RESEARCH ARTICLE OPEN ACCESS

Advanced Image Pattern Searching Duplicate Records Using Pixel Matching

J. Sunitha John [1] R.J. Pragathi [2]

[1] Assistant Professor Computer Science Department, Mangayarkarasi College of Arts & science for women, Madurai, Tamil Nadu - India.

[2] Student, PG final Year, Mangayarkarasi College of Arts & science for women, Madurai, Tamil Nadu - India.

ABSTRACT

With many potential practical applications, content-based image retrieval (CBIR) has attracted substantial attention during the past few years. Hence the authors implemented an advanced image pattern searching algorithm. In the proposed system, the images are searched based on the image patterns such as pixels, resolution, height and width of the image. To carry out the image searching process semi supervising algorithm is used. In order to carry out the search among multiple images, the parameters such as pixel height and width, resolution of the input image was compared with the images in the database. If there is a match between the two images, the matched image will be extracted and displayed as output. On selecting this extracted image one can add hidden text character, gray scale and color change property.

Keywords: Semi supervised Algorithm, Image matching algorithm, merging image with text.

I. INTRODUCTION

The main objective of the paper is to carry out image pattern matching, to extract duplicate records using pixel matching algorithm. It plays a major role in the wide area of image processing. Various algorithms have been proposed to identify the duplicate copies of the images. But even then there exists various drawbacks in carrying out the process. On considering all these, the authors proposed an approach based on pixel matching, following a semi supervised learning approach.

Semi-supervised learning is a learning paradigm concerned with the study of computers and natural systems such as humans learning pattern in the presence of both labeled and unlabeled data. During ancient days, learning can be carried out with supervised or unsupervised methods. The main goal of the authors was to analyze the performance of their system by combining both supervised and unsupervised methods. It plays a major role in data mining and machine learning. It has the capability to use the unlabeled data to improve the supervised learning approach. The proposed approach using this procedure provides better result in finding duplicate images.

II. LITERATURE REVIEW

In 2004**,** the authors have examined different types of imageforgery and their detection techniques. They mainly focused onpixel based image forgery detection techniques. All the approaches and the methodologies discussed in this paper have the capacity to recognize fraud, but few algorithms were not viable regarding identifying actual forged region. Some algorithms have a time complexityproblem. So, there is a need to develop an effective and accurate image forgery detection algorithm.

In 2017, K. K. Thyagharajan made a review on the computer vision-based approaches and feature extraction methods for the detection of near duplicate images. They also discussed the main challenges in that field and how other researchers addressed those challenges. Their review would provide research directions to the fellow researchers who are interested to work in this field.

In 2017, Zhili Zhou, et al, proposed a fast and accurate near-duplicate elimination approach for visual sensor networks. First, a coarse-to-fine clusteringmethod based on a combination of global feature and local feature was proposed to cluster near-duplicate images.Then in each near-duplicate group, they adopted PageRank algorithm to analyze the contextual relevance among images toselect and reserve seed image and remove the others. Their experimental results proved that their proposed approachachieved better performance in the aspects

of both efficiency and accuracy compared with the state-of-the-artapproaches.

In 2010 ,Bedade Jayshree, et al, proposed a method to extract patches from the given image and represented them by changeable length signature.The signature was further verified and validated in a near duplicate natural image detection which made a decisionabout whether two images are near duplicates or not. The near duplicate image retrieval aims at retrieving relevantor same images from image database which are similar to query image. The similarity between two images wasfound by using Earth Move Distance Algorithm.

In 2010, Xunyu Pan, et al, described a new region duplication detection method that is robust to distortions of the duplicated regions. Their method starts by estimatingthe transform between matched scale invariant feature transform (SIFT) key points, which are insensitive to geometrical and illuminationistortions, and then finds all pixels within the duplicatedregions after discounting the estimated transforms. Their proposed method showed effective detection on an automatically synthesizedforgery image database with duplicated and distorted regions. They further demonstrated its practical performance with several challengingforgery images created with state-of-the-art tools.

In 2012, Xunyu Pan, et al**,** described a new detection method based on matching imageSIFT features. The robustness of the SIFT features withregards to local transforms renders that this method was able to detectgeneral region duplications with efficient computation. Theeffectiveness of their method was demonstrated with experimentalresults, both qualitatively and quantitatively in terms of thedetection accuracy and the false positive rate.

In 2020, Jiayi Ma, et al**,** conducted a systematic review and analysis forclassical and latest techniques. Following the feature-based image matching pipeline, they first introduced feature detection,description, and matching techniques from handcrafted methods to trainable ones and provided an analysis of the developmentof these methods in theory and practice. Secondly, they briefly introduced several typical image matching-based applicationsfor a comprehensive understanding of the significance of image matching. In addition, they also provided a comprehensiveand objective comparison of these classical and latest techniques through extensive experiments on representative datasets. Finally, they concluded that the current status of image matching technologies delivered insightful discussions and prospects for future works. Their survey can serve as a reference for researchers and engineers in image matchingand related fields.

In 2013, MohdDilshad Ansari, et al**,** discussed about various pixel-based techniquesfor image forgery detection, mainly copy-move and splicing techniques.All the methods and approaches discussed in thispaper are able to detect forgery. But some algorithmsare not effective in terms of detecting actual forgedregion.

In 2017, Diaa M. Uliyan, et al, made astudyon image region duplication forgery detection algorithm based on the angular radial partitioning and Harris key-points. Two standard databases have beenused: image data manipulation and MICC-F220 for experimentation. Their experiment results demonstrated that the proposedtechnique can detect rotated regions in multiples of 30 degrees and can detect region duplicationwith different scaling factors from 0.8, to 1.2. Experimental results are presented to confirm theeffectiveness of detecting region duplication that has undergone other changes, such as Gaussiannoise, and JPEG compression.

In 2016, Li Liu, et al, proposed a variable-length signature fornear-duplicate image matching. An image is representedby a signature, the length of which varies with respect tothe number of patches in the image. A new visual descriptor, probabilistic center-symmetric local binary pattern, was proposedto characterize the appearance of each image patch. Beyond eachindividual patch, the spatial relationships among the patches were captured. In order to compute the similarity between twoimages, they utilized the earth mover's distance which is goodat handling variable-length signatures. The proposed image signature is evaluated in two different applications, i.e., near duplicate document image retrieval and near-duplicate natural image detection. Their promising experimental results demonstrated the validity and effectiveness of the proposed variable-length signature.

III. EXISTING METHOD

In the existing system, images are searched by the text or by the image which is given in the input field. Based on the image or text which is given in the input field, images are displayed in the output. Searching was carried out based on text search and the images were listed out. The approach was very slow, time consuming and less efficient. On considering all these issues, the authors proposed a new method based on pixel matching.

IV. PROPOSED METHOD

In order to overcome the issues of the existing method, the proposed method was organized to provide high efficiency. Pixel based searching algorithm was implemented that considers the resolution, pixel height and width. The approach follows a semi supervised learning algorithm. Once the image was extracted, certain property such as, text hiding, gray scale and color changing property was imposed on the image to differentiate the duplicate property of the image. The proposed approach was found to be highly efficient, fast and accurate.

Various steps involved in the proposed method was as follows,

- In the first step, the image to be searched was fed as input. Multiple images can be added to the database to carryout the searching process.
- In the second step, pattern matching is carried out with the semi supervised matching algorithm. The corresponding image was searched and displayed as output. The algorithm performs the comparison of the input image and the image stored in the database. The comparison was based on the image resolution, height and width of the pixels.
- In the third step, replacement of the extracted image was carried out by adding text hiding property to the image. To the extracted image, some hidden text or color changing property was imposed on the image and transferred to the server.

V. RESULTS AND DISCUSSION

The proposed method was found to be very effective in extracting the duplicate image and merging it with the image.

As a first step, the image to be searched was fed as input into the system. It was compared with the image in the database and extracted as shown in figure 1.

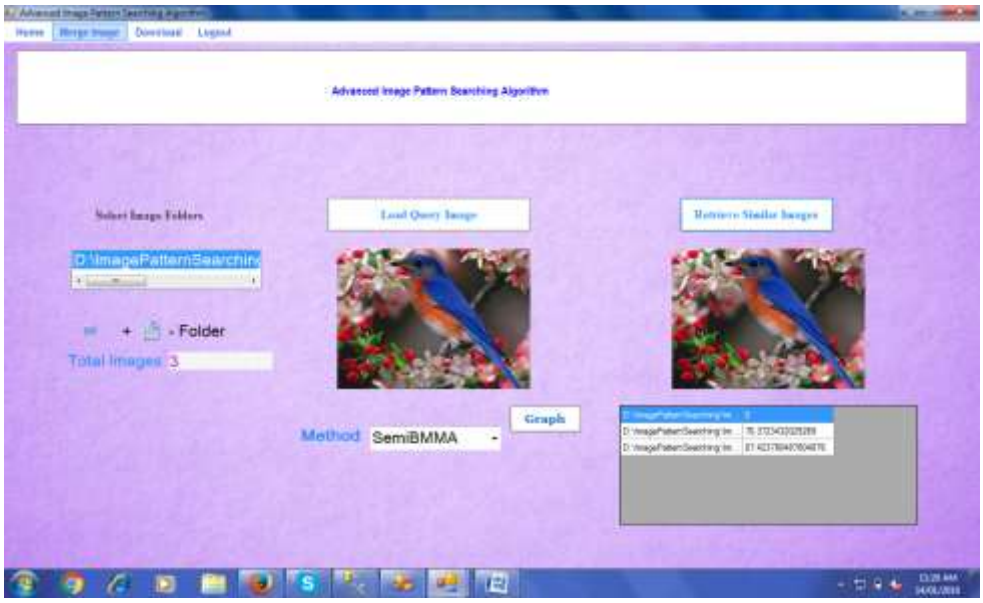

Figure 1: Input image to be searched and the extracted image from the database.

Merging operation is then carried over the extracted image using, some images, and text and color property and is shown in figure 2. It is then transferred to the server. The merged file can then be downloaded and viewed using the secret key.

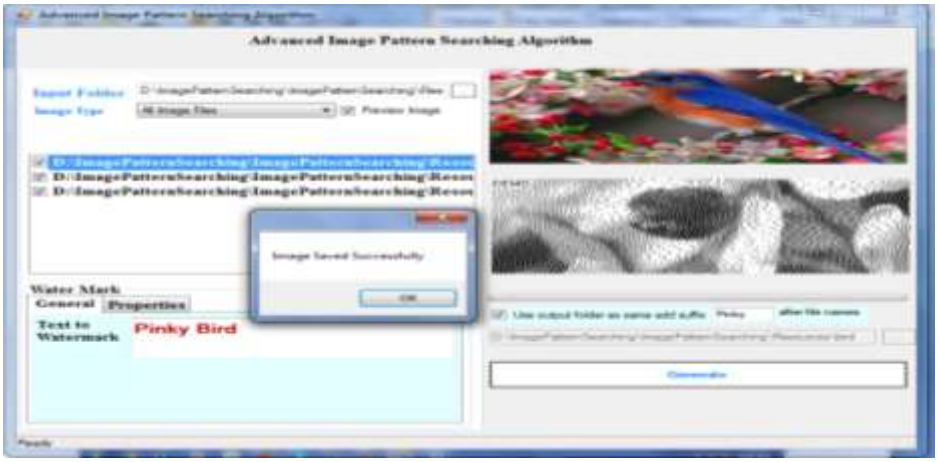

Figure 2: Merged image with text and image using pattern matching algorithm.

The merged file can then be downloaded and viewed using the secret key and is shown in figure 3.

Figure 3: The downloaded image with secret key.

The property of the input image and the merged image was as shown in the below figure 4.

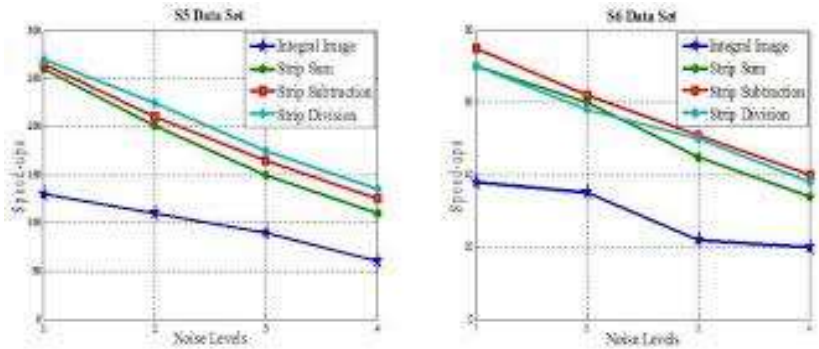

Figure 4: Graph showing the change in deviation from the input image and the modified image.

VI. CONCLUSION

The proposed method was found to be very effective in extracting the duplicate image and merging it with color property, text and other images. The semi supervised learning algorithm was very efficient in finding the duplicate image from the given dataset. As a future work, the proposed approach can be extended and implemented in Android app.

REFERENCES

[1] H. Farid, "A survey of image forgery detection," IEEE Signal Process. Mag., Vol. 26, no. 2, pp. 16-25, Mar. 2009

[2] R. E. J. Granty, T. S. Aditya, and S. Madhu, "Survey on passive methods of image tampering detection," IEEE International Conference on Communication and Computational Intelligence, in (INCOCCI), 2010, pp. 431-436.

[3] H. Huang, W. Guo, and Y. Zhang, "Detection of copy-move forgery in digital images using SIFT algorithm," Pacific-Asia Workshop on Computational Intelligence and Industrial Application, Vol. 2, pp. 272-6, Dec. 2008.

[4]. Z. Qu, and G. Qiu, "Detect digital image splicing with visual cues," Lect. Notes Comput. Sci., Vol. 5806, pp. 247-326, Jan. 2009.

[5] X. Kang, and S. Wei, "Identifying tampered regions using singular value decomposition in digital image forensics," International Conference on Computer Science and Software Engineering, 2008, Vol. 3, pp. 926-930.

[6] S. D. Lin et al., "An integrated technique for splicing and copymove forgery image detection," IEEE 4th International Congress on Image and Signal Processing (CISP), 2011, Vol. 2, pp. 1086-1090.

[7] G. H. Li, Q. Wu, D. Tu, and S. J. Sun, "A sorted neighborhood approach for detecting duplicated regions in image forgeries based on DWT and SVD," Proceedings of IEEE International Conference on Multimedia and Expo, Beijing, Jul. 2007, pp. 1750-1753.

[8] J. A. Redi, W. Taktak, and J.-L.Dugelay, "Digital image forensics: A booklet for beginners," Multimedia Tool Appl., Vol. 51, no. 1, pp. 133-62, Jan. 2011.

[9]] V. Tyagi, "Detection of forgery in images stored in digital form," Project report submitted to DRDO, New Delhi, 2010.

[10] J. Wang, G. Liu, Z. Zhang, Y. Dai, and Z. Wang, "Fast and robust forensics for image region-duplication forgery," ActaAutomaticaSinica, Vol. 35, no. 12, pp. 1488-95, Dec. 2009.