



How to detect malicious use of video/photographic content online? Insights from a NATO Stratcom competition

Ricardo Ribeiro and Daniel Canedo
{rfribeiro, danielduartecanedo}@ua.pt

Institute of Electronics and Informatics Engineering of Aveiro (IEETA)
Department of Electronics and Telecommunications, University of Aveiro, Portugal
<http://wiki.ieeta.pt/wiki>



universidade de aveiro
theoria poiesis praxis



instituto de engenharia electrónica
e telemática de aveiro

Outline

- Our team
- Motivation
- Our aim
- Image forgery detection
- Scenes/objects recognition
- Fake news detection
- Future Work



Our team

- Ricardo Ribeiro, MSc
- Daniel Canedo, MSc
- Alina Trifan, PhD
- Prof. José L. Oliveira
- Prof. António Neves



Motivation

TIME

U.S. POLITICS WORLD TECH ENTERTAINMENT

WORLD • PAKISTAN

Fake Images of Acquitted Christian Woman Leaving Pakistan Prompt Death Threats



You M

Student
march
Hungar

by France
Sponsored

by Bookin
Sponsored

by TIME

.. .

INDEPENDENT

NEWS POLITICS VOICES FINAL SAY SPORT CULTURE VIDEO INDY/LIFE INDYBEST LONG READS AIDSFREE INDY100 VOUCHERS

Threatin: Los Angeles band reportedly faked Facebook fanbase to book UK shows

DSPT
DATA SCIENCE PORTUGAL

BBC

Sign in

News

Sport

Weather

Shop

Reel

Travel

More

Search



NEWS

Home Video World UK Business Tech Science Stories Entertainment & Arts Health World News TV More

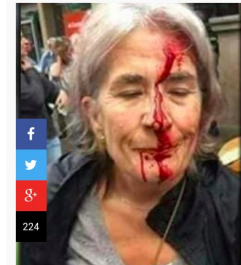
World Africa Asia Australia Europe Latin America Middle East US & Canada

Burned to death because of a rumour on WhatsApp

24

THE OBSERVERS

Film, Verify, Share



Faites circuler
que les gilets
jaunes sont
entré à l'Elysée.



Four ways to verify images linked to France's 'Yellow Vest' protests



JOIN US? SUBSCRIBE / REGISTER

LOGIN



Q

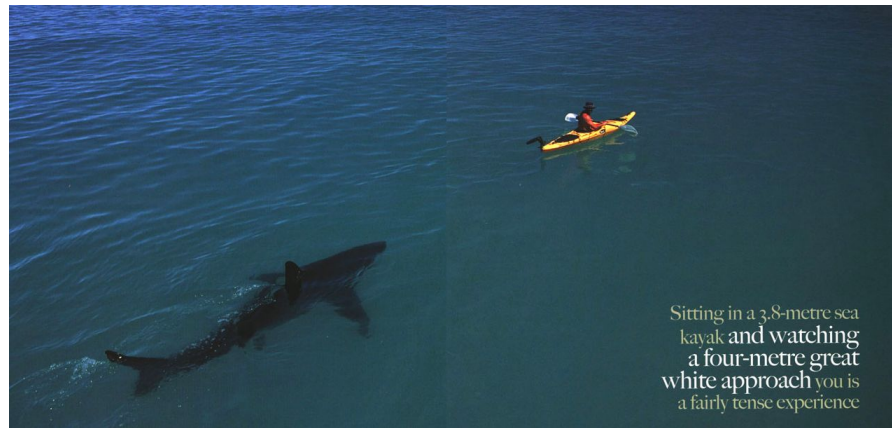


universidade de aveiro
theoria poiesis praxis



instituto de engenharia electrónica
e telemática de aveiro

Our aim



Original image - Africa Geographic Magazine, 2005

Our aim



Forgery detection



Extracted metadata



Associated text analysis



Combined analysis

Malicious



Harmless

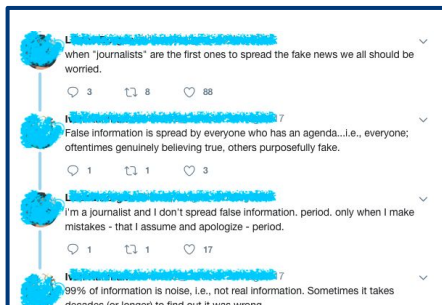
Methodology breakdown



Forgery detection



Extracted metadata



Associated text analysis



Combined analysis

Malicious

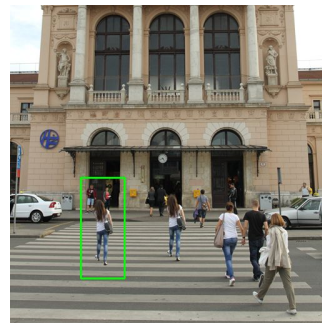


Harmless

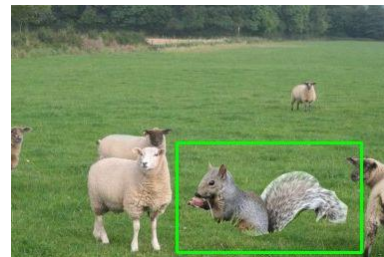
Forgery detection



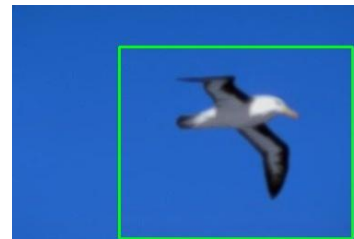
Copy-move



Splicing



Blur



How do we solve it?

→ Based on state of the art approaches for forgery detection (**Deep Learning**), we are able to detect:

- **Blur, JPEG compression, resampling,...**
- **Copy-move and splicing.**

Challenges:

→ Building a **model** that is able to:

- **Identify** if certain forgery makes the image/video malicious/extremist.
- **Protect** images/videos from manipulation.

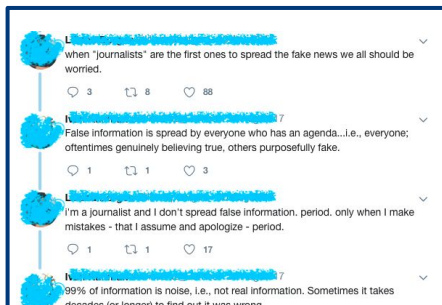
Methodology breakdown



Forgery detection



Extracted metadata



Associated text analysis



Combined analysis

Malicious



Harmless

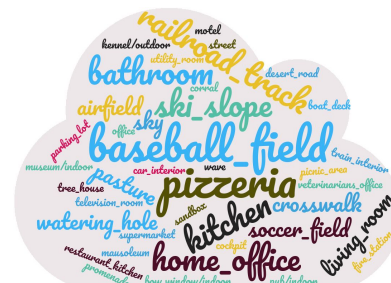
Recognition of scenes, objects and text



Scenes

Objects

Text



How do we solve it?

- Based on state of the art approaches for **image analysis**, we are able to:
- **Recognize** scenes, objects and text.
 - **Return text** containing every detection.

Challenges:

- Building a **model** that is able to:
- **Recognize** malicious/extremist scenes, objects and text in images/video.
 - **Organize** the text output into pre-defined **ontologies** and **concept hierarchies**.

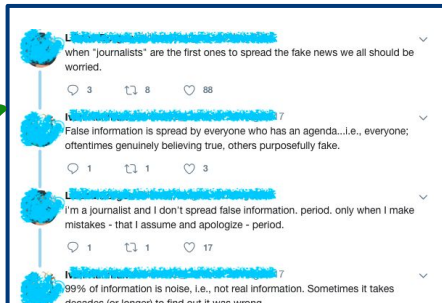
Methodology breakdown



Forgery detection



Extracted metadata



Associated text analysis



Combined analysis

Malicious



Harmless



Dataset

Statement: “The last quarter, it was just announced, **our gross domestic product was below zero**. Who ever heard of this? **It is never below zero**.”

Speaker: Donald Trump

Context: presidential announcement speech

Label: Pants on Fire

Justification: According to Bureau of Economic Analysis and National Bureau of Economic Research, the growth in **the gross domestic product has been below zero 42 times over 68 years**.

Statement: “Building a wall on the U.S. - Mexico border **will take literally years**.”

Speaker: Rick Perry, Governor of Texas

Context: Radio interview

Label: True

Justification: The U.S. - Mexico border length is around **3 145 km**. Indeed a possible construction of the wall would **take more than one year**, which is equivalent with “**literally years**”.

*Wang, William Yang. Liar, Liar Pants on Fire: A new benchmark dataset for fake news detection. *arXiv preprint arXiv:1705.00648* (2017).

How do we solve it?

- We considered this a text mining challenge:
- **machine learning**
 - **information retrieval**

Model	Accuracy
CNNs (text + all)*	0.274
Multinomial Naïve Bayes	0.230

Challenges:

- Understanding truthfulness of social media and news content:
- might **not be solvable relying solely on standard ML.**
 - requires a **prior database** of historical facts and general knowledge.

*Best result obtained by the authors of the dataset. For this result they have used all meta-data (subject, speaker, job, state, party, context, history score)

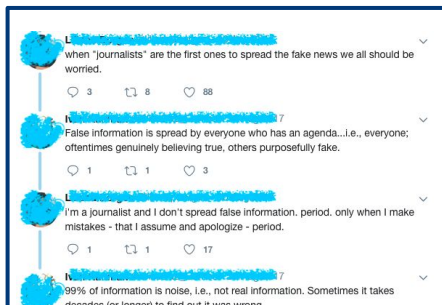
Future work



Forgery detection



Extracted metadata



Associated text analysis



Malicious



Harmless

Acknowledgments



universidade de aveiro
theoria poiesis praxis



instituto de engenharia electrónica
e telemática de aveiro



This work was supported in part by the Integrated Programme of SR&TD “SOCA” (Ref. CENTRO-01-0145-FEDER-000010), co-funded by Centro 2020 program, Portugal 2020, European Union, through the European Regional Development Fund.

