Towards learned color representations for image splicing detection

Benjamin Hadwiger*, Daniele Baracchi†, Alessandro Piva†, Christian Riess*

* IT Security Infrastructures Lab,
Friedrich-Alexander University Erlangen-Nürnberg (FAU)

† Dipartimento di Ingegneria dell'Informazione,
Università degli Studi di Firenze (UNIFI)

ICASSP, Brighton, May 2019

Did these events really occur?

Images based on
MS COCO Database

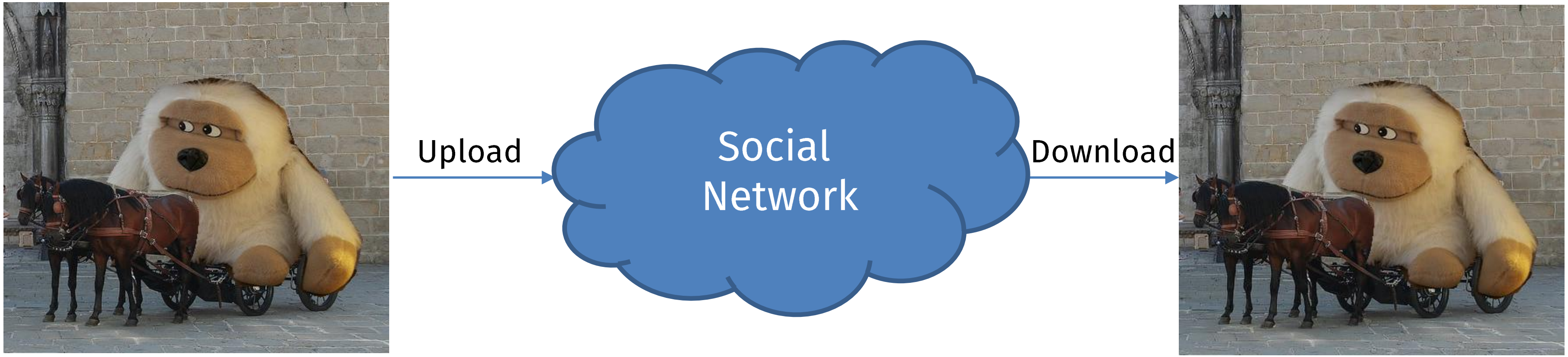
Multimedia Forensics – Goals and Methods

Important goal of multimedia forensics:
Determine authenticity of images

Typical approaches:
Exploit high frequent (HF) image statistics, e.g.

- Camera fingerprint
- Noise statistics
- Compression artifacts
- Resampling artifacts

The effect of Social Networks



The effect of Social Networks

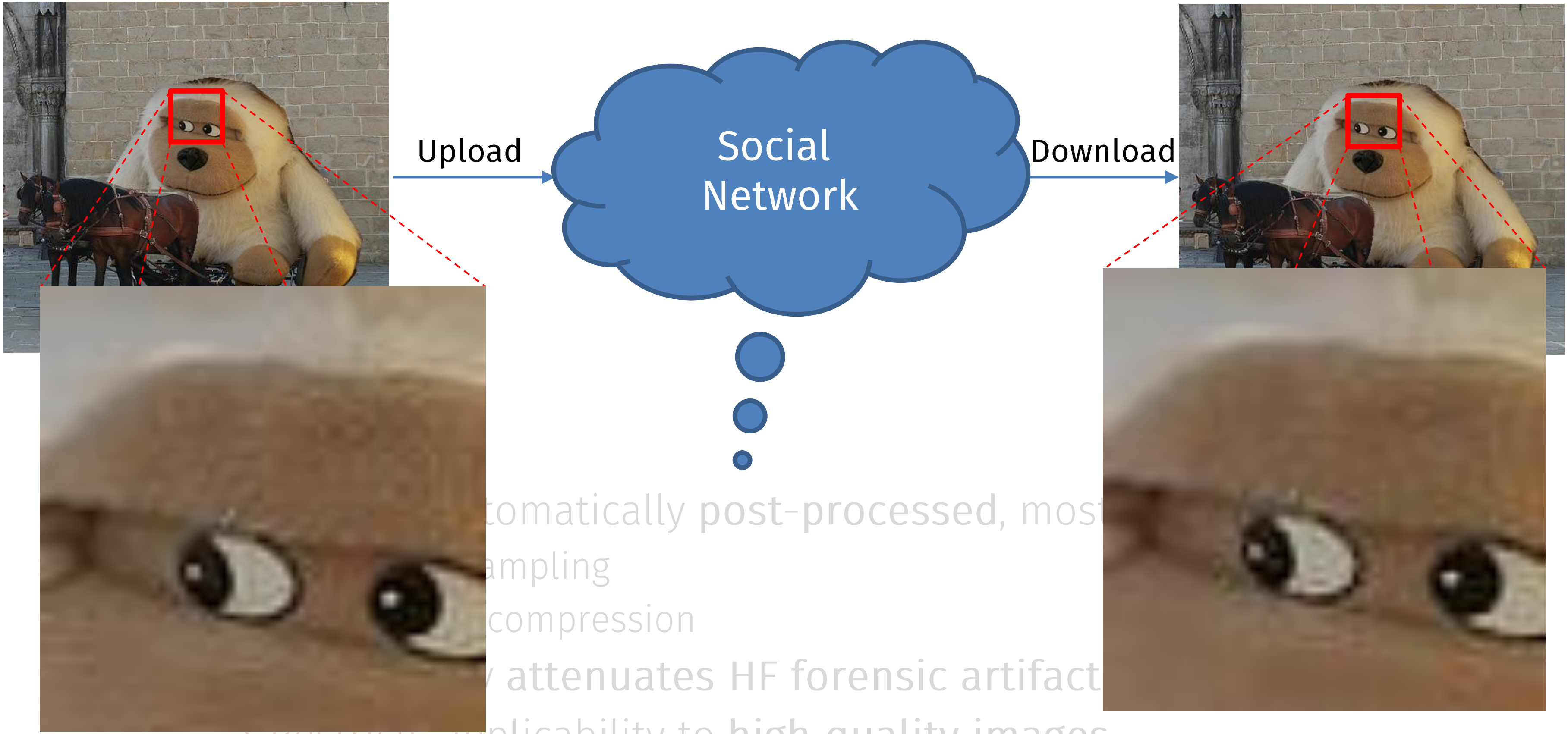
Images are automatically post-processed, most notably:

- Downsampling
- JPEG recompression

→ Significantly attenuates HF forensic artifacts

→ Restricts applicability to high quality images

The effect of Social Networks



…tomatically post-processed, most…

…ampling

…compression

…attenuates HF forensic artifact…

→ Restricts applicability to high quality images

Towards robust manipulation detection

Can we detect manipulations independent of the image quality?

Towards robust manipulation detection

Can we detect manipulations independent of the image quality?

Physics-based cues are often remarkably robust against post-processing

We explore a novel cue based on the color formation of an image

Towards robust manipulation detection

Can we detect manipulations **independent of the image quality**?

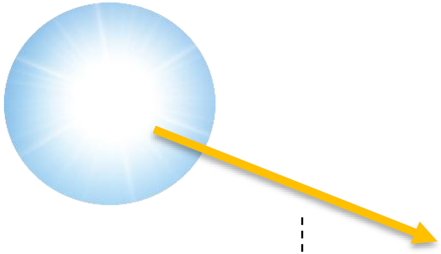
Physics-based cues are often **remarkably robust** against post-processing

We explore a **novel cue** based on the **color formation** of an image

Images based on MIT-Adobe 5k Database

Background: color image formation

Light source

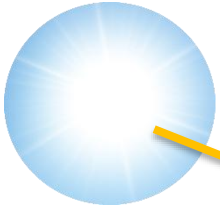


Spectral illuminant density: $e(\lambda)$

$e(\lambda)$

Background: color image formation

Light source



Spectral illuminant density: $e(\lambda)$

Scene

Spectral reflectance: $r(\lambda)$

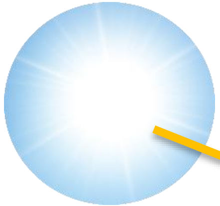
$e(\lambda)$

$e(\lambda) \cdot r(\lambda)$

Background: color image formation

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

Light source



Spectral illuminant density: $\mathrm{e}(\lambda)$

$\mathrm{e}(\lambda)$

Scene

Spectral reflectance: $\mathrm{r}(\lambda)$

$\mathrm{e}(\lambda) \cdot \mathrm{r}(\lambda)$

Camera

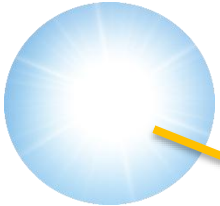
In-camera processing

- Camera sensitivity: $\vec{\mathrm{c}}\,(\lambda)$
- White balancing
  Color transformation
  etc.

$\left.\begin{array}{l}\text{White balancing}\\\text{Color transformation}\\\text{etc.}\end{array}\right\}\ \Omega$

Background: color image formation

Light source



Spectral illuminant density: $e(\lambda)$

$e(\lambda)$

Scene

Spectral reflectance: $r(\lambda)$

$e(\lambda) \cdot r(\lambda)$

Camera

In-camera processing

- Camera sensitivity: $\vec{c}\,(\lambda)$
- White balancing
  Color transformation
  etc.

$\left.\begin{array}{l}\text{White balancing}\\ \text{Color transformation}\\ \text{etc.}\end{array}\right\}\ \Omega$

$$\Omega\left(\int_{\Lambda} e(\lambda) \cdot r(\lambda) \cdot \vec{c}(\lambda)\, d\lambda\right)$$

Color image

Proposed method: Idea

$$\vec{\mathrm{I}} = \Omega\left(\int_{\Lambda} \mathrm{e}(\lambda) \cdot \mathbf{r}(\boldsymbol{\lambda}) \cdot \vec{\mathrm{c}}(\lambda)\, \mathrm{d}\lambda\right)$$

e: illuminant sp. density
Ω: in-camera processing
$\vec{\mathrm{c}}$: sp. camera sensitivity
r: spectral reflectance
$\vec{\mathrm{I}}$: image intensity

How can we control the spectral reflectance $\mathbf{r}(\boldsymbol{\lambda})$?

Proposed method: Idea

$$\vec{I} = \Omega\left(\int_{\Lambda} e(\lambda) \cdot \mathbf{r}(\boldsymbol{\lambda}) \cdot \vec{c}(\lambda)\, d\lambda\right)$$

e: illuminant sp. density
Ω: in-camera processing
$\vec{c}$: sp. camera sensitivity
r: spectral reflectance
$\vec{I}$: image intensity

How can we control the spectral reflectance $\mathbf{r}(\boldsymbol{\lambda})$?

Place Macbeth ColorChecker in image

Image source: NUS Database

Proposed method: Idea

$$\vec{I} = \Omega\left(\int_{\Lambda} e(\lambda) \cdot \mathbf{r}(\boldsymbol{\lambda}) \cdot \vec{c}(\lambda)\, d\lambda\right)$$

e: illuminant sp. density
Ω: in-camera processing
$\vec{c}$: sp. camera sensitivity
r: spectral reflectance
$\vec{I}$: image intensity

How can we control the spectral reflectance $\mathbf{r}(\boldsymbol{\lambda})$?

Image source: NUS Database

Place Macbeth ColorChecker in image

↓

Observed colors of patches characterize imaging conditions $\mathbf{e}$, Ω and $\mathbf{c}$

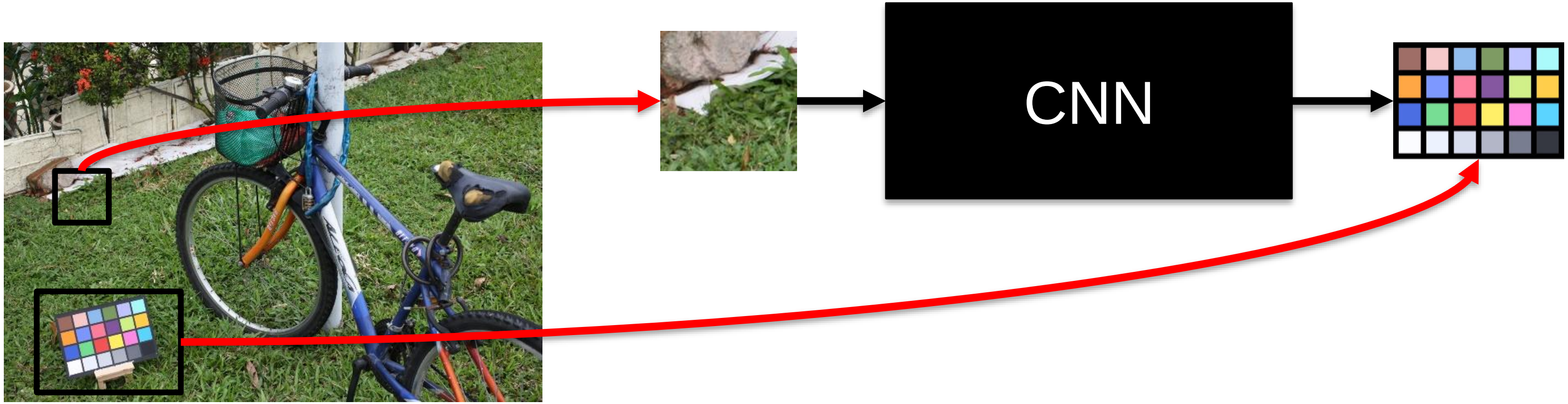
↓

Assume consistency of $\mathbf{e}$, Ω and $\mathbf{c}$ in pristine image

Proposed method: Learning the color descriptor

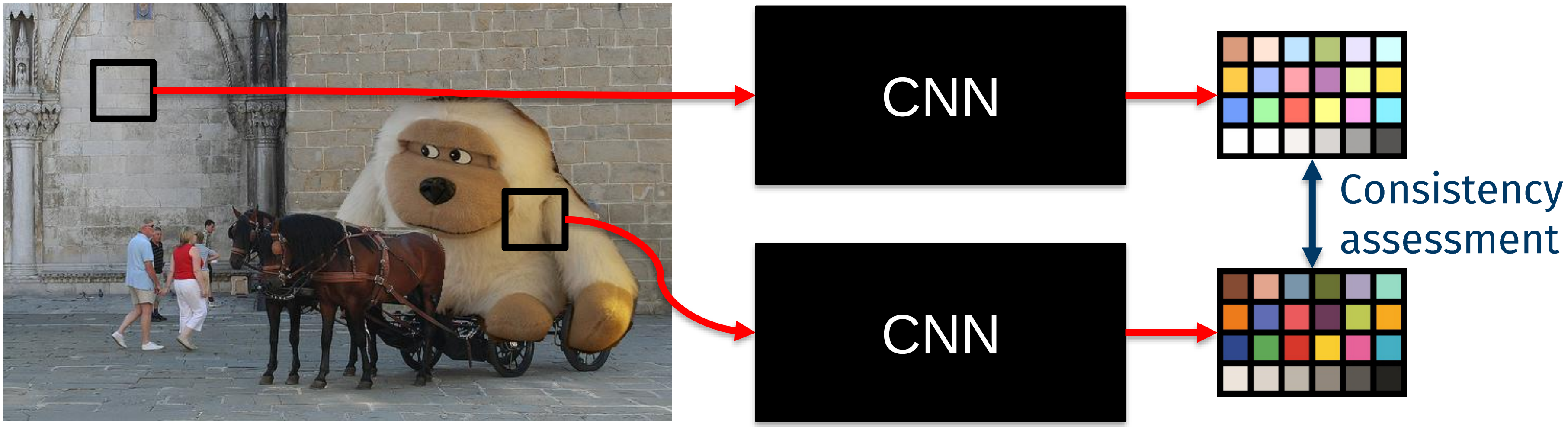
FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

1. Train a CNN to locally estimate the observed colors
   - → Covariant with respect to imaging conditions
   - → Invariant with respect to reflectance



Proposed method: Consistency assessment

2. Classify consistency of local estimates



Discriminability of the learned color descriptor

How well do the learned color features characterize the image provenance of a patch?

Discriminability of the learned color descriptor

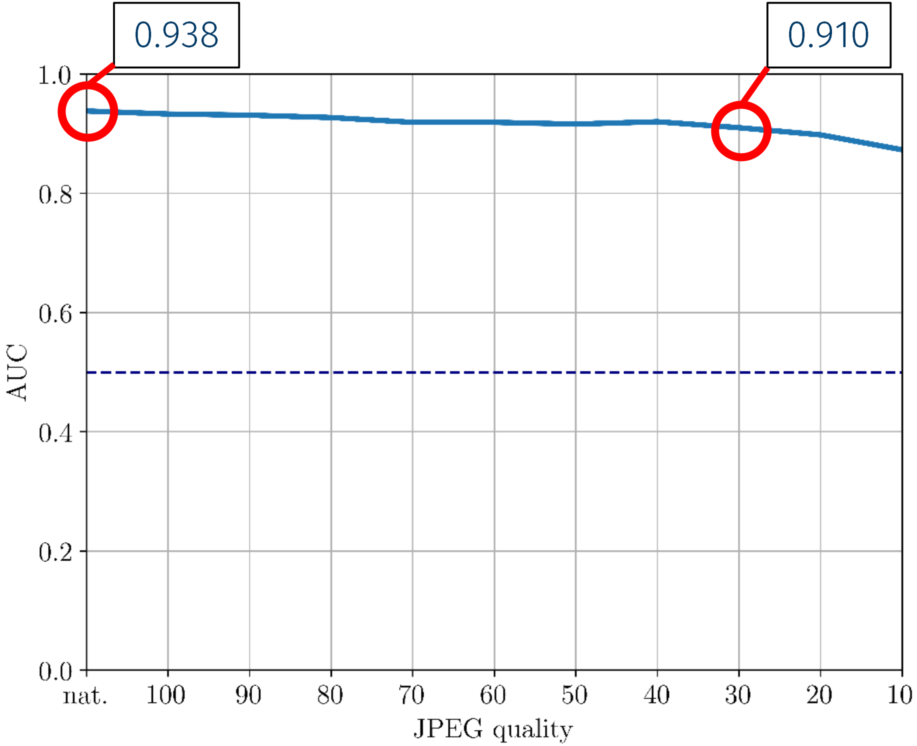
How well do the learned color features characterize the image provenance of a patch?

- Extract non-overlapping patches from test images
- Randomly split patches into training / test set
- Train a Random Forest to classify image provenance
- Repeat for increasingly stronger compressions

Discriminability of the learned color descriptor

How well do the learned color features characterize the image provenance of a patch?

- Extract non-overlapping patches from test images
- Randomly split patches into training / test set
- Train a Random Forest to classify image provenance
- Repeat for increasingly stronger compressions



Consistency assessment with the color descriptor

A more realistic scenario: How well can we assess whether different regions in an image are consistent?

Consistency assessment with the color descriptor

A more realistic scenario: How well can we assess whether different regions in an image are consistent?

- Create test splices by combining images from the same scene from different cameras
- Assess patch-based consistency of reference and target region
- Repeat for increasingly stronger compressions

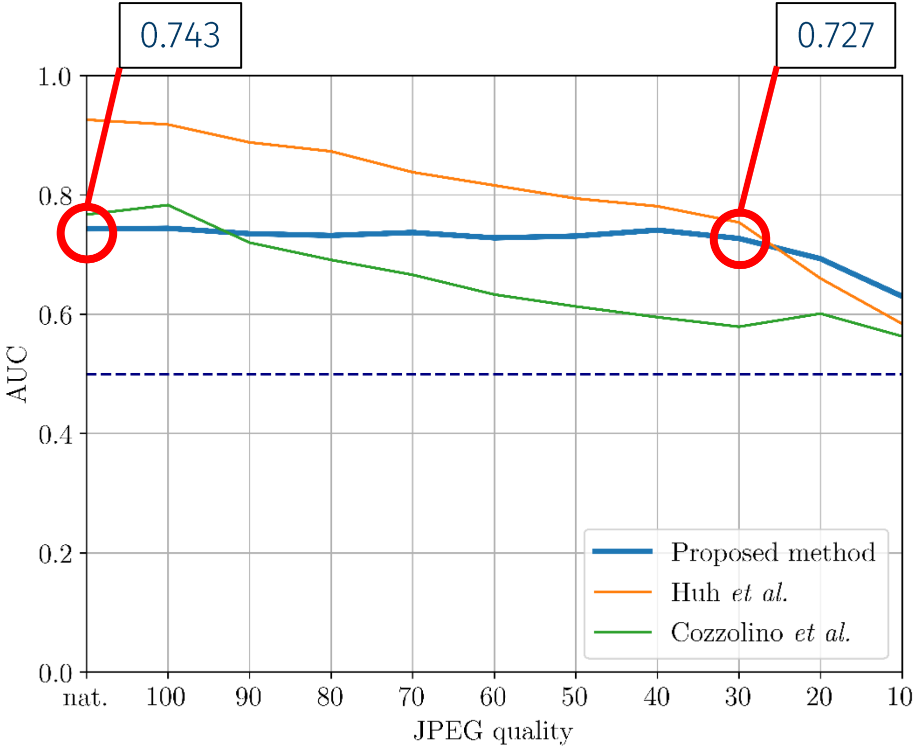
Image based on Dresden Image Database

Consistency assessment with the color descriptor

A more realistic scenario: How well can we assess whether different regions in an image are consistent?

- Create test splices by combining images from the same scene from different cameras
- Assess patch-based consistency of reference and target region
- Repeat for increasingly stronger compressions

Image based on Dresden Image Database



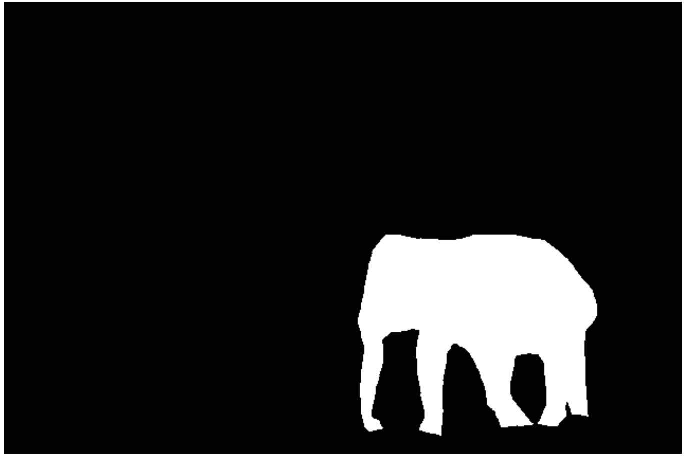
Huh et al.: *"Fighting Fake News: Image Splice Detection via Learned Self-Consistency"*, ECCV '18
Cozzolino et al.: *"Splicebuster: A new blind image splicing detector"*, WIFS '15

Qualitative Results

Qualitative Results

Qualitative Results (*cont.*)

Qualitative Results (*cont.*)

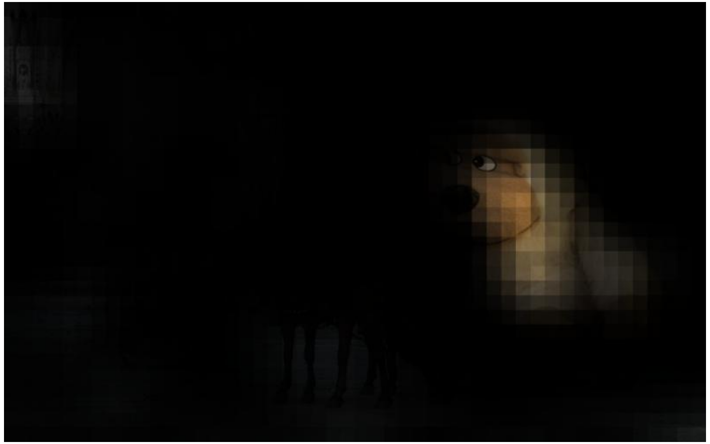


Qualitative Results (*cont.*)

Qualitative Results (*cont.*)

Results (*cont.*)

Qualitative Results (*cont.*)



Summary

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

Conclusion

- We presented a novel cue based on color image formation
- We demonstrated remarkable robustness against JPEG compression
- Promising to work in low quality settings

Ongoing work

- Incorporate prior knowledge on camera
- Perform consistency assessment using Siamese network

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

Thank you!

Questions?