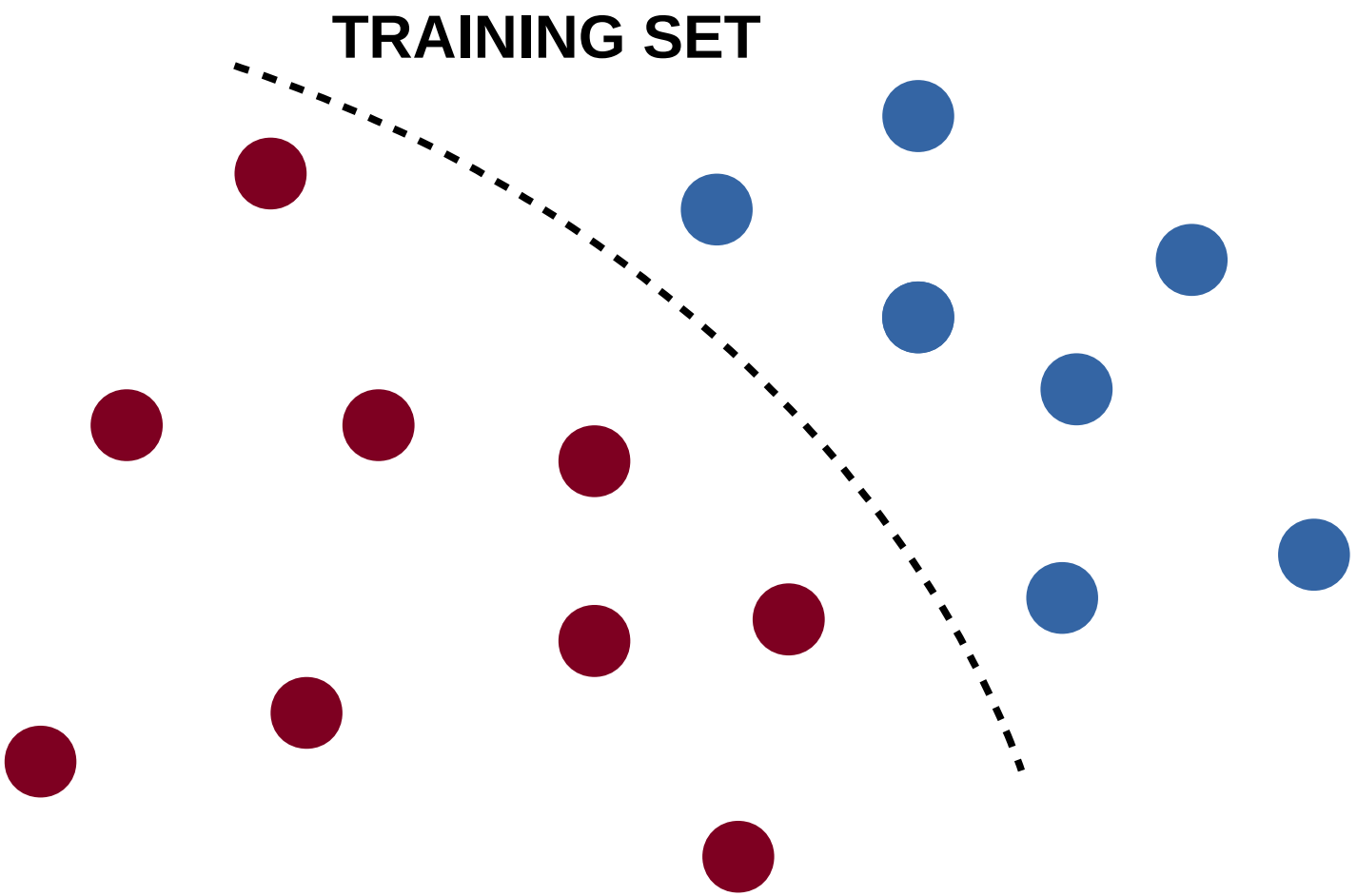
Unsupervised Steganalysis based on Artificial Training Sets

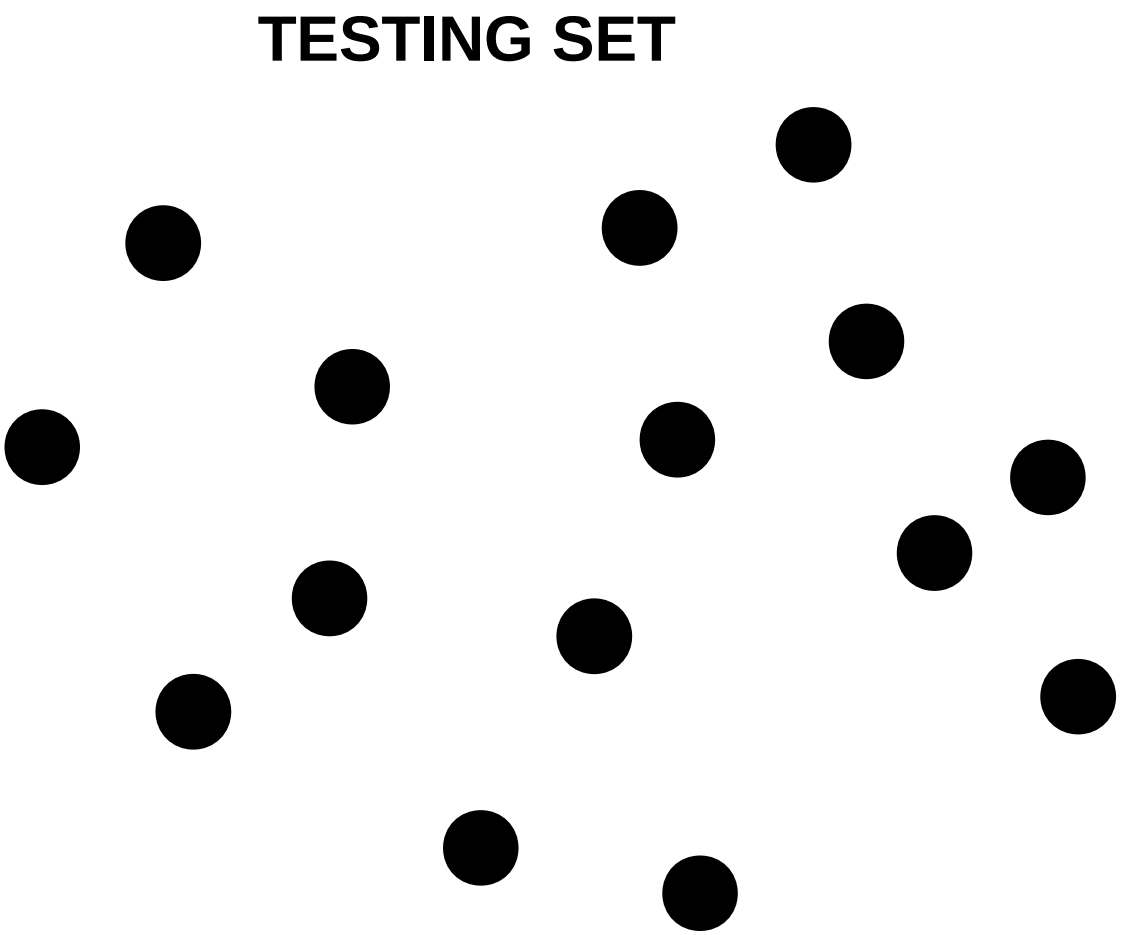
Daniel Lerch Hostalot
Davíd Megías Jiménez

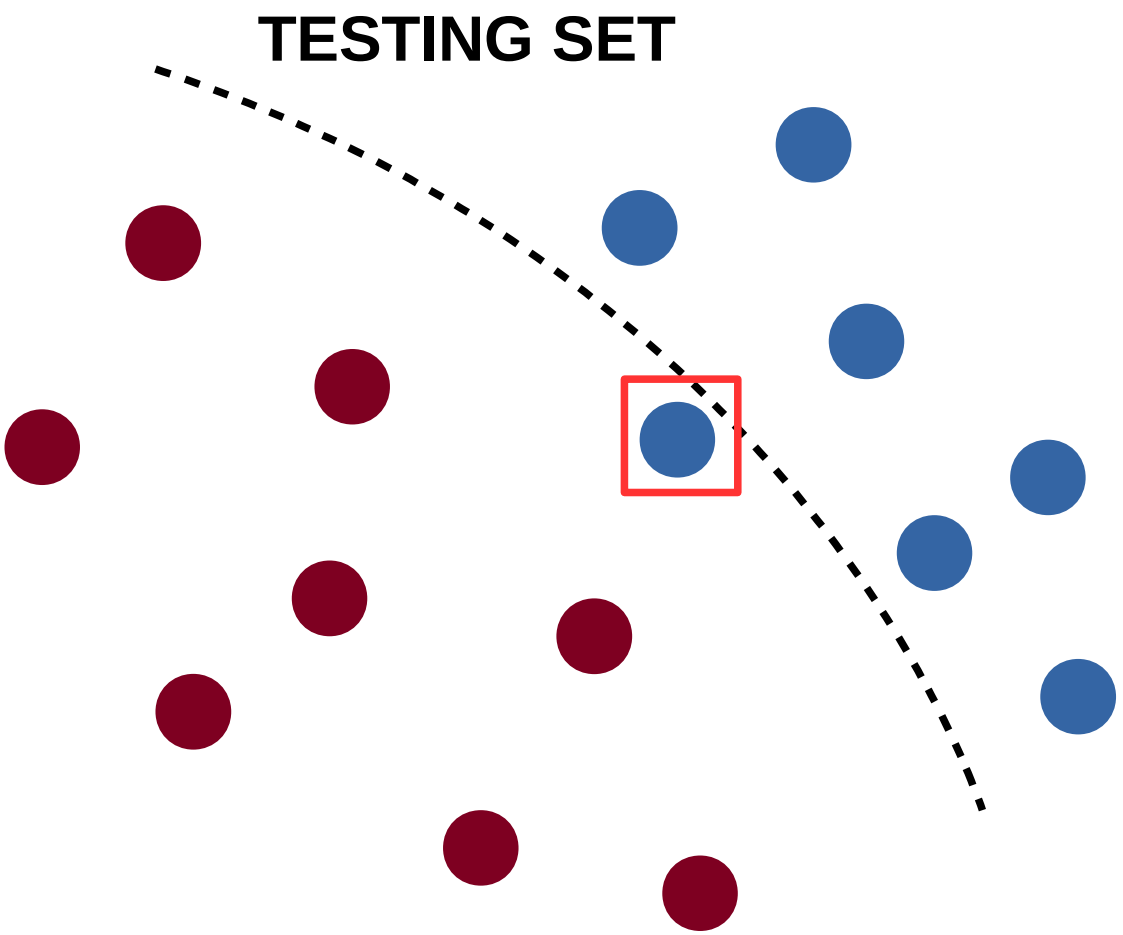
UOC Research Week 2016


STEGANOGRAPHY
00101011110
00111011110
01101001110
11100101110
00100101111
00100101000
STEGANALYSIS
Information Hidden?

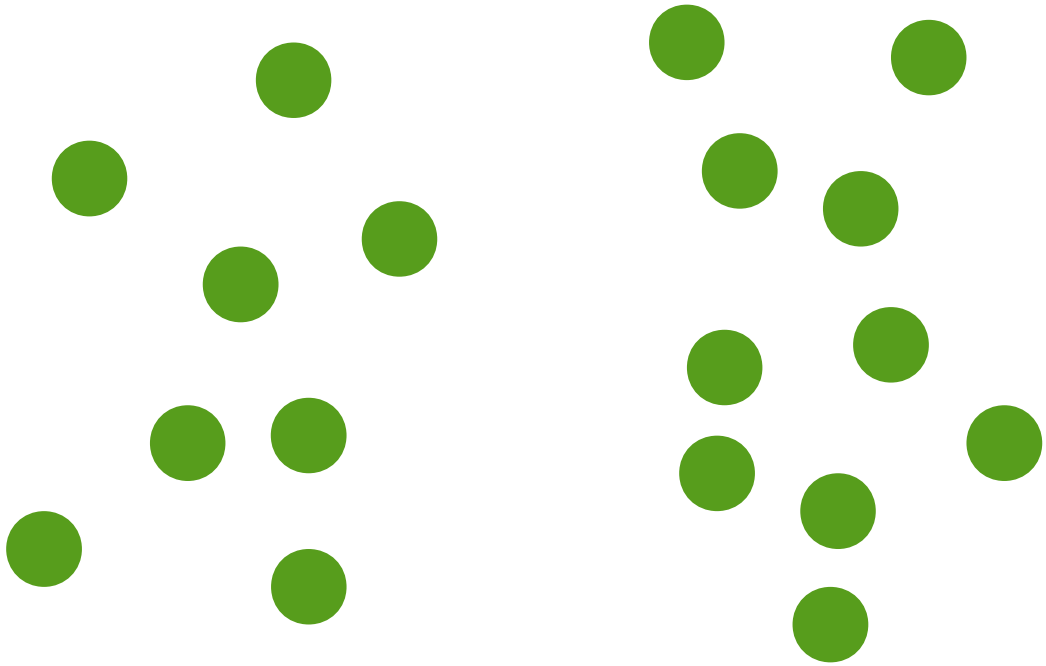

TRAINING SET
COVER IMAGES
STEGO IMAGES

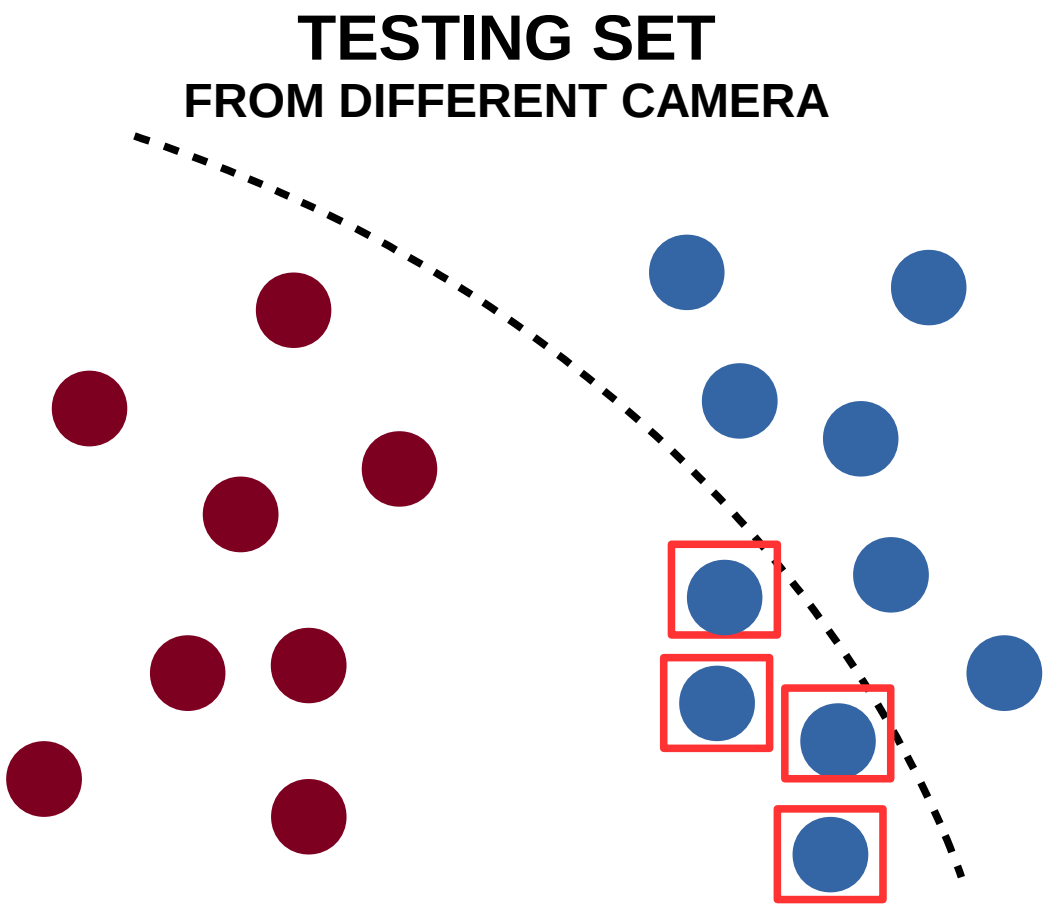

TRAINING SET


TESTING SET

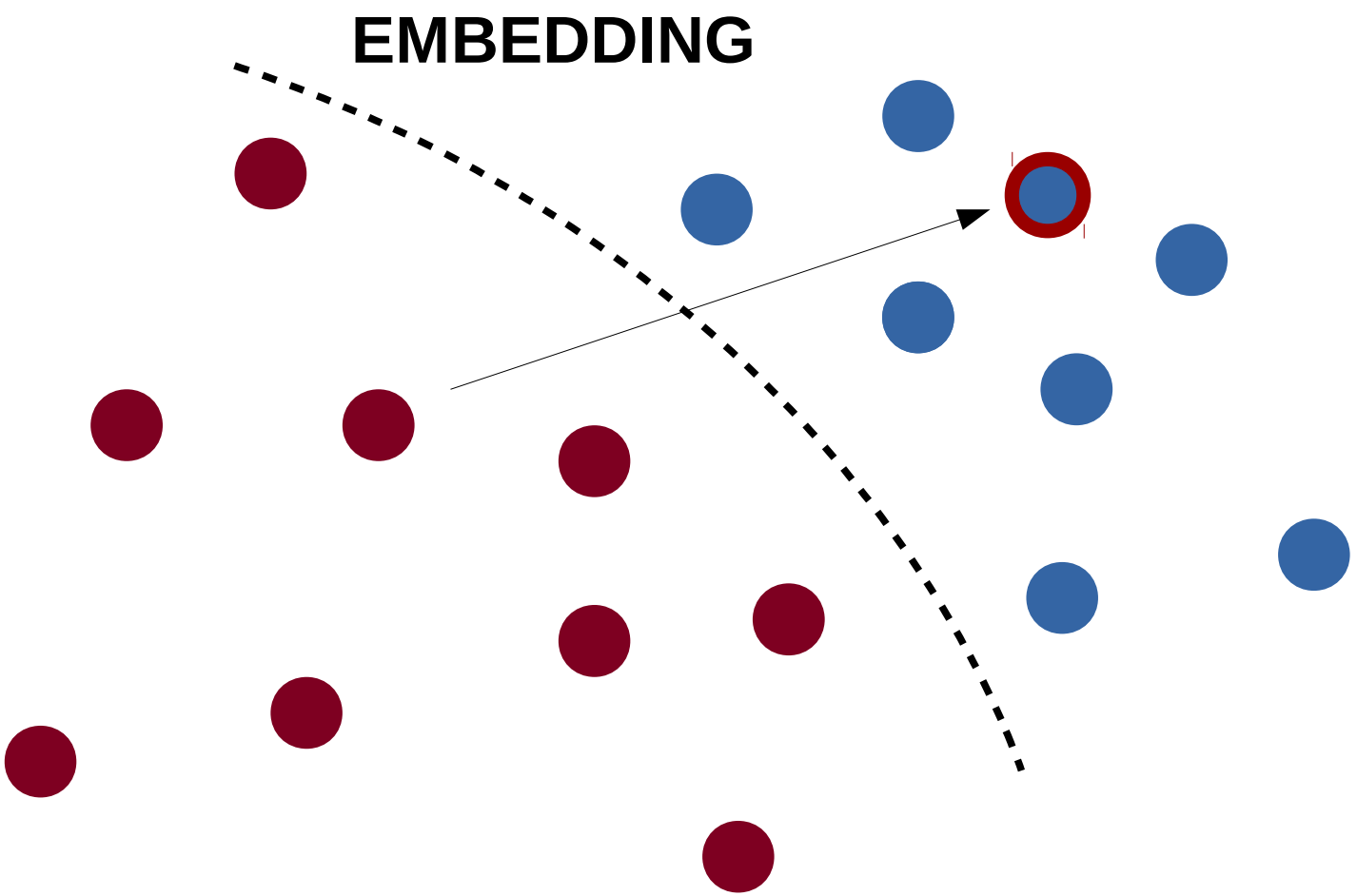

TESTING SET

TESTING SET
FROM DIFFERENT CAMERA

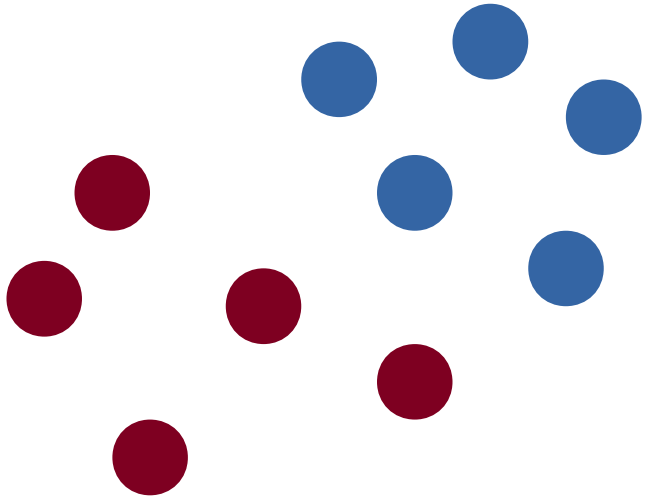



TESTING SET
FROM DIFFERENT CAMERA
COVER SOURCE MISMATCH (CSM)
PROBLEM

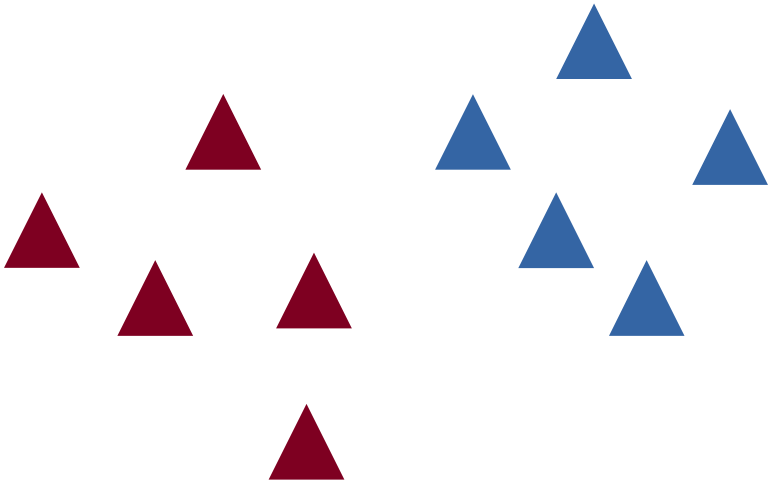
ATS: Basic Idea



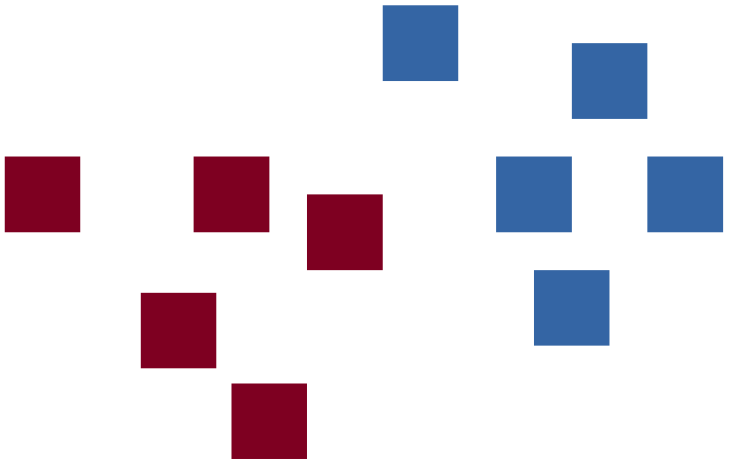
ATS: Set A – Testing Set



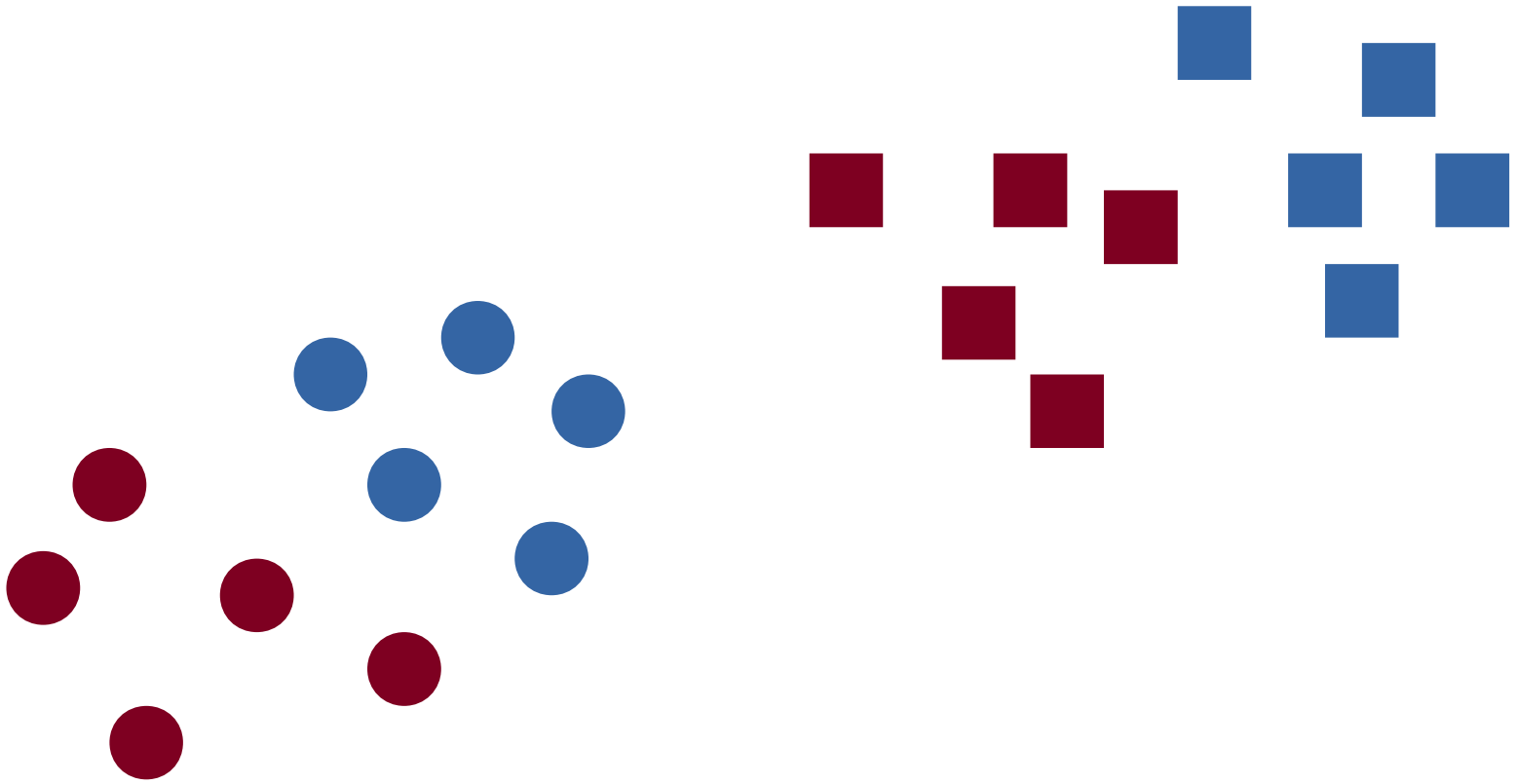
ATS: Set B – First Embedding



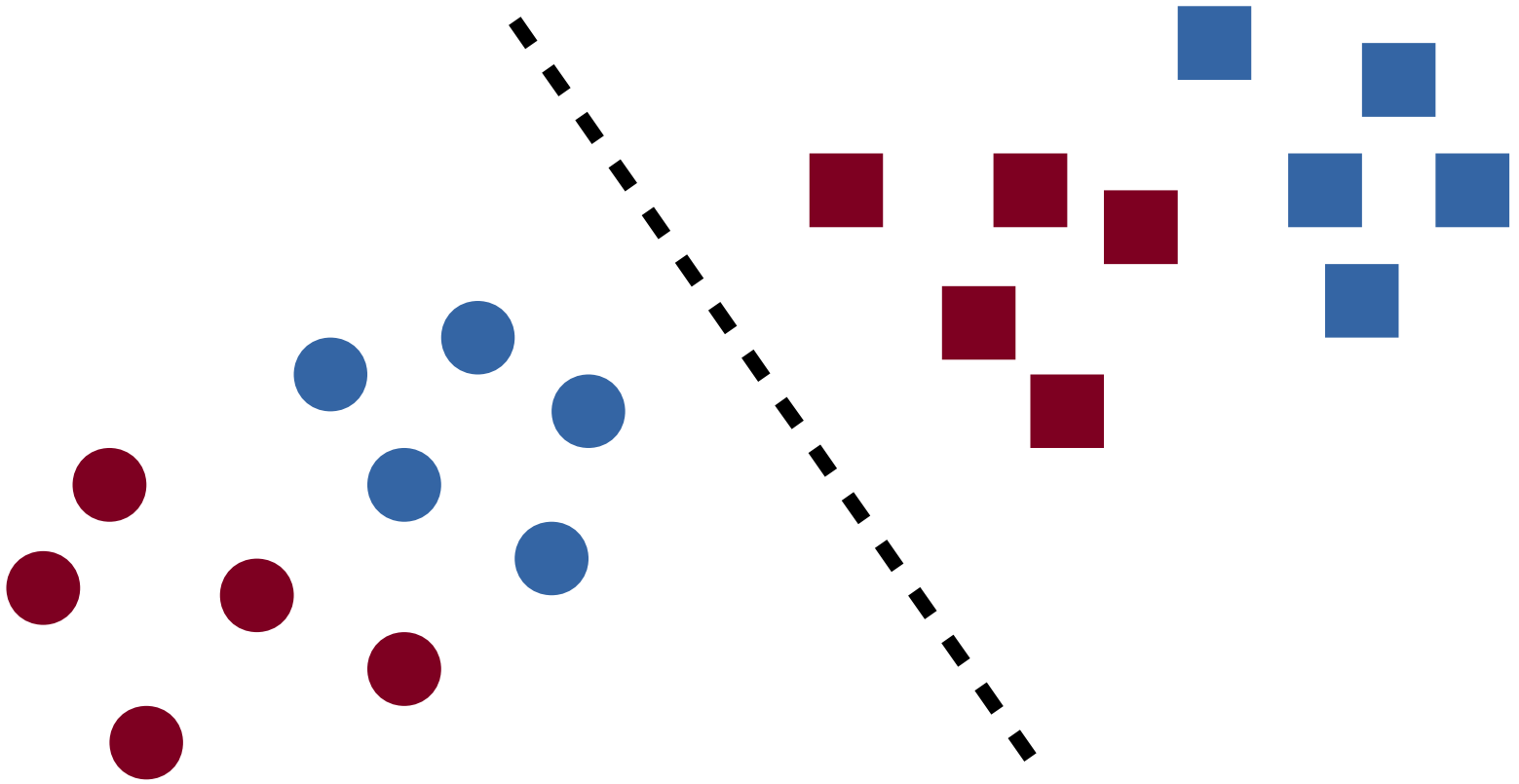
ATS: Set C – Second Embedding

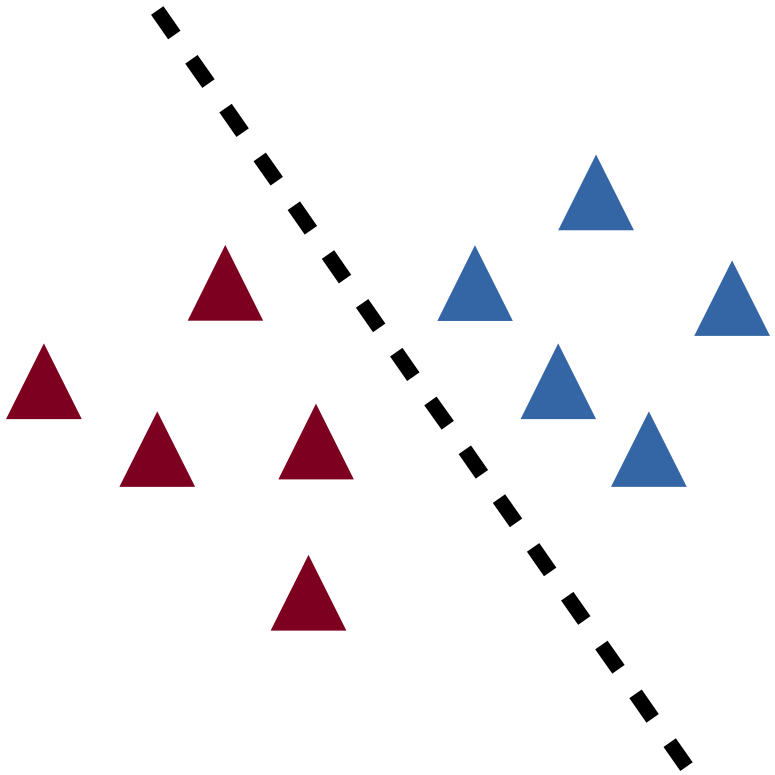


ATS: Classification Using A and C

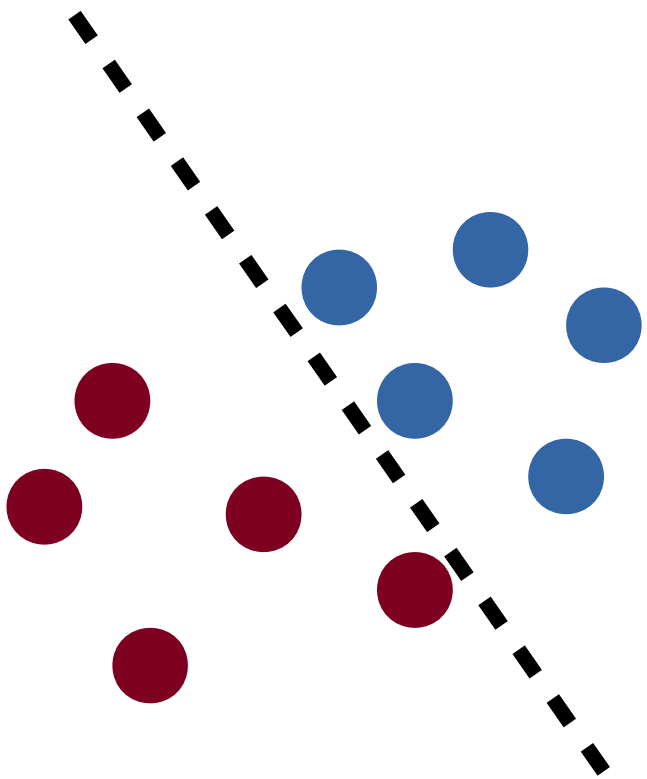


ATS: Training Using A and C





ATS: Bijection between A and B



ATS: Results

Comparative analysis between supervised classification and ATS

Test DB	LSB matching				HUGO			
	0.25 bpp		0.10 bpp		0.40 bpp		0.20 bpp	
	SUP	ATS	SUP	ATS	SUP	ATS	SUP	ATS
BOSS	0.96	**0.98**	0.90	**0.94**	0.78	**0.87**	0.67	**0.79**
ESO	0.37	**0.94**	0.50	**0.98**	0.50	**0.83**	0.45	**0.86**
CALP	0.51	**0.96**	0.57	**0.94**	0.48	**0.90**	0.48	**0.84**
INTE	0.47	**0.98**	0.50	**0.98**	0.51	**0.95**	0.51	**0.95**
NRCS	0.49	**0.68**	0.55	**0.61**	0.49	**0.63**	**0.50**	0.46
ALBN	0.57	**0.99**	0.65	**0.98**	0.50	**0.97**	0.50	**0.91**
NOAA	0.35	**1.00**	0.50	**0.98**	0.50	**1.00**	0.44	**0.96**

Conclusions:

- Tested with the algorithms LSBM, HUGO and WOW.
- Tested with the DBs: BOSS, ESO, CALP, INTE, NRCS, ALBN, NOAA.
- Tested with different cover/stego ratios.
- Avoids the Cover Source Mismatch problem
- More accurate than supervised classification.

Future work:

- Can ATS be applied to other fields?

Unsupervised Steganalysis based on Artificial Training Sets

Thank you for your attention