Comparative Study of DL-based Methods Performance for Camera Model Identification with Multiple Databases

Alexandre Berthet, Jean-Luc Dugelay; EURECOM; Sophia Antipolis, France

Abstract

Camera identification is an important topic in the field of digital image forensics. There are three levels of classification: brand, model, and device. Studies in the literature are mainly focused on camera model identification. These studies are increasingly based on deep learning (DL). The methods based on deep learning are dedicated to three main goals: basic (only model) - triple (brand, model and device) - open-set (known and unknown cameras) classifications. Unlike other areas of image processing such as face recognition, most of these methods are only evaluated on a single database (Dresden) while a few others are publicly available. The available databases have a diversity in terms of camera content and distribution that is unique to each of them and makes the use of a single database questionable. Therefore, we conducted extensive tests with different public databases (Dresden, SOCRatES, and Forchheim) that combine enough features to perform a viable comparison of DL-based methods for camera model identification. In addition, the different classifications (basic, triple, open-set) pose a disparity problem preventing comparisons. We therefore decided to focus only on the basic camera model identification. We also use transfer learning (specifically fine-tuning) to perform our comparative study across databases.

Introduction

At the beginning of the century, thanks to the development of digital devices (cameras, mobile phones, etc.), the access to images and videos increased to the point of becoming an important communication channel. In the meanwhile, the modification of digital images has become easier thanks to free and accessible image editors. These modifications can be applied to enhance the quality of an image, but it can sometimes be malicious. In some applications like trial or police investigation, images represent crucial proof and their authenticity has to be proven. Camera identification, a field of digital image forensics (DIF) [1], is a solution to image tampering detection by identifying the source camera of an image. From a considered image, the principle is to identify the associated camera among a group of cameras. There are three levels of classification: the brand, the model and the camera itself (device). Most of the work focuses on identifying the model of the camera. Several techniques have been found, using artifacts left over from the acquisition of a digital image [2]. Notably, the sensor pattern noise (SPN) [3] or the photo response non-uniformity (PRNU) [4] have first permitted establishing the pattern of cameras. Then, with the democratization of deep learning, the performances were improved in particular thanks to convolutional neural networks (CNN). A CNN takes an image as input, which is first analyzed by the feature extractor that generates an output map, which is then used in the classification part. The feature extractor consists mainly of convolutional layers that examine the image patterns from low- to high-level, and pooling layers that reduce the size of the feature map. The output map is then analyzed by fully connected layers to assign the image to a class (camera model). Among these deep learning methods, there are three classification approaches: i) basic - only the model; ii) triple - brand, model and device; iii) open-set - known (in the training set) and unknown (outside the training set) cameras. Regardless of the approach, these DL methods need databases of camera or smartphone images to learn and to be effective. Despite the availability of such databases, the Dresden image database [20] is often the only one used in most publications (see Fig. 1). In some cases, private datasets are used as a second dataset, but they cannot be reused because they are not publicly available. This evaluation process contrasts with other areas of image processing, such as face recognition, where methods are typically trained on a specific database and evaluated on other databases.

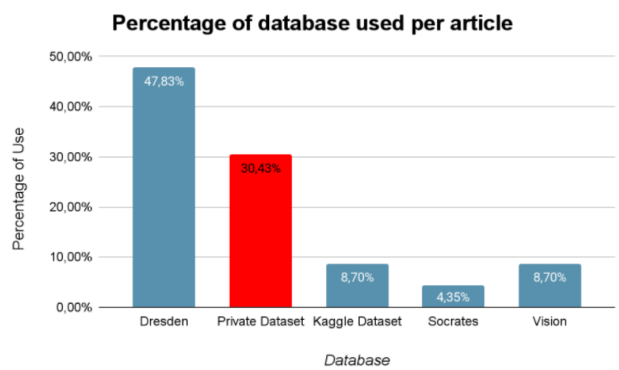

Figure 1: Statistics made among 13 articles linked to cameras: identification, extraction of pattern, etc. (blue) publicly available database (red) private datasets (not publicly available)

The purpose of our manuscript is to provide a comprehensive comparative study to evaluate state of the art (SOTA) methods for camera model identification, which includes two difficulties. The first one is the diversity of classification among the methods. This disparity prevents any comparison, even when the same database is used. To conduct our comparative study independently of this problem, we decided to focus on the basic camera model identification, without taking the additional results (brand, unknown, etc.) into account. The second problem is the lack of database diversity that is prevalent in the literature. Evaluating DL-based methods with a single database (i.e., Dresden) is not consistent with the usual evaluation process in image processing and computer vision. The results from such an evaluation are not reliable with respect to the diversity of the databases and their different characteristics. This problem comes from the digital fingerprint

Table 1: Performance of SOTA methods on Dresden for different applications of camera identification.

Applications	References	Brand		Model		Device	
		Nbr. cameras	Accuracy	Nbr. cameras	Accuracy	Nbr. cameras	Accuracy
Basic classification	[6]			18	92.83%		
	[7]			12 14 33	98% 97.09% 91.9%		
	[14]			26	98.58%		
Triple classification	[8]	13 19	99.12% 97.73%	27	94.73%	74	45.81%
	[9]	13	99.4%	27	96.1%	74	47.5%
	[18]	13	99.6%	27	97.2%	74	52.4%
Open-set Classification	[10]			25	94%		
	[13]			25	97.74%		

of cameras, which is used for identification and is unique to each device. Therefore, the classification of cameras from a database not used to train the network is more difficult. In fact, the network learns to extract a fingerprint from an image, but cannot classify fingerprints that it has not analyzed. Thus, while deep learning methods have been shown to be effective for camera pattern classification, there is a data dependency problem specific to this topic. Therefore, we conducted our comparative study on three complementary and publicly available databases (Dresden [20], SOCRatES [21] andd Forchheim [22]) that have as many features as possible: all camera types, a diverse distribution and a different number of cameras. Our evaluation will focus on identifying the basic camera model. The protocol of evaluation is based on transfer learning, an area of deep learning, and especially on the fine-tuning of the network to evaluate the methods on each database. We have applied our evaluation on four methods from the literature, based on various architectures and realized for various applications. Essentially, this article provides identification accuracy for these methods on an equal footing for diverse databases.

The section Introduction presents the camera identification and issues raised by our literature review, as well as the objective of this paper. In the section Addressed problem, we further explain the camera model classification problems that motivated us to propose this paper. In the section Camera model identification methods, we present the SOTA methods that we replicated for our performance evaluation. The section Proposed Evaluation presents the databases and properly explains the protocol of our evaluation. This section also presents the results of our performance evaluation. Finally, the section Conclusion concludes on the effectiveness of each database as well as the robustness of each of the methods studied.

Addressed problem

Despite the abundance of articles regarding camera model identification, there are some issues in the literature. The key problem remains the difficulty of comparing SOTA methods. This problem stems from the protocol of evaluation, which differs from one method to another, making it impossible to compare their performance. Moreover, as we showed in Fig. 1, most of the works are evaluated on Dresden. This lack of diversity in the databases leads to evaluation that are not reliable and complete. In this section, we illustrate these two issues through articles in the literature.

Protocols Diversity

The first issue that makes the comparison of SOTA methods difficult is the diversity of protocols. On one hand, this diversity represents a valuable aspect, as it contributes to the improvement and enrichment of camera model identification. But, on the other hand, this variety of protocols reduces the possibilities of comparing methods and evaluating on an equal footing their performance. Several approaches exist for camera model identification, and three main applications can be particularly identified from the literature: basic classification, triple classification, and open set classification. i) The basic scenario consists in identifying the camera according to a label (brand, model, or device) and is the most specific and widely discussed application. Most of the methods for this scenario classify camera models as in [6], [7]. ii) The triple classification is close to the basic scenario, as the goal is to identify the camera according to a label. But, as the name implies, it classifies cameras according to the brand, the model and the device, as in [8, 9]. The triple classification aims to identify cameras based on all labels and is more global. iii) Open set classification is special and addresses the robustness problem. The goal of this approach is to evaluate the generalization of a method by classifying cameras not seen by the network (labeled as unknown) as in [10]. As the table 1 shows, the comparison is hard to index because of the various applications and their difference of evaluation. Even among methods using the same protocol of evaluation, there is still a problem of number of cameras used for experiments. In [7], a series of experiments is conducted to underline the problem of data dependency by increasing the number of camera models in the dataset. The evaluation targets a basic model classification along three experiments with an increasing number of camera models (12 - 14 - 33). In [6], the problem of unknown cameras, which means that the network has never learned their features, is addressed with an additional experience. The issue of unknown cameras is also addressed in [10] with the open set scenario. This paper presents a Siamese network to classify pairs of images as known or unknown cameras. For this application, the evaluation was performed with a series of experiments based on three subsets according to an image pair: 1) only known camera models; 2) one known and one unknown model; 3) only unknown camera models. The protocol of evaluation was performed with 65 camera model. The aspect that stands out in the SOTA publications is the disparity of applications and thus of protocol for evaluating their method.

Database dependency

The open set classification is particularly interesting, as it addresses the problem of robustness. The evaluation of this aspect is challenging because each camera has its own fingerprint, which corresponds to the combination of artifacts generated during the creation of a digital image. Therefore, the methods in the literature are often dedicated to the cameras of the database used for their training. This database dependency, illustrated with the use of Dresden in almost 50% of the literature, represents the second issue to overcome for comparing SOTA methods. In fact, most methods use the Dresden Image database [20], which is a standard in terms of camera identification, containing 27 camera models. This overuse represents an advantage for performance evaluation because the methods can be compared on comparable terms. However, except for a few articles using an equivalent number of camera models (i.e., 27), such as [8], [9] for the triple classification, most methods classify disparate numbers of camera models. This difference is highlighted in the previous subsection as well as in the Tab. 1). In addition, using a single database for performance evaluation poses another problem related to robustness. In another area of image processing, the standard method for evaluating methods requires the use of multiple databases. Until now, the performance evaluation has almost always been conducted on the same database as the training part and most often on a single database (N.B. Dresden). This is understandable because the identification of the camera is done through its fingerprint, which is unique to each camera. Therefore, evaluating a method on another database than the one used for training is too difficult and leads to a decrease in performance. However, it allows proving the robustness of a method from one database to another.

From these particularities of the camera identification literature, we decided to propose our own comparative study to overcome the issues raised previously. The first aspect of our proposition is to provide a standard protocol to evaluate the performance of camera model identification methods. This standard is based on three databases (detailed in) to fix the problem of data-dependence: the Dresden Image Database [20], SOCRatES [21] and the Forchheim Image Database [22]. Our protocol of evaluation is fully explained later in the section Proposed Evaluation. However, to conduct a complete comparison for camera model identification, we have replicated four methods from the literature, which have their own particularities, and that are detailed in the next section.

Camera model identification methods

We have selected four SOTA methods that have proven to be effective in terms of performance. There are several protocols discussed in the literature that we detailed earlier in the section Addressed problem. In particular, to cover the full diversity of camera model identification, we took at least one method per application. In addition, we opted for different architectures in order to perform a comprehensive evaluation. To be more specific, we chose to focus on the constrained convolutional network [15] first introduced for forgery detection. This network is relatively new and has been an important improvement for preprocessing modules in deep learning methods for digital image forensics [11]. We also discussed the Siamese network, which is currently used in the literature, especially to overcome the robustness problem. The other architectures are respectively based on an improved preprocessing module and a particular layer block inspired by residual neural networks [12]. Therefore, in this section, we describe the methods we selected for our performance evaluation.

The article [13] presents a constrained convolutional network for unknown camera identification, which they introduced in a first place for forgery detection. In this architecture, the first layer is transformed from a convolutional layer to a preprocessing module. The goal is to extract the desired artifacts inside the network to achieve an end-to-end architecture. In fact, in many digital image forensic approaches, a preprocessing module is included to isolate artifacts that are overshadowed by the image content. In this network, artifacts are therefore extracted by constraining the weights of the first convolutional layer. First, these weights are randomly initialized and then forced during the back propagation of the network to learn the prediction error filters. The weights w of each filter K are forced as follows: the central value is set to -1 while the sum of the remaining pixels is set to 1 (see Eq. 1).

$$\begin{cases} w_k(0,0) = -1 \\ \sum_{l,m\neq 0} w_k(l,m) = 1 \end{cases} \tag{1}$$

The identification of unknown cameras has been conducted on Dresden [20] and a personal dataset. They defined the set T of known cameras with two types: "known" for training and "known unknown" (not used for training) to be classified. The set T consists of 10 different known camera models from Dresden and 15 known unknown camera models from their personal dataset. They also defined a set T' of "unknown" cameras, which represent cameras to be classified that are not in the base set T. The set T' is made of 15 new camera models from their personal dataset. The evaluation has been performed with a protocol of three experiments: i) They used the set T for the first evaluation. ii) They used the combination of both set T and T'. iii) The final experiment is a subset composed of "known" cameras from the set T and "unknown" cameras from the set T'. The architecture is divided into three parts: 1) the preprocessing module with the constrained convolutional layer [15]; 2) the feature extractor with four convolutional blocks consisting of a convolutional layer, batch normalization, hyperbolic tangent (tanh) activation, and pooling (max and average for the last one); 3) classification with 3 fully connected layers with tanh activation and soft max for the last one. Their results for each experiment are the following: 1) 99.38%; 2) 98.57%; 3) 97.74%.

In another of their article [14], they also proposed a network based on this constrained convolution network [15]. The only change in this architecture, compared to the previous one, is the preprocessing module. They decided to augment their model feature maps by combining two layers of artifact extraction: the constrained convolutional layer and a residual median filter. The goal was to evaluate the robustness of this artifact fusion for camera model classification. They conducted a series of experiments to compare their approach with other methods such as the high-pass filter (HPF) based CNN and the classical constrained convolutional network. They evaluated the robustness of these methods with resampling (120%, 90%, 50%), JPEG (QF=90), resampling+JPEG and normal images. In most cases, the proposed method showed the most efficient performance and, except for one test (resampling 50% + JPEG QF=90), the accuracy was always higher than 90%.

The generalization of camera pattern identification is addressed with a new approach in [16]. The objective of this work was to propose a method to determine whether two images come from the same camera without knowing their forensic traces. Their method is based on the Siamese network architecture: a feature extractor used twice in parallel to produce deep features from two image patches and a similarity network to compare them. For the feature extractor, they used only the constrained convolutional network of [17]. It is mainly the same as before, except for the classification made only of two fully connected layers. The similarity network is composed of three parts: 1) a first fully connected layer for each branch; 2) then, each branch as well as their product passes through an artificial layer f_{inter} (Eq. 2); 3) a concatenation layer, a fully connected layer and a sigmoid activation to obtain a similarity score. For their experiments, they additionally dealt with aspects of "known" and "unknown" camera models to clearly evaluate the generalizability of their method. Accuracy was calculated for three distinct cases: with known camera models only (95.93%), with unknown and known camera models (93.72%), and with unknown camera models only (92.41%). They also evaluated the best configuration (patch size, architecture, etc.) and the effects of other parameters (recompression, unknown marks) for their method.

$$f_{inter}(X) = \phi(\sum_{i=1}^{N} w_{k,i} \quad f_i(X) + b_k) \tag{2}$$

In [18], multi-classification is addressed by identifying the camera according to the triple classification (device, model and device). The method relies on a preprocessing module based on domain knowledge and ResNet blocks. The ResNet [12] is known for its shortened layers that improve diversity by allowing more exploration of the feature space. A ResNet block is defined in [18] as the merging of features obtained in parallel by two consecutive 3 convolutional layers. The architecture is developed in four parts: the preprocessing module and three successive sections for each classification (brand, then model and device). Each section consists of three consecutive ResNet blocks followed by a classification part composed of a global average pooling, a fully connected layer and a soft max layer. The preprocessing module consists of three parts: 1) a multi-scale HPF to obtain three different residuals; 2) a convolutional layer and a ResNet block applied on the 4 elements (HPF and input); 3) a concatenation layer to obtain the final output. They achieved 97.1% accuracy for camera model classification.

Methods	Preprocessing	Features extractor	Classification
Bayar [13]	Const. conv.	4 conv. blocks	3 f.-c.
Bayar [14]	Const. conv + MFR	4 conv. blocks	3 f.-c.
Mayer [16]	Const. conv	4 conv. blocks + 2 f.-c.	Similary network
Ding [18]	multi-scale HPF	3 ResNet blocks	3 class. blocks

Table 2: Details of method architecture, according to the preprocessing, the features extractor and the classification.

The table 2 summarizes completely the purpose of our choice. On one hand, we can observe four completely diverse methods, even if three of them are using the constrained convolutional layer for preprocessing. However, it is challenging to estimate which one is the most robust and the most efficient. Therefore, we have conducted a comprehensive evaluation through our protocol based on three databases that we explained in the next section.

Proposed Evaluation

Databases

Until now, the methods of the literature have proved a lack of databases diversity that we explained in the section Addressed problem. Notably, most of the methods have used only the Dresden Image Database. In opposition, we decided to exploit three databases in our evaluation, dedicated to camera model identification. We therefore used Dresden [20] because it is a standard in the literature. In addition, we exploited SOCRatES [21] and the Forchheim Image Database [22] that are dedicated to smartphone cameras. In this section, we detail the content of each database.

The *Dresden Image Database* [20] is perhaps the most popular database in the field of digital image forensics. This database is principally dedicated to camera identification, but it is also employed for forgery detection (with application of manipulations to the images). It is composed of more than 14,000 images of various indoor and outdoor scenes that were captured by 74 cameras (14 different brands and 27 models) to establish their characteristics perfectly. As Gloe *et al.* detailed in their article, developing forensic methods for camera model identification requires numerous images per camera of the same scene to conduct a complete comparison of their forensics traces.

SOCRatES [21]: SOurce Camera REcognition on Smartphones is an image and video database especially designed for source digital camera recognition on mobile devices. SOCRatES is currently one of the databases for source digital camera identification with the largest number of different devices. It is made up of about 9,700 images and 1,000 videos captured with 103 different smartphones (one unclassified) of 15 different brands and about 60 different models. The acquisition has been performed in uncontrolled conditions. To collect the database, many people were involved and asked to use their personal smartphone to collect a set of pictures. Instructions were transmitted to the participants, and they collected the set of pictures in complete autonomy. The reason behind this choice is, on the one hand, to collect a database of heterogeneous pictures and to maximize the number of devices employed, and, on the other hand, to carefully replicate the real scenario of application of the techniques that will use this database as benchmark.

The *Forchheim Image Database* [22] is also a database dedicated to source camera identification based on smartphones. The Forchheim is one of the wider databases available for source camera identification, with more than 23,000 images of 143 scenes capture by 27 different smartphone cameras (9 brands and 25 models). All the images have been captured in or near the town of Forchheim in Germany (hence the name) and each camera presents one image per scene. Moreover, each image is providing in six diverse qualities to assert of a disparate quality: the original camera-native version and five copies from social networks.

We chose these databases for their particularities in terms of variety - cameras (Dresden) and smartphones (SOCRatES and Forchheim) - and composition with different ratios of models per camera: 36.48% for 74 cameras in [20], 60.78% for 102 cameras in [21] and 92.59% for 27 cameras in [22]. With three databases, we want to plainly observe the robustness of the methods with our evaluation.

Robustness Protocol

The protocol for conducting our comparative study uses transfer learning, a field of deep learning. In particular, network refinement facilitates the evaluation of methods from one database to another (N.B. Dresden, SOCRatES and Forchheim). The protocol is divided into two parts: 1) creating a reference network for each database; 2) transferring the architecture and weights of the reference networks to obtain the transferred networks.

The first step of our protocol is to create a reference network for each possible combination of method and database. Our study is based on three databases (Dresden, SOCRatES and Forchheim) and four methods (Bayar, BayarMFR, Mayer and Ding), so we obtained a total of 12 networks for the first step. On one hand, this part provides guidance on the effectiveness of a method depending on the database, but also on the most difficult database. On the other hand, this first step allows us to obtain the reference networks necessary for the second step.

The second step requires refining (i.e., re-training) the transferred network. The principle of transfer learning is to build a new network for evaluating on a database *B2*, by transferring the architecture and weights of a network pre-trained on a database *B1*. There are alternative approaches depending on the baseline (DB1) and the transferring databases (DB2):

1. Full network training (*B1*, different *B2* and *B2* large)
2. Partial training of the network: a few layers (*B1*, similar *B2* and large *B2*) or several layers (*B1*, different *B2* and small *B2*)
3. Driven only the classification part of the network (*B1*, *B2* similar and *B2* small)

In our case, we decided to refine only the classification part to adapt a reference network to a new database. A transferred network is created for each database from the reference network. Therefore, we obtained 3 networks per methods to assess on each database (see table 3). The goal of this second step is to analyze the robustness of each method, but also to evaluate which database is the most efficient to create the reference network.

Reference Database (DB1)	Transferred Database (DB2)	Total
Dresden	Socrates (D-S)	**3 networks**
	Forchheim (D-F)	
Socrates	Dresden (S-D)	**3 networks**
	Forchheim (S-F)	
Forchheim	Socrates (F-S)	**3 networks**
	Forchheim-Dresden (F-D)	

Table 3: Scheme of the steps of the protocol with reference and transferred databases.

Experimental Study

The experimental study was conducted on the presented method with the previously detailed databases. To evaluate these methods with our protocol, we divided each database into three datasets for training, validation and testing (80:10:10). Then, from each image, we extracted patches of size 128×128 and applied much preprocessing to fit them to the input of the networks. In particular, we applied successive Gaussian filters on the patches to obtain four different inputs, as depicted in [18]. For the constrained convolutional network of [13], the preprocessing is included in the first, but we had to apply a median filter on the patches to create the second input of the improved constrained convolutional network proposed in [14]. For the method in [16], we created pairs of patches because it is a Siamese network that requires two inputs. Finally, we selected 2.5M, 0.59M, and 1.16M patches from Dresden, SOCRatES, and Forchheim, respectively, with a 60:20:20 distribution for training, validation, and testing. Once the datasets were defined, we trained the selected methods according to our protocol, i.e., in two parts: the creation of reference networks and the transferred networks. For the creation of baseline networks, we trained the models for 30 epochs with the hyper-parameters specific to each method. Then, we trained the transferred networks for 10 epochs to limit the adaptation to the new database, with the hyper-parameters specific to each method too.

Network	*Bayar* [13]	*Bayar-MFR* [14]	*Mayer* [16]	*Ding* [18]
Evaluation on Dresden				
Dresden	91.78%	92.11%	**96.17%**	84.66%
S-D	90.20%	92.19%	86.01%	**96.68%**
F-D	89.08%	91.68%	80.42%	**97.30%**
Evaluation on Socrates				
Socrates	75.20%	75.05%	**92.31%**	68.08%
D-S	79.43%	**80.98%**	75.37%	74.95%
F-S	73.97%	77.21%	81.82%	**87.45%**
Evaluation on Forchheim				
Forchheim	56.14%	57.65%	80.18%	**82.89%**
D-F	58.44%	60.63%	66.67%	**69.08%**
S-F	57.01%	58.56%	79.90%	**81.14%**

Table 4: Results of each network according to the database of evaluation and the method used.

The results obtained with the accuracy metrics give a complete overview of the camera model identification (see table 4). In contrast to the table 1, the methods can be compared with each other regardless of their approach. Therefore, the comparative study is possible, and some conclusions and discussions can be made. On the one hand, [18] seems to be the best method when changing the database between training and testing (6 out of 9 best results). On the other hand, [16] obtains the best accuracy when training and testing are performed on the same database. Methods [14] and [13] show stability in performance regardless of the training database, but their results are less convincing than those in table 1. On the other hand, Dresden appears to be the easiest database to handle (overall accuracy of 90.69%). These results confirm that assessing methods on a single database is irrelevant and that other databases are required.

Conclusion

In this paper, we propose a comparative evaluation for identifying camera models from three databases: the Dresden image database, SOCRatES and the Forchheim image database. To date, this paper is the first to evaluate methods on multiple publicly available databases. This evaluation takes into account two crucial problems in the literature: the use of a single database and the disparity in approaches between methods. The comparative study is therefore possible across various databases and across method approaches. Moreover, the results show that the method based on the triple classification [18] outperforms the others in most cases.

References

[1] Redi, J.A., Taktak, W. and Dugelay, JL. Digital image forensics: a booklet for beginners. Multimed Tools Appl 51, 133–162 (2011). https://doi.org/10.1007/s11042-010-0620-1

[2] Farid, H. "Image forgery detection." IEEE Signal Processing Magazine 26 (2009): 16-25.

[3] C. Li, "Source Camera Identification Using Enhanced Sensor Pattern Noise," in IEEE Transactions on Information Forensics and Security, vol. 5, no. 2, pp. 280-287, June 2010, doi: 10.1109/TIFS.2010.2046268.

[4] T. Filler, J. Fridrich and M. Goljan, "Using sensor pattern noise for camera model identification," 2008 15th IEEE International Conference on Image Processing, San Diego, CA, USA, 2008, pp. 1296-1299, doi: 10.1109/ICIP.2008.4712000.

[5] Guodong Guo, Na Zhang, "A survey on deep learning based face recognition", Computer Vision and Image Understanding, Volume 189, 2019, 102805, ISSN 1077-3142, https://doi.org/10.1016/j.cviu.2019.102805.

[6] L. Bondi, L. Baroffio, D. Güera, P. Bestagini, E. J. Delp and S. Tubaro, "First Steps Toward Camera Model Identification With Convolutional Neural Networks," in IEEE Signal Processing Letters, vol. 24, no. 3, pp. 259-263, March 2017, doi: 10.1109/LSP.2016.2641006.

[7] A. Tuama, F. Comby and M. Chaumont, "Camera model identification with the use of deep convolutional neural networks," 2016 IEEE International Workshop on Information Forensics and Security (WIFS), Abu Dhabi, 2016, pp. 1-6, doi: 10.1109/WIFS.2016.7823908.

[8] Y. Chen, Y. Huang and X. Ding, "Camera model identification with residual neural network," 2017 IEEE International Conference on Image Processing (ICIP), Beijing, 2017, pp. 4337-4341, doi: 10.1109/ICIP.2017.8297101.

[9] M. Zhao, B. Wang, F. Wei, M. Zhu and X. Sui, "Source Camera Identification Based on Coupling Coding and Adaptive Filter," in IEEE Access, vol. 8, pp. 54431-54440, 2020, doi: 10.1109/ACCESS.2019.2959627.

[10] O. Mayer and M. C. Stamm, "Learned Forensic Source Similarity for Unknown Camera Models," 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, 2018, pp. 2012-2016, doi: 10.1109/ICASSP.2018.8462585.

[11] A. Berthet and J. -L. Dugelay, "A review of data preprocessing modules in digital image forensics methods using deep learning," 2020 IEEE International Conference on Visual Communications and Image Processing (VCIP), 2020, pp. 281-284, doi: 10.1109/VCIP49819.2020.9301880.

[12] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

[13] B. Bayar and M. C. Stamm, "Towards Open Set Camera Model Identification Using a Deep Learning Framework," 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 2018, pp. 2007-2011, doi: 10.1109/ICASSP.2018.8462383.

[14] B. Bayar and M. C. Stamm, "Augmented convolutional feature maps for robust CNN-based camera model identification," 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 2017, pp. 4098-4102, doi: 10.1109/ICIP.2017.8297053.

[15] Belhassen Bayar and Matthew C. Stamm. 2016. A Deep Learning Approach to Universal Image Manipulation Detection Using a New Convolutional Layer. In Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec '16). Association for Computing Machinery, New York, NY, USA, 5–10. DOI:https://doi.org/10.1145/2909827.2930786

[16] O. Mayer and M. C. Stamm, "Forensic Similarity for Digital Images," in IEEE Transactions on Information Forensics and Security, vol. 15, pp. 1331-1346, 2020, doi: 10.1109/TIFS.2019.2924552.

[17] B. Bayar and M. C. Stamm, "Constrained Convolutional Neural Networks: A New Approach Towards General Purpose Image Manipulation Detection," in IEEE Transactions on Information Forensics and Security, vol. 13, no. 11, pp. 2691-2706, Nov. 2018, doi: 10.1109/TIFS.2018.2825953.

[18] X. Ding, Y. Chen, Z. Tang and Y. Huang, "Camera Identification Based on Domain Knowledge-Driven Deep Multi-Task Learning," in IEEE Access, vol. 7, pp. 25878-25890, 2019, doi: 10.1109/ACCESS.2019.2897360.

[19] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 770-778.

[20] Thomas Gloe and Rainer Böhme. 2010. The 'Dresden Image Database' for benchmarking digital image forensics. In Proceedings of the 2010 ACM Symposium on Applied Computing (SAC '10). Association for Computing Machinery, New York, NY, USA, 1584–1590.

[21] Galdi, C., Hartung, F., and Dugelay, J. L. (2019, February). SOCRatES: A Database of Realistic Data for SOurce Camera REcognition on Smartphones. In ICPRAM (pp. 648-655).

[22] Hadwiger, Benjamin and C. Riess. "The Forchheim Image Database for Camera Identification in the Wild." ICPR Workshops (2020).

Author Biography

Alexandre Berthet is currently PhD student in Digital Security at EURECOM – Sophia Antipolis since September 2019, with Professor Jean-Luc Dugelay as supervisor, working on DEFACTO project, whose subject is "Digital Image Forensics with Deep Learning".

Jean-Luc DUGELAY obtained his PhD in Information Technology from the University of Rennes in 1992. His thesis work was undertaken at CCETT (France Télécom Research) at Rennes between 1989 and 1992. He then joined EURECOM in Sophia Antipolis where he is now a Professor in the Department of Digital Security. His current work focuses on the domain of multimedia image processing, in particular activities in security (image forensics, biometrics and video surveillance, mini drones), and facial image processing. His research group is involved in several national projects and European projects. He is a fellow member of IEEE, a fellow member of IAPR, and an elected member of the EURASIP BoG. Jean-Luc Dugelay is (or was) associate editor of several international journals (IEEE Trans. on IP, IEEE Trans. on MM) and is the founding Editor-in-Chief of the EURASIP journal on Image and Video Processing (SpringerOpen).