New Perspectives on Digital Media Integrity

Walter J. Scheirer

Department of Computer Science and Engineering
University of Notre Dame

“The battle for the mind of North America
will be fought in the video arena”

Brian Oblivion, *Videodrome*

Computer Vision @ND

Videodrome (1983)

COMPUTER VISION @ND

Deep Fakes

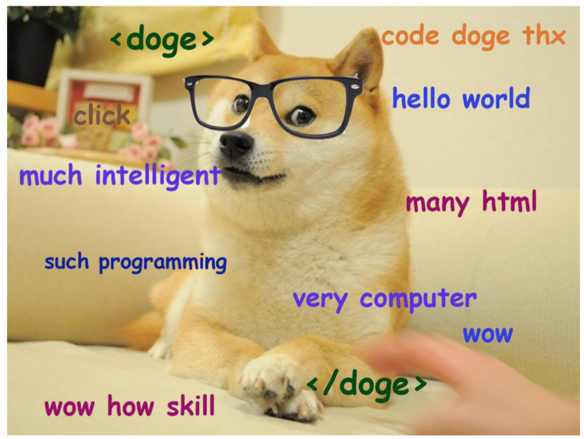
COMPUTER VISION @ND

ISIS Execution Videos (Fake)

https://www.csmonitor.com/World/Global-News/2015/0124/ISIS-video-Is-this-Japanese-beheading-real-or-fake

The Intersection of Politics and Entertainment: Internet Memes

What is a Meme?

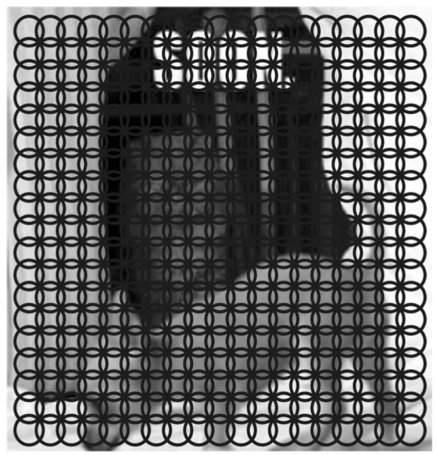


Forstall and Scheirer, *Quantitative Intertextuality*, Springer-Nature 2019

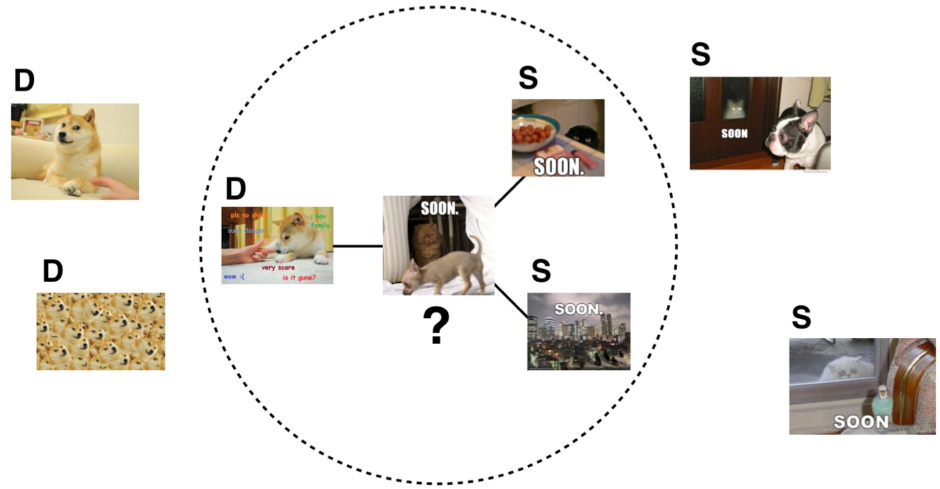
Memes, Visual Style, and Semantic Analysis

A naive approach:

Original



Dense grid of HOG descriptors



K-Nearest Neighbors

Related Work

Memes as Cultural Artifacts

L. Shifman, *Memes in Digital Culture*, MIT Press, 2014.

Visual Style in Computer Vision

J. Jupp and J. Gero, “Let’s look at style: Visual and spatial representation and reasoning in design,” In: S. Argamon, K. Burns, S. Dubnov (eds.) The Structure of Style: Algorithmic Approaches to Understanding Manner and Meaning. Springer-Verlag, 2010.

Q. Wu, H. Cai, P. Hall, “Learning graphs to model visual objects across different depictive styles,” ECCV, 2014.

Y. Lee, A. Efros, M. Hebert, “Style-aware mid-level representation for discovering visual connections in space and time,” IEEE ICCV, 2013.

Empirical Effects of Manipulated Media

Shen et al., “Fake images: The effects of source, intermediary, and digital media literacy on contextual assessment of image credibility online,” New Media & Society, 21 (2), 2018.

COMPUTER VISION @ND

Memes and Political Interference

Brazil

Legend

Bolsonaro's actual fraudulent employee: his brain

You better get used to him

The girl you like

Her father

Her brother

The guy she likes

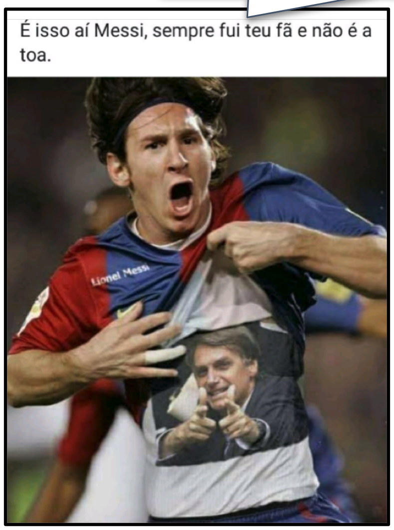
Her ex

You

COMPUTER VISION @ND

Brazil

That's it, Messi, I'm not your supporter for nothing...



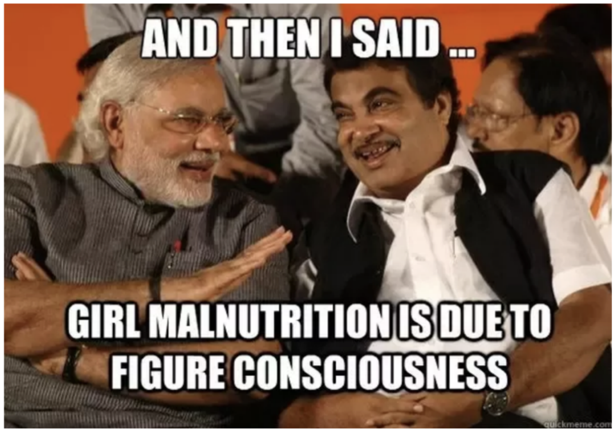
Manipulation to claim support

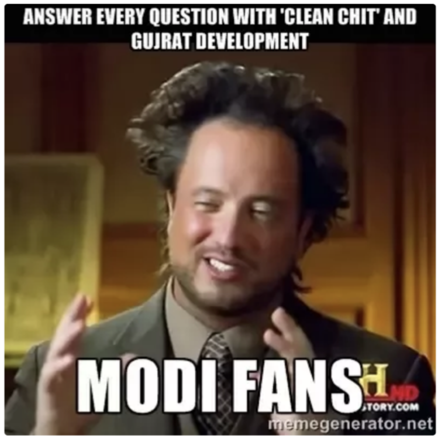
Original

A Federal University 'class'
We don't want to pay for it!

Image repurposing

India



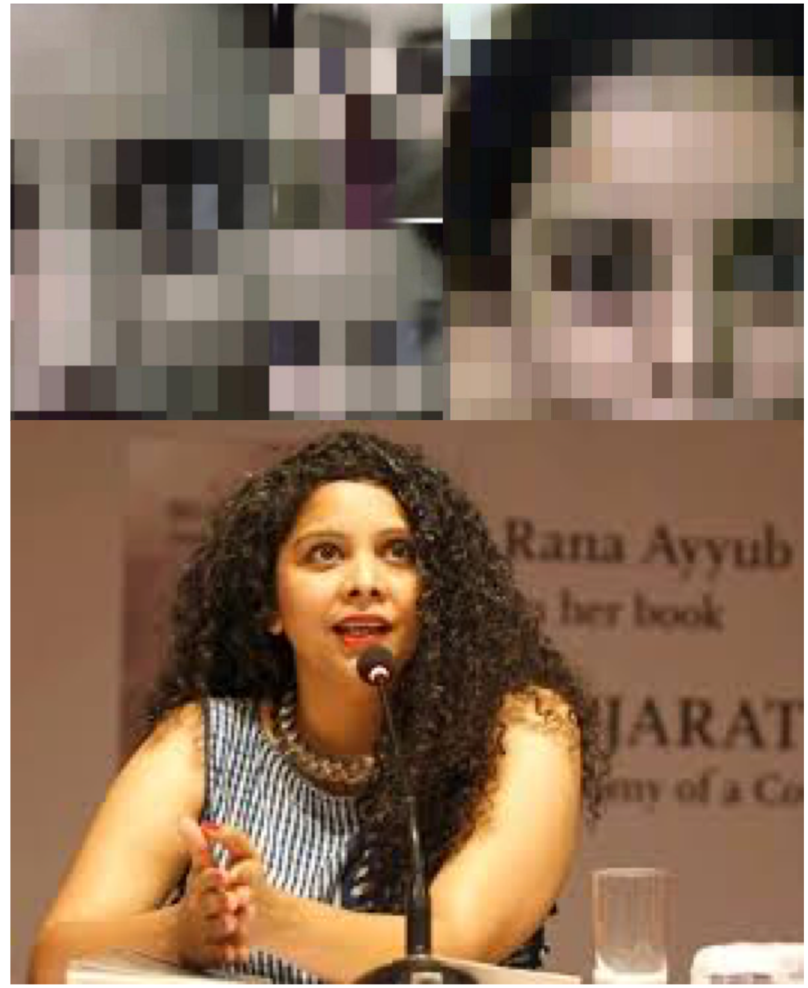


COMPUTER VISION @ND

Deep Fakes as a Political Tool

2018: BJP supporters target investigative journalist Rana Ayyub with a doctored pornographic video containing her face

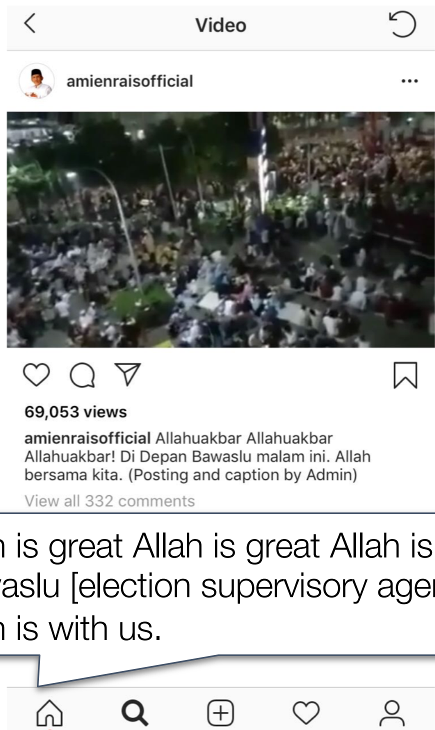
https://www.huffingtonpost.co.uk/entry/deepfake-porn_uk_5bf2c126e4b0f32bd58ba316



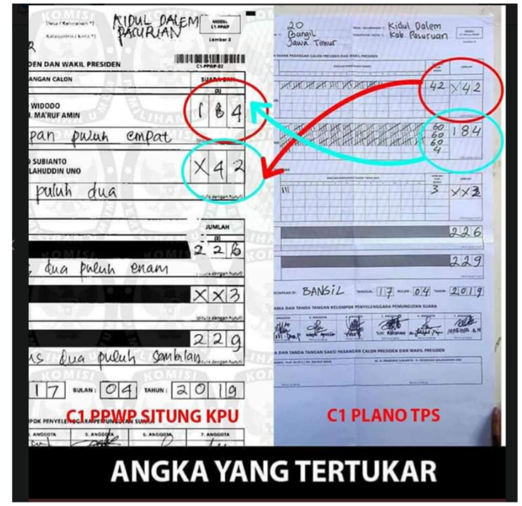
Credit: Huffington Post and Facebook

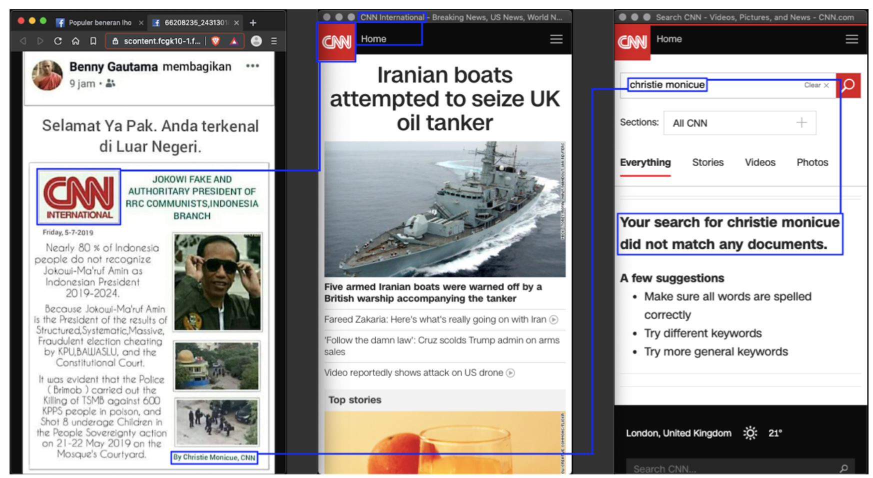
COMPUTER VISION @ND

Indonesia



Allah is great Allah is great Allah is great! at Bawaslu [election supervisory agency] tonight. Allah is with us.





COMPUTER VISION AND

Case Study: 2019 Indonesian Elections

Indonesian Presidential Election Overview

A voting station in Samarinda, East Kalimantan ⓒ BY-SA 4.0 Ezagren

- Direct general election
- Fourth presidential election
- Rematch of 2014 (five year terms)
- Joko Widodo re-elected with over 55% of the vote

COMPUTER VISION @ND

The Candidates

Joko “Jokowi” Widodo

- Incumbent
- PDI-P candidate (center-left party)
- Appealed to younger voters
- “Man of the people”

Prabowo Subianto

- Challenger
- Gerindra candidate (right populist party)
- Strongly Islamic
- Military background (Ties to Suharto dictatorship)

COMPUTER VISION @ND

Indonesian Election Aftermath

Prabowo claims a count fraud

Protests/Riots in Jakarta May 21-22, 8 dead

Indonesia restricts access to social media

https://www.straitstimes.com/asia/se-asia/old-age-poor-health-caused-deaths-of-poll-administrators-indonesia-government

COMPUTER VISION @ND

Survey of the Meme-space

Millions of images collected during the course of the election

Twitter
- ~170,000 images from 14 different sources

Instagram
- ~1.9 million images from 20 different sources

COMPUTER VISION @ND

The Data

Image sources identified by our partner in Indonesia, CekFakta

Images span from May 31st, 2018 to May 31st, 2019

Collected from 2 users and 12 popular hashtags (Twitter)

Computer Vision @ND

COMPUTER VISION @ND

Buzzer Accounts

Social media propaganda

Firms use fake accounts to generate “buzz”

Buzzers use upwards of 250 fake accounts each

https://www.reuters.com/article/us-indonesia-election-socialmedia-insigh/in-indonesia-facebook-and-twitter-are-buzzer-battlegrounds-as-elections-loom-idUSKBN1QU0AS

COMPUTER VISION @ND

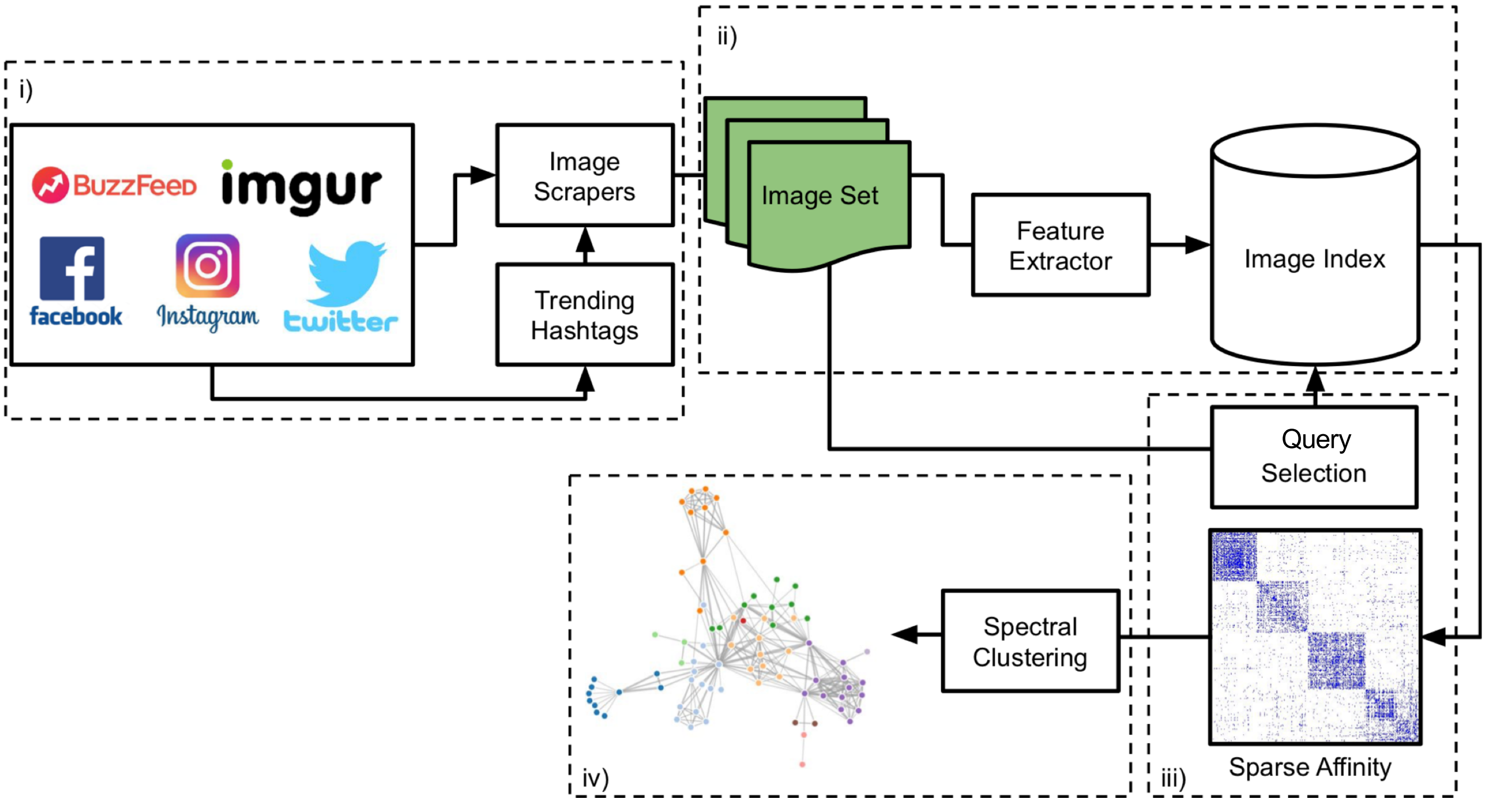
Finding Trending Memes: Motif Mining

Objectives for an early warning system

- Real-time processing pipeline with content flagging
- Assess meme landscape with respect to political actors
- Allow humans to gauge the threat of violence
- Scale to the order of billions of images per day
- Match composite images to other composite and donor images

COMPUTER VISION @ND

Motif Mining Pipeline



COMPUTER VISION @ND

Query Selection: Detecting Edited Objects

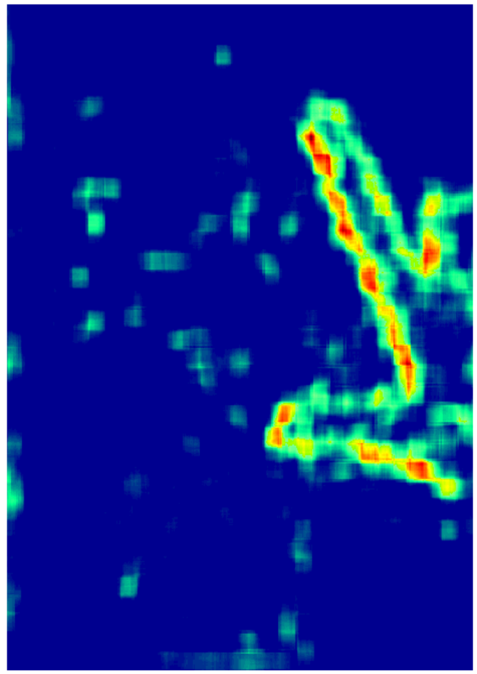
Use meme images to search for related meme content

- Option 1: Randomly sample images from known buzzers
- Option 2: Detect new memes via the identification of edits

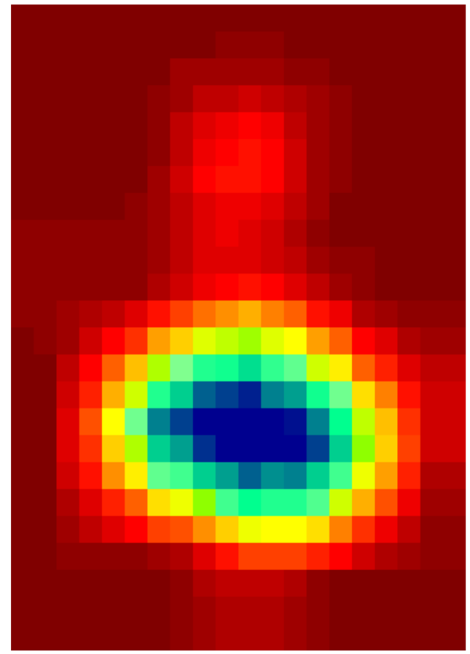
http://reveal-mklab.iti.gr/reveal/

COMPUTER VISION @ND

Detecting Image Edits



Farid 2009

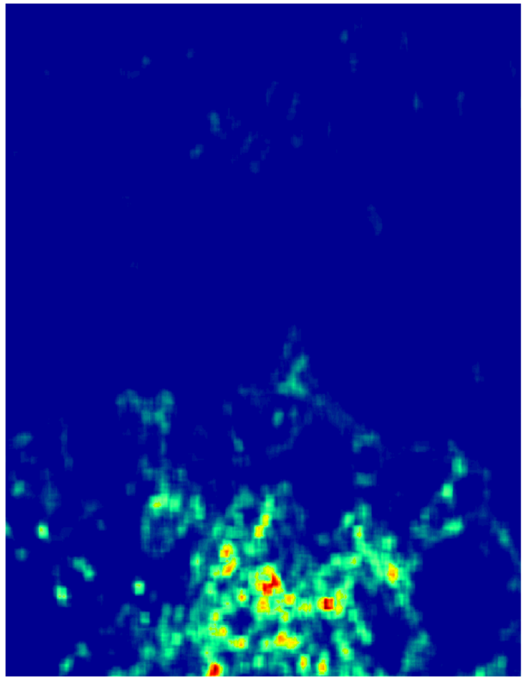


Iakovidou et al. 2018

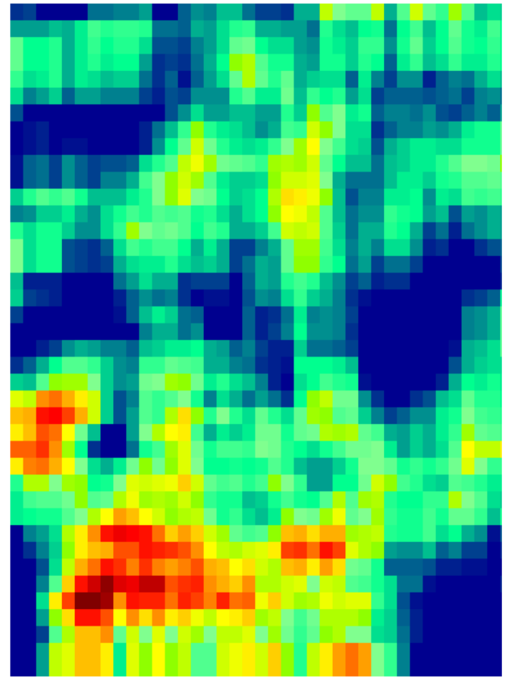
COMPUTER VISION @ND

H. Farid, *Exposing digital forgeries from JPEG ghosts*. IEEE T-IFS, 2009.
Iakovidou et al. *Content-aware detection of JPEG grid inconsistencies for intuitive image forensics*. JVCIR, 2018.

Detecting Image Edits



Farid 2009

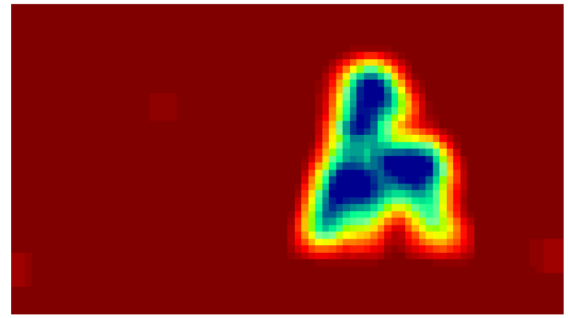


Weihai et al. 2009

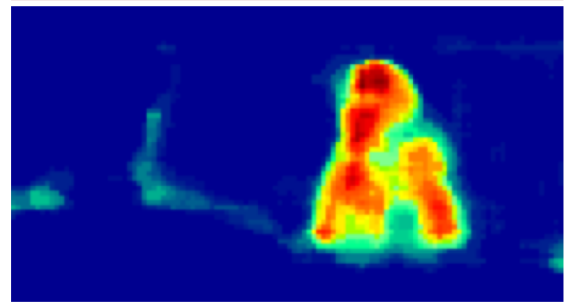
H. Farid, *Exposing digital forgeries from JPEG ghosts*. IEEE T-IFS, 2009.
Weihai et al., *Passive detection of doctored JPEG image via block artifact grid extraction*. Signal Processing, 2009.

COMPUTER VISION @ND

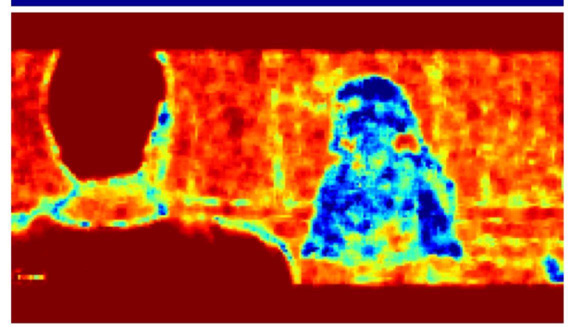
Detecting Deep Fakes



Iakovidou et al. 2018



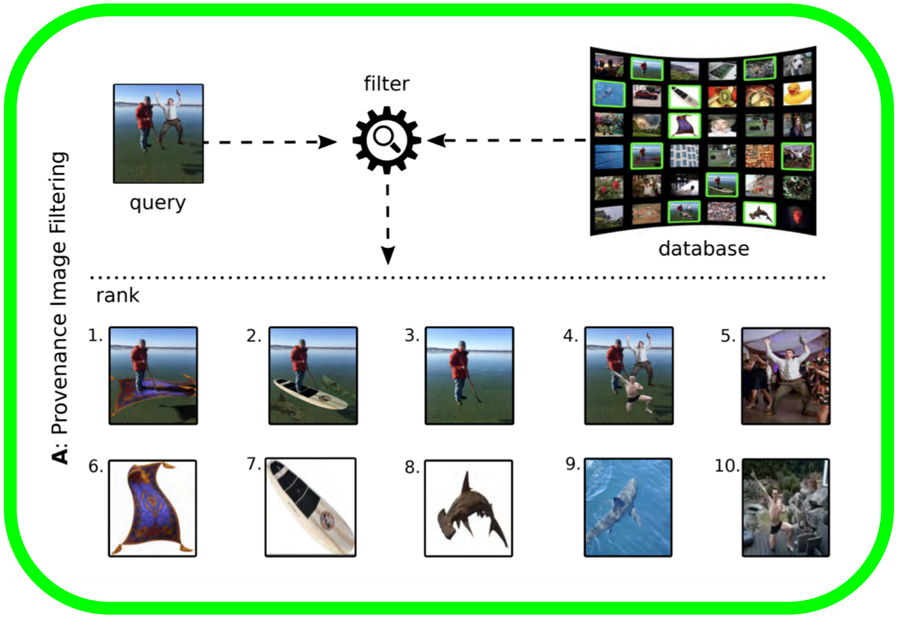
Mahdian et al. 2009

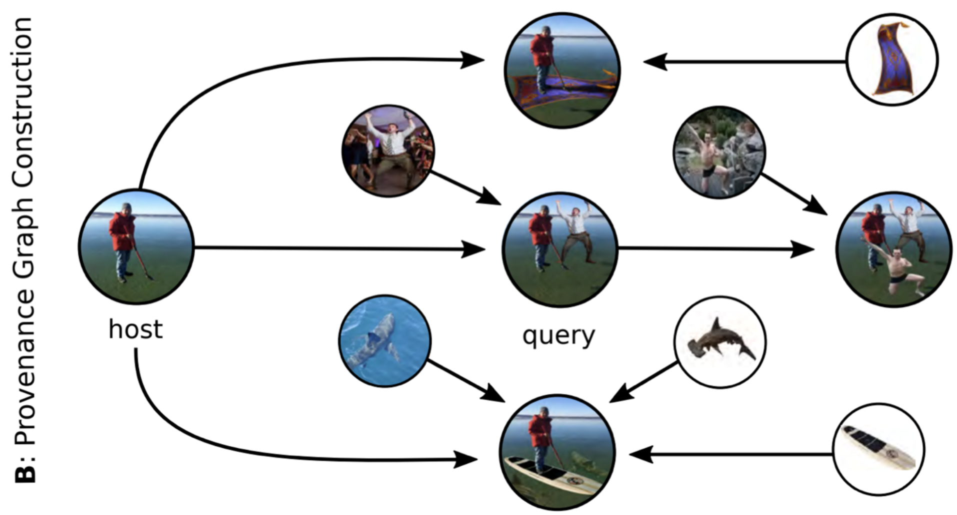


Weihai et al. 2009

Iakovidou et al. *Content-aware detection of JPEG grid inconsistencies for intuitive image forensics*. JVCIR, 2018
Mahdian et al. *Using noise inconsistencies for blind image forensics*. Elsever IVC, 2009
Weihai et al. *Passive detection of doctored JPEG image via block artifact grid extraction*. Signal Processing, 2009

Adapting a provenance pipeline to motif mining

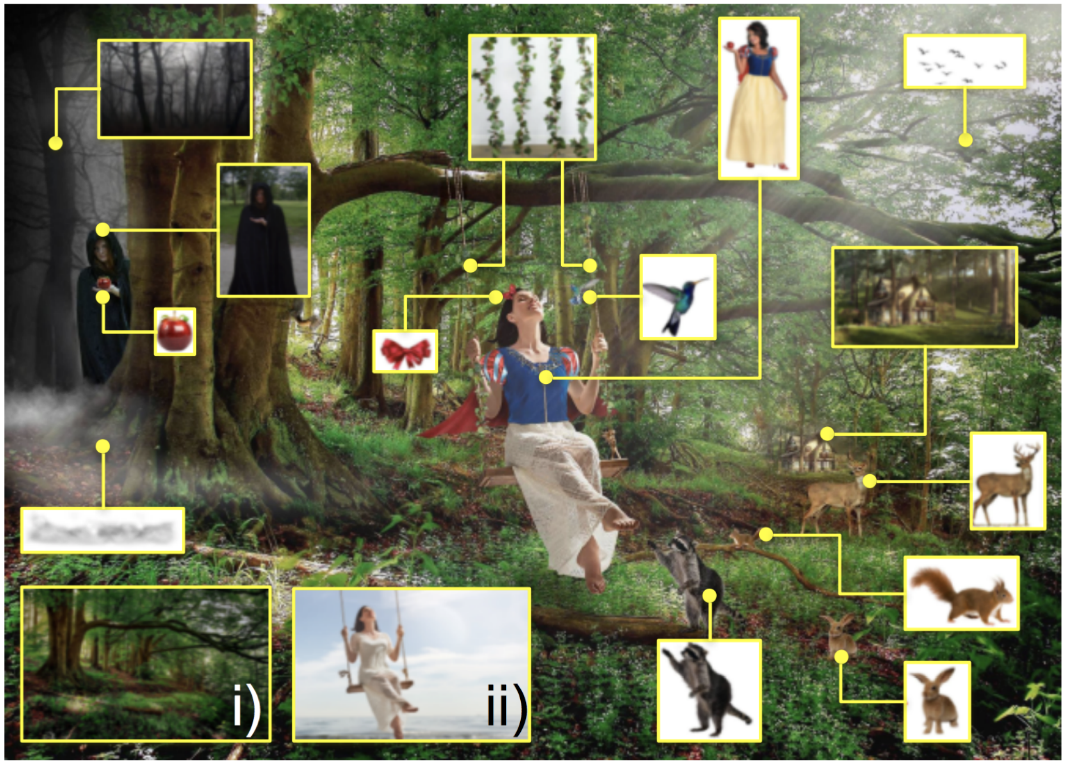




Moreira, D., et al. *Image Provenance Analysis at Scale*. IEEE T-IP, 2018

COMPUTER VISION @ND

Add the ability to localize and match small objects



J. Brogan et al. *Needle in a Haystack: A Framework for Seeking Small Objects in Big Datasets*.
Arxiv preprint (https://arxiv.org/abs/1903.10019), 2019

COMPUTER VISION @ND

Add a spectral clustering stage

Computer Vision @ND

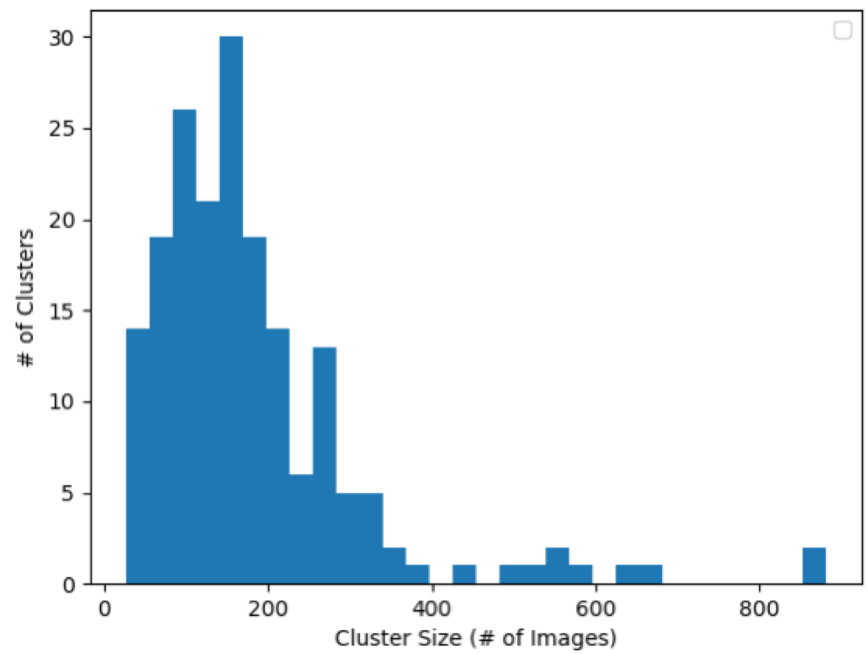
Motif Mining (Twitter Dataset)

- 174,328 images scraped from Twitter
- Image Index was built in 6.2 hours on a single GPU machine with 12 CPU cores
- 10,000 Random queries were performed for matrix generation in 6.5 hours
- **No deep learning was used**

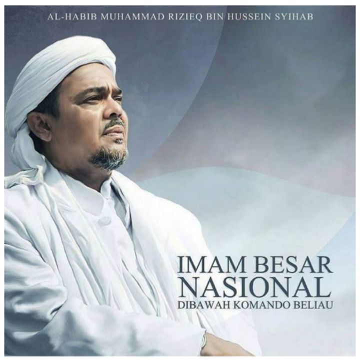
COMPUTER VISION @ND

Motif Mining (Twitter Dataset)

- 174328 total images
- 197 computed clusters
- Mean cluster size: 895 images
- Median cluster size: 156
- Minimum cluster size: 26
- Maximum cluster size: 36304



Automatic Motif Detection



Automatic Motif Detection

COMPUTER
VISION @ND

Automatic Motif Detection

Some Questions...

Q1: What capabilities do we still need for meme analysis?

Q2: What aspects of semantic analysis can improve the process?

Q3: What is a good process for data collection?

Q4: Who outside of the forensics community should we engage with?