Limiting Factors in Smartphone-Based Cross-Sensor Microstructure Material Classification

Johannes Schuiki, Christof Kauba, Heinz Hofbauer and Andreas Uhl

Department of Artificial Intelligence and Human Interfaces, University of Salzburg, Jakob-Haringer-Straße 2, 5020 Salzburg, Austria

November 27, 2023

PARIS LODRON UNIVERSITÄT SALZBURG

WAVELAB
Multimedia Signal Processing and Security Lab

Introduction & Motivation

- Counterfeit products are a problem
  - Causing economic damage to the original manufacturers
  - Especially in medical field they can directly influence patients' health

  $\rightarrow$ Product authentication:
  - Extrinsic: based on QR codes, external markers
  - Intrinsic: based on the product's properties, no external markers etc. needed

- Manufacturers do not favour extrinsic authenticity methods due to extra production costs

- Intrinsic product authentication is preferred

Previous Work

Goal: Classify manufacturer of dental ceramics using images captured with smartphone cameras (sensors)[1].

Findings:

- Intra-sensor ✓ works well

  Train/enroll on sensor A → Test/evaluate on sensor A

- Cross-sensor ✗ results unsatisfying

  Train/enroll on sensor A → Test/evaluate on sensor B

- PRNU (sensor noise[2]) was eliminated as a reason

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

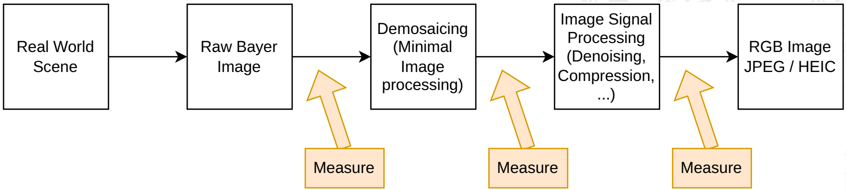
[2] Lukas, J., Fridrich, J. J., & Goljan, M. (2008).Digital camera identification from sensor pattern noise.. *IEEE Transactions on Information Forensics and Security*

Aim of this research

- **Goal:** Determine limiting factors why cross-sensor material classification yields inferior results

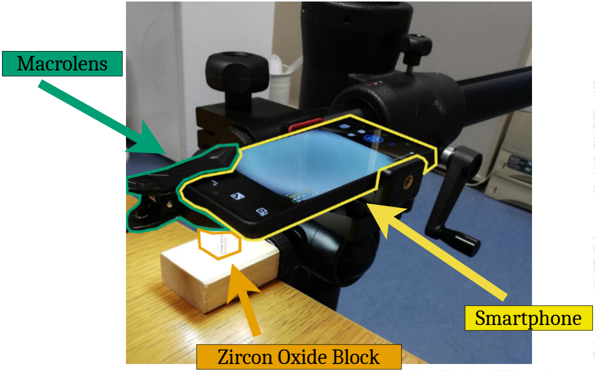
- **Assumption:** Signal of material property is much lower than device inherent signals & artefacts from image processing pipeline.

- Imaging Pipeline



Dental Ceramic Data Acquisition

- 7 different smartphones
- Zircon oxide blocks from 3 manufacturers
- Acquisition of images in RAW mode via app:
  - Android: OpenCamera (https://opencamera.org.uk/)
  - iPhone: Halide Mark II (https://halide.cam/)
- Macro lens with built-in illumination

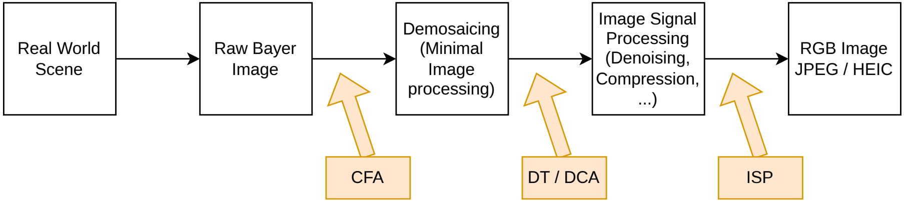


Smartphones and their imaging sensor resolution

Smartphone	Image Resolution	Scaling Factor
Google Pixel 4a (GP)	3024 x 4032	0.686
Huawei P20 Lite (H20)	3456 x 4608	0.600
Huawei P30 Pro (H30)	2736 x 3648	0.758
iPhone 11 (i11)	3024 x 4032	0.686
iPhone 13 Pro (i13)	3024 x 4032	0.686
Samsung Galaxy A52 (SG)	3468 x 4624	0.598
Xiaomi Mi A3 (XM)	3000 x 4000	0.691

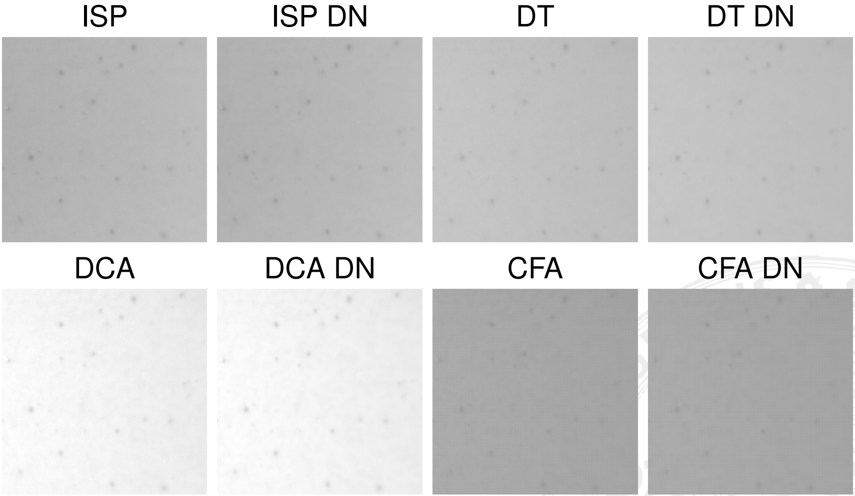
576 patches per smartphone for Ivoclar Vivadent (Manufacturer 1), 216 for Dentsply Sirona (Manufacturer 2) and 360 for 3M (Manufacturer 3).

Dental Ceramic Data Legend - Abbreviations



- **ISP** "Image Signal Processing" Pipeline: JPEG and HEIC files from smartphones.
- **DT** Color filter array image demosaiced using *darktable-cli*.
- **DCA** Color filter array demosaiced using *dcraw -a*: Average the whole image for white balance.
- **CFA** Color filter array extracted using *dcraw -d*: Document mode (no color, no interpolation)
- **DN** Apply *bm3d* [3] denoising filter on the whole image.

[3] Dabov, K., Foi, A., Katkovnik, V., & Egiazarian, K. (2007). Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on Image Processing*


ISP
ISP DN
DT
DT DN
DCA
DCA DN
CFA
CFA DN

Texture Classification Toolchain

- Same as in [4], originally proposed in ”Textures in the Wild” [5]
- Different feature extraction schemes:
  - Dense SIFT
  - Dense Micro-block Difference
  - LBP
  - Local Phase Quantization
  - Weber Pattern

  ↑ Followed by a PCA based dimensionality reduction, a Fisher Vector encoding and finally an SVM based classification
  - Rotation invariant LBP
- For brevity, only **SIFT** results shown in results

---

[4] Kauba, C., Debiasi, L., Schraml, R., & Uhl, A. (2016).Towards drug counterfeit detection using package paperboard classification. *Advances in Multimedia Information Processing – Proceedings of the 17th Pacific-Rim Conference on Multimedia*

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

Experiment Guideline

1) Test ISP as main factor: Use raw images

   Train/enroll on 1 sensor $\rightarrow$ Test/evaluate on 1 other sensor

Table: Average ceramic classification accuracy in inter-sensor setup using SIFT features and DCA images.

Train \ Test	GP	H20	H30	i11	i13	SG	XM
GP	–	0.837	0.852	0.707	0.832	0.607	0.997
H20	0.605	–	0.515	0.333	0.352	0.455	0.465
H30	0.846	0.716	–	0.676	0.854	0.751	0.906
i11	0.480	0.663	0.381	–	0.864	0.380	0.613
i13	0.813	0.579	0.362	0.819	–	0.344	0.552
SG	0.831	0.668	0.603	0.675	0.963	–	0.943
XM	0.982	0.804	0.451	0.787	0.903	0.358	–

Experiment Guideline

1) ~~Test ISP as main factor: Use raw images~~

   ~~Train/enroll on 1 sensor → Test/evaluate on 1 other sensor~~

   Does not solve problem!

2) Evaluate effect of color filter array
   - Baseline: **Intra-Sensor**

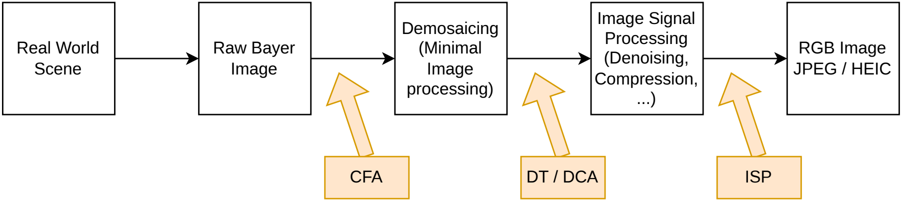
     Train/enroll on sensor A → Test/evaluate on sensor A

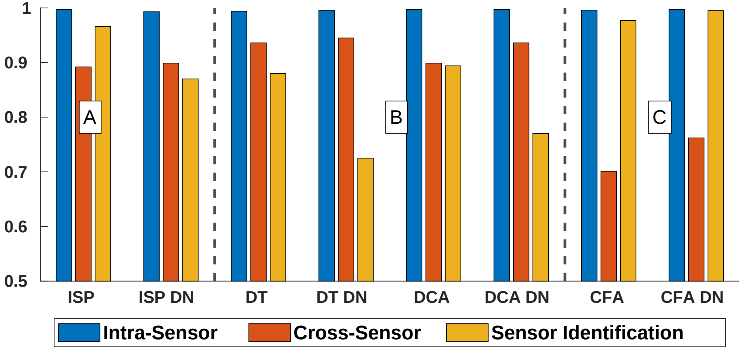
   - **Cross-Sensor**: Leave one out cross validation

     Train/enroll on all but sensor A → Test/evaluate on sensor A

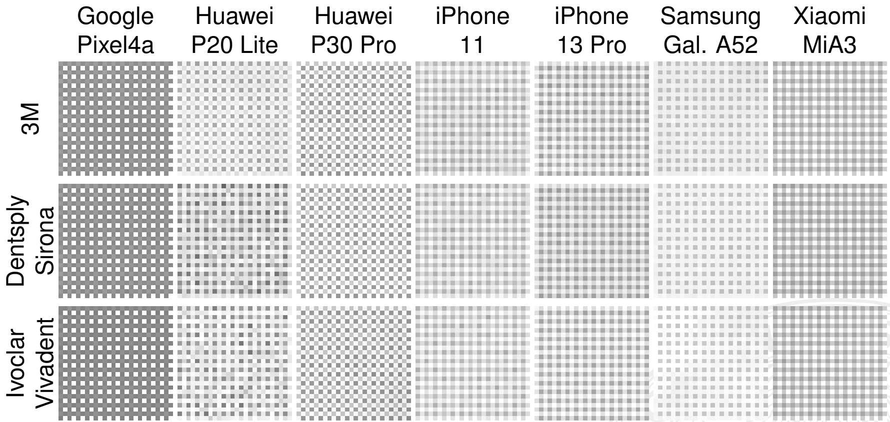
   - **Sensor Identification**

Results (SIFT)





Classification accuracy results CFA unscaled



	Texture Classification		Sensor Identification	
	CFA	CFA DN	CFA	CFA DN
SIFT	0.453	0.736	1.000	1.000

Conclusion & Future Work

Conclusion:

- While raw (DT) tends to slightly increase cross-sensor accuracy, using only the mosaiced image (CFA) greatly reduces the accuracy
- Second, the CFA and the processing pipeline (ISP) signals can be used for sensor identification
- This suggests: CFA mainly interferes with the low-amplitude texture signal of the material

Future Work:

- Further investigations by employing smartphone pairs (multiple devices of same model)
- Try to remove CFA artefacts through deep learning (e.g. domain adaptation)

Thank You!
Q & A