AUTHOR VERSION COMPILED ON January 25, 2025

© IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.
DOI: 10.1109/IWBF62628.2024.10593842

On the Influence of Texture Strength and Camera Signal on Micro-Texture Surface Classification

Johannes Schuiki ♦ Christof Kauba ♦ Heinz Hofbauer ♦ Andreas Uhl

Department of Artificial Intelligence and Human Interfaces, University of Salzburg, Austria

Abstract

The digital cameras contained in nearly all modern smartphone devices can be utilised for a broad range of applications, e.g. biometric recognition and extrinsic product authentication which are well-established and work reliably. However, previous research revealed undeniable problems with micro-structure surface classification if different smartphone devices are involved (cross-device scenario). This work is a step further towards understanding the underlying reasons by employing different training strategies (training on general texture patches and testing on the dental ceramic ones), analysing the texture properties (information entropy) and testing the performance of pairs and triples of the same make and model of smartphones compared to arbitrary device pairs.

Contents

This research was partially funded by the Salzburg State Government within the Science and Innovation Strategy Salzburg 2025 (WISS 2025) under the project AIIV-Salzburg (Artificial Intelligence in Industrial Vision), project no 20102-F2100737-FPR.

AUTHOR VERSION COMPILED ON January 25, 2025

1 Introduction

Nowadays smartphone devices are utilised for a variety of different tasks. Thanks to the digital cameras contained in every modern smartphone device they can also be employed as biometric capturing devices for personal authentication [1] but also for various kinds of classification purposes, e.g. detecting paving materials in urban environments [2] or wood type identification [3]. The classification of object surfaces based on their micro-structure texture in combination with smartphone devices to capture the images of the objects' surfaces (to alleviate the need for special microscopes or other optical devices) is an essential part in various practical applications with mobile product authentication being only one example.

There are several works showing that intrinsic micro-structure based material classification is feasible using commodity smartphone devices, e.g. in [4]–[6] based on physical non-cloneable functions (PUFs), relying on the surface micro-structure. Sun et al. [7] utilised an efficient micro-structure orientation estimation technique, modelling the entire propagation path of the light. Schraml et al. [8] presented an approach for drug packaging authentication using micro-texture images in an open set scenario. While for most applications (biometric authentication, extrinsic product authentication, object detection and classification) a cross-device scenario (more than one different smartphone device used for training/enrollment and classification/authentication) does not impact the performance, cross-device intrinsic micro-structure based material classification has hardly been investigated so far. There is some evidence that the material classification performance is reduced in the cross-device setting e.g. in [9], mentioning scaling issues (different dpi resolutions) as well as artefacts introduced during the image processing pipeline as main reasons for the non-satisfactory performance.

Motivated by those works, our previous work [10] investigated the applicability of surface micro-structure images captured with commodity off-the-shelf smartphone devices in combination with a clip-on macro lens to perform intrinsic product authentication for zircon oxide blocks (commonly used for dental ceramics). The results confirmed that the micro-structure based surface classification works well in the intra-device (training and evaluation only on the same smartphone device) scenario but revealed undeniable problems in practical applications if more than one type of smartphone device is employed. While the PRNU [11] could be ruled out as a reason, a device-inherent signal, stemming from the image sensor itself as well as an additional signal from the image processing tool-chain (ISP) were identified as possible reasons for the inferior cross-device performance. Our follow up work [12] attempted to shed some more light on the limiting factors and their underlying reasons in cross-device micro-structure material classification. The initial zircon oxide block dataset was extended (7 smartphone devices instead of the original 4) and in addition to the JPG images, samples were captured in raw mode (to rule out artefacts stemming from the ISP). Moreover, a denoising filter was applied to the images in order to remove the PRNU noise signature. The experiments confirmed that there is a device-inherent signal, which while being stronger for the ISP processed JPG images, is even present in the raw images and can partly be attributed to the different physical structure of the image sensors in different smartphone models (the arrangement of the single pixels - colour filter array). Hence, the device-inherent signal is assumed to be composed of an ISP depended part and a sensor hardware-related part, both not necessarily independent (the ISP works on the images containing the hardware-related signal). Additional device identification experiments (instead of classifying the different surface structures, the capturing devices are classified) showed that this signal is strong enough to perform reliable device identification (with an accuracy of 99%) using the zircon oxide block images.

For the domain of natural images, recent works tried to separate the image content from the device-specific artefacts introduced during the imaging process: Cozzolino and Verdoliva [13] employed a convolutional neural network (CNN) architecture to extract a noise residual coined *noiseprint* which serves as a model-specific fingerprint. Chen et al. [14] adopted a Siamese network architecture to remove artefacts in natural images introduced by the capturing device as an anti-forensics tool. Manisha et al. [15] utilized a CNN architecture to explore the existence of device specific artefacts other than PRNU.

The main aim of this work is to further investigate the influence and characteristics of this device-inherent signal for the domain of micro-texture surfaces, especially the interdependence of the image content and the device-inherent signal. This is done by employing different types of textures (two additional texture datasets) and a train-test strategy for device identification, where training is performed on general texture data and testing is

AUTHOR VERSION COMPILED ON January 25, 2025

done on zircon oxide block samples (and vice versa). The ceramic samples serve as a representative of surfaces with little structure and low texture similar to medical packaging [8] and paper structure [4], i.e. the results apply for other surfaces of similar type analogously. Moreover, the individual samples are categorised according to their image entropy and the different categories are evaluated separately to analyse the influence of certain texture/surface properties on the classification performance. In contrast to the PRNU [11] which is instance/device specific, our previous results indicate that the device-inherent signal could be rather model than instance specific. Hence, there should be a significant performance difference for device identification if pairs of the same device model are employed during training and testing (lower classification performance) compared to pairs of different devices. In order to verify this assumption, the zircon oxide block dataset is extended to 19 devices in total, by capturing sample data with four pairs (same manufacturer and model) and one triple of devices. Of course, the other two texture datasets are also captured with all 19 devices. To further verify the assumption that the device-inherent signal is model specific, the material classification experiments are conducted for the pairs/triples of the same devices and compared to arbitrary sets of devices. If the assumption holds, the material classification performance should be influenced in the opposite way than the device identification one, i.e. higher performance for the pairs/triples of the same model compared to the arbitrary sets.

The rest of the work is organised as follows: Section 2 describes the datasets and their acquisition. The experimental set-up, the utilised classification approach and the evaluation results are explained in Section 3. Finally the conclusion and an outlook on future work is provided in Section 4.

2 Database

The acquired dataset used in the experimental part of this study can be divided into three subsets / texture types:

Ceramic The ceramic database comprises of images acquired from the top side of zircon oxide blocks (see Fig. 1) produced by three different manufacturers: 10 blocks produced by Ivoclar Vivadent, 6 by Dentsply and 16 by 3M. Examples of the already cropped ceramic patches can be seen in Fig. 2.

Figure 1: Unprocessed Zircon Oxide Ceramic Blocks.

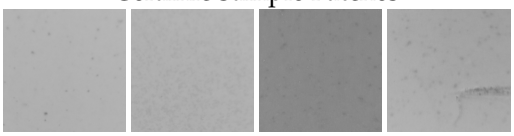

Figure 2: Examples of ceramic patches.

Zinterhof Zinterhof sequence [16] are a simple form of texture with an even spread of uniformly distributed points in a uniform square and a special form of Weyl sequences. In dimension $d = 2$ with a uniform distribution in the union square $U_s = \cup_{i=1}^{d}[0,1[$. Weyl sequences are defined as $x_n = (\{n\theta_1\}, \ldots, \{n\theta_d\})\ n = 1, 2, 3, \ldots$, where $\{a\}$ is the fractional part of a. Weyl sequences are uniformly distributed if and only if θ_i are independent irrational numbers. For Zinterhof sequences we set $\theta_i = e^{1/i}$, and consequently the sequence of points to $z_n = (\{ne^1\}, \{ne^{1/2}\})\ n = 1, 2, \ldots, N$, where N is set depending on the number of points we require. The resulting images, with the number of points N to produce them, can be seen in Fig. 3. The textures were then printed using a laser printer.

Brodatz The Brodatz Texture Database [17] consists of 112 texture images. Six textures were chosen which can be seen on Fig. 4.

For every ceramic block and every printed texture, four images, one on every corner, were captured using a macro lens *Agritix WIDK-24X01 Xylorix Wood Identification Tool* clipped onto a smartphone's camera. The macro lens has a circular illumination ring which was used during the acquisition to suppress

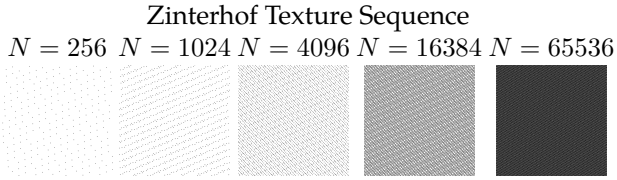

Figure 3: Zinterhof textures: produced by painting the points in a 600x600 pixel image. N indicates the points per image.

AUTHOR VERSION COMPILED ON January 25, 2025

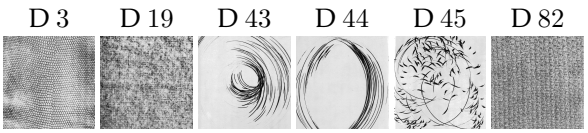

Figure 4: The six Brodatz textures used in this work.

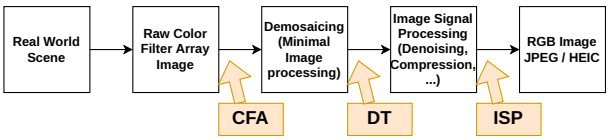

Figure 5: Imaging pipeline with measurement points for CFA, DT and ISP.

Table 1: Smartphones and their imaging sensor resolution.

Smartphone		Image Resolution	Scaling Factor
Google Pixel 1		3036 × 4048	0.683
Google Pixel 4a*	× 3	3024 × 4032	0.686
Huawei P20 Lite*		3456 × 4608	0.600
Huawei P30 Pro*		2736 × 3648	0.758
iPhone 11*		3024 × 4032	0.686
iPhone 13	× 2	3024 × 4032	0.686
iPhone 13 Pro*		3024 × 4032	0.686
iPhone X		3024 × 4032	0.686
iPhone XS	× 2	3024 × 4032	0.686
LG Nexus 5	× 2	2448 × 3264	0.847
Nokia 6.1	× 2	3456 × 4608	0.600
Samsung Galaxy A52*		3468 × 4624	0.598
Xiaomi Mi A3*		3000 × 4000	0.691

the influence of varying lighting conditions and enhance the visibility of the surface micro-structure.

In addition to the described textures, checkerboard pattern images were acquired, which revealed that the fields of view between the various devices are similar but the resolutions are only coarsely matching. Therefore, a resolution normalization is applied by down-scaling the images to a fixed resolution of 2074 × 2765. Sensor resolutions and corresponding scaling factors for each smartphone are listed in Table 1. For tiling, images are first rotated to landscape orientation. Afterwards, the images are converted to gray-scale and nine patches of size 512 × 512 are cropped from the center from every image.

Results from previous experiments [10] suggested that the smartphone inherent ISP greatly influences the cross-sensor performance. Hence, all images were captured in raw in addition to the "normal" (ISP) image. Usually, the standard smartphone camera application does not allow to capture raw images. Hence, suitable camera applications were employed: OpenCamera for devices running Android and Halide Mark II for the iPhone devices. Two applications are used for conversion from the raw data, i.e. the colour filter array image, to an RGB image: Darktable (DT) and dcraw together with the additional parameter *-d* to omit demosaicing, hence coined CFA (for colour filter array) in the later experiments. A schematic depiction of the imageing pipeline and corresponding measurement points for CFA, DT and ISP is shown in Fig. 5. In order to remove PRNU and other sensor artefacts, ISP images additionally underwent denoising by application of the BM3D [18] denoising filter. Results in Section 3 generated using denoised images are denoted with DN.

3 Experimental Evaluation

Two distinct feature extraction techniques are employed in this study: (i) Dense SIFT [19] (SIFT descriptors applied on a fixed-spaced grid), which are subsequently encoded using improved Fisher vector encoding [20] in a similar way as originally proposed in [21], and (ii) a local binary pattern (LBP) feature descriptor [22] made rotational invariant by using a number of rotational shifts of the pattern equal to the number of points and taking the minimum. A support vector machine is utilized to classify the encoded features.

The experiments are based on well-established classic texture features (SIFT and LBP) rather than on deep-learning techniques for two reasons: i) better interpretability and more straightforward evaluation of the classic texture features and ii) a lack of sufficient training data for a deep-learning approach due to the rather small sample datasets.

The accuracy, defined as the proportion of correct classifications, is used as an evaluation metric to quantify the classification performance. The reported value in the following experiments is the average accuracy, i.e. the arithmetic mean over all accuracies for a particular experiment. The experiments can be separated into two main parts: The first part (*Experiment A - Experiment E*) evaluates the influence of the image content (surface texture, image entropy) on the device inherent signal utilised for camera classification, where only seven devices are used, indicated by * in Table 1. In the second part (*Experiment F* and *Experiment G*) the assumption is tested if the device-inherent camera signal is device or model specific by utilizing more devices from similar models. Therefore additional samples with all 19 devices

AUTHOR VERSION COMPILED ON January 25, 2025

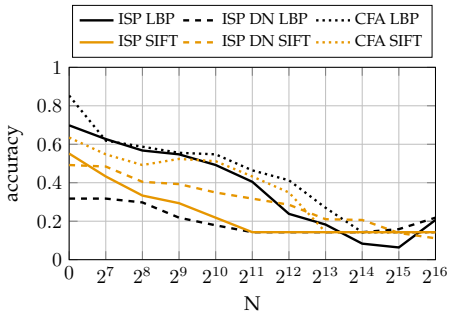

Figure 6: Experiment A: train on ceramic samples and test on Zinterhof sequences

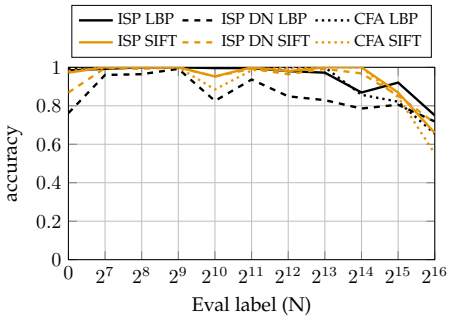

Figure 7: Experiment B: Zinterhof sequences - take given label for evaluation, train model on the rest in an x-fold manner.

listed in Table 1 are used.

Experiment A A classifier is trained to classify devices using ceramic images, which are low in texture. Zinterhof patterns, which have increasingly more image texture as the number of dots N increases, are used as test data. The results in Fig. 6 show the same trend for ISP, ISP DN and CFA for both, SIFT and LBP, indicating an inverse relation of N and ability to classify devices, i.e. the classification accuracy decreases with increasing image texture which is as expected. More image content interferes with the device inherent signal, suppressing it and leading to diminished ability to separate devices. The accuracy raises again towards the $N = 2^{16}$ which can be interpreted as black being just inverted white (i.e. no texture as the black spots cover the whole image).

Experiment B Here each N of the Zinterhof sequence is used in turn as evaluation set and the classifier is trained on the remaining sequences (instead of the ceramic ones). This should test if training on a range of textures can alleviate the trend from *Experiment A*. The results are depicted in Fig. 7. In general they suggest that employing more variety in texture during training, which corresponds to learning different "magnitudes" of device inherent signal, indeed improves the ability to separate devices as the classification accuracy is improved compared to *Experiment A*. However, for a higher N it seems to be more difficult as the accuracy decreases, especially for CFA. This might be due to the bigger difference in the number of black dots between the textures (the scale of N is log). ISP DN exhibits the same general behaviour too, but performs worse than ISP, indicating that with the higher texture content, the denoising is able to filter out the parts of the device inherent signal which are essential for device classification.

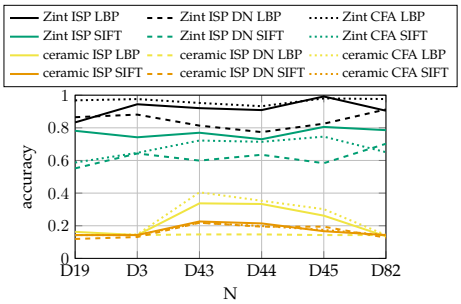

Figure 8: Experiment C: classifier trained on either Zinterhof or ceramic and evaluated on Brodatz textures (per label)

Experiment C This experiment evaluates the generalisation capability. Classifiers are trained using ceramic and Zinterhof pattern samples, respectively. Brodatz texture images, which are meant to represent random texture in the wild, are used for evaluation. This shall answer the question of how much texture is required during training such that the classifier works on all possible input images. The assumption based on *Experiment A* is that the lower the image texture during training, the lower the accuracy of the classifier. The results depicted in Fig. 8 confirm this assumption. Training on the ceramic samples achieves the lowest camera classification accuracy throughout all Brodatz samples, no matter if ISP, ISP DN or CFA samples are utilised, while training on the Zinterhof samples generalises better, resulting in a higher accuracy, especially for LBP (with SIFT still working 3-4 times better than trained on the ceramic samples). In contrast to *Experiment B*, here the CFA samples are slightly better than the ISP ones.

Experiment D The results of *Experiment A* suggested that the more image texture content, the less accurate the camera classification works if trained

AUTHOR VERSION COMPILED ON January 25, 2025

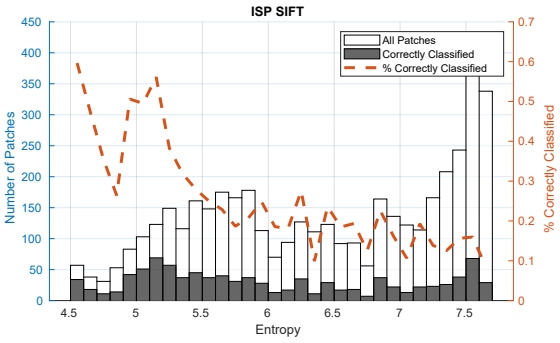

Figure 9: Experiment D: Patch entropy analysis, train on Ceramic, test on the rest, results for ISP SIFT

on the ceramic samples. In this experiment the amount of image texture is quantified using the image entropy to verify the earlier findings. The classifier is once again trained on the ceramic patches and evaluated on the rest (Zinterhof, Brodatz and checkerboard pattern), while the evaluation patches are sorted based on the image entropy. The device classification accuracy results in Fig. 9 only includes patches within a certain entropy range for each bin. Except for the spike at the beginning, the earlier findings can be confirmed (c.f. Fig. 6), i.e. the accuracy decreases with increasing image entropy. Hence, the higher the image entropy, the more the camera inherent signal is suppressed. The results are only shown for ISP SIFT as the other combinations follow the same trend.

Experiment E Based on the findings derived from *Experiment C*, the opposite question is if training on more diverse image texture improves the camera classification accuracy, where the assumption is that it does. Hence, in contrast to *Experiment C*, the classifier is trained on the Zinterhof patches and evaluated on the ceramic samples (exhibiting low image texture). The results listed in Table 2 are controversial. For LBP, CFA confirms the assumption, while ISP DN and ISP should have a much higher accuracy, thus, invalidating the assumption that more diverse texture during training (or training on Zinterhof samples) generalizes well. SIFT follows a similar trend. However, results once more indicate that the camera inherent signal is composed of an ISP as well as a image sensor component, with the ISP component being highly influenced by the image content while CFA remains relatively stable (otherwise the performance of ISP and ISP DN should be similar to CFA).

Table 2: Experiment E: Accuracy of classification of model trained on Zinterhof images applied to ceramic images.

Source	Accuracy	Source	Accuracy
CFA LBP	0.963	CFA SIFT	0.800
ISP LBP	0.477	ISP SIFT	0.625
ISP DN LBP	0.271	ISP DN SIFT	0.321

Table 3: Experiment F: device classification performance (accuracy) for the different features on the four defined subsets.

Source	**all**	**all uid**	**NDF**	**NDF uid**
ISP LBP	0.959	0.793	0.992	0.641
ISP DN LBP	0.789	0.606	0.962	0.537
CFA LBP	0.952	0.832	0.999	0.787
DT LBP	0.877	0.700	0.990	0.661
ISP SIFT	0.878	0.881	0.994	0.826
ISP DN SIFT	0.823	0.812	0.984	0.809
CFA SIFT	0.873	0.939	0.999	0.955
DT SIFT	0.805	0.775	0.974	0.765

Experiment F This experiment tests if the device-inherent signal tends to be model or instance/device specific by training and evaluating a camera identification classifier on four different subsets (of the ceramic samples) in a leave-one-out cross-fold validation: **all** simply uses all the 19 available devices, **NDF** (non-duplicates filtered) is a subset where only devices are included which have more then one instance, e.g., two different iPhones 13 from two different users. Both subsets are further subdivided into **uid**, which keeps the user id for the device (or device id), i.e., the iPhone 13 from user A and the iPhone 13 of user B are separate classes (iPhone13-A and iPhone13-B are the target classes), and the non-uid where both, iPhone13-A and iPhone13-B belong to the same single class (iPhone13 as target class). If the device-inherent signal was instance specific, there should be no considerable difference between the four subsets, while if the contrary is true, **NDF** is expected to achieve the highest device classification performance.

The results given in Table 3 confirm that the device classification works well for the **NDF** subset across all image sources. It still works well for the **all** but inferior to the **NDF** one (note that there is a lower set of classes with higher number of samples per class in **NDF**). For the subsets where the user id is kept (**all uid** and **NDF uid**) the device classi-

AUTHOR VERSION COMPILED ON January 25, 2025

Table 4: Experiment F: Dental ceramic material classification performance (accuracy) for the different feature sources with the four sets given.

Source	all	all uid	NDF	NDF uid
ISP LBP	0.898	0.933	0.833	0.977
ISP DN LBP	0.907	0.933	0.875	0.968
CFA LBP	0.969	0.887	0.500	0.945
DT LBP	0.963	0.974	0.856	0.983
ISP SIFT	0.982	0.990	0.996	0.999
ISP DN SIFT	0.980	0.981	0.993	0.999
CFA SIFT	0.980	0.988	0.957	0.999
DT SIFT	0.995	0.998	0.994	0.999

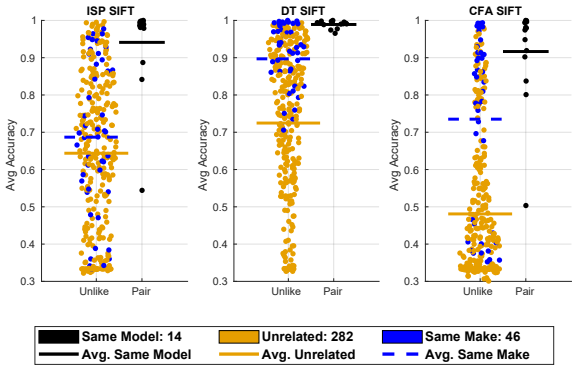

Figure 10: Exp. G: Pairwise cross-sensor ceramic classification.

fication performance drops. The **all uid** subset performance drops less than the **NDF uid** subset due to the number of classes with only a single device available in the **all** subset, which of course perform equally with and without **uid**. This trend is the same throughout all image sources (ISP, ISP DN, CFA and DT). The performance drop is a clear indicator that the source/signal of the features used for classification is model specific, i.e. the device inherent signal is rather model than instance specific. However, even **NDF uid** is far away from random guessing, i.e. while the majority of the camera inherent signal is model specific there seems to be a part which is specific to the individual device.

To further validate the assumption that the main part of the device-inherent signal is model specific, the same experiment is conducted for material classification. Here the classes are reversed as the different dental ceramic materials are the target classes: The cross-fold validation puts a device either in the training or in the evaluation set. If the device-inherent signal was instance specific, there should be distinct "data" (images of the same model but different user id should be different) in the training and test sets for the **uid** subsets. If this assumption does not hold, there is the same "data" in training and test sets, leading to an increase in accuracy (more pronounced in the pair of **NDF** sets). The results provided in Table 4 reflect the assumed accuracy increase in the **uid** sets due to individual devices of the same model being similar for LBP but not SIFT. Again the increase in the pair **NDF** sets is more pronounced than in the **all** set due to the smaller number of classes.

The device as well as the material classification results confirm (for all image sources), that the main part of the device-inherent signal is *model* specific. However, there is also a smaller part of the signal which is specific to each *individual* device, enabling device classification even between instances of the same model (c.f. **NDF uid** in Table 3).

Experiment G As a further validation of the ceramic material classification results in *Experiment F*, here a one to one cross-sensor train-test evaluation is performed: All ceramic samples captured with exactly one of the 19 devices are used as training set while all ceramic samples captured with a different one are used as test set. If the assumption of *Experiment F* still holds, the train-test pairs of the same model should achieve a higher performance than uncorrelated pairs. The results are given in Fig. 10. Each dot represents one train-test device pair with different colours, black for pairs of the same model, blue for pairs of the same make and orange for uncorrelated pairs (note that of course the uncorrelated ones have the highest number of pairs). The assumption that the main part of the device-inherent signal is model specific is confirmed, as the pairs from the same model achieve the highest performance across all the three tested image sources, followed by the pairs of the same make while the uncorrelated pairs perform worst.

4 Conclusion

This work analysed the influence of the image texture on the device-inherent signal, utilised for camera classification, as well as the question if the nature of this signal is instance/device or model specific. The experimental results confirmed that the device-inherent signal is strongly interdependent with the texture/image content and weak in nature. If the training and testing is performed on different textures, there is a significant drop in the device clas-

AUTHOR VERSION COMPILED ON January 25, 2025

sification accuracy. The higher the image entropy the more the signal is suppressed and the lower the classification accuracy. The device as well as material classification results on the pairs and triples of devices from the same model revealed that the main part of the signal is rather model than instance/device specific (on sets of devices from the same model device identification performs worse while material classification performs better than on arbitrary sets).

References

[1] A. Das, C. Galdi, H. Han, R. Ramachandra, J.-L. Dugelay, and A. Dantcheva, "Recent advances in biometric technology for mobile devices", in *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, 2018.

[2] S. Jain and M. Gruteser, "Recognizing textures with mobile cameras for pedestrian safety applications", *IEEE Transactions on Mobile Computing*, vol. 18, no. 8, 2019.

[3] X. J. Tang, Y. H. Tay, N. A. Siam, and S. C. Lim, "Mywood-id: Automated macroscopic wood identification system using smartphone and macro-lens", in *Proceedings of the 2018 International Conference on Computational Intelligence and Intelligent Systems*, ser. CIIS 2018, Association for Computing Machinery, 2018.

[4] S. Voloshynovskiy, M. Diephuis, F. Beekhof, O. Koval, and B. Keel, "Towards reproducible results in authentication based on physical non-cloneable functions: The forensic authentication microstructure optical set (famos)", in *2012 IEEE International Workshop on Information Forensics and Security (WIFS)*, 2012.

[5] M. Diephuis, S. Voloshynovskiy, and T. Holotyak, "Sketchprint: Physical object microstructure identification using mobile phones", in *2015 23rd European Signal Processing Conf. (EUSIPCO)*, 2015.

[6] O. Taran, J. Tutt, T. Holotyak, R. Chaban, S. Bonev, and S. Voloshynovskiy, "Mobile authentication of copy detection patterns: How critical is to know fakes?", in *2021 IEEE International Workshop on Information Forensics and Security (WIFS)*, 2021.

[7] Y. Sun, X. Liao, and J. Liu, "An efficient paper anti-counterfeiting method based on microstructure orientation estimation", in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021.

[8] R. Schraml, L. Debiasi, C. Kauba, and A. Uhl, "On the feasibility of classification-based product package authentication", in *IEEE Workshop on Information Forensics and Security (WIFS'17)*, 2017.

[9] R. Schraml, L. Debiasi, and A. Uhl, "Real or fake: Mobile device drug packaging authentication", in *Proceedings of the 6th ACM Workshop on Information Hiding and Multimedia Security*, ser. IH&MMSec '18, Association for Computing Machinery, 2018.

[10] J. Schuiki, C. Kauba, H. Hofbauer, and A. Uhl, "Cross-sensor micro-texture material classification and smartphone acquisition do not go well together", in *Proc. of the 11th International Workshop on Biometrics and Forensics (IWBF'23)*, 2023.

[11] J. Lukas, J. J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise.", *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, 2008.

[12] J. Schuiki, C. Kauba, H. Hofbauer, and A. Uhl, "Limiting factors in smartphone-based cross-sensor microstructure material classification", in *22th Int. Workshop on Digital-forensics and Watermarking (IWDW'23) (to appear)*, ser. Springer LNCS, vol. 14511, 2023.

[13] D. Cozzolino and L. Verdoliva, "Noiseprint: A cnn-based camera model fingerprint", *IEEE Transactions on Information Forensics and Security*, vol. 15, 2020.

[14] C. Chen, Z. Xiong, X. Liu, and F. Wu, "Camera trace erasing", in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[15] Manisha, C.-T. Li, X. Lin, and K. A. Kotegar, "Beyond prnu: Learning robust device-specific fingerprint for source camera identification", *Sensors*, vol. 22, no. 20, 2022.

[16] H. Hofbauer, A. Uhl, and P. Zinterhof, "Zinterhof sequences in grid-based numerical integration", in *Monte Carlo and Quasi-Monte Carlo Methods 2006*, Springer-Verlag, 2008.

[17] P. Brodatz, *Textures; a photographic album for artists and designers*, English. Dover Publications New York, 1966.

AUTHOR VERSION COMPILED ON January 25, 2025

[18] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering", *IEEE Transactions on Image Processing*, vol. 16, no. 8, 2007.

[19] D. G. Lowe, "Distinctive image features from scale-invariant keypoints", *Int. J. Comput. Vision*, vol. 60, no. 2, 2004.

[20] F. Perronnin, J. Sánchez, and T. Mensink, "Improving the Fisher kernel for large-scale image classification", in *Computer Vision – ECCV 2010*, Springer Berlin Heidelberg, 2010.

[21] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, and A. Vedaldi, "Describing textures in the wild", in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014.

[22] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution Gray-Scale and rotation invariant texture classification with local binary patterns", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, 2002.