

# How We Learned to Stop Worrying about Content and Love the Metadata

Thomas Gloe<sup>a</sup>

Matthias Kirchner<sup>b,a</sup>

Christian Riess<sup>c,a</sup>

<sup>a</sup> dence GmbH, Dresden, Germany

<sup>b</sup> University of Münster, Germany

<sup>c</sup> University of Erlangen-Nuremberg, Germany

# The Truth about the Challenge

(all images taken from the dataset)

German officials received a sealed envelope ...

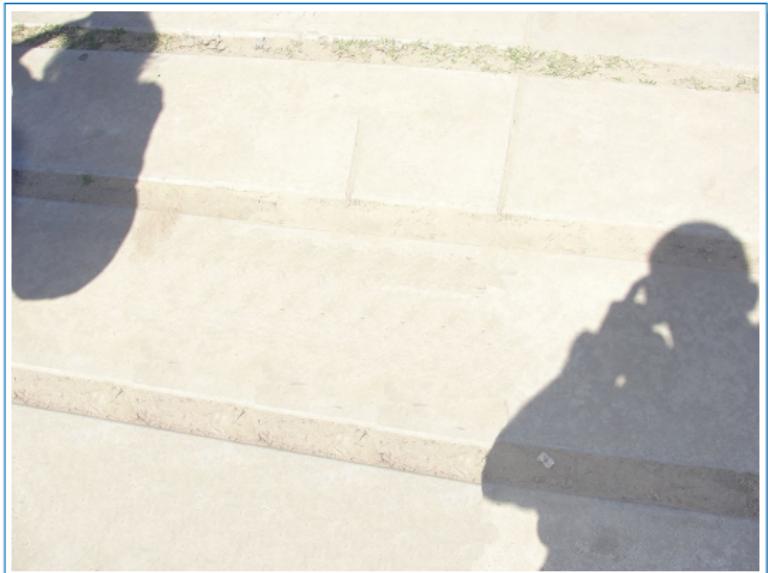


63e39d849d4ed857f136248a1845788e.png

# The Truth about the Challenge

(all images taken from the dataset)

... from an [unknown source](#), ...



8866aa058e3c1ec17d4ed43e429a0be4.png

# The Truth about the Challenge

(all images taken from the dataset)

... apparently flown in ...

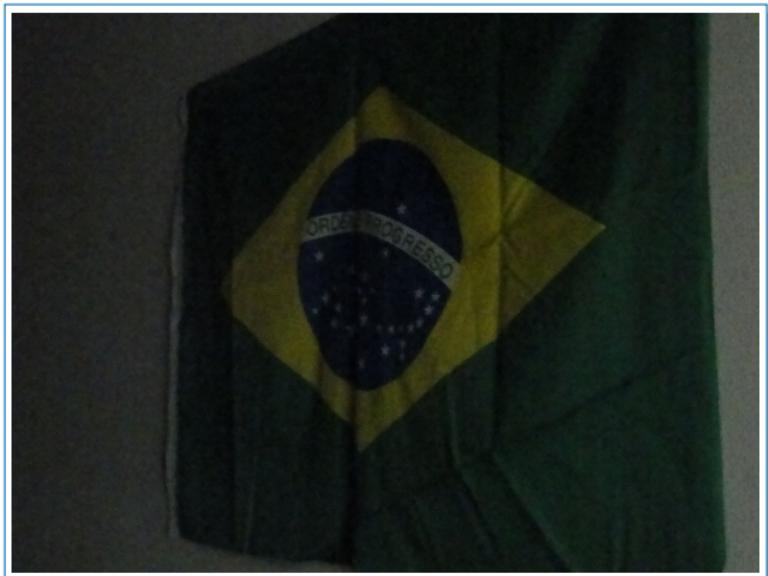


0657a892b04b62e7c5276bb487b728a3.png

# The Truth about the Challenge

(all images taken from the dataset)

... from Brazil ...



72c0c6740bdb07f422e1d261bf023b74.png

# The Truth about the Challenge

(all images taken from the dataset)

... to Germany.

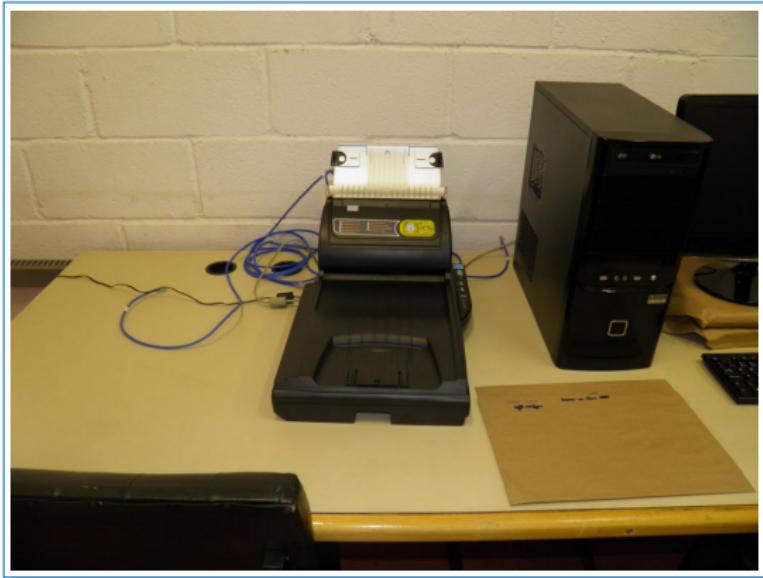


b46123b02139d593c1f842bc46ddc634.png

# The Truth about the Challenge

(all images taken from the dataset)

The envelope ...



1085577e4101f576514286b1f2255fd4.png

# The Truth about the Challenge

(all images taken from the dataset)

... contained a CD ...



887d94aa141d94481775fb3b03c794a6.png

# The Truth about the Challenge

(all images taken from the dataset)

... with loads of [digital images](#).



dd99d9140a3a93673f959b1a4b81c676.png

# The Truth about the Challenge

(all images taken from the dataset)

A letter explained that Brazilian researchers ...



867493635e57b7d6d19fd5e7c38e11b7.png

# The Truth about the Challenge

(all images taken from the dataset)

... were working ...



051ef83611bbaa7632bfefceb8f9a44a.png

# The Truth about the Challenge

(all images taken from the dataset)

... on ways how to make the Seleção ...



5147ac806357670e6bb24045b5dd6c7c.png

# The Truth about the Challenge

(all images taken from the dataset)

... win the forthcoming soccer world cup in Brazil.



ab59e5fdefc229a4a07592c4376e2ffc.png

## The Truth about the Challenge

(all images taken from the dataset)

They had developed complex techniques ...



a2d3b7163d67b60036c207d3bc31638f.png

# The Truth about the Challenge

(all images taken from the dataset)

... to defeat their [German](#) competitors ...

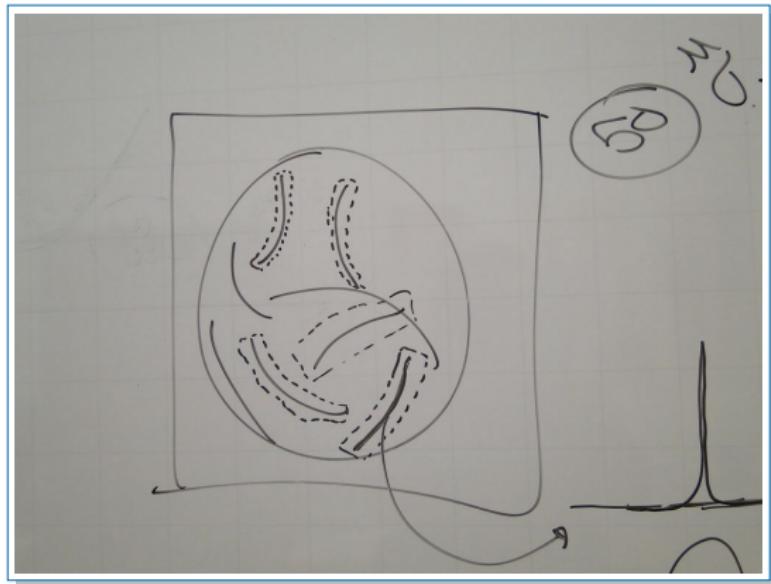


b46123b02139d593c1f842bc46ddc634.png

# The Truth about the Challenge

(all images taken from the dataset)

... with secret weapons.



0d3e4b46139d791d9f99b1272a7d454b.png

# The Truth about the Challenge

(all images taken from the dataset)

The Germans realized that it was time to react, ...



4419c64cdc21148336909c2dedf4d155.png

# The Truth about the Challenge

(all images taken from the dataset)

... and to find out whether this **threat** was real.



1277f286708944ab898b65d57dce2f49.png

# The Truth about the Challenge

(all images taken from the dataset)

They put lots of money onto the table.



1679b831fd157ddc848910949bb691e0.png

# The Truth about the Challenge

(all images taken from the dataset)

“Use all your [compute power](#) ...



294ec5a02a8c2d3ed8dcc4aa1b08e6cb.png

# The Truth about the Challenge

(all images taken from the dataset)

...and [time](#) to analyze these images”, they said to us.



271b57d5018f06c8df01761db37b286a.png

# The Truth about the Challenge

(all images taken from the dataset)

We rather wanted to spend money and time [differently](#).



b46123b02139d593c1f842bc46ddc634.png



b46123b02139d593c1f842bc46ddc634.png

# The Truth about the Challenge

(all images taken from the dataset)

So we came up with a master plan.



93751b603412ddcfb0dea231809e4561.png

# Metadata and File Format Forensics



0ed24a267d71f3c23582d481c42d93f4.png



2aa8cb7dcff33fdf6a0642eee0d06139.png

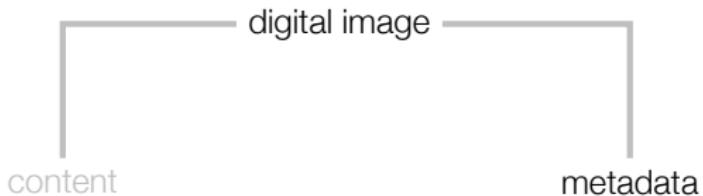
# Metadata and File Format Forensics



next presentations!  
(also Italians compete  
for the World Cup)

- file name
- compression parameters
- JPEG file structure
- EXIF data

# Metadata and File Format Forensics



next presentations!

- file name      — randomized
- compression parameters
- JPEG file structure
- EXIF data

# Metadata and File Format Forensics



next presentations!

- file name
  - compression parameters
  - JPEG file structure
  - EXIF data
- randomized
- } downsampling and conversion to PNG format

# Metadata and File Format Forensics

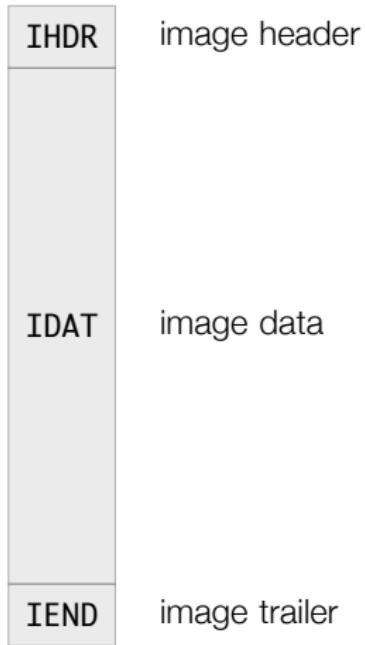


next presentations!

- file name — randomized
- compression parameters }  
parameters
- JPEG file structure }  
structure
- EXIF data — not supported

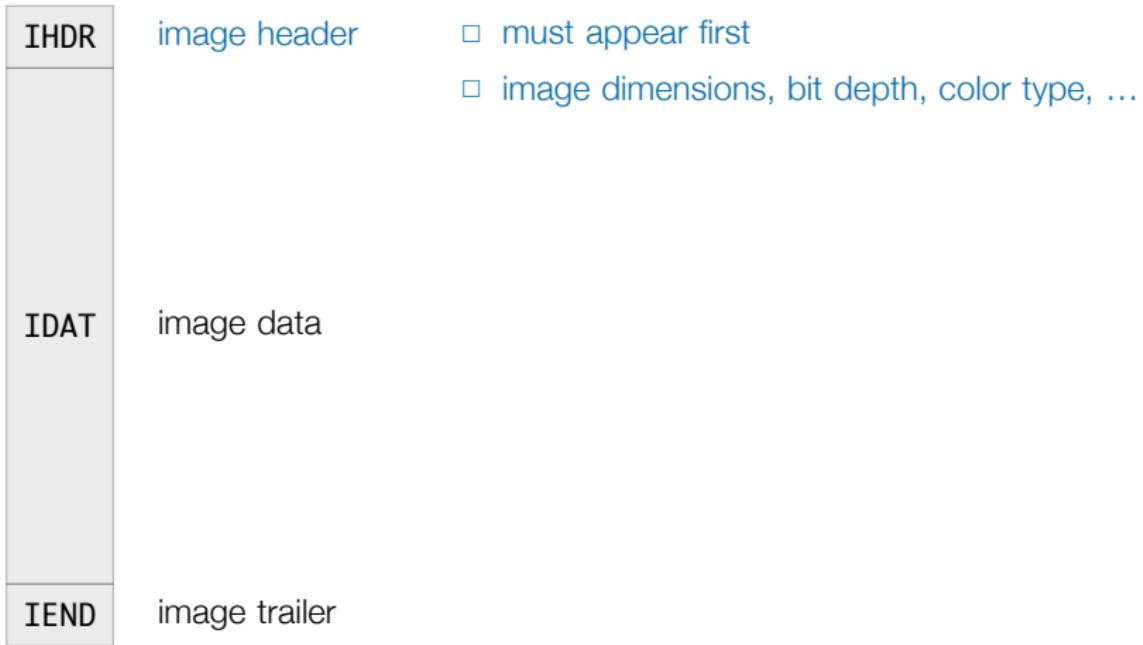
# PNG File Structure

- PNG files are organized in chunks.



# PNG File Structure

- PNG files are organized in chunks.



# PNG File Structure

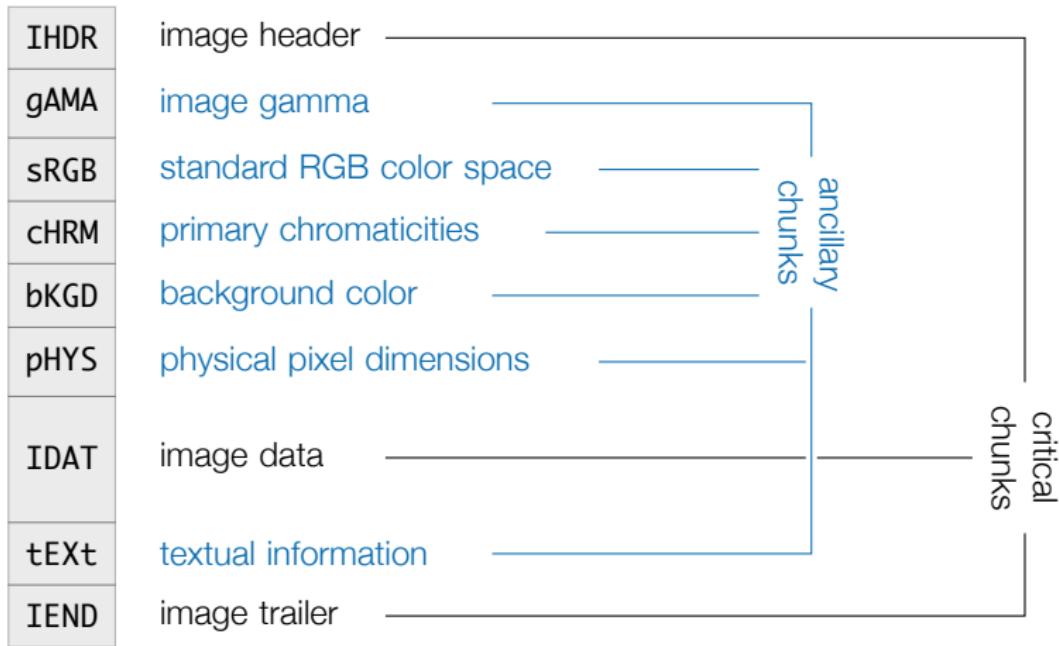
- PNG files are organized in chunks.



ab005cd25e71d13f3ff86bb9f394c556a.png

# PNG File Structure

- PNG files are organized in chunks.



# Example Files

chunk	description	entry	image #1
IHDR	image header	width	1024
		height	768
		bit depth	8
		color type	2
		compression method	0
		filter method	0
		interlace method	0
gAMA	image gamma		2.2
sRGB	standard RGB color space		perceptual
cHRM	primary chromaticities	white point x	0.31270
		white point y	0.32900
		red x	0.64000
		red y	0.33000
		green x	0.30000
		green y	0.60000
		blue x	0.15000
		blue y	0.06000
bKGD	background color		255 255 255
pHYS	physical pixel dimensions	pixels per unit x	7086
		pixels per unit y	7086
		unit specifier	meter
IDAT	image data		
tEXt	textual information	date:create	2013-05-01T17:07:36-03:00
tEXt	textual information	date:modify	2013-04-16T15:42:00-03:00
tEXt	textual information	jpeg:colorspace	2
tEXt	textual information	jpeg:sampling-factor	1x1,1x1,1x1
IEEND	image trailer		

# Example Files

chunk	description	entry	image #1	image #2
IHDR	image header	width	1024	1024
		height	768	645
		bit depth	8	8
		color type	2	6
		compression method	0	0
		filter method	0	0
		interlace method	0	0
gAMA	image gamma		2.2	2.2
sRGB	standard RGB color space		perceptual	perceptual
cHRM	primary chromaticities	white point x	0.31270	0.31269
		white point y	0.32900	0.32899
		red x	0.64000	0.63999
		red y	0.33000	0.33000
		green x	0.30000	0.21000
		green y	0.60000	0.71000
		blue x	0.15000	0.14999
		blue y	0.06000	0.59999
bKGD	background color		255 255 255	255 255 255
pHYS	physical pixel dimensions	pixels per unit x	7086	9449
		pixels per unit y	7086	9449
		unit specifier	meter	meter
IDAT	image data			
tEXt	textual information	date:create	2013-05-01T17:07:36-03:00	2013-05-01T17:07:36-03:00
tEXt	textual information	date:modify	2013-04-16T15:42:00-03:00	2013-04-16T16:29:15-03:00
tEXt	textual information	jpeg:colorspace	2	NA
tEXt	textual information	jpeg:sampling-factor	1x1,1x1,1x1	NA
IEEND	image trailer			

# Example Files

chunk	description	entry	image #1	image #2
IHDR	image header	width	1024	1024
		height	768	645
		bit depth	8	8
		color type	2	6
		compression method	0	0
		filter method	0	0
		interlace method	0	0
gAMA	image gamma		2.2	2.2
sRGB	standard RGB color space		perceptual	perceptual
cHRM	primary chromaticities	white point x	0.31270	0.31269
		white point y	0.32900	0.32899
		red x	0.64000	0.63999
		red y	0.33000	0.33000
		green x	0.30000	0.21000
		green y	0.60000	0.71000
		blue x	0.15000	0.14999
		blue y	0.06000	0.59999
bKGD	background color		255 255 255	255 255 255
pHYS	physical pixel dimensions	pixels per unit x	7086	9449
		pixels per unit y	7086	9449
		unit specifier	meter	meter
IDAT	image data			
tEXt	textual information	date:create	2013-05-01T17:07:36-03:00	2013-05-01T17:07:36-03:00
tEXt	textual information	date:modify	2013-04-16T15:42:00-03:00	2013-04-16T16:29:15-03:00
tEXt	textual information	jpeg:colorspace	2	NA
tEXt	textual information	jpeg:sampling-factor	1x1,1x1,1x1	NA
IEEND	image trailer			

camera resolution?

camera settings?

# Example Files

chunk	description	entry	image #1	image #2
IHDR	image header	width	1024	1024
		height	768	645
		bit depth	8	8
		color type	2	6
		compression method	0	0
		filter method	0	0
		interlace method	0	0
gAMA	image gamma		2.2	2.2
sRGB	standard RGB color space		perceptual	perceptual
cHRM	primary chromaticities	white point x	0.31270	0.31269
		white point y	0.32900	0.32899
		red x	0.64000	0.63999
		red y	0.33000	0.33000
		green x	0.30000	0.21000
		green y	0.60000	0.71000
		blue x	0.15000	0.14999
		blue y	0.06000	0.59999
bKGD	background color		255 255 255	255 255 255
pHYS	physical pixel dimensions	pixels per unit x	7086	9449
		pixels per unit y	7086	9449
		unit specifier	meter	meter
IDAT	image data			
tEXt	textual information	date:create	2013-05-01T17:07:36-03:00	2013-05-01T17:07:36-03:00
tEXt	textual information	date:modify	2013-04-16T15:42:00-03:00	2013-04-16T16:29:15-03:00
tEXt	textual information	jpeg:colorspace	2	NA
tEXt	textual information	jpeg:sampling-factor	1x1,1x1,1x1	NA
IEEND	image trailer			

camera resolution?

alpha matting?

camera settings?

# Example Files

chunk	description	entry	image #1	image #2
IHDR	image header	width	1024	1024
		height	768	645
		bit depth	8	8
		color type	2	6
		compression method	0	0
		filter method	0	0
		interlace method	0	0
gAMA	image gamma		2.2	2.2
sRGB	standard RGB color space		perceptual	perceptual
cHRM	primary chromaticities	white point x	0.31270	0.31269
		white point y	0.32900	0.32899
		red x	0.64000	0.63999
		red y	0.33000	0.33000
		green x	0.30000	0.21000
		green y	0.60000	0.71000
		blue x	0.15000	0.14999
		blue y	0.06000	0.59999
bKGD	background color		255 255 255	255 255 255
pHYS	physical pixel dimensions	pixels per unit x	7086	9449
		pixels per unit y	7086	9449
		unit specifier	meter	meter
IDAT	image data			
tEXt	textual information	date:create	2013-05-01T17:07:36-03:00	2013-05-01T17:07:36-03:00
tEXt	textual information	date:modify	2013-04-16T15:42:00-03:00	2013-04-16T16:29:15-03:00
tEXt	textual information	jpeg:colorspace	2	NA
tEXt	textual information	jpeg:sampling-factor	1x1,1x1,1x1	NA
IEEND	image trailer			

camera resolution?

alpha matting?

camera settings?

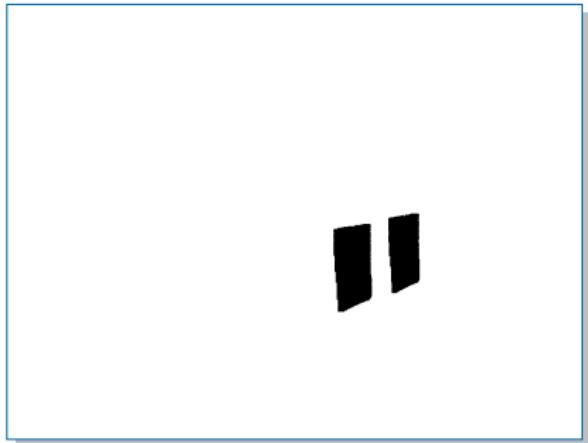
JPEG? Hm ...

# Color Type

- IHDR Color Type 2: RGB



7359205ed162a1f63a4537be468f9a71.png



# Color Type

- IHDR Color Type 2: RGB
- IHDR Color Type 6: RGB with alpha channel



010543abfb0db1e9aa1b24604336e0c.png

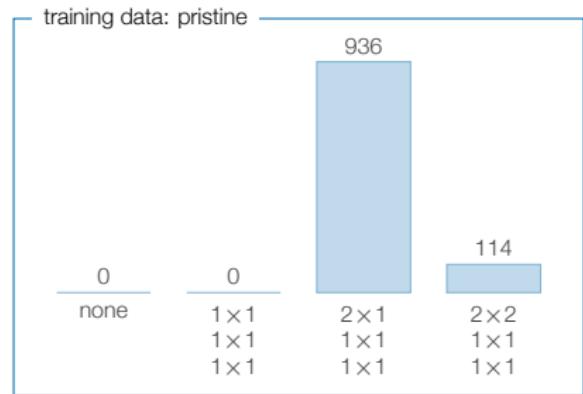


# JPEG Sampling Factor

- four instances of JPEG sampling factor chunks in the training set
  - ▷ missing
  - ▷  $1 \times 1, 1 \times 1, 1 \times 1$
  - ▷  $2 \times 1, 1 \times 1, 1 \times 1$
  - ▷  $2 \times 2, 1 \times 1, 1 \times 1$ 
    - ↙ luminance subsampling
    - ↘ chrominance subsampling

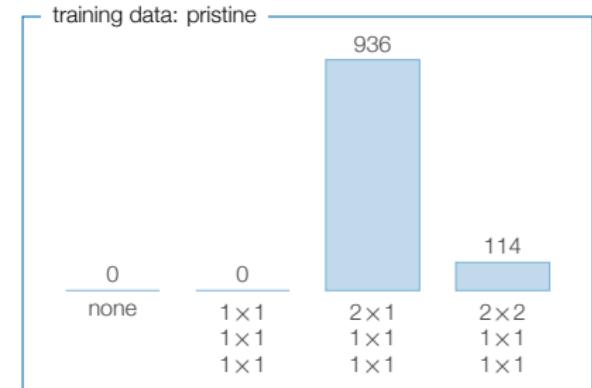
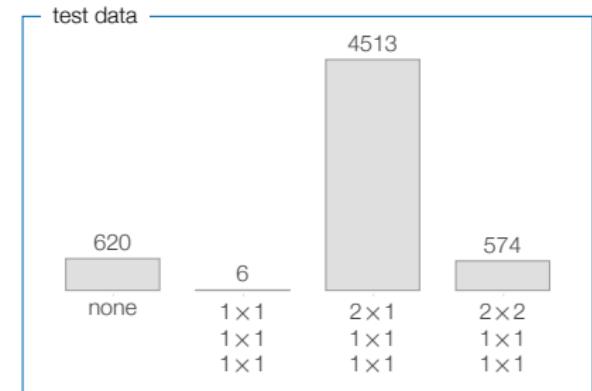
# JPEG Sampling Factor

- four instances of JPEG sampling factor chunks in the training set
  - ▷ missing
  - ▷  $1 \times 1, 1 \times 1, 1 \times 1$
  - ▷  $2 \times 1, 1 \times 1, 1 \times 1$
  - ▷  $2 \times 2, 1 \times 1, 1 \times 1$ 
    - ↙ luminance subsampling
    - ↘ chrominance subsampling



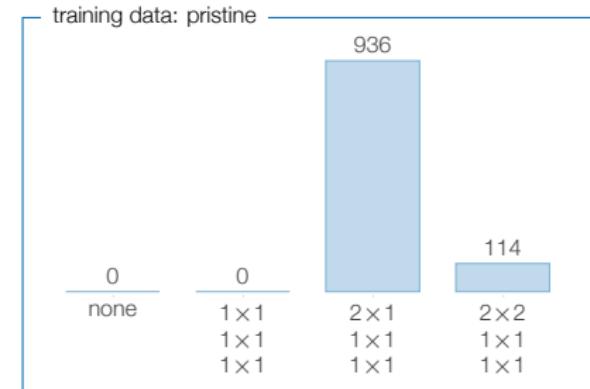
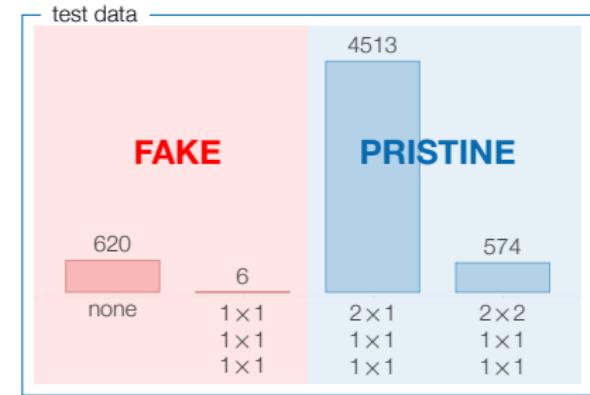
# JPEG Sampling Factor

- four instances of JPEG sampling factor chunks in the training set
- same instances in the test set



# JPEG Sampling Factor

- four instances of JPEG sampling factor chunks in the training set
- same instances in the test set
- Hypothesis:** All images in the test set with the 'jpeg:sampling-factor' field set to '1x1,1x1,1x1' or missing are manipulated.



# Visual Cross-Checking



67bd605bfb967e11d6fe481be8e21269.png

test images with  
sampling factor  
 $1 \times 1, 1 \times 1, 1 \times 1$

# Visual Cross-Checking



a5af9017a60e5f2206f69486694e3210.png

test images with  
sampling factor  
 $1 \times 1, 1 \times 1, 1 \times 1$

# Visual Cross-Checking



a5e5a9a9110a293f63e8cd10e908d489.png

test images with  
sampling factor  
 $1 \times 1, 1 \times 1, 1 \times 1$

# Visual Cross-Checking



a30ef361e8eba3e176989dcb7bb95056.png

test images with  
sampling factor  
 $1 \times 1, 1 \times 1, 1 \times 1$

# Visual Cross-Checking



a3d2ed17f91f008ea41951795545eba7.png

test images with  
sampling factor  
 $1 \times 1, 1 \times 1, 1 \times 1$



# Visual Cross-Checking

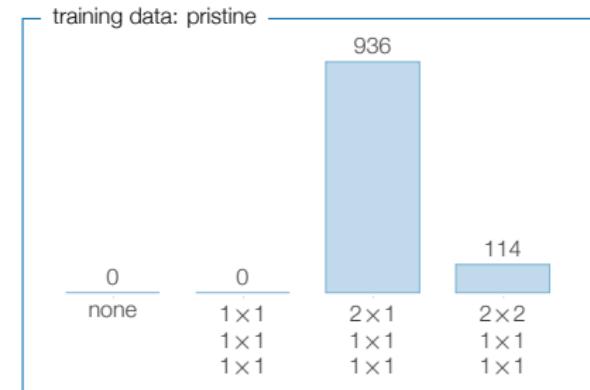
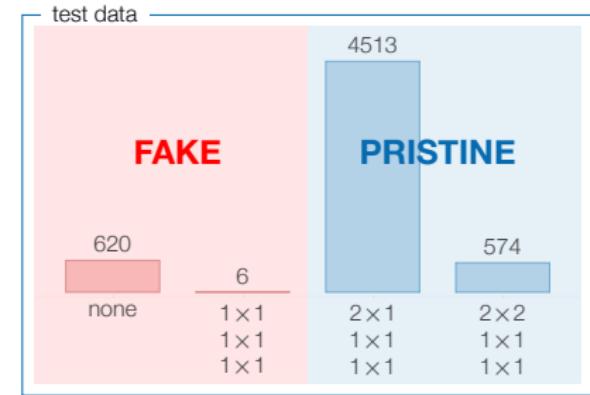


a79e69e6960588007f647bcd6a599f66.png

test images with  
sampling factor  
 $1 \times 1, 1 \times 1, 1 \times 1$

# JPEG Sampling Factor

- four instances of JPEG sampling factor chunks in the training set
- same instances in the test set
- Hypothesis:** All images in the test set with the ‘jpeg:sampling-factor’ field set to ‘1x1,1x1,1x1’ or missing are manipulated.



# Result

The screenshot shows a web browser window for the "Submission System" at the URL [ifc.recod.ic.unicamp.br/fc.submission/#dashboard](http://ifc.recod.ic.unicamp.br/fc.submission/#dashboard). The page displays results for the "1st Image Forensics Challenge".

The dashboard includes a sidebar with navigation links: Dashboard, My profile, My team, Phase #1, and Phase #2. The main content area shows the "Top teams (Phase #1)" table:

#	Leader	Team	Score
1	Thomas Glocé Matthias Kirchner Christian Riess	detector	1.000

# Summary

- Metadata—if present—is a forensically relevant part of a digital image file, independent of the file format.

# Summary

- Metadata—if present—is a forensically relevant part of a digital image file, independent of the file format.
- Different software tools yield different selections and sequences of PNG chunks.
- Database was most likely generated with ImageMagick's `convert`.

# Summary

- Metadata—if present—is a forensically relevant part of a digital image file, independent of the file format.
- Different software tools yield different selections and sequences of PNG chunks.
- Database was most likely generated with ImageMagick's `convert`.
- Option `-strip` suppresses all ancillary PNG chunks.

# Summary

- Metadata—if present—is a forensically relevant part of a digital image file, independent of the file format.
- Different software tools yield different selections and sequences of PNG chunks.
- Database was most likely generated with ImageMagick's `convert`.
- Option `-strip` suppresses all ancillary PNG chunks.



**Thank You!**