Cross-Sensor Micro-Texture Material Classification and Smartphone Acquisition do not go well together

Johannes Schuiki, Christof Kauba, Heinz Hofbauer and Andreas Uhl

University of Salzburg

Department of Artificial Intelligence and Human Interfaces

Jakob-Haringer-Str. 2, 5020 Salzburg, AUSTRIA

{jschuiki, ckauba, hofbauer, uhl}@cs.sbg.ac.at

Abstract—Intrinsic, non-invasive product authentication is still an important topic as it does not generate additional costs during the production process. This topic is of specific interest for medical products as non-genuine products can directly effect the patients' health. This work investigates micro-texture classification as a mean of proving the authenticity of zircon oxide blocks (for dental implants). Samples of three different manufacturers were acquired using four smartphone devices with a clip-on macro lens. In addition, an existing drug packaging material database was utilized. While the intra-sensor micro-texture classification worked well, the cross-sensor classification results were less promising. In an attempt to track down the limiting factors, intrinsic sensor features usually used in device identification were investigated as well.

Index Terms—micro texture, texture classification, smartphone imaging, cross sensor, material classification, product authentication

I. Introduction

The act of verifying a product's origin is known as product authentication. There are various scenarios where someone desires to perform product authentication. Two of the more prominent ones are classifying a product to be either genuine or counterfeit, and assuring that a product stems from a certain manufacturer. Such a proof of origin can be as simple as a signed physical document, which can be easily forged, up to individual, hard-to-forge features that are embedded in each single product. The wide-spread use of mobile devices with imaging sensors over the last decades led to multiple new approaches for product authentication:

Cai et al. [1] proposed a deep learning based texture authentication framework that extracts texture features from leather products which are encoded into a 2D barcode. These labels should then be attached to the products in an non-detachable manner and a user can verify the authenticity of a product by capturing the barcode and compare it to texture features generated on-cite using a pre-trained deep network.

Another framework that combines micro-texture features in combination with a QR code for anti-counterfeiting was proposed by Yan et al. [2]. Visual features and the QR code are registered on the assembly line and stored in a cloud. For verification, a mobile phone then compares acquired features with preregistered records online.

However, using attachable labels, barcodes or any form of external embedded features as anti-counterfeiting methods can be cumbersome to implement into a running production chain. Furthermore, some approaches additionally demand for a back-end database infrastructure, which is often not feasible form a financial perspective. Thus, the preferred approach from a cost perspective is to achieve product authentication based solely on constant but discriminative intrinsic features of the product or its packaging material. A further requirement is that the product authenticity can be established in a non-invasive way (without altering the product).

In [3], [4], micro-texture images of drug package material was used for product authentication in an open set scenario. Further, [4] showed that the capability of classifying manufacturers using micro-texture structures is greatly reduced when testing a cross-sensor scenario, i.e. train a classifier using data from one sensor and test on data captured with another sensor. It was suggested that in order for micro-texture cross-sensor product authentication to work, probably scaling between the various imaging sensors needs to be considered.

However, there are other types of artefacts that might cause a decrease in cross-sensor material classification accuracy. Such artefacts include optical lens distortions, device-inherent processing steps from the image signal processor (ISP, i.e. the processing unit that maps raw sensor images to rgb images), or pixel response non-uniformity (PRNU, often simply coined sensor noise). Such artefacts were successfully applied for source camera identification in natural scene images [5]–[7]. More details on such artefacts are given in Section II.

Establishing a proof of authenticity is of special interest for medical products as it has a direct impact on the patients' health. The current gold standard is to extract a material sample with a subsequent chemical analysis, which is intrusive, time and cost consuming.

The main goal of this study is to answer the question if the authenticity of zircon oxide blocks, commonly used for dental ceramics, can be established based on the micro-texture of their surface in a non-intrusive way using off-the-shelf smartphone devices. The contribution of this study can be summarized in the following points:

- Establishing a database of dental zircon oxide block samples from three different manufacturers captured by multiple different smartphone devices.

- Evaluation of material texture based classification performance in an intra-sensor setting on the established as well as an existing database including micro-texture structures from drug packaging material.
- Evaluation of the classification performance in a cross-sensor setting for both databases.
- Assessment of the impact of varying scale factors in a cross-sensor setting as proposed by [4].
- Evaluation of sensor identification performance
- Investigating the effect of PRNU for sensor identification

The remainder of this research is structured as follows: Section II describes the feature extraction and classification pipeline, followed by a detailed description of the involved databases used within this work and also introduces the experiments in Section III. Finally, Section IV summarizes the findings of this article.

II. EXPERIMENTAL SETUP

This section first explains the feature extraction and classification pipeline, followed by a detailed description of the two databases included in this study and finally gives an overview of the experiments in Section III.

A. Feature Extraction & Classification

In total, five distinct techniques for feature extraction are employed in this study that are briefly introduced in the following. All of the algorithms were successfully used in classical texture classification, image tampering detection and paper identification [8]: Dense SIFT [9] (SIFT descriptors applied in a grid), Dense Micro-block Difference (DMD) [10], Local Binary Pattern (LBP) [11], Weber Pattern (WP) [12] and Local Phase Quantization (LPQ) [13].

For all experiments in this study, feature vectors are encoded using improved Fisher vector encoding [14] in a similar way as described in [15]. Therefore, extracted features from the previous step undergo soft-quantization using a Gaussian mixture model and dimensionality reduction using principle component analysis. As a classifier for the encoded features, a support vector machine with linear kernel is utilized. Experiments within this study are accomplished by randomly splitting the data into training and testing subsets such that every subset contains half of the data. Since no hyper-parameter optimization is done, no validation set is needed. Because the databases explained in the upcoming section are unbalanced in terms of micro-texture patches per class, the class with the minimum patches in a particular setting dictates the numbers of patches involved for an experiment. Doing so, classification bias caused by the amount of available data is avoided.

To quantify the outcome of the classification experiments throughout this study, the accuracy is used as an evaluation metric. Accuracy is defined as the proportion of correct classifications. The reported value in the experiments in Section III is the average accuracy, i.e. the arithmetic mean over all the accuracies for a particular experiment.

B. Databases

Hereafter, the two texture databases used in the later parts of this research are introduced. After dividing the images from both databases into multiple patches, no further pre-processing, such as contrast enhancement, is employed in order to preserve possible artefacts generated by the sensor or the device-inherent image signal processing pipeline.

1) Drug Package Material (Drug Data): The first database included in this study is a subset of a larger drug package authentication database introduced by Schraml et al. [4]. The drug package authentication database comprises of images captured from packaging material of pharmaceutical products. The database includes different types of packaging material: cardboard (M1), blister top-side (M2) and blister bottom-side (M3). For some of the drug packages, images were captured using three different devices: A Canon 70D (S1), a Samsung S5 Mini (S2) and an iPhone 5 (S3). Drug packages captured with less than three devices are not included in the experiments in this research. Hence, packaging material data of six different drugs are available. Due to the fact that some of the blisters were already opened at the time of image acquisition and also side flaps from the cardboard provide a smaller texture region, the available data consists of hand-cropped areas of different size. Therefore, all the bigger areas are cut into patches of size 200 by 200 pixels. Table I lists the available number of patches per drug.

Table I: Number of available patches per sensor (S1-S3), material (M1-M3) and drugs (D1-D6). The rotated text lists the manufacturer and product names of the corresponding drug.

		Lambacher Thrombo ASS	Gynial Bilinda	Gedeon Richter Floumizin	Pelpharma Peliette	Kwizda Pharma Delia	Ratiopharm Mexalen
		D1	D2	D3	D4	D5	D6
M 1	S1	4852	310	144	11982	721	—
	S2	800	1200	800	1600	1600	1600
	S3	800	1200	800	1600	1600	1600
M 2	S1	202	59	82	472	54	—
	S2	780	1104	727	1582	1580	1600
	S3	721	1073	723	1432	1407	1524
M 3	S1	546	33	19	335	61	—
	S2	710	922	588	1389	1466	501
	S3	656	985	667	1235	1546	435

2) Zircon Oxide Blocks (Ceramic Data): The second database used in the experiments was captured within the scope of this study. Here, images were acquired from the top side of zircon oxide blocks from three different manufacturers (C1-C3). The acquisition setup is depicted in Figure 1. Five images were captured per block, namely four on the corners and one in the center. This database includes images captured with four different smartphones: A Huawei P20 Lite (S4), a Samsung Galaxy S8 (S5), a Samsung Galaxy Tab S6 (S6) and a Samsung Galaxy XCover 4 (S7). In order to make the material texture visible, a macro lens *Agritix WIDK-*

24X01 Xylorix Wood Identification Tool [1] was clipped onto the imaging sensor of the devices. It was found that although the resolution of the devices varies, the captured field of view is roughly similar. Therefore, all images underwent down-scaling to a fixed resolution. Afterwards, nine patches of size 512 by 512 pixels were cropped from the center as can be seen in Figure 2. Sample patches can be seen in Figure 3.
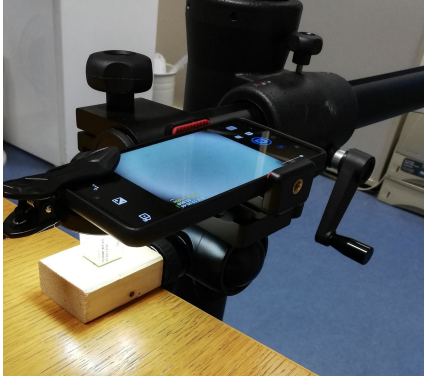

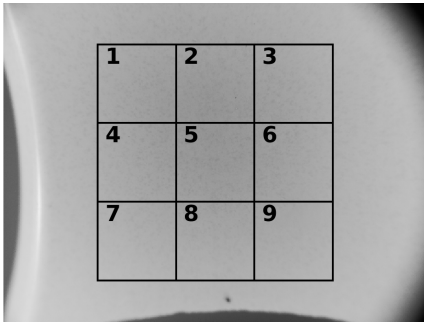Figure 1: Acquisition setup for the zircon oxide blocks, smartphone with clip-on macro lens.

Figure 2: Image of a corner on the top-side of a zircon oxide block with patch annotations.

In Figure 2, one can see a pincushion distortion (visible through curved lines on the left and bottom) introduced by the macro lens. The black areas in the right corners are caused by the macro lens not perfectly aligning with the imaging sensor on the smartphone. By choosing the downsizing factor and patch size in a way that a margin from the patches to all image borders remains, these two types of artifacts can be evaded in the classification experiments. The used macro lens has a circular ring of light emitting diodes around the center which was switched on during all image acquisitions because preliminary tests suggested that this setting produces images with more texture visible. Table II lists the available number of patches per ceramic and imaging sensor.

C. Experiment Design

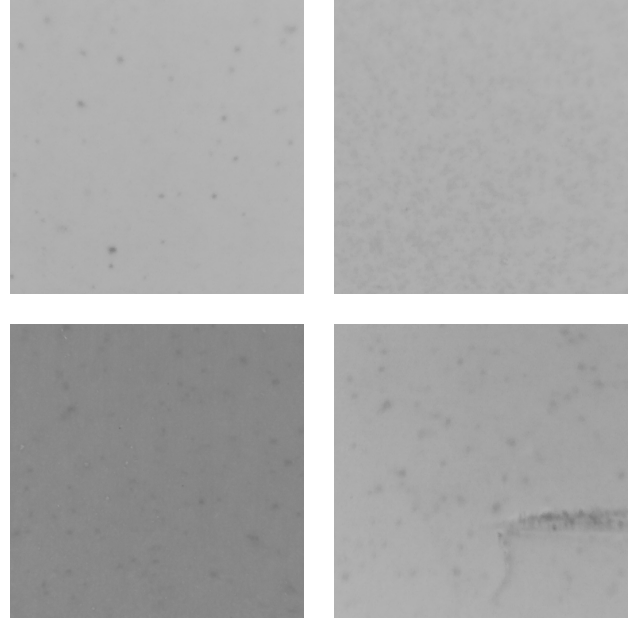This section aims to provide a brief description regarding the experiments in Section III.

Figure 3: Zircon oxide block sample patches.

Table II: Number of available patches per sensor (S4-S7) and ceramic (C1-C3). The rotated text lists the manufacturer names of the corresponding ceramic.

	Ivoclar Vivadent	Dentslpy Sirona	3M
	C1	C2	C3
S4	720	270	450
S5	720	270	450
S6	720	270	450
S7	720	270	450

As a first step, experiments are carried out to verify that material classification can be achieved in an intra-sensor setup.

Afterwards, the cross-sensor material classification scenario is evaluated. For the drug packaging data, as suggested in [4], it is investigated whether employing different scales would reveal any improvements. A different scale here means that, prior to cropping images into patches of size 200 by 200 pixels, resizing by a factor is applied. Through combining different train- and test-scales, as visualized in Table III, various effective scale factors ($\frac{train}{test}$), ranging from 0.5 to 3.0, are covered. Note that this is only done for the drug package data because the ceramic patching strategy already includes a step to unify the resolution.

For the ceramic database, the impact of applying a "leave one sensor out" strategy (i.e. train on patches from three sensors, test on patches from the remaining sensor) is investigated. Doing so, more training data variability should be available and therefore a better performance is expected.

Next, it is tested whether sensor identification can be achieved. To do so, "sensors" and "classes" (manufacturers) are switched for the experiments. The idea is to investigate

Table III: Train-test scale combinations and resulting scale-factor.

		Test				
		0.50	0.75	1.00	1.25	1.50
Train	0.75	1.50	1.00	0.75	0.60	0.50
	1.00	2.00	1.33	1.00	0.80	0.67
	1.25	2.50	1.67	1.25	1.00	0.83
	1.50	3.00	2.00	1.50	1.20	1.00

whether artefacts caused by the imaging device are predominantly present in the images such that they overrule the micro-texture structures. Sensor or device inherent artefacts can be caused by various sources. Three prominent artefact sources are briefly explained in the following:

1) *Image Signal Processor (ISP):* The so called ISP is the processing unit responsible for converting the raw color filter array (Bayer pattern) to an rgb image. This processing pipeline includes denoising, demosaicing, white balance, color transformation and compression (mostly into JPEG format).
2) *Lens Distortion:* Since for both databases included in this study, the data was captured using optical magnification (an additional lens), assumably the lens also introduces distortion artefacts that change from device to device. When looking at Figure 2, one can see an example of a pincushion distortion on the left an bottom border. The marco lens has a small positional variability everytime it is attached for an image acquisition session. Therefore, one can not rule out the possibility that images corresponding to the same smartphone acquisition session have smaller distortions by which they could be identified.
3) *Photo Response Non-Uniformity (PRNU):* Due to imperfections in the manufacturing process, imaging sensors have a sort-of fingerprint which is measurable as a noise pattern that is applied to every image acquisition. PRNU does also work with JPEG quality 90 [7], which is the standard case for smartphone acquisitions. Therefore, it can also be considered as a sensor inherent artefact within the context of this work.

The last experiment in this study focuses on the PRNU sensor noise. The patching strategy that can be seen in Figure 2 remains constant throughout the entire ceramic data set. Since both, sensor noise and also patches with a certain number, are "bound" to a fixed location in the image, meaning that a certain patch will always contain the same sensor area, an experiment can be designed such that location dependent parts of the PRNU impact can be ruled out as a contributing factor for sensor identification performance: For training, only patches numbered 1-8 are used. For testing, the remaining patches (those at the position number 9) are then used for testing. Doing so, it is assured that the sensor noise in patches stemming from area number 9 is not used during training.

III. RESULTS AND DISCUSSION

The following section includes the results for the experiments motivated in Section II-C. Note that some values in Tables IV and VII are missing for feature extraction schemes LBP and WP. These schemes produce one feature vector per image and thus too few feature vectors are available for the Fisher encoding step in cases where only a small amout of data is available.

A. Intra-Sensor Material Classification

For the first experiment, material classification performance is assessed for both micro-texture databases in an intra-sensor setup. The average classification accuracy is displayed in Table IV. The results suggest that, in general, materials stemming from different manufacturers provide enough distinctiveness to classify them correctly when using SIFT or DMD feature extraction. With LBP, WP and LPQ having their highest achieved average accuracy value at 91%, 85% and 55%, respectively, the classification results for these feature extractions schemes can be labelled unsatisfactory, even in an intra-sensor setup.

Table IV: Average material classification accuracy in intra-sensor setup.

		SIFT	DMD	LBP	WP	LPQ
M 1	S1	0.86	0.87	0.75	0.61	0.35
	S2	0.98	0.97	0.87	0.71	0.30
	S3	0.85	0.85	0.48	0.49	0.28
M 2	S1	1.00	0.99	—	—	0.25
	S2	1.00	0.99	0.91	0.75	0.49
	S3	0.99	0.99	0.88	0.72	0.40
M 3	S1	1.00	0.97	—	—	0.15
	S2	0.99	0.98	0.80	0.58	0.35
	S3	0.98	0.97	0.59	0.53	0.21
Ceramic	S4	0.99	0.98	0.68	0.66	0.54
	S5	0.98	0.96	0.62	0.68	0.41
	S6	0.98	0.95	0.74	0.62	0.55
	S7	0.99	0.99	0.86	0.85	0.55

B. Cross-Sensor Material Classification

The cross-sensor experiments are divided into three parts. First, exemplary results for the ceramic data using SIFT features are reported in Table V. An accuracy value beyond 90% was only reached once. Interestingly, Train SensorX - Test SensorY does not necessarily yield the same results as Train SensorY - Test SensorX. It is also worth mentioning that an average accuracy of around 33% would indicate that the classifier is guessing, since only three ceramic manufacturers are included in the database.

Table V: Cross-sensor material classification on ceramic data using SIFT.

		Test			
		S4	S5	S6	S7
Train	S4	—	0.71	0.84	0.87
	S5	0.33	—	0.34	0.40
	S6	0.83	0.60	—	0.64
	S7	0.34	0.91	0.40	—

It is also interesting to see that, although 3 out of 4 devices used for capturing the ceramic data are Samsung devices, cross-sensor material classification between those devices does not necessarily work any better than between a Samsung and the Huawei device. One would assume that either the ISP processing steps or the imaging sensor would be rather similar when developed by the same manufacturer.

Second, Figure 4 reports results for cross-sensor experiments on the drug packaging database. Plots show exemplary results for the material blister top-side (M2) together with SIFT features. The reported numbers correspond to the scaling scheme described in Table III. Regarding experiments including sensor S1, it can be concluded that scaling does not improve the results in a way that they could be labelled satisfactory. Experiments from sensor S2 to S3 or vice versa yield accuracy numbers up to 88% for the cases where the samples had similar scale. However, as with the ceramic cross-sensor experiments, results are noticeably lower compared to the intra-sensor case.

The third cross-sensor classification experiment evaluates, whether a leave one sensor out strategy would increase the classification performance. When comparing the numbers in Table VI for SIFT to the single sensor to sensor performances in Table V, one can observe that the numbers appear to be more stable but nevertheless are still far from the intra-sensor case.

Table VI: Leave one sensor out cross validation for material classification.

	SIFT	DMD	LBP	WP	LPQ
Leave Out S4	0.77	0.85	0.36	0.43	0.20
Leave Out S5	0.66	0.49	0.34	0.61	0.32
Leave Out S6	0.84	0.50	0.40	0.47	0.33
Leave Out S7	0.46	0.57	0.33	0.70	0.37

C. Sensor Identification

Next, it is evaluated whether sensor classification can be performed. To do so, the sensors and classes are switched. Results are reported in Table VII. An interesting observation can be made when looking at the numbers for LBP, WP and LPQ and comparing them to the intra-sensor material classification approach. It seems that, although these feature extraction schemes were not able to correctly classify the materials, they are however able to identify sensors. This indicates that smartphone-inherent artefacts tend to be more present than the micro-texture material information on the product.

D. Impact of sensor noise (PRNU) (Ceramic Data)

As described in Section II-C, in order to assess the impact of PRNU, patches on position 1-8 are used for training, while patches in position 9 are used for performance evaluation. To actually see a possible impact of PRNU, one needs to compare

Table VII: Average sensor classification accuracy per feature extraction scheme and material class.

		SIFT	DMD	LBP	WP	LPQ
M 1	D1	1.00	1.00	0.99	1.00	0.66
	D2	1.00	1.00	0.98	0.97	0.98
	D3	1.00	1.00	—	—	0.94
	D4	1.00	1.00	1.00	0.99	0.45
	D5	1.00	1.00	1.00	0.99	0.49
	D6	1.00	1.00	1.00	0.99	0.98
M 2	D1	0.99	0.99	0.48	0.87	0.78
	D2	0.97	0.97	—	—	0.77
	D3	0.99	0.99	—	—	0.78
	D4	1.00	0.99	0.94	0.91	0.74
	D5	1.00	1.00	—	—	0.82
	D6	1.00	1.00	1.00	0.99	1.00
M 3	D1	1.00	1.00	0.92	0.88	0.75
	D2	1.00	1.00	—	—	0.68
	D3	0.92	0.92	—	—	0.40
	D4	0.99	0.99	0.90	0.83	0.83
	D5	1.00	1.00	—	—	0.86
	D6	1.00	1.00	0.26	0.86	0.82
Ceramic	C1	1.00	0.99	0.96	0.86	0.72
	C2	1.00	1.00	1.00	0.96	0.97
	C3	1.00	1.00	0.99	0.94	0.81

the results reported in Table VIII with the previous results from Table VII. If the PRNU sensor noise would be a factor that highly impacts the sensor identification experiments, then one would expect a decrease in accuracy. Since this is not the case, the location dependent parts of the PRNU can be ruled out as the main reason why sensor identification seems to work significantly better than material classification.

Table VIII: Average sensor classification accuracy per feature extraction scheme and material class, when using patch 1-8 for training and patch 9 for testing.

		SIFT	DMD	LBP	WP	LPQ
Ceramic	C1	1.00	1.00	0.99	0.90	0.73
	C2	1.00	1.00	1.00	0.95	0.99
	C3	1.00	1.00	1.00	0.95	0.89

IV. CONCLUSION

This study conducted various classification experiments on two micro-texture databases. Intra-sensor material classification could successfully be achieved using classical (i.e. non deep learning) texture classification methods as long as there was only one capturing device involved. However, cross-sensor material classification, for the most part, yields unsatisfying results. Further investigations considered a suggestion from a related work, where it was recommended to try the cross-sensor experiment using varying scales. The experiments showed, that the different scales are not the main reason for the poor performance of the cross-sensor classification.

Additional experiments to further track down the limiting factors for the cross-sensor classification performance involving commonly used device-intrinsic sensor features, in particular the PRNU, showed that the location dependent parts of the PRNU can be ruled out as a reason. In fact it turned out that sensor identification on the two databases works even better than the desired material classification.
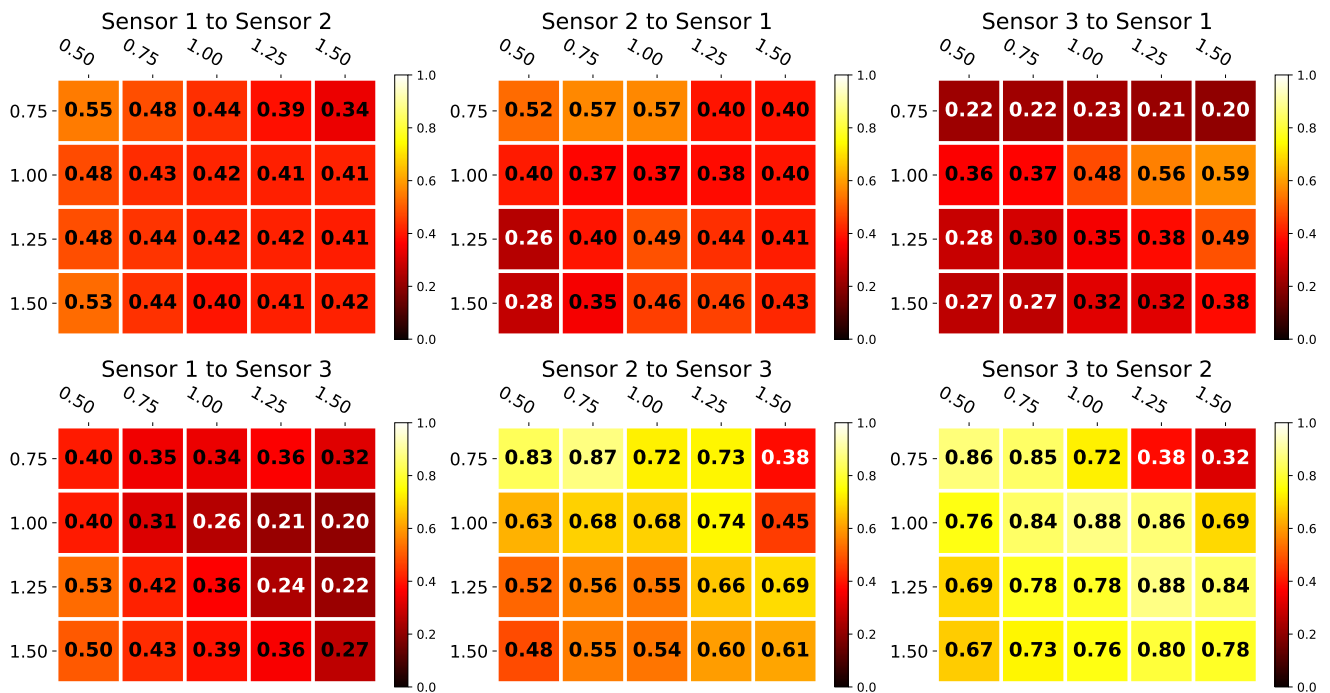
Figure 4: Cross-sensor drug material classification - different scale factors on the drug data (M2, blister top-side) using SIFT.

As mentioned earlier, there are also other artefacts introduced, e.g. optical artefacts due to the lenses or ISP pipelines. In our future work these additional intrinsic artefacts present in the captured sample images will be considered in order to answer the question why sensor identification reaches higher accuracies than cross-sensor material classification. Experiments based on denoising filters to remove the PRNU influence as well as captured raw images to reduce the influence of the image processing toolchain will be conducted. Once the causes have been identified, suitable countermeasures can be applied in order to improve the cross-sensor material classification performance.

REFERENCES

[1] Sudao Cai, Lin Zhao, and Changsheng Chen, "Open-set product authentication based on deep texture verification," in *Image and Graphics*, Yuxin Peng, Shi-Min Hu, Moncef Gabbouj, Kun Zhou, Michael Elad, and Kun Xu, Eds., Cham, 2021, pp. 114–125, Springer International Publishing.

[2] Yulong Yan, Zhuo Zou, Hui Xie, Yu Gao, and Lirong Zheng, "An IoT-based anti-counterfeiting system using visual features on QR code," *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6789–6799, 2021.

[3] Rudolf Schraml, Luca Debiasi, Christof Kauba, and Andreas Uhl, "On the feasibility of classification-based product package authentication," in *IEEE Workshop on Information Forensics and Security (WIFS'17)*, Rennes, FR, December 2017, p. 6.

[4] Rudolf Schraml, Luca Debiasi, and Andreas Uhl, "Real or fake: Mobile device drug packaging authentication," in *Proceedings of the 6th ACM Workshop on Information Hiding and Multimedia Security*, New York, NY, USA, 2018, IH&MMSec '18, p. 121–126, Association for Computing Machinery.

[5] S. Bayram, H. Sencar, N. Memon, and I. Avcibas, "Source camera identification based on CFA interpolation," in *IEEE International Conference on Image Processing 2005*, 2005, vol. 3, pp. III–69.

[6] Kai San Choi, Edmund Y. Lam, and Kenneth K. Y. Wong, "Automatic source camera identification using the intrinsic lens radial distortion," *Opt. Express*, vol. 14, no. 24, pp. 11551–11565, Nov 2006.

[7] Jan Lukas, Jessica J. Fridrich, and Miroslav Goljan, "Digital camera identification from sensor pattern noise.," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 205–214, 2006.

[8] Christof Kauba, Luca Debiasi, Rudolf Schraml, and Andreas Uhl, "Towards drug counterfeit detection using package paperboard classification," in *Advances in Multimedia Information Processing – Proceedings of the 17th Pacific-Rim Conference on Multimedia (PCM'16)*, Xi'an, CHINA, 2016, vol. 9917 of *Springer LNCS*, pp. 136–146.

[9] David G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[10] Rakesh Mehta and Karen Egiazarian, *Texture Classification Using Dense Micro-block Difference (DMD)*, pp. 643–658, Lecture Notes in Computer Science. Springer-Verlag, Berlin, 2015.

[11] T. Ojala, M. Pietikainen, and D. Harwood, "Performance evaluation of texture measures with classification based on kullback discrimination of distributions," in *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, Oct. 1994, vol. 1, pp. 582 –585 vol.1.

[12] Ghulam Muhammad, "Multi-scale local texture descriptor for image forgery detection," in *Industrial Technology (ICIT), 2013 IEEE International Conference on*, 2013, pp. 1146–1151.

[13] V. Ojansivu, E. Rahtu, and J. Heikkila, "Rotation invariant local phase quantization for blur insensitive texture analysis," in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, Dec 2008, pp. 1–4.

[14] Florent Perronnin, Jorge Sánchez, and Thomas Mensink, "Improving the Fisher kernel for large-scale image classification," in *Computer Vision – ECCV 2010*, Kostas Daniilidis, Petros Maragos, and Nikos Paragios, Eds., Berlin, Heidelberg, 2010, pp. 143–156, Springer Berlin Heidelberg.

[15] Mircea Cimpoi, Subhransu Maji, Iasonas Kokkinos, Sammy Mohamed, and Andrea Vedaldi, "Describing textures in the wild," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3606–3613.