

See discussions, stats, and author profiles for this publication at: [https://www.researchgate.net/publication/346898668](https://www.researchgate.net/publication/346898668_Plagiarism_Detection_of_Images?enrichId=rgreq-b9d980ddc24a81f0c2ef4257de43b4ac-XXX&enrichSource=Y292ZXJQYWdlOzM0Njg5ODY2ODtBUzoxMDc4MDk2NTM0OTk5MDQwQDE2MzQwNDk2NTI1NTU%3D&el=1_x_2&_esc=publicationCoverPdf)

[Plagiarism Detection of Images](https://www.researchgate.net/publication/346898668_Plagiarism_Detection_of_Images?enrichId=rgreq-b9d980ddc24a81f0c2ef4257de43b4ac-XXX&enrichSource=Y292ZXJQYWdlOzM0Njg5ODY2ODtBUzoxMDc4MDk2NTM0OTk5MDQwQDE2MzQwNDk2NTI1NTU%3D&el=1_x_3&_esc=publicationCoverPdf)

**Conference Paper** · September 2020

DOI: 10.1109/SCOReD50371.2020.9250940

CITATIONS

2

READS

2,016

**3 authors**, including:

[Othman Omran Khalifa](https://www.researchgate.net/profile/Othman-Khalifa?enrichId=rgreq-b9d980ddc24a81f0c2ef4257de43b4ac-XXX&enrichSource=Y292ZXJQYWdlOzM0Njg5ODY2ODtBUzoxMDc4MDk2NTM0OTk5MDQwQDE2MzQwNDk2NTI1NTU%3D&el=1_x_5&_esc=publicationCoverPdf)

[International Islamic University Malaysia](https://www.researchgate.net/institution/International-Islamic-University-Malaysia?enrichId=rgreq-b9d980ddc24a81f0c2ef4257de43b4ac-XXX&enrichSource=Y292ZXJQYWdlOzM0Njg5ODY2ODtBUzoxMDc4MDk2NTM0OTk5MDQwQDE2MzQwNDk2NTI1NTU%3D&el=1_x_6&_esc=publicationCoverPdf)

**501** PUBLICATIONS **3,436** CITATIONS

[Diaa Eldein Mustafa Ahmed](https://www.researchgate.net/profile/Diaa-Eldein-Ahmed?enrichId=rgreq-b9d980ddc24a81f0c2ef4257de43b4ac-XXX&enrichSource=Y292ZXJQYWdlOzM0Njg5ODY2ODtBUzoxMDc4MDk2NTM0OTk5MDQwQDE2MzQwNDk2NTI1NTU%3D&el=1_x_5&_esc=publicationCoverPdf)

University of Science and Technology (UST) - Sudan

**10** PUBLICATIONS **81** CITATIONS

[SEE PROFILE](https://www.researchgate.net/profile/Othman-Khalifa?enrichId=rgreq-b9d980ddc24a81f0c2ef4257de43b4ac-XXX&enrichSource=Y292ZXJQYWdlOzM0Njg5ODY2ODtBUzoxMDc4MDk2NTM0OTk5MDQwQDE2MzQwNDk2NTI1NTU%3D&el=1_x_7&_esc=publicationCoverPdf)

[SEE PROFILE](https://www.researchgate.net/profile/Diaa-Eldein-Ahmed?enrichId=rgreq-b9d980ddc24a81f0c2ef4257de43b4ac-XXX&enrichSource=Y292ZXJQYWdlOzM0Njg5ODY2ODtBUzoxMDc4MDk2NTM0OTk5MDQwQDE2MzQwNDk2NTI1NTU%3D&el=1_x_7&_esc=publicationCoverPdf)

**Some of the authors of this publication are also working on these related projects:**

Available Transfer Capability Determination hybrid wind power system usin Probabilistic Colocation Method [View project](https://www.researchgate.net/project/Available-Transfer-Capability-Determination-hybrid-wind-power-system-usin-Probabilistic-Colocation-Method?enrichId=rgreq-b9d980ddc24a81f0c2ef4257de43b4ac-XXX&enrichSource=Y292ZXJQYWdlOzM0Njg5ODY2ODtBUzoxMDc4MDk2NTM0OTk5MDQwQDE2MzQwNDk2NTI1NTU%3D&el=1_x_9&_esc=publicationCoverPdf)



Healthcare, Medical and Emergency [View project](https://www.researchgate.net/project/Healthcare-Medical-and-Emergency?enrichId=rgreq-b9d980ddc24a81f0c2ef4257de43b4ac-XXX&enrichSource=Y292ZXJQYWdlOzM0Njg5ODY2ODtBUzoxMDc4MDk2NTM0OTk5MDQwQDE2MzQwNDk2NTI1NTU%3D&el=1_x_9&_esc=publicationCoverPdf)

All content following this page was uploaded by [Diaa Eldein Mustafa Ahmed](https://www.researchgate.net/profile/Diaa-Eldein-Ahmed?enrichId=rgreq-b9d980ddc24a81f0c2ef4257de43b4ac-XXX&enrichSource=Y292ZXJQYWdlOzM0Njg5ODY2ODtBUzoxMDc4MDk2NTM0OTk5MDQwQDE2MzQwNDk2NTI1NTU%3D&el=1_x_10&_esc=publicationCoverPdf) on 12 October 2021.

The user has requested enhancement of the downloaded file.

Plagiarism Detection of Images

Amirul S. Bin Ibrahin Electrical and Computer Engineering,

International Islamic University Malaysia

Kual Lumpure, Malaysia

[Amirul.syahiran.ibrahim@gmail.com](mailto:Amirul.syahiran.ibrahim@gmail.com)

Othman O. Khalifa Electrical and Computer Engineering,

International Islamic University Malaysia

Kual Lumpure, Malaysia [khalifa@iium.edu.my](mailto:khalifa@iium.edu.my)

Diaa Eldein M. Ahmed Electrical and Computer Engineering,

Sudan University of Scienec and

Technology, Sudan [diamamed@gmail.com](mailto:diamamed@gmail.com)

***Abstract*—Plagiarism is when someone takes another author’s works, thoughts, ideas, etc. without proper referencing and claim it as his/her own works. Plagiarism detection is the process to find the plagiarism within a work or documents. With the advance of modern technology, it makes it easier for people to search for information and plagiarize the work of others. Although the effort and ideas for an image-based plagiarism detection has been increasing over the years, flaws are still presence in the current systems. This paper proposes a new system that can cover those flaws. It consists three stages: the pre-processing, feature extraction and comparison stage. The results showed in an ascending order of similarity index and true and false. However, the accuracy is 100% in case of unedited images and variated in other operations such as flipped, rotated, greyscales and cropped**

***Keywords—Image Plagiarism, Image Retrieval, Feature extraction.***

1. INTRODUCTION

Plagiarism basically means the wrongful stealing of an author’s work, thoughts, ideas, etc. and claiming it as your own original work. Plagiarism is considered as deceit and a breach of ethics. In academics, students that are caught with plagiarism are exposed to various level of penalties and punishment and may even lead to expulsion. Plagiarism in itself cannot be considered as a crime but as copyright violation. In the academics and other industries that are sensitive to copyright infringement, plagiarism is grave misconduct in integrity. The law cannot and usually will not punish plagiarism, but it is up to the institution on how to handle it once it happens [1]. Plagiarism detection is usually split into two which is text-based plagiarism detection and image-based plagiarism detection. For text-based plagiarism detection there are currently five techniques that is used most often in different fields. These techniques are Fingerprinting, String Matching, Bag of Words, Citation Analysis and Stylometry. String Matching is mostly used in computer science where it compares the documents words for words. Bag of words represents the documents in one or two vectors for comparison. Citation analysis is mainly used in scientific texts because it only compares the citation and reference of the documents. Stylometry check the author’s unique writing style for detection author’s ownership [2]. For image-based plagiarism detection, there is no commonly used techniques like the text-based plagiarism detection, but they usually share the same processes and steps. When we say plagiarism checking or detection we usually mean checking only the text in the file or document for plagiarism. Most of the times when you check your documents or files for plagiarism through a plagiarism checker software they will check for images and then discard them. This is one of the fatal flaws that the current system is facing. In the field of research, images and flowchart can carry vital piece of information that can easily be plagiarized if the flaw in the system is there.

1. PLAGIARISM

From the Oxford Dictionary, plagiarism means the act of taking someone else's work, ideas, thought, etc. and claiming it as your own work. The word plagiarism comes from the Latin word Plagiarus which means kidnapper, plunderer or seducer. The word Plagium which means kidnapper is derived from the word Plaga which means to capture or trap. Modern days the word plagiarism means to plagiarize. The process of checking a work or documents for plagiarism is called Plagiarism Detection or Checker [3].

The history of plagiarism first began from religious texts where most of them were authorless, so it is copied extensively and merged into later works. At the mid-1600, it is very common for there to be accusation of plagiarism for every creative field [3]. In the year 1709, the first copyright law was passed but it has more to do with protecting the publisher’s right than the author’s, but another law was passed soon after that to protect author’s right. James Boswell, who is also known as the biographer for Samuel Johnson, was a lawyer that opposed how long the copyright of the author lasted which at that time ended up to 21 years [4]. In the beginning of the 19th century, the laws for copyright is pretty like what we have today. The only difference is the issue of enforcing those laws across the borders. Most European country sign an agreement to prevent book piracy except for America which signs it at the year 1891 [4].

1. PLAGIARISM DETECTION TECHNIQUE

*A.Text-Based Plagiarism Detection Technique*

The main technique used for text-based plagiarism detection is, Fingerprinting, String Matching, Bag of Words, Citation Analysis and Stylometry. The most used technique is the Fingerprinting technique where the system will select a set of multiple substrings from the documents and the sets signifies the fingerprints which is made up of the elements called minutae. The plagiarism checking is done by taking the fingerprints of the documents and comparing them to a precomputed index of reference collection [2].

String matching is mainly used in the computer science field where the system will compare word for word on each document. This system detect plagiarism in a pair with the original documents and with the collection of references. Although plentiful methods have been proposed but this system is still computationally expensive making it unfit for large number of documents. For Bag of Word technique, it used a vector space retrieval representation where the documents are represented with one or two vectors to be used for similarity comparison. The system can use the regular cosine similarity measure or further advance similarity comparison technique. Citation analysis more widely used for checking plagiarism in scientific text because

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

978-1-7281-9317-5/20/$31.00 ©2020 IEEE

183

it is the only technique that does not use textual analysis but examines the citation and references of the documents to recognize comparable pattern. This technique is still considerably new, so it is not ready for commercial used yet [2]. The last technique, Stylometry detects plagiarism by checking the writer’s unique writing style so it is more widely used for checking the original owner attribute. The system compares the stylometric models for different text segment that are stylistically different from others.

*B. Image-Based Plagiarism Detection Technique*

Image-based plagiarism detection is less used compared to text-based plagiarism detection, so it does not have a widely popular technique that is used everywhere. Instead, they are still researching for a good method to detect plagiarism in images. Here we will discuss a few of the methods that was proposed every year.

Method by Popescu and Farid [5] proposed a system by using the Principal Part Analysis (PCA). This system divides the numerous tiny sized blocks into vectors which is then organized lexicographically before matching them. The main drawback for this system is that if the image quality is too low then the accuracy will fall as well.

Mahdian and Saic [6] apply a blur movement variant to signify the image region so that the images will not be degraded from blurring and noises. This system begins by tilting the image with selected size blocks and defined it with blur invariants. The drawback from this method is that the computation time for this system is very long.

An experiment conducted by Wang et al [7] used a copy-move plagiarism detection system by applying the victimization Hu moments to cut down the computation time of the system. This system divides the image into numerous sized blocks and then applying the Hu moments on the block they computed the Eigen value.

A more enhanced method has been proposed by Zimba and Xingming [8] for the copy-move detection system. The system starts by converting the image into a gray scale image and then applying DWT. The image is then divided into overlapping blocks and then PCA is done to each block. This method cuts down the computation time of the system compare to the PCA method by reducing the size of the image. Bravo-Solorio and Nandi [9] proposed a system to detect reflection, scaling and rotation of an images. The drawback is that their methods produced a lot of matches which needed to be further improve. A system that had been proposed by Sridevi et al [10] used the copy-move detection system in parallel which makes the system unable to used methods that has a requirement of long computation time. The only disadvantage with this system is that it cannot detect colored images.

works by comparing the shape of the query image with those in the database as shown in Figure 1.

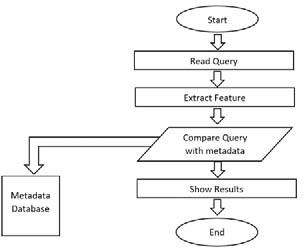
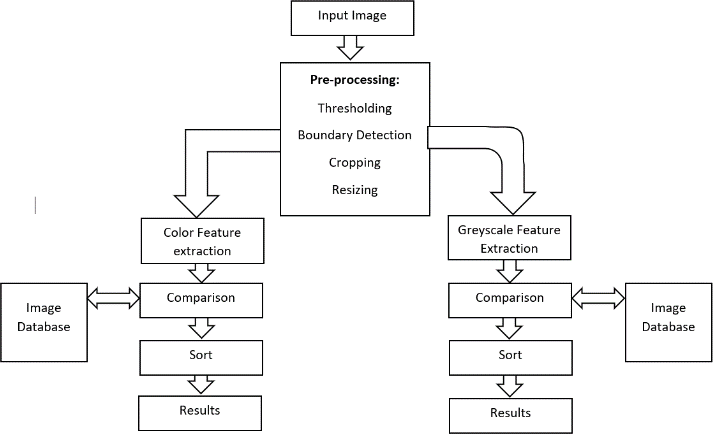


Figure 1: Flowchart of shape-based flowchart detection [11]

This method begins by pre-processing the flowchart to get maximum result, reducing error and better accuracy in plagiarism detection. The pre-processing has 3 stages which is thinning the images, removing connected lines and removing texts. The database consists of 2 sets. The 1st set is for storing figures and the 2nd set is for storing preprocess figure. Each of these figures has an ID which makes it easier to recall in the comparison stages. The database generally only consists of 4 basic shapes in the flowchart which is circle (connector), square (process), diamond (decision) and oval (start/stop). The next stage is the metadata extraction. Here we extract information from the figures in the database into a metadata table to speed up the process of the comparison. Each metadata entries contain the number of shapes that a figure in the database has. The technique used to determine the shaped of an object is the area detection technique where we measure the distance between the centers of the object to its boundaries. Using the distance obtained and inserting it into a mathematical formula the corresponding shape can be obtained.

*B. Content Based Image Retrieval Technique (Cbir)*

This system was proposed by Prajakta Mahendra Ovhal [12] which used the CBIR technique. In CBIR, the features of the image such as shape, and color is extracted and then compared to the feature that was saved in the database.

1. RELATED WORK

The technique and system for image-based plagiarism detection system is always improving. Below are some of the recently proposed system from around the world that can enhanced the current image-based plagiarism detection technique.

*A. Shape Based Flowchart Detection*

Based on the research article by Senosy Arrish et al [11] this system focuses entirely on flowchart images. This system

Figure IV: Flowchart of CBIR [12]

This system mainly consists of 3 stages which are the pre-process stage, feature extraction stage and comparison stage. In the pre-process stage, the images must undergo gray scaling, thresholding, boundary detection and cropping to make it more readable and freer from error. Next getting the image boundary, the original image is crop with

the image boundary which will only leave us with the shape of the original image. Everything else is discarded. Cropping is done to make the image more compact and thus processing it requires less time increasing the system efficiency.

The feature extraction stage has 2 parts which is color feature extraction and shape feature extraction. Color feature extraction is done in 2 ways by describing the color space and the color descriptor. This system uses the HSV color space which is more suited to represent computer graphic. Shape is a boundary that gives an object its form. Some of the techniques used for shape feature extraction are Mass, Centroid, Mean, Variance, Aspect ratio and Moment invariant. In the comparison stage, the system uses the Euclidean distance formula to calculate the similarity between 2 images. Once the comparison between the images is done, the results that will be displayed to the user will be sorted using merge sort in decreasing order of similarity.

1. *Perceptual Hash Technique*

Based on the system proposed by Vipul Bajaj et al [13] which used the perceptual hash technique. This technique calculates the hash value of an image and compare it to the database using a distance function formula.

This system has 3 main functions, first starts by decreasing the size of the image to simplify the calculation. After that is to convert the image into gray scale image for further simplification. Then it computes the discrete cosine transform (DCT) which is used to separate the images into frequency and scalar. For this system it creates a 32x32 DCT and the next step reduce that DCT into 8x8 DCT. Next is to compute the mean DCT where it completely excludes solid colors from the hash description. The system further reduces the DCT size by setting the 64(8x8) hash bit into 0 or 1 for each bit depending on whether the bit is below or above the average value. Lastly by setting the 64 bits into 64-bit integer the PHash value for the image is constructed.

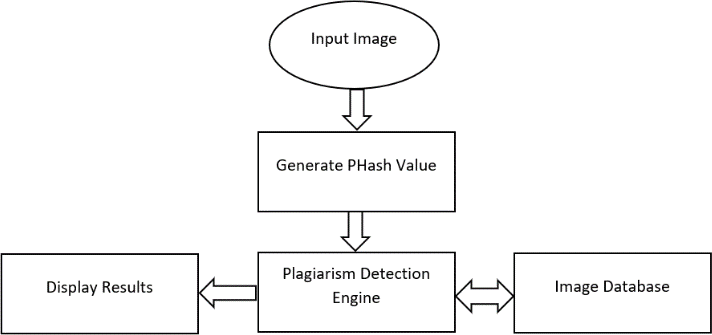


Figure 3: Flowchart of Perceptual Hash [13]

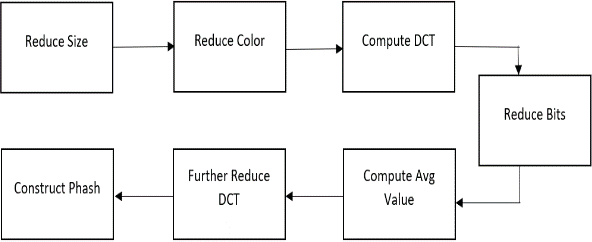


Figure 4: Workflow of Phash Generation [13]

1. *Hierarchal Near Duplicate Retrieval*

Based on the system proposed by Siddharth Srivastava et al [14], this system is more suited for natural images compared to other classes of image.

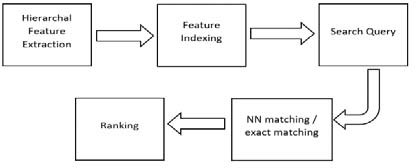


Figure 5: Flowchart of Hierarchal Near Duplicate [14]

Here, the proposed system is very similar to the previous technique in that it also uses the Perceptual Hash for feature detection but what differ this system with the previous one is that this system also uses the SIFT technique. This system first constructs and then stores the Phash value for all the images in the database. After that is to calculate and store the SIFT features and change the images into a bag of visual word representation and store the resultant word and histogram into the database using the LSH. For the displayed results, the system calculates in the distance between histogram instead of exact matching of key feature.

1. *Image Processing Approeach For Flowchart*

Based on the article by Jithin S Kuruvila et al [15]. This system converts the flowchart into a directed graph which makes it able to detect the shape of the flowchart even if the positioning has been changed.

This system consists of 4 main parts which is the pre- processing, shape detection, graph creation and comparison module. For the pre-processing part, it starts by converting the input image into a binary image and then it is applied with Canny Edge Detection algorithm to find the edges of the image. The Canny Edge Detection is used which is more stable and accurate in its result.

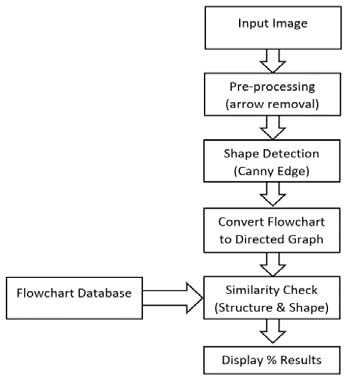


Figure 6: Flowchart for Image Processing Approach [15]

Next is to find the intensity gradient of the image and then to apply edge thinning technique to the image for getting rid of unwanted response to the edge detection. After that to find the potential edges, double threshold is applied and is then finalize by suppressing all the weak edges and edged not connected to any strong edges. Next in the pre- processing step is to remove the arrows in the flowchart.

1. *Artificial Neural Network (ANN)*

Based on the paper by Behnam Hadi and Kargar [16]. This system used the ANN to compare the similarity rate of a paper.

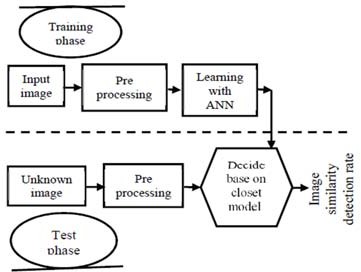


Figure 8: Flowchart for Artificial Neural Network [16]

This system is split into 2 phases which is the training phase and the testing phase. The training phase consists of the feature interpretation step and the learning step. In the feature interpretation step the first thing the system does is identify whether the images are a flowchart or a non-flowchart.

For the testing phase, the system will identify whether the images are a flowchart or non-flowchart. The system will ignore and discard images that are not flowchart. For the flowchart images, the system first pre-processes them to remove noise and then the feature extraction is done. In the pre-processing step, the system does thinning operation, deleting the texts, equalize the image histogram, and image resizing. After the feature extraction is done then the system will do the plagiarism check by comparing it with the closest created model in the learning step of the training phase.

1. PROPOSED SOLUTION AND IMPLEMENTATION

The methodology of this work t is divided into 2 parts as shown in figure 9. This work stars with the training phase where the pre-processing the images, extracting the image features and then save the extracted feature in the database. This step is repeated until enough images in the database has been reached. The next part of the project is the testing phase where we edited some of the images in the database using some common technique used to pass through image detection software, pre-process the image, extract the image feature, compare it to the database and then display the results

in a descending order of similarity. In the pre-processing stage we extract both the global and local feature of the image while for feature extraction we extract the images color, texture and shape feature and then the comparison is done using the Euclidean Distance formula.

For the input images, the only restriction is that the dimensions of the images must be either be 364x236 or 236x364. The software will not accept any images outside of that dimension and will gives an error. Besides that, the user can freely input any images they want but for the purpose of testing the software in detecting similarities, we only picked images that have already been in the database as sample images and then edited them slightly to test the accuracy of the software.

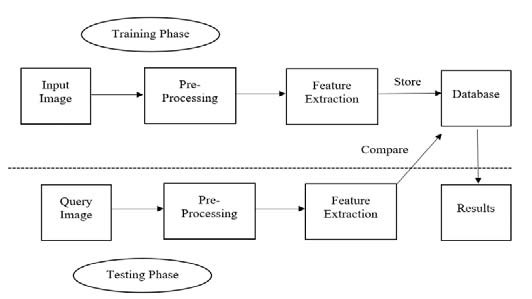
The database used in this work is gotten online where all the images of the database is used specifically to test image retrieval software. The database contains varied categories of images starting from buildings, animals, vehicle, plants, objects and sceneries. Each of these categories contain around 100 images each so the size of the database can be considered big to test the accuracy of the software.

1. *Pre-Processing*

The processing stage is used to remove the noise contain in the input image so that the feature extraction can be done faster, smoother or more accurate. Features are sometimes called descriptor. The global descriptor described the color and texture features of the entire image. To extract the global feature of an image we used the color, texture and shape descriptor of the project. The texture feature is not always precise since it is calculated from the entire image. The local descriptor focuses more on a certain region or key points of an image. To extract the local feature of the images, this work used the Bag-Of-Visual-Words (BOVW) which is inspired by the bag-of-words model used for text-based plagiarism checker. Then, find similar images using the frequency histogram.

1. *Feature Extraction*

The features extracted from images and saving it to the databases is color using the RGB and HSV color space, texture using Tamura texture and then shape using canny edge algorithm.

Figure 9: Block Diagram of the proposed System

1. RESULT ANALYSIS

The results are divided into two parts. The first part shows the development of the GUI of the project which shows some feature extraction options that the user can enable or disable, and the second part shows the results obtained from editing the images. For a CBIR image plagiarism checker using Matlab, one of the most important aspect is the graphic user interface (GUI) as shown in figure 10 &11. The GUI enable easy selection of the query image, the ability to enable and disable some feature extraction process and being able to show the image processing results in an easy to understand manner. The GUI also displays the chosen query image with a slider under it to choose how many results that wants to be shown. There is a progress bar to show the user how much longer will it take to finish the process and a display to show how long the process takes to finish. The results are displayed in an ascending order of similarity index. The smaller the similarity index the more similar the query image is to the result image as shown in Table 1.

1. *Query Image Results*

These are the results obtained from the software after the sample query images was edited using common techniques to bypass plagiarism checker which is flipping, rotating, cropping and gray scaling. The results shown in the GUI for each image while the table will show the similarity index of up to 10 results and whether it is the original image or not. This part will show the similarity index results of the original image before editing them.

1. *Results Accuracy*

The table 2 below will show the accuracy of the software where we divide the number of sample image that got their original image in the first results with total number of sample images used.

Table 2: Accuracy Image Results

|  |  |
| --- | --- |
| **Editing Technique** | Accuracy (%) |
| Unedited | 100 |
| Flipped | 100 |
| Cropped | 60 |
| Rotated | 80 |
| Gray Scale | 20 |

1. *Result Discussion*

There are 3 main feature extraction in this work which is color, texture and shape feature. The user can freely disable or enable one or two of the feature extractions to exclude the feature from being extracted in the query image. Changing the feature extraction option will cause variation in the results but for this project’s results, all of the feature options were enabled. In the color feature extraction, this project used the RGB and HSV color space to get the color histogram of the image. The texture feature extraction used the Tamura texture feature because it extracts features that are close to how human detect texture. For the shape feature extraction, the canny edged algorithm is used to get the edges of the image and store them. Euclidean formula is used in order to calculate how similar the query image is with those in the database. The software then displays results with the smallest similarity index first and then show the other results with an ascending order of similarity index. The results evaluation is based on how close the original image is to be the 1st results after the sample images have been through some editing. Table 1 shows the similarity index of the sample images without any editing done to them. As you can see, all of the 1st results show the original image, so it indicates that the software is working properly when no editing is done to them. At table 2, the results of the flipped sample images show that the original image is still at the 1st results which means that this project can detect flipped image without problems. In table 4.3, the sample images were cropped so that it only shows the important part of the image and the results display that only 3 of the sample images have their original image as the 1st result. Table 4.4 shows the result of rotating the sample images by 90 degree, which only 4 of the rotated images shows the original image at the 1st result while the other rotated image’s original image is in the 2nd result. Rotation of the sample images at different degree will show different result but for this project we will only show the results for 90- degree rotation. The last edited sample image was by gray scaling the sample images, which only 1 of the sample images shows the original image at the 1st results. The gray scale image shows the poorest accuracy in detecting the original image when comparing to the other edited image.

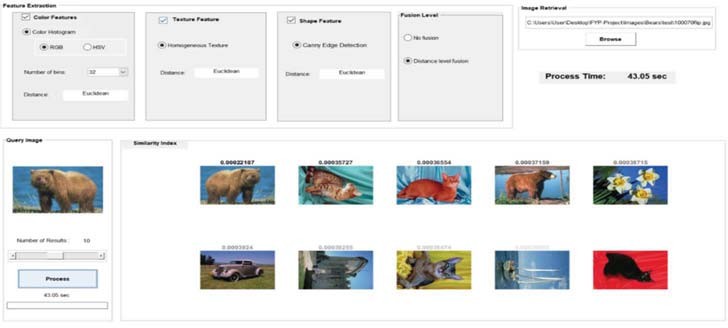
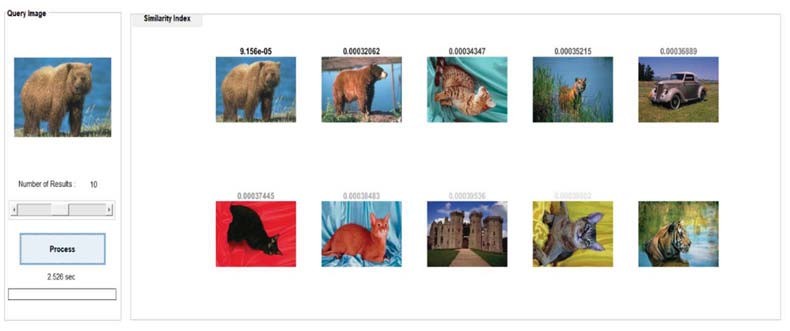


Figure 10: Graphical User Interface (GUI)

Figure 11: Database Image Query Table 1: Database Image Results

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Query Images | 1st Result Similarity Index (Original Image) | 2nd Result Similarity Index (Original Image) | 3rd Result Similarity Index (Original Image) | 4th Result Similarity Index (Original Image) | 5th Result Similarity Index (Original Image) |
|  | 9.156e-5 (True) | 0.00032062 (False) | 0.00034347  (False) | 0.00035215  (False) | 0.00036889  (False) |
|  | 8.1426e-5  (True) | 0.00026695  (False) | 0.0026928  (False) | 0.00030176  (False) | 0.0003082  (False) |
|  | 5.1257e-5  (True) | 0.00027858  (False) | 0.00029516  (False) | 0.00029737  (False) | 0.0003092  (False) |
|  | 9.1346e-5  (True) | 0.00026941  (False) | 0.00031308  (False) | 0.00032041  (False) | 0.00032461  (False) |
|  | 6.6255e-5  (True) | 0.00027033  (False) | 0.0003185  (False) | 0.00032391  (False) | 0.00032708  (False) |

1. CONCLUSION

Plagiarism detection is a well-known phenomenon in the academic arena. Copying other people is considered as a serious offence that needs to be checked. In this paper, an enhanced system to detect the Plagiarism of Images was proposed. The feature extracted from images and saving it to the databases is color using the RGB and HSV color space, texture using Tamura texture and then shape using canny edge algorithm. The results showed in an ascending order of similarity index and true and false. However, the accuracy is 100% in case of unedited images and variated in other operations such as flipped, rotated, greyscales and cropped.

**References**

[1]. Green, Stuart P, Plagiarism, Norms, and the Limits of Theft Law: Some Observations on the Use of Criminal Sanctions in Enforcing Intellectual Property Rights, *Hastings Law Journal*, 2002, 54 (1)

[2]. How Plagiarism Detection Works. (2016, May 18). Retrieved April 17,

2018, from https://[www.plagiarismtoday.com/2016/05/03/plagiarism-](http://www.plagiarismtoday.com/2016/05/03/plagiarism-) detection-works/

[3]. Vinod K.R.\*, Sandhya.S, Sathish Kumar D, Harani A, David Banji and Otilia JF Banji, Plagiarism – History, Prevention and Detection, *journal for drugs and medicines,* 2011, 3(1), 1-4

[4]. Eshgh, A. (n.d.). Copyright Timeline: A History of Copyright in the United States. Retrieved April 17, 2018, from <http://www.arl.org/focus-areas/copyright-ip/2486-copyright-timeline>

[5]. AC Popescu and H Farid, Exposing Digital Forgeries by Detecting Duplicated Image Regions, Dept Computer Science, Dartmouth College, *Hanover,* 2004, 515.

[6]. B Mahdian and S Saic, Detection of Copy–Move Plagiarism using a Method based on Blur Moment Invariants, *Forensic Science International*, 2007, 171(2), 180-189.

[7]. JW Wang, GJ Liu, Z Zhang, Y Dai and Z Wang, Fast and robust forensics for image region-duplication Plagiarism, *Acta Automatica Sinica*, 2009, 35(12), 1488-1495.

[8]. M Zimba, and S Xingming, DWT-PCA (EVD) Based Copy-move Image Plagiarism Detection, *International Journal of Digital Content Technology and its Applications*, 2011, 5(1), 251-258.

[9]. S Bravo Solorio and AK Nandi, Automated Detection and Localisation of Duplicated Regions Affected by Reflection, Rotation and Scaling, *Image Forensics Signal Processing*, 2011, 91(8), 1759-1770.

[10]. M Sridevi, C Mala and S Sandeep, Copy–Move Image Plagiarism Detection, *Journal of Computer Science and Information Technology*, 2012, 52, 19-29.

[11]. Senosy Arrish, Fadhil Noer Afif, Ahmadu Maidorawa and Naomie Salim, Shape-Based Plagiarism Detection for Flowchart Figures in Texts, *International Journal of Computer Science & Information Technology (IJCSIT)*, 2014, 6(1).

[12]. Prajakta Ovhal, Detecting Plagiarism in Images, *International Conference on Information Processing (ICIP)*, 2015.

[13]. Vipul Bajaj, Sanket Keluskar, Ravi Jaisawal and Prof. Rupali Sawant, Plagiarism Detection of Images*, International Journal of Innovative and Emerging Research in Engineering*, 2015, 2(2).

[14]. Siddharth Srivastava, Prerana Mukherjee and Brejesh Lall, imPlag: Detecting Image Plagiarism Using Hierarchical Near Duplicate Retrieval, *IEEE INDICON,* 2015*.*

[15]. Jithin S Kuruvila, Midhun Lal V L, Rejin Roy, Tomin Baby, Sangeetha Jamal, Sherly K K\*,

[16]. Flowchart Plagiarism Detection System: An Image Processing Approach, *International Conference on Advances in Computing & Communications,* 2017.

188

[View public](https://www.researchgate.net/publication/346898668)A[at](https://www.researchgate.net/publication/346898668)u[ion](https://www.researchgate.net/publication/346898668)th[sta](https://www.researchgate.net/publication/346898668)o[ts](https://www.researchgate.net/publication/346898668)rized licensed use limited to: Robert Gordon University. Downloaded on May 27,2021 at 04:33:11 UTC from IEEE Xplore. Restrictions apply.