

University of Groningen

Leveraging image noise: source camera identification and increased robustness of convolutional neural networks

Bennabhaktula, Guru Swaroop

DOI:

[10.33612/diss.843513794](https://doi.org/10.33612/diss.843513794)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2023

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Bennabhaktula, G. S. (2023). *Leveraging image noise: source camera identification and increased robustness of convolutional neural networks*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen. <https://doi.org/10.33612/diss.843513794>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Leveraging image noise:
Source camera identification and increased robustness
of convolutional neural networks

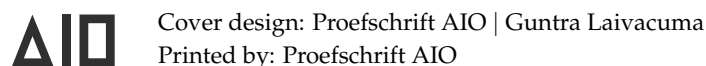
Guru Swaroop Bennabhaktula

This research has been conducted at the Information Systems group of the Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence of the University of Groningen, and at the Group for Vision and Intelligent Systems at the Department of Electrical, Systems, and Automation of the University of León.

This work was partly supported by the framework agreement between the University of León and INCIBE (Spanish National Cybersecurity Institute) under Addendum 01. This research has been partly funded with support from the European Commission under the 4NSEEK project with Grant Agreement 821966.



Cover image: The picturesque landscape is from Alt St. Johann in Switzerland, taken during the memorable SSSIHL alumni retreat (SEAM 2022). It illustrates, on the left, the manifestation of the hidden camera noise used in forensic analysis. On the right is a representation of the high-level scene content in images useful in the design of robust computer vision models.





university of
 groningen



universidad
 de León

Leveraging image noise: source camera identification and increased robustness of convolutional neural networks

PhD thesis

to obtain the degree of PhD of the
 University of Groningen
 on the authority of the
 Rector Magnificus Prof. J.M.A. Scherpen
 and in accordance with
 the decision by the College of Deans

and

to obtain the degree of PhD of the
 University of León
 on the authority of the
 Rector Magnificus Prof. J.F.G. Marín
 and in accordance with
 the decision by the College of Deans.

Double PhD degree

This thesis will be defended in public on
 Monday 18 December 2023 at 14.30 hours

by

Guru Swaroop Bennabhaktula

born on 13 August 1991
 in Visakhapatnam, India

Supervisors

Prof. D. Karastoyanova

Prof. E. Alegre

Co-supervisor

Dr. G. Azzopardi

Assessment Committee

Prof. M. Castejon Limas

Prof. M. Biehl

Prof. A. Telea

Prof. D. Meuwly

*Dedicated with Love to
Bhagawan Sri Sathya Sai Baba*

Contents

Table of contents in Spanish	xi
Acknowledgements	xv
List of figures	xix
List of tables	xxii
List of acronyms	xxiii
Spanish abstract	xxv
1 Introduction	3
1.1 Research questions and contributions	5
1.2 Outline	7
Part I Source Camera Identification	11
2 Device-based image matching for camera device identification from images	13
2.1 Introduction	13
2.2 Related work	14
2.2.1 Traditional approaches	16
2.2.2 Approaches based on deep learning	17
2.3 Proposed approach	18
2.3.1 Learning phase I	19

2.3.2	Learning phase II	20
2.4	Preliminary experiments and results	21
2.4.1	Data set	21
2.4.2	Experiments	22
2.5	Discussion and future work	23
2.6	Conclusions	24
3	Camera model identification using homogeneous patches	27
3.1	Introduction	27
3.2	Related works	29
3.3	Methodology	33
3.3.1	Overview	34
3.3.2	Homogeneous patch selection and extraction	34
3.3.3	Patch classification	37
3.3.4	Majority voting	40
3.3.5	Hierarchical classification scheme	40
3.4	Experiments	41
3.4.1	Data set	41
3.4.2	Data balancing	42
3.4.3	Data set split	43
3.4.4	Brand classification	43
3.4.5	Model classification	47
3.4.6	Hierarchical classification	49
3.5	Discussion	51
3.5.1	Hierarchical versus flat approach	52
3.5.2	Future work	53
3.6	Conclusion	54
4	Camera identification from videos using constrained convolutions	57
4.1	Introduction	57
4.2	Related work	59
4.2.1	Model-based techniques	59
4.2.2	Data-driven technologies	60
4.3	Methodology	61
4.3.1	ConstrainedNet	61
4.3.2	Frame extraction	64
4.3.3	Voting procedure	64
4.4	Experiments and results	64
4.4.1	Data set	64

Contents

4.4.2	Frame-based device identification	66
4.5	Discussion	69
4.6	Conclusion	71
5	Source camera device identification from videos	73
5.1	Introduction	73
5.2	Related works	75
5.3	Methodology	80
5.3.1	Frame extraction	80
5.3.2	Convolutional neural networks	80
5.3.3	Video-level predictions	83
5.4	Experiments	83
5.4.1	Data set - VISION	83
5.4.2	Data set - QUFVD	85
5.4.3	ConvNet training	86
5.4.4	ConvNet evaluation	86
5.5	Discussion	91
5.5.1	Video compression	91
5.5.2	Number of frames per video	92
5.5.3	Pre-training ConvNet	94
5.5.4	Camera model identification	94
5.5.5	Counter-productive learning strategies	95
5.5.6	Prediction time efficiency	96
5.5.7	Future work	97
5.6	Conclusion	97
 Part II Improved robustness of Convolutional Neural Networks		 101
6	Improving the robustness of ConvNet using data pre-processing	103
6.1	Introduction	103
6.2	Related works	104
6.3	Methods	106
6.3.1	Overview	106
6.3.2	CORF operator with push-pull inhibition	106
6.3.3	AlexNet	108
6.3.4	Image perturbations	109
6.4	Experiments and results	110
6.4.1	Data set	110

6.4.2	Experiments	111
6.5	Discussion	112
6.6	Conclusion	113
7	PushPull-Net: Inhibition-driven ResNet robust to image corruptions	115
7.1	Introduction	115
7.2	Related works	118
7.3	Approach	120
7.3.1	The Push Pull unit	120
7.3.2	Characteristics of Push Pull convolutions	121
7.3.3	Embedding PushPull-Conv in ConvNets	123
7.4	Experiments and results	124
7.4.1	Evaluation metrics	125
7.4.2	Experimental setup	125
7.4.3	Experiments	126
7.4.4	Comparison with the state-of-the-art	128
7.5	Discussion	130
7.5.1	Clean error vs robustness	130
7.5.2	Robustness to high-frequency corruptions	131
7.5.3	Image retrieval and future work	132
7.6	Conclusion	132
7.A	Negative inhibition values	133
7.B	Robustness transferability to other applications	133
7.C	Absolute scores	136
8	Conclusion	141
8.1	Evaluation of research questions	141
8.2	Summary	145
8.3	Outlook	147
8.A	Summary of thesis in Spanish	150
	Samenvatting	153
	Bibliography	157
	Research Activities	171
	Curriculum Vitae	173

Tabla de contenido en español

Expresiones de gratitud	xv
Lista de Figuras	xix
Lista de tablas	xxii
Lista de acrónimos	xxiii
Resumen español	xxv
1 Introducción	3
1.1 Preguntas de investigación y contribuciones	5
1.2 Esquema	7
2 Coincidencia de imágenes basada en dispositivos para identificación de dispositivos de cámara a partir de imágenes	13
2.1 Introducción	13
2.2 Trabajo relacionado	14
2.2.1 Enfoques tradicionales	16
2.2.2 Enfoques basados en Deep Learning	17
2.3 Enfoque propuesto	18
2.3.1 Aprendizaje Fase I	19
2.3.2 Aprendizaje Fase II	20
2.4 Experimentos preliminares y resultados	21
2.4.1 Conjunto de datos	21
2.4.2 Experiments	22
2.5 Discusión y trabajo futuro	23

2.6	Conclusiones	24
3	Identificación del modelo de cámara mediante parches homogéneos	27
3.1	Introducción	27
3.2	Obras relacionadas	29
3.3	Metodología	33
3.3.1	Descripción general	34
3.3.2	Selección y extracción de parches homogéneos	34
3.3.3	Clasificación de parches	37
3.3.4	Votación mayoritaria	40
3.3.5	Esquema de clasificación jerárquica	40
3.4	Experimentos	41
3.4.1	Conjunto de datos	41
3.4.2	Equilibrio de datos	42
3.4.3	División del conjunto de datos	43
3.4.4	Clasificación de marca	43
3.4.5	Clasificación del modelo	47
3.4.6	Clasificación jerárquica	49
3.5	Discusión	51
3.5.1	Enfoque jerárquico versus plano	52
3.5.2	Trabajo futuro	53
3.6	Conclusión	54
4	Identificación de cámara a partir de videos usando circunvoluciones restringidas	57
4.1	Introducción	57
4.2	Trabajo relacionado	59
4.2.1	Técnicas basadas en modelos	59
4.2.2	Tecnologías basadas en datos	60
4.3	Metodología	61
4.3.1	Red restringida	61
4.3.2	Extracción de cuadros	64
4.3.3	Procedimiento de votación	64
4.4	Experimentos y Resultados	64
4.4.1	Conjunto de datos	64
4.4.2	Identificación de dispositivos basada en marcos	66
4.5	Discusión	69
4.6	Conclusión	71

5	Identificación del dispositivo de la cámara de origen a partir de videos	73
5.1	Introducción	73
5.2	Obras relacionadas	75
5.3	Metodología	80
5.3.1	Extracción de cuadros	80
5.3.2	Redes neuronales convolucionales	80
5.3.3	Predicciones a nivel de video	83
5.4	Experimentos	83
5.4.1	Conjunto de datos - VISIÓN	83
5.4.2	Data Set - QUFVD	85
5.4.3	Capacitación de ConvNet	86
5.4.4	Evaluación de ConvNet	86
5.5	Discusión	91
5.5.1	Compresión de video	91
5.5.2	Número de cuadros por video	92
5.5.3	Pre-entrenamiento ConvNet	94
5.5.4	Identificación del modelo de cámara	94
5.5.5	Estrategias de aprendizaje contraproducentes	95
5.5.6	Eficiencia del tiempo de predicción	96
5.5.7	Trabajo futuro	97
5.6	Conclusión	97
6	Mejorar la solidez de ConvNet mediante el preprocesamiento de datos	103
6.1	Introducción	103
6.2	Trabajos relacionados	104
6.3	Métodos	106
6.3.1	Descripción general	106
6.3.2	Operador CORF con inhibición push-pull	106
6.3.3	AlexNet	108
6.3.4	Perturbaciones de imagen	109
6.4	Experimentos y resultados	110
6.4.1	Conjunto de datos	110
6.4.2	Experimentos	111
6.5	Discusión	112
6.6	Conclusión	113
7	PushPull-Net: ResNet impulsada por inhibición resistente a las corrupciones	115
7.1	Introducción	115

7.2	Trabajos relacionados	118
7.3	Enfoque	120
7.3.1	La unidad Push Pull	120
7.3.2	Características de las circunvoluciones Push Pull	121
7.3.3	Incorporación de PushPull-Conv en ConvNets	123
7.4	Experimentos y resultados	124
7.4.1	Métricas de evaluación	125
7.4.2	Montaje experimental	125
7.4.3	Experimentos	126
7.4.4	Comparación con el estado del arte	128
7.5	Discusión	130
7.5.1	Error limpio vs robustez	130
7.5.2	Robustez a corrupciones de alta frecuencia	131
7.5.3	Recuperación de imágenes y trabajo futuro	132
7.6	Conclusión	132
7.A	Valores de inhibición negativa	133
7.B	Transferibilidad de la robustez a otras aplicaciones	133
7.C	Puntajes absolutos	136
8	Conclusión	141
8.1	Evaluación de preguntas de investigación	141
8.2	Resumen	145
8.3	Outlook	147
8.A	Resumen de la tesis en Español	150
	Samenvatting	153
	Bibliografía	157
	Actividades de investigación	171
	Currículum vitae	173

Acknowledgements

The journey leading to the preparation of this thesis has been both challenging and fun. It goes without saying this was made possible only by the support and guidance of family, friends, colleagues, well-wishers, and most importantly my beloved God. I am immensely grateful to everyone.

Firstly, I would like to express my gratitude to my mother and father for always supporting me and playing a vital role in my upbringing. In spite of being far away in India, my father was always available to have a conversation with me whenever I reached out to him. I look up to him for his qualities of forbearance, hard work, sincerity, and love. These helped me cultivate a part of my personality that helped me immensely during my PhD studies and beyond.

Teachers are the pillars of this society. I consider myself very fortunate to have been guided by selfless teachers who inculcated noble values along with secular education. Particularly, I am grateful to all the teachers from the Sri Sathya Sai Educational Institutions in India, with whom I had the opportunity to interact and learn. I am indebted to all of them. I would especially like to thank Prof. Raghunatha Sarma and Prof. Pallav Kumar Baruah for their support in guiding and inspiring me to pursue a Ph.D. degree. I am also grateful to the supervisors of my Master's thesis Dr. Srikanth Khanna and Prof. Venkatachalam Chandrasekaran who have been instrumental in helping me develop personal and research skills necessary for my career.

Coming to the Ph.D. thesis itself, the most important people behind its success are, of course, my supervisors. I had the opportunity to be supervised by Prof. Dimka Karastoyanova and Dr. George Azzopardi from the University of Groningen, the Netherlands, and Prof. Enrique Alegre from the University of León, Spain. It was a valuable learning experience for me both professionally and personally. George played the role of a daily supervisor and his involvement and interest in the project was beyond my expectation. We had a lot of interesting interactions on several ideas related to our project and sometimes these were

virtually late in the evenings while working from home. I am sure this wouldn't have happened without the understanding of George's wife Charmaine, and their daughters Alaia and Amira. I thank you all for your support. I would like to express my deepest sense of gratitude to all my supervisors for the opportunities they have provided me for my growth.

My stay at Groningen was pleasant due to my colleagues from the Information Systems Group and Artificial Intelligence group. I would like to thank Abolfazl, Ali Reza, Mohammed Alghazwi, Arash, Xueyi, Nafiseh, Estefania, Ahmad, Sreejita, and Anusha. Likewise, I would like to thank my colleagues from the GVIS research group in Léon. I must mention that it is because of you that my stay in Léon became possible and comfortable due to your assistance with Spanish and local arrangements. I would like to thank Surajit, Eduardo, Victor, Laura, Rubel, David, Pablo, Aitor, Andres, Fran, Javier, and Manuel. I also cherish the collaborations with students in their Master's and Bachelor studies. Thanks to Derrick, Joey, Adrain, and Thijs. I would also like to thank Peter van der wal for reviewing the Dutch versions of my thesis.

Family and friends are very important to your well-being. I was lucky enough to have the company of my cousin Pushpa in Groningen. Along with Sukhmander, we all enjoyed preparing nice dinners, praying, and having a good time together. Incidentally, I was also lucky to have the company of my cousin Vishal in Léon. Thanks to Ali Reza and Pushpa for kindly agreeing to be my paranymphs. Although the rest of my family was in India, they were in constant touch and never let me feel that I was physically far away. I thank you all for your love and care.

I would fail in my duty if I did not thank all the administrative and support staff at both the Universities of Groningen and Léon. They did a phenomenal job in taking care of several things right from the issue of contracts, arranging residence permits, and finally getting the double doctorate agreement signed. I would like to especially thank the HPC group at Groningen for providing access to the Peregrine supercomputer and the team at Léon for similar access to computing resources.

Finally, I would like to express my deepest sense of gratitude to Bhagawan Sri Sathya Sai Baba. I remember His words, rephrased, "A Ph.D. is of no use to the nation if it cannot be applied for the welfare of the society". I am grateful to Him for providing me with a topic that has a societal impact. I know that I am merely an instrument and He is the real doer. I am ever grateful for His guidance and for giving me this experience.

Guru Swaroop Bennabhaktula
Groningen
November 22, 2023

List of Figures

Figure 1.1	A high-level pipeline of the proposed approaches for images and videos	4
Figure 1.2	A high-level pipeline of the proposed approaches for improved robustness of Convolutional Neural Networks (ConvNets).	4
Figure 2.1	Topology of digital camera sensor noise	15
Figure 2.2	Proposed workflow of device-based image matching	18
Figure 2.3	The proposed neural network architecture of the Similarity Network.	20
Figure 2.4	Similarity matrix for the 31 camera devices in the test set	21
Figure 3.1	The proposed flow for hierarchical classification of homogeneous patches	30
Figure 3.2	A typical image generation pipeline inside digital cameras	30
Figure 3.3	Categorization of camera processing noise embedded within every image	31
Figure 3.4	A high-level pipeline of the proposed methodology for the classification of homogeneous patches	34
Figure 3.5	An illustration of homogeneous patch selection based on standard deviation	36
Figure 3.6	An illustration of the patch pre-processing step	37
Figure 3.7	Distribution of an average number of homogeneous, non-homogeneous, and saturated patches in the Dresden data set	38
Figure 3.8	The proposed ConvNet architecture for camera model identification using homogeneous patches	39
Figure 3.9	An illustration of the hierarchical patch balancing	44

Figure 3.10	Learning plots of the trained models depicting model convergence during training	45
Figure 3.11	Confusion matrix for brand-level classification	46
Figure 3.12	Confusion matrix for device-level classification	48
Figure 3.13	Confusion matrix for model-level classification	50
Figure 3.14	A plot of image-level accuracy for an increasing number of homogeneous patches	51
Figure 4.1	Image acquisition pipeline inside digital cameras	59
Figure 4.2	A high-level overview of the proposed methodology for video camera identification with a constrained ConvNet	62
Figure 4.3	Architecture of the proposed ConstrainedNet	63
Figure 4.4	Distribution of duration of videos in the VISION data set	66
Figure 4.5	Test set accuracy plot for every epoch	68
Figure 4.6	Confusion matrices for flat, indoor, and outdoor scenarios	68
Figure 4.7	Confusion matrices of the flat scenario for native, WhatsApp, and YouTube compression	69
Figure 5.1	Video generation pipeline in digital cameras	76
Figure 5.2	Classification of noise in imaging sensors	78
Figure 5.3	Proposed pipeline for camera identification from videos with sophisticated ConvNets	78
Figure 5.4	Learning plots for the trained ConvNet models	87
Figure 5.5	Confusion matrices obtained on classifying I-frames using MobileNet on the VISION data set	90
Figure 5.6	Confusion matrices obtained on classifying I-frames using MobileNet on the QUFVD data set	90
Figure 6.1	The proposed application pipeline.	106
Figure 6.2	Application pipeline depicting training and deployment scenarios	106
Figure 6.3	Combination of Receptive Fields (CORF) computation model of a simple cell	107
Figure 6.4	CORF model with push-pull inhibition	108
Figure 6.5	Robustness of the push-pull CORF delineation operator to Gaussian noise	108
Figure 6.6	AlexNet architecture	109
Figure 6.7	Sample images from Fashion MNIST data set	110
Figure 6.8	Learning plots for model training	111
Figure 6.9	Impact of Gaussian and uniform additive noise on the trained ConvNet models	112

Figure 7.1	An illustration of the proposed approach for using PushPull convolutions for improved robustness	116
Figure 7.2	The proposed push-pull computation unit	120
Figure 7.3	Illustration of selected push-pull kernels from ResNet50	122
Figure 7.4	Illustration of push-pull behaviour with a simulated input stimulus corrupted by Gaussian noise	123
Figure 7.5	Averaged Fourier spectrum determined from the kernels of the conv1 layer of ResNet50 (with positive inhibition)	123
Figure 7.6	Illustration of image corruptions from the ImageNet-C data set	124
Figure 7.7	Distribution of learnt inhibition strength α for PushPull-based ResNets trained on ImageNet.	128
Figure 7.8	A comparison of ResNet18 models trained with various configurations of PushPull-Conv	130
Figure 7.9	Averaged Fourier spectrum determined from the kernels of conv1 layer of ResNet50 (with negative inhibition)	135

List of Tables

Table 2.1	The proposed ConvNet architecture of the signature network	19
Table 3.1	List of camera models considered for our experiments from the Dresden data set	38
Table 3.2	Classification results in terms of accuracy	46
Table 3.3	Classification results in terms of macro F1 score	48
Table 3.4	Reduction in error rate from the patch- to image-level predictions .	49
Table 3.5	Comparison with the state-of-the-art for camera model identification on the Dresden data set	50
Table 3.6	Hierarchical versus flat approach	53
Table 4.1	List of camera devices considered for experiments	65
Table 5.1	List of camera devices considered for experiments	83
Table 5.2	Classification accuracy of the proposed methods on the VISION data set	88
Table 5.3	Results on MobileNet for video identification	89
Table 5.4	State-of-the-art comparison on the QUFVD data set	91
Table 5.5	Results for various learning strategies on the VISION data set	92
Table 5.6	Results with varying number of frames per video	93
Table 5.7	Results with varying number of I-frames per video	94
Table 5.8	Runtime performance of the proposed method	96
Table 7.1	Summary of results on various data sets for robust image classification using push-pull convolutions	127

Table 7.2 Comparison with the state-of-the-art on CIFAR-10 and ImageNet-1K data sets for robust image classification	129
Table 7.3 Ablation study with different configurations of push-pull	131
Table 7.4 Results per low-, mid-, and high-frequency corruptions	131
Table 7.5 Summary of results - image retrieval	134
Table 7.6 Absolute mCE scores - image classification	137
Table 7.7 Absolute mCE scores - image retrieval	138
Table 7.8 Absolute mCE scores - comparison with state-of-the-art	139

List of acronyms

- 4NSEEK** Forensic Against Sexual Exploitation of Children
- AWGN** Additive White Gaussian Noise
- BN** Batch Normalization
- CCD** Charge Coupled Device
- CFA** Color Filter Array
- CNN** Convolutional Neural Network
- Conv** Convolutional
- ConvNet** Convolutional Neural Network
- CORF** Combination of Receptive Fields
- EXIF** Exchangeable Image File
- FC** Fully Connected
- FPN** Fixed Pattern Noise
- FN** False Negatives
- FP** False Positives
- LEA** Law Enforcement Agency
- MaxPool** Max-Pooling
- MISLNet** ConvNet architecture proposed by Bayar and Stamm (2018a)

PNU Pixel Non-Uniformity noise

PRNU Photo Response Non-Uniform noise

QUFVD Qatar University Forensic Video Database

ReLU Rectified Linear Unit

RGB Red, Green, and Blue image-colour channels

SCI Source Camera Identification

SGD Stochastic Gradient Descent

SPN Sensor Pattern Noise

SVM Support Vector Machine

TP True Positives

Spanish abstract

El ruido en una imagen digital puede ser una ventaja o una desventaja dependiendo del problema con el que se esté trabajando. Analizando el ruido de la imagen, esta tesis estudia y propone soluciones a dos aplicaciones diferentes del mismo. En la primera, se explora el papel del ruido generado por los sensores de una cámara para realizar un análisis forense de medios digitales donde la presencia de ruido en una imagen resulta crucial. Este trabajo se realizó dentro del proyecto 4NSEEK (análisis forense contra el abuso sexual de menores), financiado por la UE, que aborda un importante problema social. La otra línea en la que se trabaja en esta tesis es la vulnerabilidad de los sistemas de aprendizaje profundo frente al ruido presente en la imagen de entrada, para lo que se proponen medidas para mitigar el impacto de dicho ruido.

La primera parte de la tesis aborda varias cuestiones de investigación centradas en identificar la cámara con la que se ha obtenido una imagen o un vídeo. En primer lugar, dado un par de imágenes, se pretende determinar la probabilidad de que ambas hayan sido capturadas utilizando la misma cámara. Para ello, se propone una red que consta de dos partes: la primera extrae huellas forenses de cada imagen y la segunda coteja un par de dichas huellas para establecer la similitud con otra imagen determinando así la cámara con la que se obtuvo. Hasta donde tenemos constancia, esta es la primera propuesta que aborda el problema de identificar qué dispositivo obtuvo una imagen determinada, mostrándose con los resultados obtenidos el potencial de esta línea de investigación. En segundo lugar, se pretende conocer si algunas regiones de una imagen tienen una mayor presencia de ruido de cámara que otras. Para responder a esta pregunta, se examinaron diferentes regiones de una imagen y se planteó la hipótesis de que las regiones homogéneas de una imagen son las menos distorsionadas en comparación con las regiones con detalles de alto nivel de la escena, como bordes y patrones de textura.

Posteriormente, propusimos un sistema jerárquico para identificar las cámaras basado en parches homogéneos. La metodología propuesta alcanzó una precisión de clasificación del 99,01% en el conjunto de datos públicos de Dresde para la identificación de modelos de cámara. Además, el esquema de clasificación jerárquica propuesto hace que el sistema sea modular y facilita que se amplíe eficientemente con la incorporación de nuevos dispositivos de cámara. Nos basamos en el trabajo de Bayar and Stamm (2018a) que propuso una convolución restringida de un solo canal para suprimir el contenido de la escena, ampliando la solución propuesta en este trabajo a una solución multicanal. El tercer problema que propusimos resolver fue determinar en qué configuraciones de redes neuronales convolucionales (ConvNets) son eficaces las convoluciones restringidas. Se comprobó empíricamente que las convoluciones restringidas sólo son eficaces con ConvNets poco profundas y no con ConvNets profundas. Además, en el análisis forense de vídeos, se examinó qué fotogramas de vídeo conservaban más rastros forenses. Se abordó el problema consistente en extraer rastros forenses de vídeos comprimidos como los de YouTube y WhatsApp para observar que los fotogramas I retienen mejores rastros forenses en comparación con los demás. Posteriormente, se observó que, incluso con las compresiones de WhatsApp y YouTube, el sistema es capaz de identificar la cámara de origen con la misma precisión que en los vídeos nativos.

La segunda parte de la tesis se centra en hacer que los sistemas de aprendizaje profundo sean robustos frente a la corrupción de imágenes no vistas previamente. Alejándose del enfoque tradicional, que consiste en utilizar la aumentación de datos para mejorar la robustez, el objetivo que se tuvo fue explorar métodos alternativos para lograr la generalización fuera de la distribución original. En el primer enfoque, se propone un nuevo paso de preprocesamiento de datos, en el que se utilizan mapas de delineación CORF en lugar de la imagen tradicional como entrada a la ConvNet. Esto mejoró significativamente la robustez del modelo frente al ruido aleatorio uniforme y gaussiano, con una precisión comparable a la obtenida en imágenes de prueba limpias. Posteriormente, se propone una novedosa modificación intrínseca de la arquitectura de ConvNet para hacerla robusta frente a distintos tipos de corrupción de imágenes. En concreto, se propone una convolución basada en push-pull para sustituir la primera capa convolucional de una ConvNet. Los amplios experimentos realizados con los conjuntos de datos ImageNet-C y CIFAR10-C demuestran la eficacia de este enfoque.

