https://doi.org/10.2352/ISSN.2470-1173.2021.8.IMAWM-279

Visual document tracking and blockchain technologies in mobile world

M. Allouche, M. Ljubojevic, M. Mitrea
Telecom SudParis, ARTEMIS Department, SAMOVAR Laboratory
9, rue Charles Fourier, 91011 Evry, mihai.mitrea@telecom-sudparis.eu

Abstract

As during the last decade the limit between professional and personal usage of smartphones gradually disappeared, the present study is devoted to the tracking of the visual documents scanned by a personal mobile phone for some professional reasons. By a visual document we assume a composition of text, graphics and images corresponding to various physical-world documents (invoices, calls for tenders, legal contracts, etc.). As the scanning (capturing) conditions cannot be reproductible, the main issue is to unambiguously and securely identify various digital representations for a same physical document. A second issue is related to the inherent constraints in resources made available for such a task in the mobile/embedded environment. To jointly solve these issues, a solution based on coupling the blockchain technologies to the visual fingerprinting principles is advanced. The novel elements thus brought to light relate to (1) the coupling of fingerprint and blockchain solutions, (2) the unitary smart contracts generation and management (with illustrations for the Tezos blockchain) and (3) an on-chain / off-chain work balancing solution for coping to the mobile world constraints. The experimental results obtained on a database of more than 10 000 visual documents resulted in F1 score equal up to 0.98 while being compatible with low-resources computing environments (Raspberry Pi).

1. Problem statement

Digitization is currently considered as the "industrial revolution of our age" and the main enabler for ensuring background compatibility with the "old world" is the dematerialization of the documents. As this process is still expected to span over several years, there is a particular need for ensuring a unitary management system for visual documents (a composition of text, graphics, images corresponding to various physical world documents like invoices, calls for tenders, legal contracts, etc.), irrespective of their creation source (native digital, scanned by a professional device or by a mobile phone camera).

The tracking of visual content can be accommodated by two existing methodologies, namely blockchain and fingerprinting. A blockchain is an information storage technology, ensuring trust in the tracking and the authentication of the binary data exchanged in a decentralized, peer-to-peer network: even the smallest (1 bit) modification in a message can be identified. As such, a bit sensitivity property is incompatible with the digital document tracking (where multiple digital representations can be associated to a same semantic), the blockchain principles should be coupled to the visual fingerprints (which are compact and salient visual content features computed from the content itself and which can uniquely identify duplicated and/or replicated versions of it in a reference database). Yet, state-of-the-art blockchain solutions are known as being computational prohibitive in mobile and/or embedded environments.

Under this framework, the objective of the present paper is to unambiguously and securely identify various digital representations for a same physical document by ensuring effective coupling of visual fingerprinting to blockchain methodologies.

2. State of the art

The state-of-the-art section is structured at two levels: first, the basic notions related to the blockchain are presented, then the image fingerprinting framework is introduced.

2.1 Blockchain basic concepts

Blockchain relates to the peer-to-peer networks and provides a universal dataset that every actor can trust on, even though they might not know or trust each other. It provides a shared and trusted ledger of transactions, where immutable and encrypted copies of information are stored on every node in the network. Those characteristics look appealing for our use case to track digital documents.

Be Alice and Bob, two out of the N users in a network, exchanging peer-to-peer messages in an authenticated and decentralized way. The messages sent by Alice to Bob are first authenticated by one or a group of users in the network (called the *miners*), according to a protocol everybody agreed on. Once authenticated, the message is stored (together with the Alice's and authenticator's IDs) into a ledger which is distributed to all the users. Assuming a malicious user or a group of users wants to modify local copy of the ledger, the other users are able to identify this modification and to fix it.

The main blockchain concepts are:

- **transaction:** Transactions constitute the data exchanges between users. Each transaction is signed by the sender private key. This signature guarantees the security of transactions. This is to prevent any changes to these transactions during its transmission.
- **block:** A block is a record in the blockchain that contains the transactions confirmed. Each pending transaction will be added to a block. After a period, a new block containing transactions is added to the blockchain. First, this block is given to the miners to be verified and validated by a process called mining.
- **chain of blocks:** Each block in the blockchain is linked to the previous block by incorporating the value of its hash. That is, the hash of each block includes the hash of the previous block. This prevents any modification of the contents of existing blocks.
- **smart contract:** A Smart contract is a software "installed" on the blockchain that automatically executes a pre-programmed contractual commitment. It is not a

legal document in itself, but it automates the execution of a contractual commitment.

Nowadays, many of blockchain solutions coexist, like Bitcoin, Ethereum, Hyperledger or Tezos, to mentioned but a few. Each blockchain comes across with conceptual and/or technical advantages and disadvantages. As our study considers the Tezos blockchain [2], the next sub-section will be devoted to it.

2.2 Tezos main features

According to the Tezos creators, *"Tezos is a new decentralized blockchain that governs itself by establishing a true digital commonwealth".*

Tezos is a platform for secure smart contracts and decentralized applications. The important features differentiating Tezos from other smart contracting platforms relate to both new concepts and new ways of realizing concepts already advanced for other blockchains: formal verification, on-chain governance, self-amendment, and Proof-of-Stake (PoS) consensus, as described in the following sections.

2.2.1 Reliability and formal verification

Reliability of the blockchain and the contracts that run on it is always a concern. If platforms do not behave as expected, or if contracts do not behave as the authors expected, the consequences can be quite serious. Incidents of defects diminish trust in smart contracts, the system, and even the blockchain technology in general. For Tezos, the contracts are more secure and reliable thanks to the formal verification natively embedded in the platform.

2.2.2 On-chain governance

Upgrades are previously agreed processes that define how governance rules are to be dynamically changed. Without clearly defined governance mechanisms, upgrades can be disordered, inefficient, and damaging to the communities. Tezos governance process is designed to address this problem. Tezos offers a pre-defined governance mechanism to upgrade the protocol.

Tezos presents a mechanism that empowers all stakeholders to more efficiently coordinate to reach consensus on protocol upgrades. In doing so, stakeholders can more effectively conduct a decision-making process which takes all stakeholders' preferences into account.

2.2.3 Self-Amendment

Node operators of other blockchain implementations must manually upgrade the software of their nodes when conducting a network upgrade, which can be extremely difficult and becomes more and more difficult as the network grows.

Self-amendment in Tezos is designed to solve this coordination and execution problem of conducting upgrades in a decentralized network. If an amendment (upgrade) to Tezos is approved by its stakeholders via the governance mechanism, then all nodes are automatically upgraded.

2.2.4 Proof of Stake

Blockchain is just a linked list of blocks where each block contains a set of transactions. Each blockchain has a technique to create these blocks and for nodes in the blockchain network to add blocks to the chain (mining). The most popular technique is the proof of work (PoW) used in blockchains like Bitcoin and Ethereum. One of the big disadvantages of PoW is that it takes lot of computational power (waste of electricity) to earn the right to include a block in the blockchain.

Tezos uses a technique called Proof of Stake (PoS). The baker, that produces blocks, is randomly selected to do so based on its own funds staked and those that are delegated to it by other stakeholders. Randomly selected endorsers validate blocks produced by bakers.

2.3 Image fingerprinting

One way to define image fingerprints is in relation to the human fingerprints. For humans, natural identifiers are the patterns of dermal ridges on their fingertips. Although they convey very little information compared to the entire person, human fingerprints are sufficient to uniquely identify a person even if a person changes haircut, clothes, or wears a wig or a disguise.

Analogously, image fingerprints are intended to be image identifiers. Image fingerprints have to be able to uniquely identify visual contents even if the image content goes under a predefined, application dependent set of transformations. The transformations an image can undergo can be different kinds of modifications, distortions, or attacks, either malicious or mundane. The image which is transformed, modified, distorted or attacked will be denoted as a copy or a replica of an image [3].

In general, the fingerprinting function needs to have the following properties:

- **Robustness**: fingerprints resulting from degraded versions of an image should result in the same or at least similar fingerprints with respect to the fingerprint of the original image. Robustness refers to the ability to positively identify two perceptually similar objects as similar. The robustness property can be quantified by recall rate.
- **Uniqueness**: If two images are perceptually different, the fingerprints from two images should be considerably different. For pairwise independence, it is desirable that fingerprint bits have uniform distribution and are uncorrelated with each other. The uniqueness property can be quantified by precision rate.
- **Database search efficiency**: for applications with a large-scale database, fingerprints should be conducive to efficient database search (fast fingerprint computation and matching, compact form …).

3. On-chain / off-chain work balancing solution

3.1 Conventional smart contract workflow

Tezos is a smart contract platform focusing on enabling formal verification for smart contracts, that provides a mathematical proof that all deployed smart contracts do what they are designed and programmed to do. The formal verification can be considered as a tool to help smart contract programmers, but deployment of smart contract is not an easy process yet. That process milestones are presented in Figure 1.

Programming is related to writing a Michelson program that converts a logical contact into Tezos script. Formal verification is done by executing the Tezos commant "*typecheck script*" with proper parameters. The command does not apply any action on the blockchain, its output just notifies the developer about the smart contract correctness. Deployment is done by executing the Tezos command *"originate"* with the proper parameters (smart contract name, owner, transaction cost, ...). The smart contract is added to

the new block (the smart contract keys are generated) and is ready to be invoked. With Call/Execution, the action is added to the blockchain, the smart contract is executed, and the results are stored in the block.

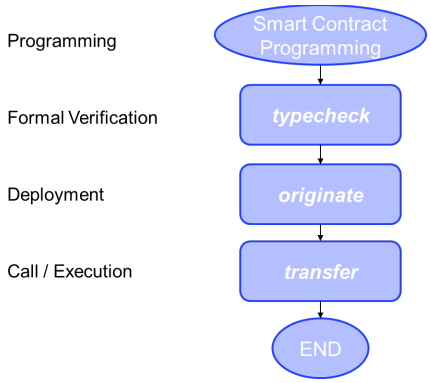


Figure 1: Smart contract workflow

The main usual blockchain and smart contracts use cases are finance related, therefore, it is important to note their limitations when confronted to the requirements of more complex use cases.

As a fundamental limitation, Michelson language is bounded to process integer values and cannot be extended (while keeping backward compatibility) so as to process floating point values.

As applicative limitations, the blockchain operation cost steady increases with the size and complexity of the smart contract.

In order to overcome these limitations, two solutions can be considered. First, the mapping of the operations from floating point to integers can be a solution for this fundamental limitation but it increases both the size and the computational complexity of the smart contracts. Consequently, we designed an architecture for ensuring an on-chain / off-chain code balancing, as described in the next sections.

3.2 New smart contract workflow for on-chain / off-chain code balancing

In order to overcome the conventional smart contract limitations presented in Figure 1, for complex applications, we designed the architecture presented in Figure 2.

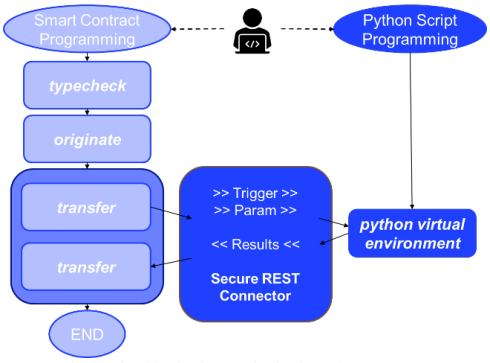


Figure 2: On-chain / off-chain code balancing

The new design ensures that the contract computational load is balanced between the on-chain and the off-chain computing. So, instead of translating all the contract articles into a smart contract that is likely to lead to fundamental and functional limitations (as explained in the section above), two computational resources are in charge to the well execution of the digital contract.

On one hand, the smart contract code representing the transactional (monetizing, legal articles) aspects of the contract preserves exactly the same workflow as presented in the section above. On the other hand, a python script, paired with the transactional script, takes in charge the heavy calculation load. Having a complete python virtual environment gives us the opportunity to implement high complexity and resource-consuming algorithm that cannot be supported in the blockchain or the data processing, which is black-boxed in a specific use case.

Of course, in such a system that mixes multiple processing environments, data communication between the entities is usually considered as a failure point. The connector serves as an entry and exit point for both on-chain and off-chain entities. First, it is a tool which scans the blockchain state, block generation and transaction made in order to notify the off-chain entity whenever an action or intervention is asked for. Second, it collects the required data (block ID, transaction ID, needed action, parameters, …) and passes them to the python script. In the python script point of view, the passed data is its entry parameters and execution invoker.

The python program decodes the entry parameters, parses them, executes the asked operations and encodes the results. The encoded result is sent back to the connector, which will check authenticity and integrity, and then it goes back to the on-chain entity.

Finally, the resulting data is added to the blockchain as a proof of accomplishment. The connector transfers the results to the smart contract that finishes its execution.

The connector has access to both on-chain and off-chain entities thanks to a REST API. The Tezos node provides a JSON/RPC interface [4] and our work was related to developing the off-chain part.

4. Fingerprinting method

Thanks to its ability to represent, in a compact way, salient content of images and to its low complexity, the discrete wavelet transform (2D-DWT) is intensively considered in many image processing applications like watermarking, compression, defect detection, etc. Actually, after a preliminary experimental study, we decided to consider for this paper a DWT-based fingerprinting method described in [3]; this method is illustrated in Figure 3 and will be detailed here-after.

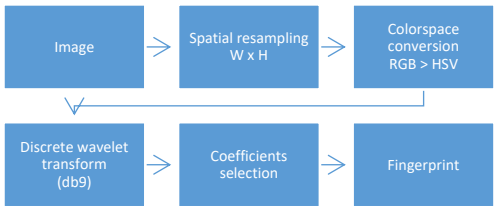


Figure 3: DWT-based fingerprint computation principle

4.1 Spatial re-sampling

The CIF format (W = 352, H = 288) was chosen in our implementation in order to decrease the computational time. The computation time of the image fingerprint is directly proportional to the size of the image, therefore the smaller the size, the shorter the processing time. Important to note is that any other resolution will not influence the stability of the fingerprinting function.

4.2 Color space conversion

The image color space is changed from RGB (or RGBA) to HSV (Hue-Saturation-Value) and only the V component is considered in the next fingerprint computation steps. The HSV color space separates the luma (the image intensity, in the H and S components) from chroma (i.e. the color information in the V component). Selecting only the V component makes the fingerprinting function invariant to color distortions or changes so increases the robustness of the method.

4.3 Daubechies wavelet transform

The fingerprint was computed in the discrete wavelet domain due to its capacity of identifying the overall salient image's visual properties and representing them through edges in the high frequency sub-bands.

Moreover, the 9/7 Daubechies wavelets were used over other DWT types like 2/2 or 4/4 due to their very fine capacity of approximating the visual content [10].

4.4 Coefficients selection

The fingerprint is computed by selecting 2D-DWT coefficients. The coefficient selection aims to collect as much information as possible in order to achieve robustness with the minimal influence on the database search efficiency. Consequently, all the 2D-DWT coefficients in each frequency sub-band (LL, LH, HL, HH) yielded by the 3rd level wavelet transform are selected.

4.5 Fingerprint matching

A match between the digital image and one of the reference fingerprints is defined by:

$$t_{test} > Z_{\alpha/2}$$

where:

$$t_{test} = \rho \times \sqrt{\frac{N-2}{1-\rho^2}} \text{ and } \alpha = 0.05,$$

$$\rho = corr(f,t) = \frac{1}{N-1}\sum \frac{(f_k(x,y) - \bar{f_k})(t_k(x,y) - \bar{t_k})}{\sigma_{fk}\sigma_{tk}}$$

f_k and t_k designate the 2D-DWT coefficients of the digital document and the reference documents respectively, in a frequency sub-band k, $\bar{t_k}$ and $\bar{f_k}$ are the mean values of the 2D-DWT coefficients in the considered frequency sub-band, while σ_{fk} , σ_{tk} are the related standard deviations, respectively.

N designates the number of 2D-DWT coefficients in every frequency sub-band k. is each of the frequency sub-bands: LL, LH, HL, HH yield by the wavelet transform. A perfect match (identity) between the query and the reference fingerprints is obtained when $|\rho| = 1$; a value $\rho = 0$ indicates no correlation between f_k and t_k.

5. Experimental results

5.1 Experimental testbed

To establish the PoC on the coupling of visual fingerprinting solutions to blockchain, under joint functional/deployment constraints, the experiments were performed on a multiplatform set-up, composed by a Raspberry Pi 3 and a PC. To ensure the ease of use of the system, a *Rest API* was added to communicate with Raspberry Pi via network. We developed a graphic user interface (GUI) that interacts with the system, as illustrated in Figure 4.

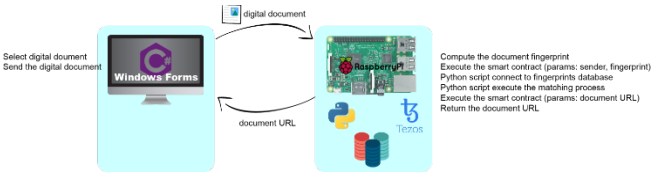


Figure 4: System components

5.2 The applicative workflow

The following applicative workflow is considered. Assume a new document is produced and scanned.

Its fingerprinting is first computed, then checked in the database by a smart contract deployed on a Tezos blockchain.

Assuming the fingerprint is already present, a detection message is sent to the user.

On the contrarily, assuming the fingerprint is not detected in the database, it is added to the database and the corresponding message is sent to the user.

Note that the computational/storage intensive operations (fingerprinting computation, matching, data-base updating) are processed by the off-chain module.

5.3 Experimental database

To evaluate the studied methods performances, we created a 11557 digital document database. The digital documents are JPEG images obtained from 25 PhD thesis defended in France in various research fields (physics, biology, philosophy, …) accessible on [11]. For illustration, some samples are shown in Figure 5.

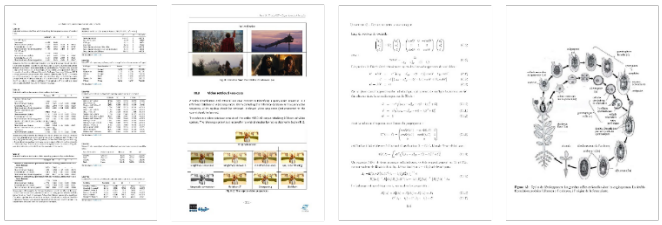


Figure 5: Original images sampled from database

We created a set of test images, which consists of images produced by various attacks on the original images from database.

Each test image is composed of two parts extracted from two database images with different ratios ([50% image1, 50% image2], [67% image1, 33% image2], [75% image1, 25% image2]) as shown in Figure 6.

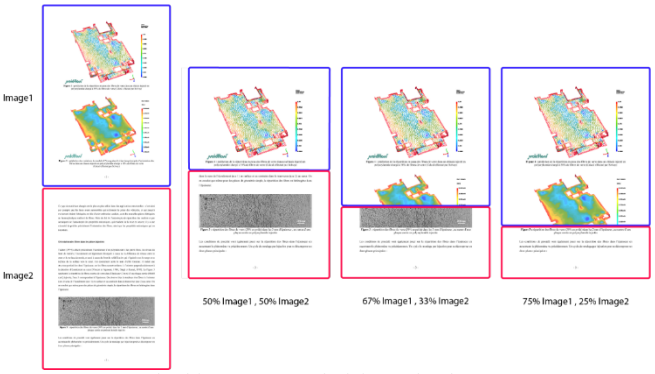


Figure 6: Attacked images sampled from database

The other types of attacks are produced using various types of distortions, such as: sharpening, Gaussian filtering, shearing, median filtering (2x2, 3x3 and 4x4), row and column removal, scaling with various factors (0.5, 0.75, 0.9, 1.1, 1.5 and 2), rotation by degree 0.25, as well as coupling shearing with Gaussian filtering, which are usual distortions made by a digital camera.

5.4 Performance evaluation criteria

We shall denote by *fp* the false positive, by *fn* the false negative, by *tp* the true positives and by *tn* the true negatives.

To evaluate the uniqueness property, two evaluation criteria are considered in the literature: the *probability of false alarm* (*Pfa*) and the *precision rate* (*Prec*) defined as [1]:

$$P_{fa} = \frac{fp}{tp+tn+fn+fp} \qquad Prec = \frac{tp}{tp+fp}$$

To evaluate the robustness property, two evaluation criteria are considered in the literature: the *probability of missed detection* (*Pmd*) and the *recall rate* (*Rec*) defined as [1]:

$$P_{md} = \frac{fn}{tp+tn+fn+fp} \qquad Rec = \frac{tp}{tp+fn}$$

To evaluate the system accuracy, the *F1 score* is considered. It is computed only in the case of correctly identified copies, as the harmonic mean of *Prec* (precision) and *Rec* (recall) rates and defined as [1]:

$$F1 = \frac{2 \times Prec \times Rec}{Prec + Rec}$$

An efficient fingerprinting function should ensure a low probability of missed detection and a low probability of false alarm, as well as a high precision rate, a high recall rate and a high F1 score.

5.5 Numerical results

Blockchain main and the most used algorithms are baking, transferring coins, new smart contract type check to approve the formal verification and smart contract deployment. Figure 7 presents the execution time of Tezos blockchain operations, with 5 Tezos nodes running, on computer and Raspberry Pi.

	PC	Raspberry Pi
Baking	0.994 s	2.216 s
Coin transfer	1.261 s	4.575 s
Type check	0.505 s	3.167 s
Smart contract deployment	1.836 s	7.263 s

Figure 7: Tezos basic operations execution time

The following figures represent results obtained by applying various attacks: F1 score, precision and recall, as well as thresholds at which those results are obtained for each type of attack. Thresholds are selected at values which give maximum F1 score.

F1 score is presented on Figure 8, and it varies from 0.9, up to 0.98 for mixed-content. Shearing combined with Gaussian filtering, which simulates most common distortions when image is taken by a mobile phone camera, gives F1 score of 0.92.

Figure 9 represents precision values for all of the attacks, which vary from 0.85 to 0.97.

Figure 10 represents recall with values from 0.93 for row and column removal, up to 1 for sharpening.

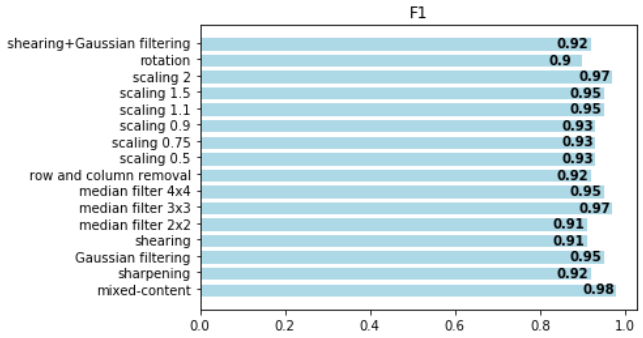


Figure 8: F1 score for various types of attacks

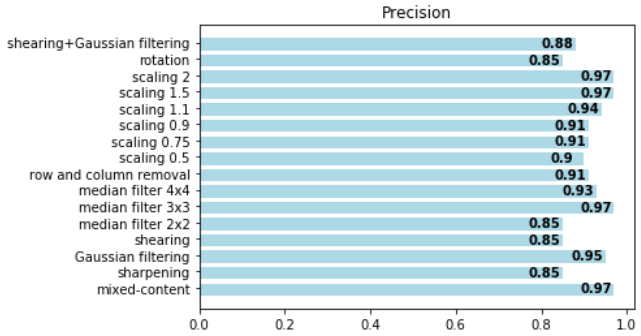


Figure 9: Precision for various types of attacks

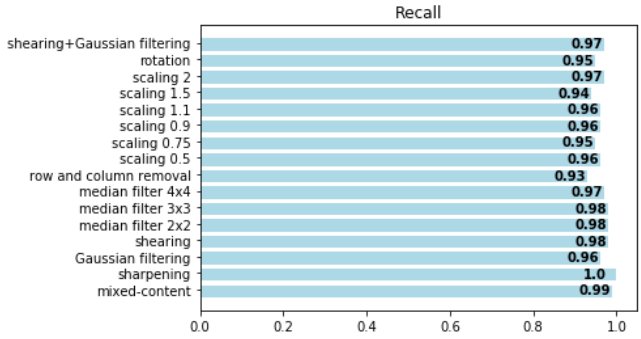


Figure 10: Recall for various types of attacks

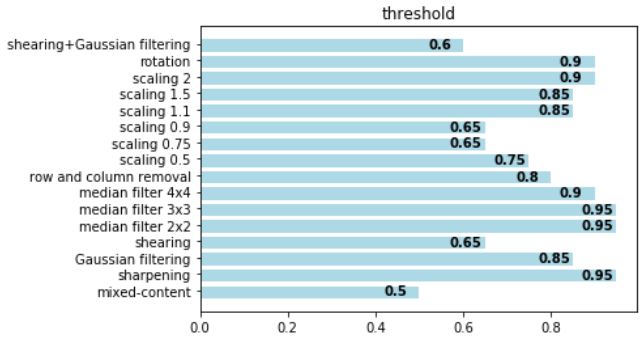


Figure 11: Thresholds for F1 score, precision and recall

6. Conclusion

The main goal of our work was to validate a proof-of-concept for coupling visual fingerprint and blockchain technologies. We

focused on analyzing the Tezos blockchain and deploying it on an embedded device. Fingerprinting method implemented in the system is based on discrete wavelet transform. Adopting the on-chain / off-chain code balancing technique that we designed and developed, we are able to execute smart contract with high computational complexity (the fingerprinting method in our case) that was not possible with the conventional smart contract workflow proposed by Tezos.

The experimental results were obtained on a database with more than 10 000 original visual documents. Measures are formalized using F1 score, which resulted up to 0.98.

References

[1] A. Garboan, "Towards camcorder recording robust video fingerprinting," 2012.

[2] "Tezos," [Online]. Available: https://tezos.com/.

[3] A. Garboan and M. Mitrea, "Live camera recording robust video fingerprinting," *Multimedia Systems,* vol. 22, no. 2, pp. 229-243, 2016.

[4] L.Goodman, *Tezos — a self-amending crypto-ledger White paper,* 2014

[5] V. Awasthi, V. Awasthi and K. Kumar Tiwari, "Finger Print Analysis Using Termination and Bifurcation," *International Journal of Emerging Technology and Advanced Engineering,* vol. 2, 2012.

[6] "liquidity," 2017. [Online]. Available: liquidity-lang.org.

[7] "SmartPy," 2019. [Online]. Available: smartpy.io.

[8] "LIGO," 2019. [Online]. Available: https://ligolang.org/.

[9] "FI," 2019. [Online]. Available: fi-code.com.

[10] D. Tay, S. Marusic, M. Palaniswami and G. Deng, "Mathematical properties of the two-parameter family of 9/7 biorthogonal filters," in *IEEE International Symposium on Circuits and Systems*, 2004.

[11] "Investopedia,"[Online].Available: https://www.investopedia.com/terms/b/blockchain.asp.

[12] "Tezos Developer Documentation," TEZOS , [Online]. Available: http://tezos.gitlab.io/developer/rpc.html.

[13] "hal," [Online]. Available: https://hal.archives-ouvertes.fr/.

[14] "Cambridge Bitcoin Electricity Consumption Index," 2018. [Online]. Available: https://www.cbeci.org/comparisons/.

[15] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," *ryptography Mailing list at https://metzdowd.com,* 2008.

[16] "technopedia," 2019. [Online]. Available: https://www.techopedia.com/definition/30119/digital-transformation

[17] "InStream," April 2018. [Online]. Available: https://instreamllc.com/document-scanning-first-step-digital-transformation-process/.

[18] "IBM," [Online]. Available: https://www.ibm.com/blogs/blockchain/2017/10/blockchain-in-healthcare-patient-benefits-and-more/.

[19] "betanews," November 2019. [Online]. Available: https://betanews.com/2019/11/11/digital-transformation-is-the-industrial-revolution-of-our-age/.

Author Biography

Mohamed Allouche received his computer engineering degree from University of Sfax at National engineering school of Sfax (2019) and MsC in embedded systems from the same school (2020). The results presented in this paper were obtained during his MS internship at Telecom SudParis. He is currently a PhD student with Vidmizer (https://vidmizer.com/en/) and Telecom SudParis.

Marina Ljubojevic received her BS in electrical engineering from the University of Belgrade (2018) and MsC in multimedia networking at Telecom Paris (2020). The results presented in this paper were obtained during her MS internship at Telecom SudParis. She is currently a PhD student with info3 (https://www.infocubed.com/) and Telecom SudParis.

Mihai Mitrea holds an HDR degree Pierre and Marie Curie University in Paris (2010) and a PhD from Politehnica University in Bucharest (2003). He is currently Associate Professor at Telecom SudParis. He is vice-president of the Cap Digital's Technical Commission on Digital Content and serves as advisor for the French delegation at ISO/IEC JTC1 SC29 (a.k.a. MPEG).

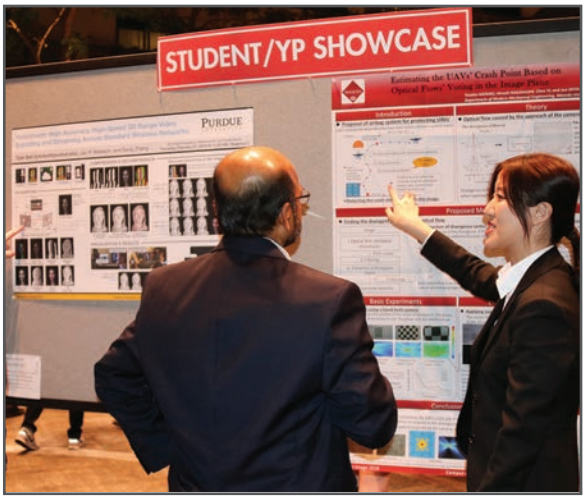
JOIN US AT THE NEXT EI!

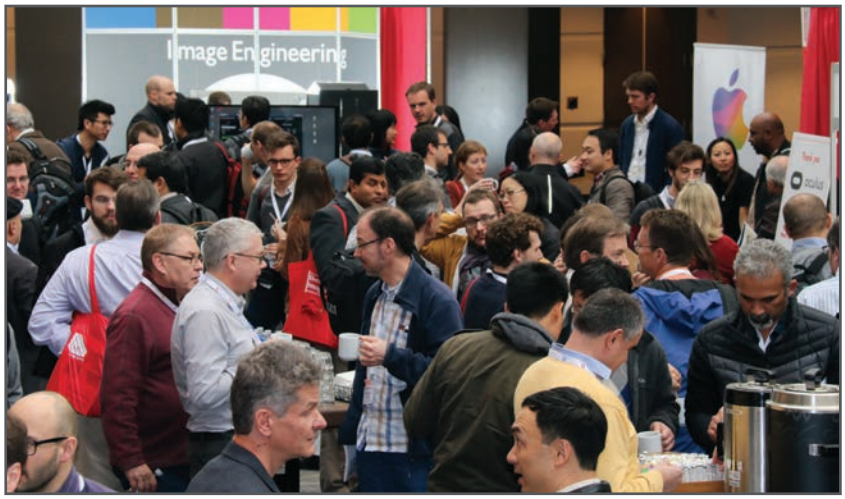
IS&T International Symposium on

Electronic Imaging

SCIENCE AND TECHNOLOGY

Imaging across applications . . . Where industry and academia meet!





• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •
• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •

www.electronicimaging.org

IS&T
imaging.org