Modeling of Image Copyright Protection using Discrete Cosine Transform Hash and Blockchain

Carles Juliandy[1], Ronsen Purba[1], Roni Yunis[2], Darwin[1]
[1]*Information Technology Master Department, STMIK Mikroskil, Medan, Indonesia*
[2]*Information System Department STMIK Mikroskil, Medan, Indonesia*

Keywords: Discrete Cosine Transform Hash, Blockchain, Copyrights Protection, Images

Abstract: Protecting copyright is an important issue because now such works as an image can be sold online for making income. With the rapid development of distribution media, a centralized management system cannot protect copyrights properly. Because now some research about image plagiarism can detect image modification as plagiarism, but cannot detect rotated image as plagiarism. Application of Discrete Cosine Transform (DCT) hash with adding looping steps can detect rotated images as plagiarism. Uploaded image we looped it rotate 22,5° and saved the hash value from DCT hash each time rotate until 180° and then compare each hash with the first hash to get the rotating plagiarism image. After that, a combination with Blockchain which is a decentralized management system is a solution to protect copyrights now, with the application of Blockchain and digital signature, making it difficult for other people to make changes to the data which is stored in the block. This research results showed that the use of DCT hash can reach accuracy until 99,67% to detect rotating image as plagiarism, and the mining time of the Blockchain with 10.000 blocks and difficulty target 5 needed 1591204,671 seconds.

1 INTRODUCTION

Protecting copyright was an important issue because copyright was an appreciation of the work and creativity of the author. Currently works such as images can be sold online for income. Copy of the media that easy to do has an impact on the modification of media that is easily done too (Cho and Jeong, 2019; Mehta, 2019; Ravindran, Zacharia & Roy, 2018). Image plagiarism was used or modified some or whole parts of the image without any permission and give credit to the author (Aghav et al. 2014; Ovhal et al. 2016). To detect an image modification, the use of a cryptographic hash can cause an avalanche effect, which was the effect where a small change of input value can lead to a drastic change of output value that can make it difficult to detect if there was any modification of the image. Perceptual hash was hashing algorithm that differently from the cryptographic hash, that can keep away from the avalanche effect, which is a small change in input value will affect some or none bit change (Mehta, 2019; Drmic *et al.*, 2017). Recently centralized management system can't protect the copyright as well as the decentralized management system. Blockchain as a decentralized management system that is immutable, integrity, traceability, and transparency can protect it better because Blockchain doesn't need a centralized server and interference from network members (Cho and Jeong, 2019; Kibet, Simon and Karume, 2018).

The used of Blockchain to protect the image copyright has been done by several researchers. Knirsch (Knirsch et al. 2018) researched the use of Blockchain and smart contracts using the concept of digital signature to provide private key and public key to the proof of possession and can be forensic evidence when claiming the copyright of the image. (Mehta, 2019) research the use of Blockchain and perceptual hash, where the Blockchain and smart contracts are used to record all transactions at the image marketplace. Perceptual hash and hamming distance are used to detect the similarity of two images. In this research, the perceptual hash can detect any modification as plagiarism, but it can't detect the modification of 90° rotation image as plagiarism.

The use of perceptual hash to detect the image similarity has been researched by some researchers before. Aghav *et al.*, 2014 researched the capability

Juliandy, C., Purba, R., Yunis, R. and Darwin, .
Modeling of Image Copyright Protection using Discrete Cosine Transform Hash and Blockchain.
DOI: 10.5220/0010304600003051
In *Proceedings of the International Conference on Culture Heritage, Education, Sustainable Tourism, and Innovation Technologies (CESIT 2020)*, pages 128-134
ISBN: 978-989-758-501-2

of perceptual hash to detect the rotation image as plagiarism. This research generated a hash value every time rotate the image 22,5° and then compare it with the real hash value of the image in the dataset. Rivas *et al.*, 2017 researched the use of perceptual hash to detect similarity images when uploaded on social media. Drmic (Drmic et al. 2017) compared each well-known algorithm in perceptual hash such as average hash, differential hash, discrete cosine transform (DCT) hash, and wavelet hash. This research showed that DCT hash is the most robust perceptual hash algorithm to detect similarity image.

In this paper, we introduced the model that not only detected the similarity of the uploaded image but can protect the copyright of the image that is already saved. The proposed model combined the use of Blockchain technology with the digital signature and perceptual hash to (1) protect and prevent the attempt to change important information such as hash value, image owner name, image name, uploader name, and image added date, (2) to detect the similarity of modification image such as gamma correction, resize, rotate, crop, and salt and pepper noise that try to upload especially to detect rotation of 90°, 180°, and 270° that previous research can't solve it, (3) with ECDSA digital signature it can be an additional security to protect and proof the possession of the received upload image.

As mention above about the proposed model, it can be explained that the contribution of this research is the use of looping steps for DCT hash to detected the plagiarism image, especially for the rotation image which cannot be detected by the previous researcher model. The improved DCT hash is combined with Blockchain technology and the digital signature to improve the security of data, so it cannot be changed once it is saved into the block.

The remaining paper is structured as follows: Section 2 provides background research related to perceptual hash and blockchain application. Section 3 presents research methodology or our approach to detect the similarity of the uploaded image and prevent the attempt to change important data about the image that is already saved before and to proof of possession with the use of ECDSA. Section 4 provided the result and discussion and Section 5 provided the conclusion of the paper.

2 RELATED WORKS

Blockchain was a decentralized management system invented by Satoshi Nakamoto in 2008 and implemented in 2009. Bitcoin is the first application that implemented this technology to handle the transaction of cryptocurrency. As a result, Bitcoin did not need a third party to validate the transaction. All transaction in Bitcoin is validated by together agreement which is called Consensus. Blockchain isn't a standalone technology, it consists of cryptography, mathematics, algorithm, economic model, combine of peer to peer network (P2P), and consensus algorithm which is agreed by everyone who joined to the network (Wang et al. 2018). The use of Blockchain is well known because it's capable to secure data inside, it prevented other people who want to change the data which is already saved in the block. But this technology still has an opportunity to be hacked, if the attack is offense more than 50% (50%+1) of network members at the same time. But this is something that almost impossible to do because it needed many resources in computing. (Lin and Liao 2017).

The related works about Blockchain and perceptual hash have been done by some researchers. In 2014 (Aghav et al. 2014) research the use of DCT hash which is one of the perceptual hash algorithms to detect rotation image modification. This research generated the hash value every 22,5° rotation and then compared it with the hash value of each image that previously saved before. Compared the hash value is using a hamming distance if the hamming distance value below the threshold image will be rejected, if the hamming distance value above the threshold it will looping rotate the image and repeat the step before until the image rotated 180° clockwise and anti-clockwise. This research model is when there is just one hamming distance value below the threshold all processes will stop and it will be considered as plagiarism image. Bhowmik and Feng, (2017) researched the use of Blockchain store the watermark of unique information that consisting of transaction history and the hash value of image which can be used to find a similarity image. The result of this research is using history transactions and the value of image hash, it can be defined as the part of the image that is edited or be changed. Knirsch *et al.*, (2018) research the use of smart contracts with digital signature to handle the claim of copyright possession by generated private key and public key. In this research private key is kept by the author of image, and the public key is used by another to verified the possession of image. But the analysis of this research is focus on evaluate the operational cost implementation of this methoed indeed of evaluate the effectiveness of this method. The conclusion of this research is if the more image

size than the more cost is needed to implement this method.

Jnoub and Klas, (2019) researched the use of Blockchain to protect the image and the copyright by register the ownership information and copyright into the block of Blockchain. In this research, the image is not stored in Blockchain, but just store tow hash value of the image which is extracted directly from the image, and the second hash value is extracted using speed up robust feature (SURF). This research efficient the storage of the Blockchain because it just stored text value. Mehta (2019) researched the use of perceptual hash to detect the similarity of the modification image. Blockchain is used in this research to record the transaction for someone who uploaded the plagiarism image and get a financial penalty because of that. In this research, there is a problem when the proposed method failed to detect the rotate image of 90° as a plagiarism image. Andi, Purba, and Yunis, (2019) researched the use of Blockchain to prevent plagiarism in a scientific publication. This research uses the combination of Blockchain, SHA-256, and ECDSA to protect the publication data, so it's nearly impossible for other people who tried to change the data. The benefit of this research is it can prevent the reviewer or someone who handles the publication to used, changed, or modified the paper which needs to be published. Because the paper is already signed with ECDSA and only can be changed with the private key.

3 RESEARCH METHODOLOGY

In this paper, we proposed the model to solve three problems: (1) the use of cryptographic hash to detect the similarity of two images can cause the avalanche effect, wherewith the small change of input value can cause a drastic change in output value, which make it impossible to detect the similarity image, (2) the implementation of perceptual hashing (DCT hash) with looping step can detect other modification as plagiarism, such as gamma correction, resize, crop, rotate, and salt and pepper noise. Which is for the rotation it failed to detect 90° rotated image as plagiarism in previous research, (3) the implementation of Blockchain to protect the copyright just used it to store the possession of image, so it difficult for the owner to prove that the copyright of the image is their image when the image is too many.

This research used the same dataset as previous research from (Mehta, 2019) the dataset is from Berkeley Segmentation Dataset (BSDS 500) which contained 500 images and then modified the dataset with 42 modifications that fitted (Mehta, 2019) dataset So at last, the dataset has 21.500 images (500 original images and 21.000 modified. The 42 modification of the dataset is consisted of:

- Rotation (in degrees): 5°, 10°,15°,20°,25° (clockwise)
- Gamma correction: 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5
- Salt and pepper noise: 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 1.00
- Crop (fixed aspect ratio mode): 5%, 10%, 15%, 20%, 25%
- Resize (size reduction - fixed aspect ratio mode): 5%, 10%, 15%, 20%, 25%

This research proposed a new model of implementation of perceptual hash and Blockchain to detect similarity and prevent the attempt to change the data in Blockchain. The perceptual hashing algorithm in this research is used the DCT hash algorithm which is the most robust perceptual hash algorithm (Drmic et al. 2017).

The DCT hash algorithm is a perceptual hashing algorithm that concludes binaries value which represented the image. This algorithm is based on cosine transform which is first it reduced the image into lower pixels than it transformed to grayscale after that it got the DCT value from that reduced and grayscale image. After that, it counts the average value of that DCT value, and that compared it one by one that DCT value with the average value to get the binary number of an image.

The proposed model is shown in Figure. 1 below

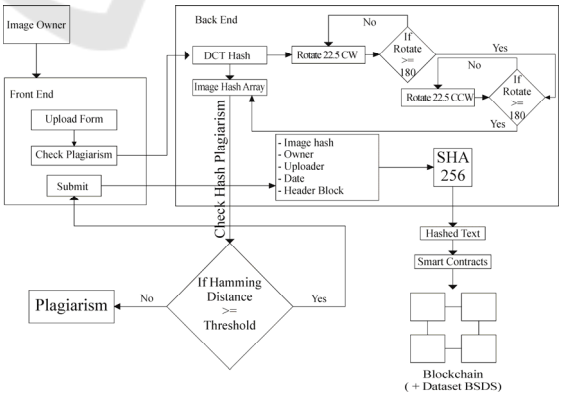

Figure 1: Proposed Model of Perceptual Hash dan Blockchain.

The proposed model consisted of two big part that can be explained below:

3.1 Front-end

In the front-end menu, the image owner or uploader will fill the upload form that consisted of the image name, image owner name, uploader name, and upload date. After filled the upload form, before submit, the uploader needed to check the plagiarism. This check will process on the back-end that will return the result if the image is considered plagiarism or not plagiarism. If the image is considered plagiarism the submit button cannot be clicked, if the image is not plagiarism the submit button will be available to click.

3.2 Back-end

On the back-end side after receiving the request to check the plagiarism, the image will convert into a hash value using the DCT hash algorithm and the hash value will store temporarily in the array. After that the looping process will start, it started to rotate the image 22,5° until 157,5° clockwise after that it will generate and store the image hash in the array that was created before every time it rotated 22,5°. After rotation is 157,5° image will rotate 22,5° anti-clockwise until 180° and generate and store it again in the array. Now we have 16 hash value of the uploaded image (1 from original, 7 from clockwise rotation, and 8 from anti-clockwise rotation). This array with all that hash value will compare with all image hash that is already in the dataset. To compared this image hash we used a hamming distance algorithm to compare bit per bit hash. The hamming distance result will compare to a threshold value that is already set. In this research, the threshold value is 14 (based on previous research by (Drmic et al. 2017)). If there is just one compared result below the threshold value it will be detected as a plagiarism image, and if all of the compare results above of threshold result in it return a value as free from the plagiarism act to the front end. After that, it can press the submit button and can be continued to submit. The submission process will store the data that already fill in the upload form include the original image hash and convert all that data with the SHA-256 algorithm to create a new block and that data will be signed with the ECDSA algorithm. The private key will be store by the image owner as proof of possession that the owner has that image. The public key will be used by another member to verify that the sign value is true and never changed before.

In this proposed model, we tried two different testing methods that are qualitative test and a quantitative test. In the qualitative test we test the function of all proposed models, from testing the DCT hash model to detect some modification images such as rotation, crop, resize, gamma correction, and salt and pepper noise. After that, we tested the Blockchain model to protect the data in the block, and next is to testing the ECDSA to verify the data that already sign. In the quantitative tested we test the accuracy of the DCT hash model to detect the rotate image as plagiarism. The testing is using 3 scenarios, try to upload 200, 400, and 600 images, and the second quantitative test is to test the mining time of the Blockchain model if there is someone who tried to change the data in the block. Our testing show that this model is capable to detect the plagiarism rotate the image and protect the image data from the possible plagiarism act.

4 RESULTS AND DISCUSSIONS

In this research, we build the application based on the proposed model using node.js. In the testing, we separate the test into 2 parts, qualitative test to show the functionality of the proposed model from the DCT hash to detect plagiarism and Blockchain with ECDSA to protect the data that already saved. The second part test is a quantitative test to show the accuracy of the DCT hash model to detect the rotate image as plagiarism and to show how the Blockchain model to handle the attempt to change the data.

For the qualitative test we try the DCT hash model with upload the original image that doesn't in the dataset before, and some modification images such as rotation, crop, resize gamma correction, and salt and pepper noise. The result showed that the functionality of the DCT hash model works well. When uploaded the original image that doesn't in the dataset before, the model can detect it as a non-plagiarism image and can continue to submit the process. The resulted testing can be seen in Figure 2.

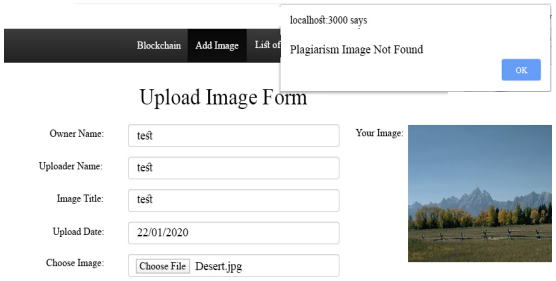

Figure 2: Successful Received Image

For the modification image, the model can detect it well as a plagiarism image, we have tried it with several modifications and all of them can be detected as plagiarism. The result of this testing can be seen in Figure 3. Below, Figure 3. is the rotate 90° image.

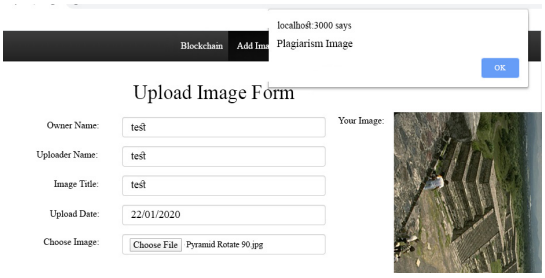

Figure 3: Rotation Plagiarism Image.

Another qualitative test is testing the capability of Blockchain and ECDSA to prevent an attempt to change the data inside. For the Blockchain model, we tried to change the block data. For example, we tried to change the name of the image that was already saved in the block. This change made the whole block after the changed block become an invalid block. This can be seen in Figure. 4 which is a valid block with green color before any changed data and in Figure. 5 below which is the invalid block became red color with changed data from image 2 to image 3.

Figure 4: Valid Block Before Change Data.

Figure 5: Invalid Block After Change Data.

This qualitative test showed that our proposed model can protect the data well. It prevented other people to change the value inside. To make it a valid block it needed to mine block per block when on the other hand it seems like something impossible for the usual computer except supercomputer.

For ECDSA testing we try to change the hash value of the image that is already signed before, after that we try to validate the signed text with the public key from the first signed. The result shows that the public key can't validate the data because the data is different when it first signed. This means that our proposed model of ECDSA can protect the signed data well and prevent it from changing. The result of ECDSA testing can be seen in Figure 5.

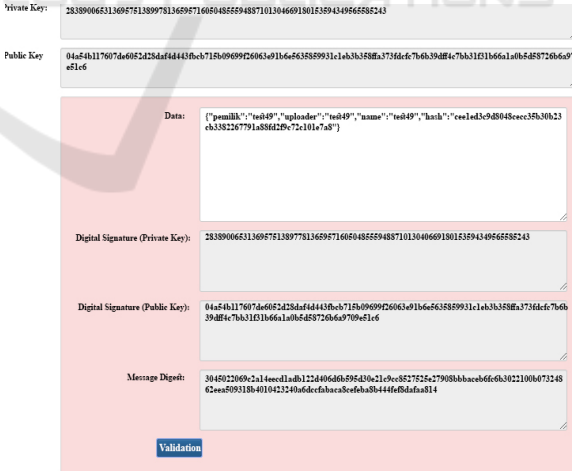

Figure 6: Fail Validation ECDSA.

For the quantitative test, we have tested the accuracy of DCT hash with 3 different scenarios. The first scenario is uploaded 200 images, the second scenario is uploaded 400 images, and the third scenario is uploaded 600 images. the

measurement of this quantitative test is using the confusion matrix with the operator

- True Positive (TP) for actually the plagiarism images and the DCT recognized it as a plagiarism image.
- True Negative (TN) for actually the not plagiarism images and the DCT recognized it as not plagiarism image.
- False Positive (FP) for actually the not plagiarism images and the DCT recognized it as a plagiarism image.
- False Negative (FN) for actually the plagiarism images and the DCT recognized it as a not plagiarism image.

For the quantitative test, we have tested the accuracy of DCT hash with 3 different scenarios. The first scenario is uploaded 200 images that consist of 50 original images that don't in the dataset before, 50 random images from the dataset and rotated it with 90°, 50 random images from the dataset and rotated it with 180°, and 50 random images from the dataset and rotated it with -90°. The detailed testing of the first scenario can be seen in Table 1 (Rc is Received and Rj is Rejected).

Table 1: First Scenario Testing Result.

Image type	Image qty	Result					
		Rc	Rj	TP	TN	FP	FN
Original Image	50	49	1	0	49	1	0
Rotated 90°	50	0	50	50	0	0	0
Rotated -90°	50	1	49	49	0	0	1
Rotated 180°	50	0	50	50	0	0	0
Total:	200			149	49	1	1

For the first scenario with a confusion matrix, it gained 49 TN and 149 TP. In percentage, it gained 99% accuracy this is because there is one image that not a plagiarism image, and the DCT hash detected it as a plagiarism image. On the other hand that is one plagiarism image and the DCT hash cannot detect it as a plagiarism image.

For the second scenario, the image type is the same as the previous scenario, but the difference is the amount of testing image. In this scenario, we use 100 images of each image type. The total image of this testing is 400 images. In this scenario, we get 100% accuracy with the confusion matrix. The result of this scenario can be seen in Table 2.

Table 2: Second Scenario Testing Result.

Image type	Image qty	Result					
		Rc	Rj	TP	TN	FP	FN
Original Image	100	100	0	0	100	0	0
Rotated 90°	100	0	100	100	0	0	0
Rotated -90°	100	0	100	100	0	0	0
Rotated 180°	100	0	100	100	0	0	0
Total:	400			300	100	0	0

For the second scenario with a confusion matrix, it gained 100 TN and 300 TP. In percentage, it gained 100% accuracy this is because all of the images detected like the images should be, plagiarism as plagiarism, and not plagiarism as a not plagiarism.

For the third scenario, the image type is the same as the previous scenario, but the difference is the amount of testing image. In this scenario, we use 150 images of each image type. The total image of this testing is 600 images. In this scenario, we get 100% accuracy with the confusion matrix. The result of this scenario can be seen in Table 3.

Table 3: Third Scenario Testing Result.

Image type	Image qty	Result					
		Rc	Rj	TP	TN	FP	FN
Original Image	150	150	0	0	150	0	0
Rotated 90°	150	0	150	150	0	0	0
Rotated -90°	150	0	150	150	0	0	0
Rotated 180°	150	0	150	150	0	0	0
Total:	600			450	150	0	0

For the third scenario with a confusion matrix, it gained 150 TN and 450 TP. In percentage, it gained 100% accuracy this is because all of the images detected like the images should be, plagiarism as plagiarism, and not plagiarism as a not plagiarism.

From three scenarios testing, we get the average accuracy of this model is

(99%+100%+100%) / 3 = 99,67%.

For the second quantitative test, we test the mining time of the Blockchain. Mining time was a time to use for mine or to find a fit public key for Blockchain private key. We tested mining time in second by using the number different of blocks from

10, 50, 500, 1000, 5000, 10000 blocks with different difficulty target (DT) from 1 to 5. The details of this testing can be seen in Table 4 below.

Table 4: Mining Time Testing Result.

Number of Blocks	Mining Time (Second)				
	DT1	DT2	DT3	DT4	DT5
10	0,2	0,6	4,7	11,341	28,578
50	0,8	3,0	6,9	15,321	39,465
500	33,0	60,8	133,8	289,561	621,59
1000	156,1	412,0	865,5	1832,5	4120,5
5000	2.962,6	7167,5	15012,3	37891,5	81952,91
10000	75643,5	160824,5	340156,8	721051,3	1591204,6

The result of this testing showed that the number of blocks and difficulty target is correlated to mining time. It meant that the more blocks and the higher the difficulty target you want to mine and use, the more time you need. For example, we can see the highest mining time to mining 10000 blocks with the difficulty target is 5, it needed about 1.591.204,671 seconds, or in the day it about 18 days more.

This result meant that Blockchain is a robust technology to protect the data that are already saved inside.

5 CONCLUSION AND FUTURE RESEARCH

Copyright is the same as giving a reward to the author for their works and idea, that why protecting the copyright of the image is important. Previously researched image copyright protection using perceptual hash and Blockchain can detect plagiarism when uploading images, but for some cases like the rotated image like 90 degrees. The proposed model with the implementation of DCT hash and Blockchain can protect the image copyright. The DCT hash can successfully detect the modification image especially the rotated image with an accuracy of 99,67%. The implementation of the Blockchain proved that is capable to protect the data that already in the block. With the mining time needed when in the largest difficulty target, it shows that it was nearly possible to change the data. For future research, it can be implemented using another algorithm that can detect the modification images easier and combined well with the implementation of Blockchain so the new model can more efficient in time consumption and more accurate to detect the modification image.

REFERENCES

Aghav, Sushila, et al. 2014. “Mitigation of Rotational Constraints in Image-Based Plagiarism Detection Using Perceptual Hash.” 2(1): 28–32.

Andi, Ronsen Purba, and Roni Yunis. 2019. “Application of Blockchain Technology to Prevent The Potential Of Plagiarism in Scientific Publication.” *Proceedings of 2019 4th International Conference on Informatics and Computing, ICIC 2019*.

Bhowmik, Deepayan, and Tian Feng. 2017. “The Multimedia Blockchain: A Distributed and Tamper-Proof Media Transaction Framework.” *International Conference on Digital Signal Processing, DSP* 2017-Augus.

Cho, Sunghyun, and Chiyoung Jeong. “A Blockchain for Media : Survey.” *2019 International Conference on Electronics, Information, and Communication (ICEIC)*: 1–2.

Drmic, Andrea, et al. 2017. “Evaluating the Robustness of Perceptual Image Hashing Algorithms.” : 995–1000.

Jnoub, Nour, and Wolfgang Klas. “Detection of Tampered Images Using Blockchain Technology.” *2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*: 70–73.

Kibet, Alex, Prof Simon, and Maina Karume. 2018. “A Synopsis of Blockchain Technology.” 7(11): 789–95.

Knirsch, Fabian, et al. 2018. “EVALUATION OF A BLOCKCHAIN-BASED PROOF-OF-POSSESSION IMPLEMENTATION Center for Secure Energy Informatics, Salzburg University of Applied Sciences, Puch / Salzburg , Austria.” 865082(865082).

Lin, Iuon Chang, and Tzu Chun Liao. 2017. “A Survey of Blockchain Security Issues and Challenges.” *International Journal of Network Security* 19(5): 653–59.

Mehta, Rishabh. “Decentralised Image Sharing and Copyright Protection Using Blockchain and Perceptual Hashes.” *2019 11th International Conference on Communication Systems & Networks (COMSNETS)* 2061: 1–6.

Ovhal, Prajakta Mahendra, et al. 2016. “Plagiarized Image Detection System Based on CBIR To Cite This Version : HAL Id : Hal-01284675 Plagiarized Image Detection System Based on CBIR.” 4(3).

Ravindran, M K, Balaji Zacharia, and Antony Roy. 2018. “Plagiarism and Copyright , Acknowledgements, Disclosure and Conflicts of Interest.” : 207–14.

Rivas, Alberto et al. 2017. “On Hashes Extraction.” 1(d): 87–94.

Wang, Huaimin et al. 2018. “Blockchain Challenges and Opportunities: A Survey.” *International Journal of Web and Grid Services* 14(4): 352.