Computer Vision Lab

Noise Modeling, Synthesis, and Classification for Generic Object Anti-Spoofing

Joel Stehouwer, Amin Jourabloo, Yaojie Liu, Xiaoming Liu
Department of Computer Science and Engineering, Michigan State University

I A R P A
BE THE FUTURE

Anti-Spoofing

Problem Statement:

Given an image of an object or scene, determine if the object or scene are genuine, or if they are presented to the camera via a spoof attack medium.

Definitions:

Sensor Noise – A high frequency pattern that is encoded, unique to each sensor.

Spoof Noise – A high frequency pattern that is encoded, unique to each spoof medium.

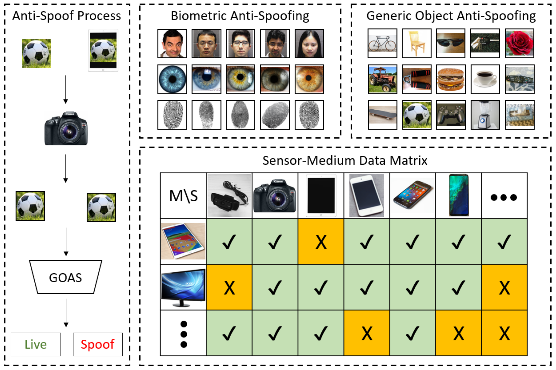
Why generic objects:

Biometric anti-spoofing systems are limited to the context in which they are developed.

The spoof noise is independent of the content in the image.

Insights and Contributions:

- ✧ Novel database for the purpose of Generic Object Anti-Spoofing and spoof noise analysis
- ✧ Patch-Based Camera Model Identification and Spoof Medium Classification
- ✧ Modeling and Synthesis of Sensor Noise, Spoof Noise, and their combination



M\S							•••
	✓	✓	X	✓	✓	✓	✓
	X	✓	✓	✓	✓	✓	X
⋮	✓	✓	✓	X	✓	X	X

Fig 1: Overview of Generic Object Anti-Spoofing and the Sensor-Medium Matrix.

Generic Object DataSet (GOSet)

The GOSet dataset was collected using various objects and backgrounds available in an office or home environment. Generic objects with varying backgrounds were chosen to show the effect of the assumption that spoof noise is independent of the content in the image.

- ❖ 308 Live Videos
- ❖ 2453 Spoof Videos
- ❖ 24 Objects
- ❖ 7 Backgrounds
- ❖ 7 Sensors
- ❖ 6 Mediums (plus live)

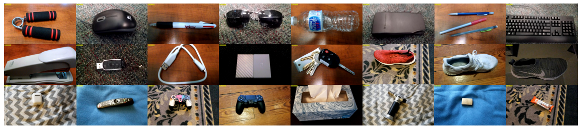


Fig 2: Example objects and backgrounds for the GOSet dataset.

Generic Object Image Synthesis

Problem Statement:

Given target sensor and medium, synthesize a novel image as if from that sensor and medium.

Specifically, given a genuine live image I captured by a camera C_1, generate I' such that I' appears to have been captured by C_2 viewing M_1.

We do this by learning noise prototypes that can be used to specify the target sensor and medium. The prototypes are convolved with the source image to produce the synthetic image.

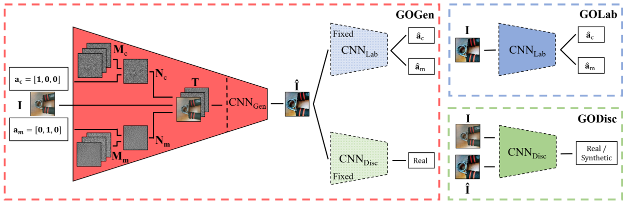


Fig 3: The proposed architecture for a machine learning computer vision system that can generate the missing elements from a sensor-medium data matrix. This system will learn to produce any sensor-medium combination, such that given a live image, it can produce an image for the desired sensor-medium combination.

Motivation:

Convolutional neural networks (CNNs), the backbone of computer vision systems, are fragile to:

- ❖ Adversarial Attacks
- ❖ Blurry Images (low resolution or motion blur)
- ❖ Extreme Illumination
- ❖ New or Different Sensors

To enhance robustness, more data is required. Data collection is time consuming and costly. If we can collect part of the data and systematically generate more, we could significantly reduce the cost.

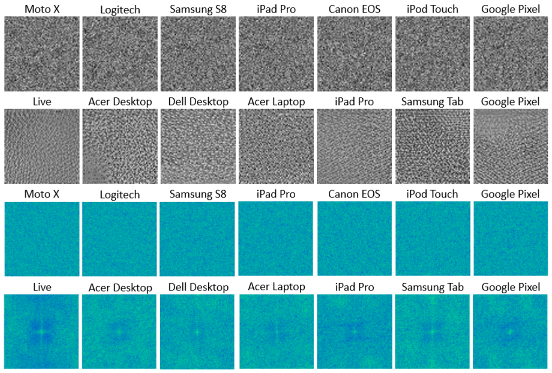


Fig 4: Noise prototypes for the various sensors and mediums in the GOSet. Shown below are the FFT spectra of the noise prototypes.

This research is performed under generous funding from IARPA. IARPA exists to promote and invest in novel research that is of potential benefit to future national intelligence.

Generic Object Anti-Spoofing

Problem Statement:

Given an image, determine if it is a genuine image or a spoof attack.

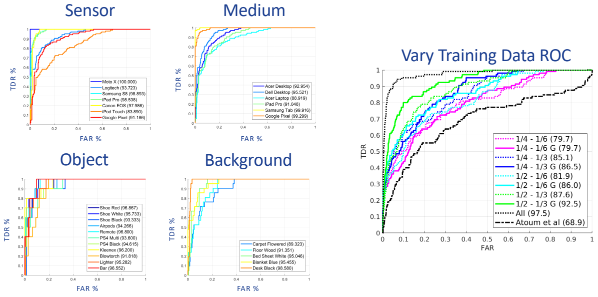


Fig 5: ROC curves for the GOLab algorithm according to image variables (left) and the amount of live and spoof, with or without synthetic, training data (right).

Generalization to Face Anti-Spoofing:

Algorithm	Train Data	OULU-NPU P1		MSU SiW	
		HTER	EER	HTER	EER
Chingovska LBP	Face	38.5	44.2	30.5	31.7
Boulkenafet Texture	Face	40.8	43.3	28.6	29.9
Boulkenafet SURF	Face	38.2	40.8	36.0	36.7
Atoum *et al.*	Face	11.8	13.3	11.0	11.2
Chingovska LBP	GOSet	44.1	46.1	42.2	42.4
Boulkenafet Texture	GOSet	34.6	36.7	44.1	44.9
Boulkenafet SURF	GOSet	45.3	45.8	47.7	48.6
Atoum *et al.*	GOSet	32.9	35.0	8.2	8.8
GOPad	GOSet	33.4	34.2	9.5	10.2
GOLab	GOSet	41.2	42.5	15.6	16.0

Tab 1: Cross-dataset testing on the face modality.

Sensor and Medium Classification

Problem Statement:

Given an image, determine the camera that was used to collect it. Also determine whether the image is genuine or if it is a spoof attack. If it is a spoof attack, determine which spoof medium was used to present the image.

GT \ Est	Moto X	Logitech	Samsung S8	iPad Pro	Canon EOS	iPod Touch	Google Pixel
Moto X	**16**	0	7	5	18	0	0
Logitech	2	**320**	0	0	0	0	3
Samsung S8	1	2	**353**	1	0	7	17
iPad Pro	6	0	42	**220**	0	3	0
Canon EOS	55	0	7	32	**68**	0	3
iPod Touch	0	0	0	0	0	**270**	0
Google Pixel	1	1	0	0	0	1	**259**

GT \ Est	Live	Acer Desktop	Dell Desktop	Acer Laptop	iPad Pro	Samsung Tab	Google Pixel
Live	**97**	7	0	0	0	1	0
Acer Desktop	50	**116**	67	36	9	45	3
Dell Desktop	31	52	**83**	59	20	77	8
Acer Laptop	58	53	4	**141**	7	3	5
iPad Pro	43	30	31	29	**107**	30	0
Samsung Tab	4	0	0	79	5	**115**	0
Google Pixel	7	54	5	12	34	20	**84**

Fig 6 and Tab 2,3: The sensors and spoof mediums that were used for data collection. The confusion matrix for the GOLab trained to identify sensor and spoof medium.