EURECOM
Sophia Antipolis

Two-stream Convolutional Neural Network for Image Source Social Network Identification

28/09/2021

Alexandre Berthet
PhD student
berthet@eurecom.fr

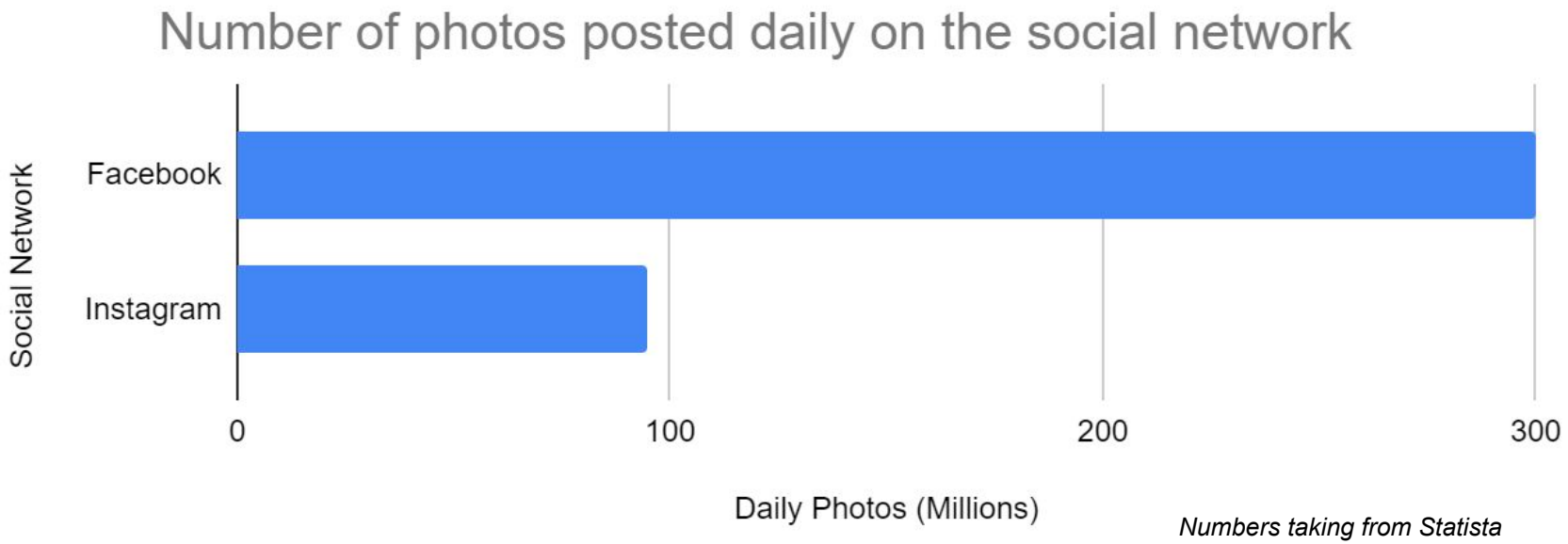
Chiara Galdi
Research associate
galdi@eurecom.fr

Jean-Luc Dugelay
Professor
dugelay@eurecom.fr

Département Sécurité numérique, EURECOM
Campus SophiaTech, 450 route des Chappes, F-06410 Biot Sophia Antipolis

Social Network (SN) Background

- **Images and social network : new way to communicate**



Numbers taking from Statista

- **Digital data on SN can be used in a malicious way**

Cyber-bullying

Psychological harassment

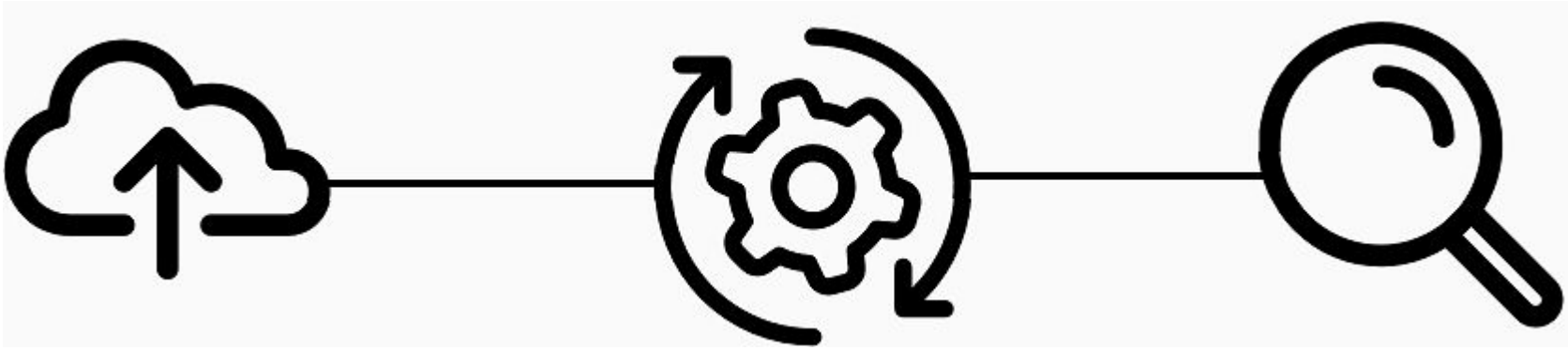
Attack personal reputation

Violence instigation

EURECOM
Sophia Antipolis

Image Source SN Identification

- **Retracing the image history**



- **Blind detection through artifacts**

~~METADATA~~

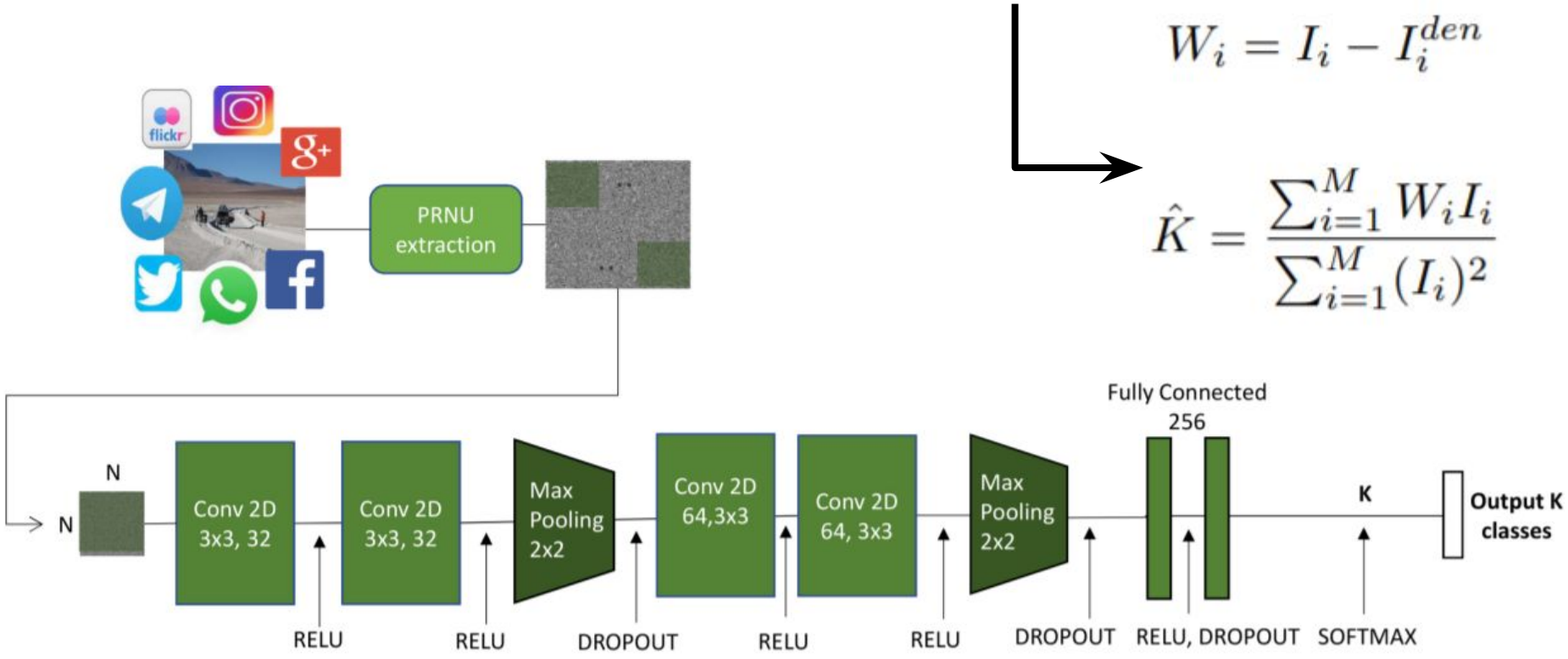
Discrete Cosine Transform (DCT)

Photo Response Non-Uniformity (PRNU)

EURECOM
Sophia Antipolis

State-of-the-art - PRNU-based method [1]

- **Convolutional Neural Network (CNN) with PRNU**

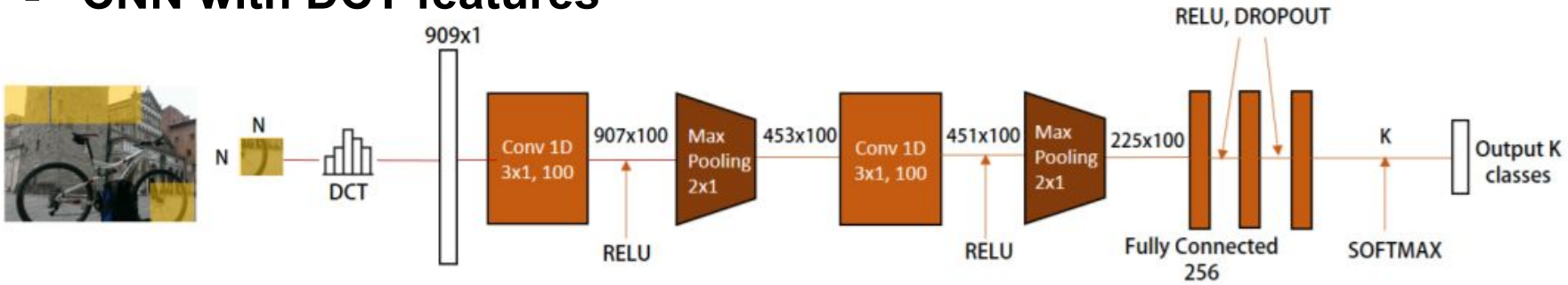


- **Each SN leaves traces that modulates PRNU**

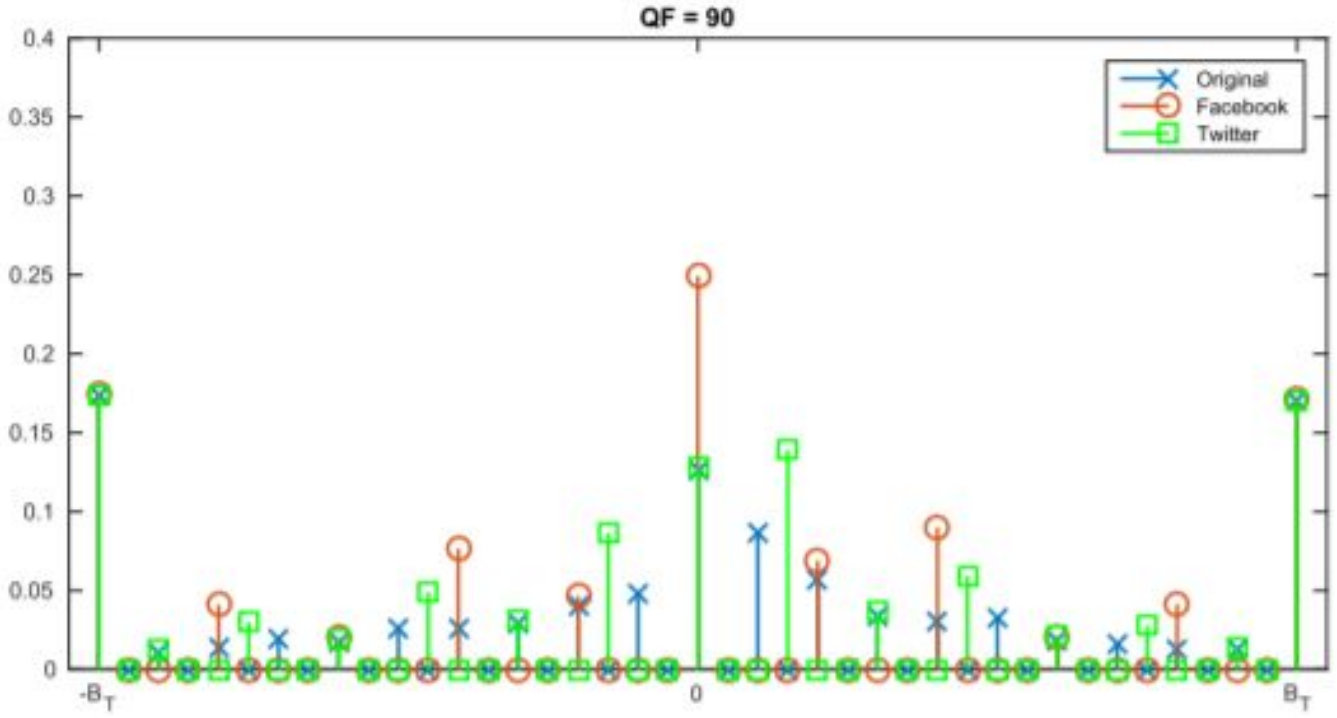
[1] R. Caldelli, I. Amerini, and C. T. Li, "Tracing images back to their social network of origin: a CNN-based approach"

State-of-the-art - JPEG-based method [2]

- **CNN with DCT features**



- **Histogram of DCT coefficients differs from one SN to another**

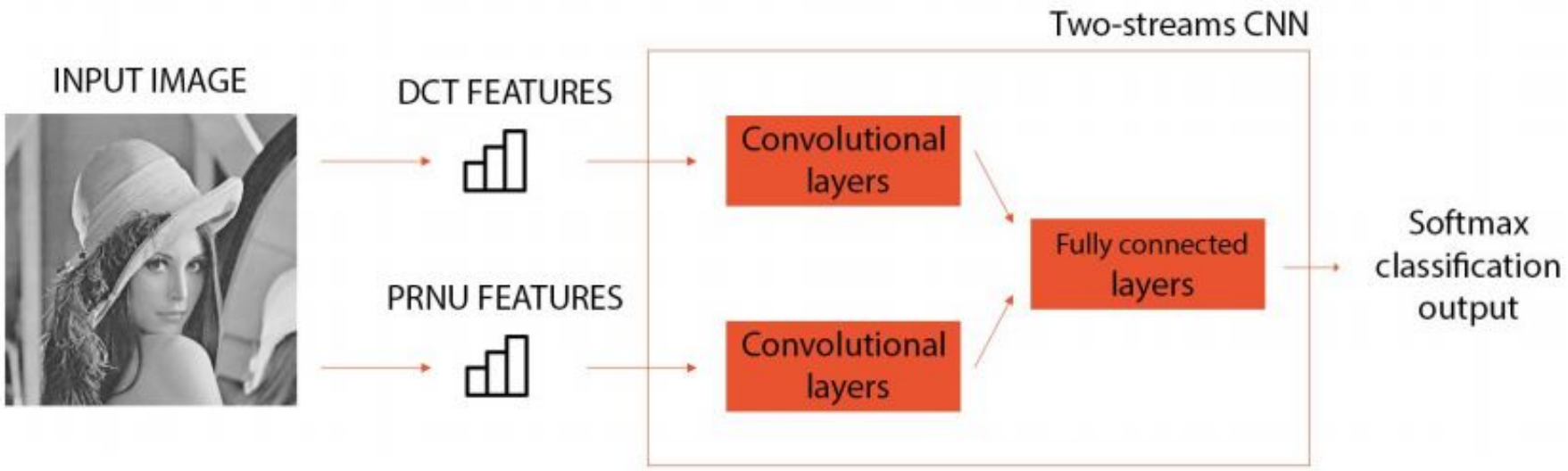


[2] I. Amerini, T. Uricchio, and R. Caldelli. Tracing images back to their social network of origin: A cnn-based approach.

EURECOM
Sophia Antipolis

Proposed Method - Two-Stream CNN

- **Two-Stream CNN: analysis of PRNU and DCT features**

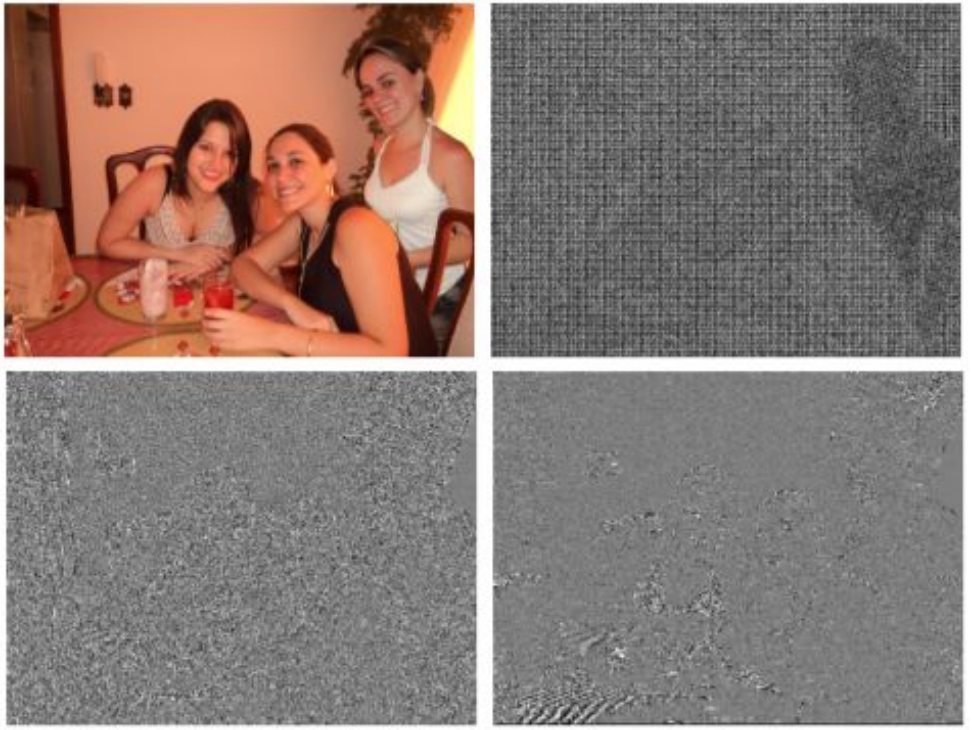


- **Combination of two domains to enhance artifacts diversity**

Proposed Method - PRNU & DCT features

- **PRNU extraction with CNN**

Use of Open Source Noiseprint [3]
Pre-trained CNN for camera model pattern extraction



- **Histogram of DCT coefficients**

New encoding method with normalized values of DCT coefficients to increase size of histograms vector

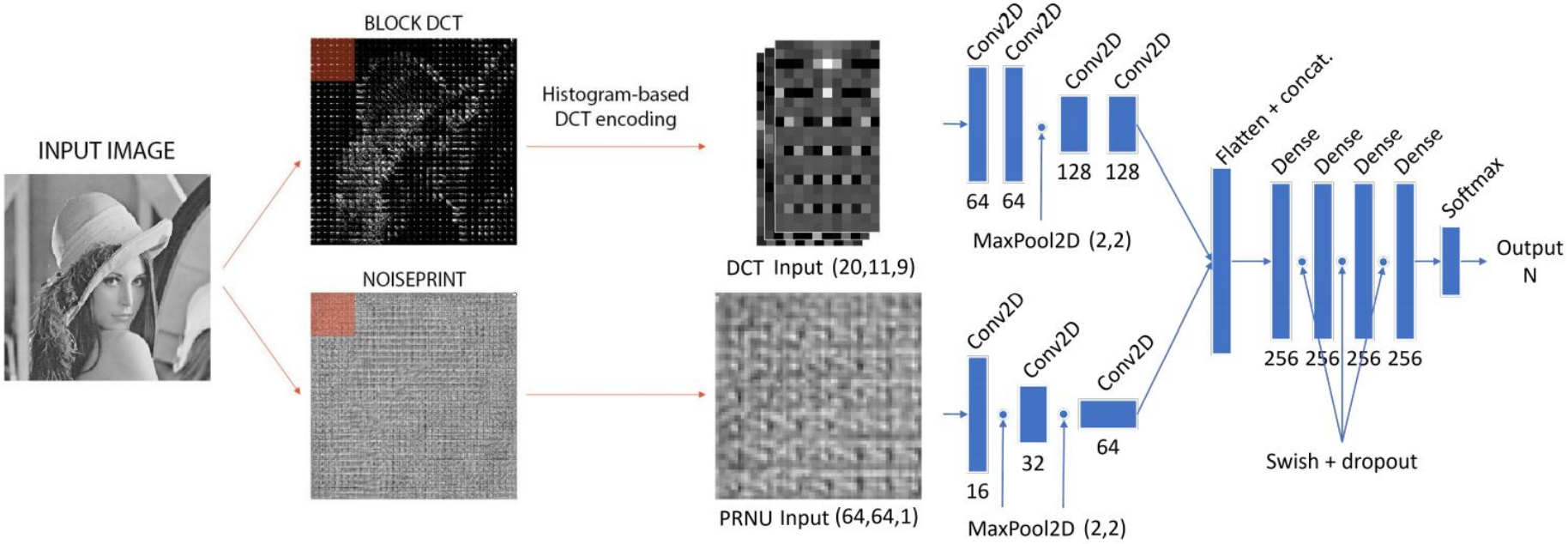
$$x_n = \frac{x}{q} - round(\frac{x}{q})$$

909 ⟶ **1980**

[3] Davide Cozzolino and Luisa Verdoliva. “Noiseprint: A cnn based camera model fingerprint”

Proposed Method - Final Structure

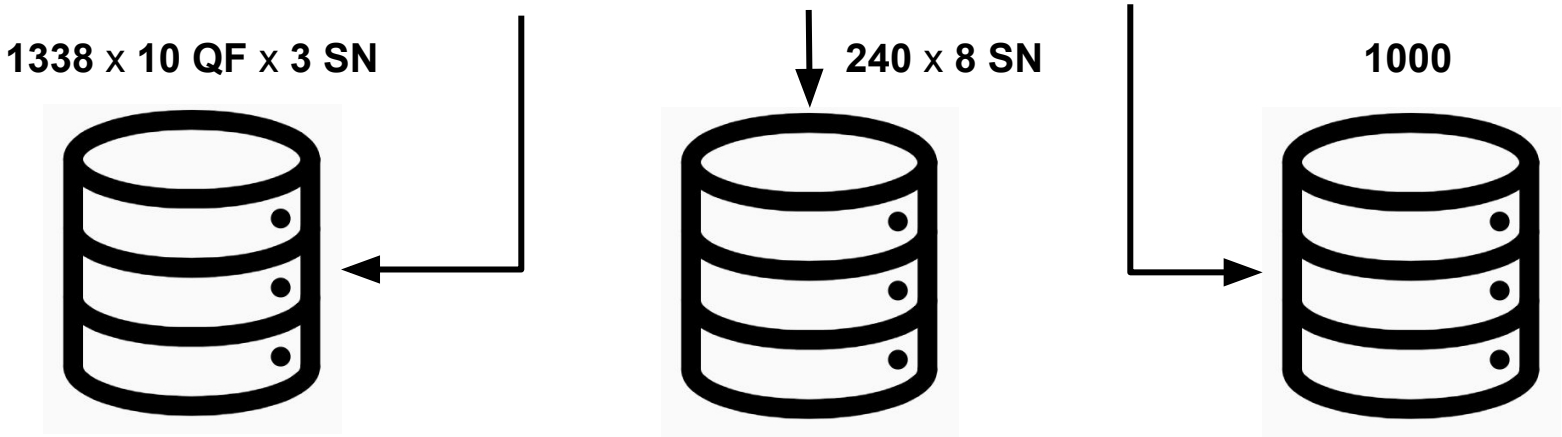
- **Input image is used for each stream and split in patches of 64 x 64**



- **Two evaluations: patch-level & image-level**

Dataset - Balancing Classes & Patches

- **Three databases: UCID social / IPLAB / Social PUBLIC**



- **Datasets balanced at image-level but not patch-level**

DUPLICATE TO OBTAIN INFINITE GENERATORS

EURECOM Sophia Antipolis

Dataset - Balancing Classes & Patches

- **Balance classes for validation and testing set**
  - Creation of a class weight factor *F* for an image of class *k* with *N* patches

$$F_k = N \cdot \frac{1}{|P_k|} = \frac{N}{|P_k|}$$

- **Unbalancing patches at image-level because of image size**
  - Associate a weight *w* to each sample from an image of class *k* with *N* patches
  - Combinate this weight to the class weight factor *F* to obtain a fully balance factor

$$w_{Nk} = \frac{|P_k|}{|I_k|} \cdot \frac{1}{N} \qquad f_k = w_{Nk} * F_k = \frac{|P_k|}{|I_k|} \cdot \frac{1}{N} \cdot \frac{N}{|P_k|} = \frac{1}{|I_k|}$$

- **Therefore, image are equally balanced at image-level**

Results - Preliminary evaluation

- **Evaluation of each single stream of our proposed method**
  - DCT-based CNN reaches same result as the state-of-the-art (SOTA) JPEG-based method, with about 20% fewer epochs
  - Noiseprint obtains 10% more accuracy than the SOTA PRNU-based method

- **Comparison of the two-stream network at patch-level with state-of-the-art methods**

UCID social	*Facebook*	*Twitter*	*Flickr*
PRNU-based SOTA	72.80 %	72.49 %	93.15 %
JPEG-based SOTA	96.15 %	**99.30 %**	99.79 %
Proposed method	**97.60 %**	**99.30 %**	**100 %**

Results - Image source SN identification

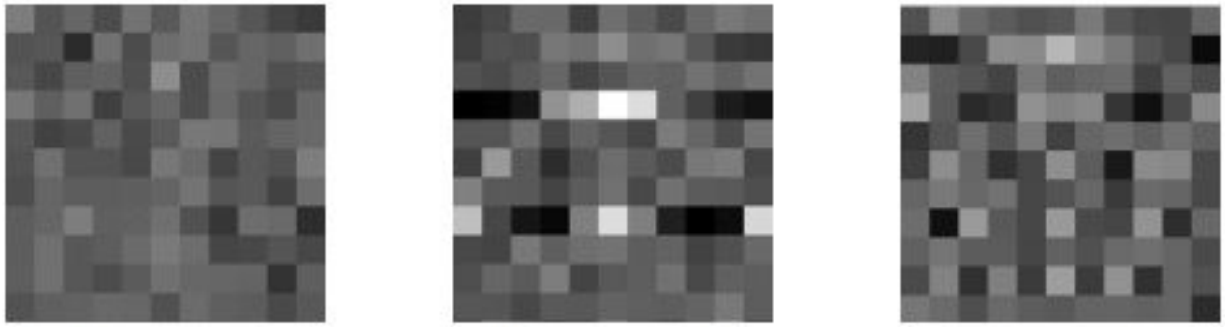
- **Comparison of the confusion matrix with JPEG-based SOTA method**

Public Social	*Our*			*JPEG*		
	Facebook	*Twitter*	*Flickr*	*Facebook*	*Twitter*	*Flickr*
Facebook	**91.21 %**	8.79 %	0 %	88.24 %	11.76 %	0 %
Twitter	0.1 %	98.67 %	1.23 %	0 %	**100 %**	0 %
Flickr	0	1.85 %	**98.15 %**	0.99 %	1.98 %	97.03 %

THANK YOU ALL FOR LISTENING

Annexes - PRNU & DCT features

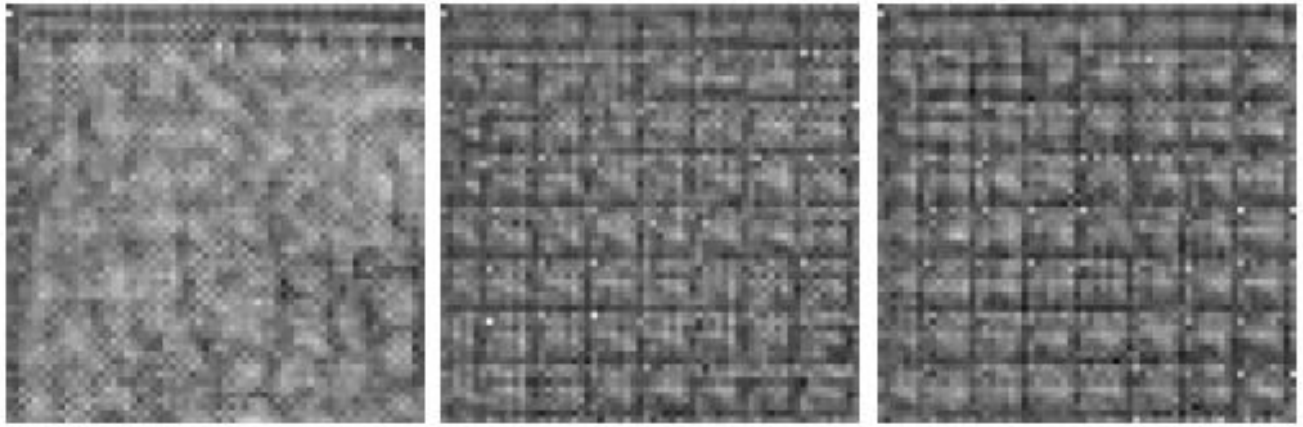
- **Comparison of JPEG & PRNU features according to SN**



(a) Uncompressed (b) Facebook (c) Flickr

DCT-block

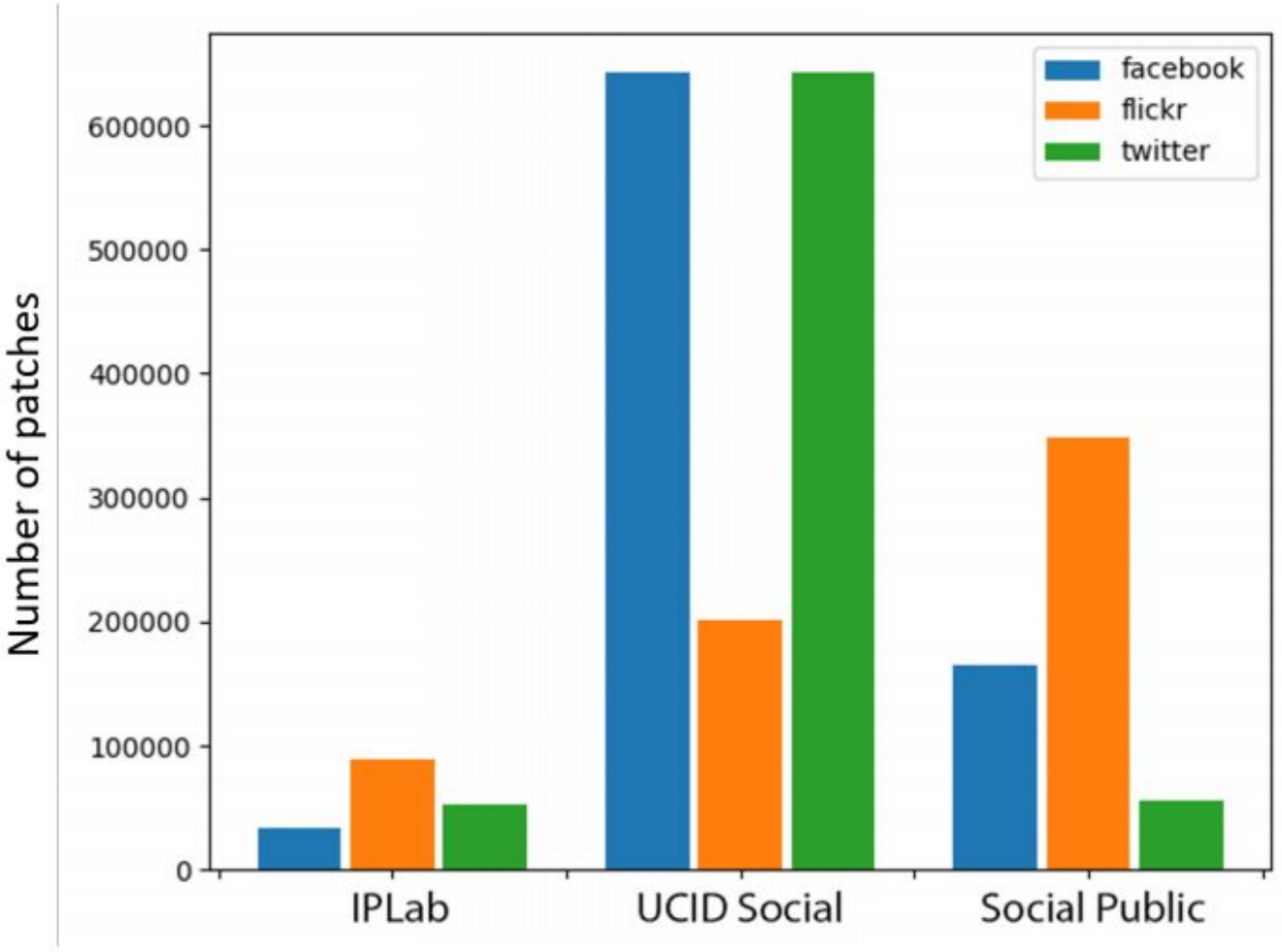
Noiseprint 64x64 patches



(a) Uncompressed (b) Facebook (c) Flickr

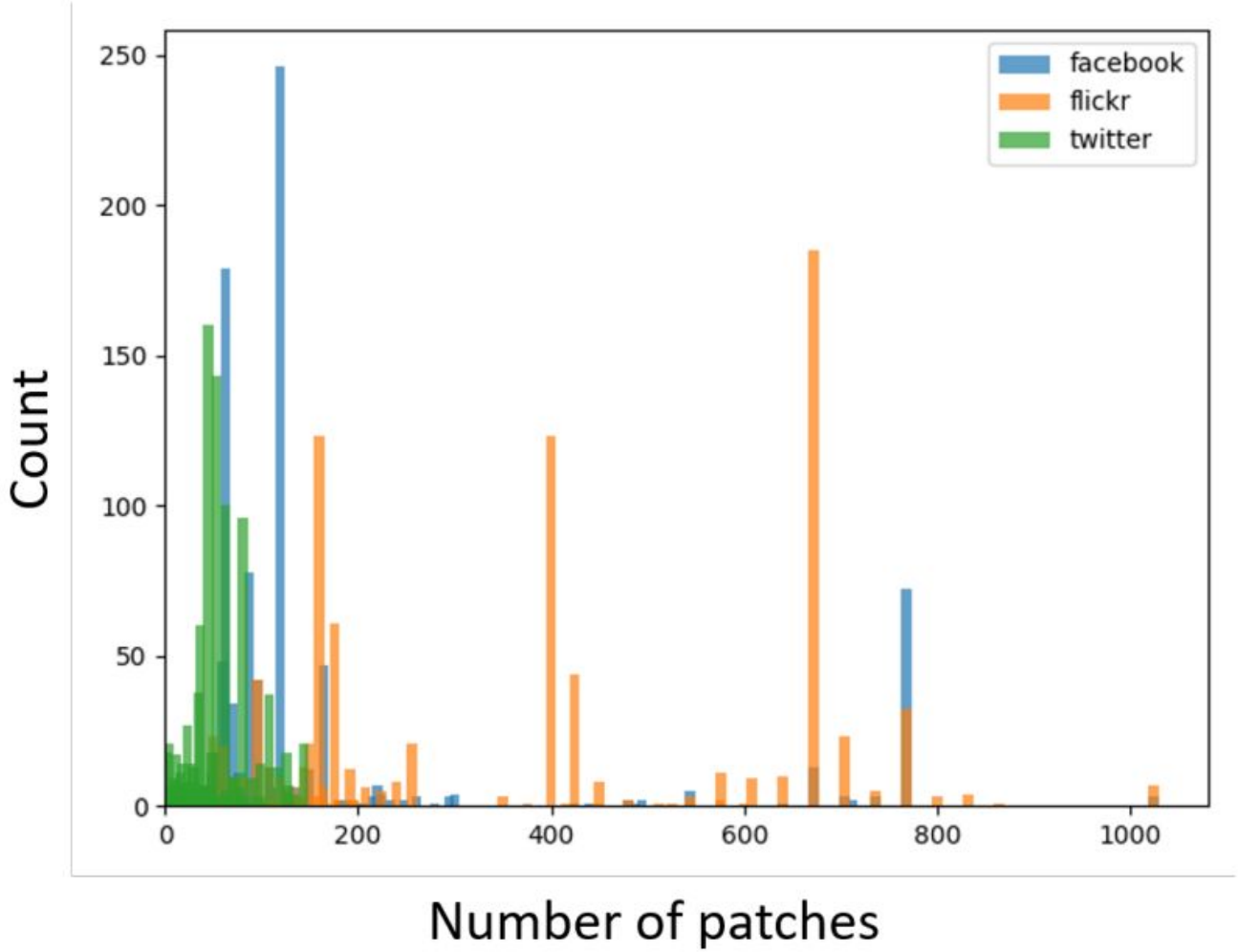
Annexes - Balancing Classes & Patches

- **Number of patches per images according to each database**



Annexes - Balancing Classes & Patches

- **Histogram of images according to their number of patches**



Annexes - Results at patch-level

- **Comparison against JPEG-based method at patch-level**

UCID Social	Our	[2]	Our	[2]	Our	[2]
	Facebook		*Flickr*		*Twitter*	
Facebook	**97.60%**	96.15%	0.20%	0.19%	2.20%	3.66%
Flickr	0.00%	0.03%	**100%**	99.79%	0.00%	0.18%
Twitter	0.67%	0.59%	0.13%	0.11%	**99.30%**	**99.30%**
Social Public	Our	[2]	Our	[2]	Our	[2]
	Facebook		*Flickr*		*Twitter*	
Facebook	**97.30%**	94.00%	2.52%	6.00%	0.18%	0.00%
Flickr	7.11%	1.76%	89.61%	**92.13%**	3.28%	6.11%
Twitter	0.10%	0.00%	3.49%	13.47%	**96.41%**	86.53%
IPLab	Our	[2]	Our	[2]	Our	[2]
	Facebook		*Flickr*		*Twitter*	
Facebook	**97.30%**	94.00%	2.52%	6.00%	0.18%	0.00%
Flickr	7.11%	1.76%	89.61%	**92.13%**	3.28%	6.11%
Twitter	0.10%	0.00%	3.49%	13.47%	**96.41%**	86.53%