Image Forgery Detection Using Machine Learning

Y. Harshitha, V. Brahmeswari, T. Siva Parvathi, Y. Wishwanth
Under The Esteemed Guidance Of S.Anil Kumar M. Tech,(Ph.D) Asst. Prof
Department Of Computer Science And Engineering ,
Bachelor Of Technology
Tirumala Engineering College
Jonnalagadda, Narasaraopet, Guntur (Dt.), A.P.

ABSTRACT

Digital Image Forgery can be done by deceiving the digital image to mask some meaningful or important data of the image. It is usually difficult to spot out the manipulated region of the original image. To sustain the uprightness and legitimacy of the image, the detection of forgery in the image is mandating. Acclimation of the modern way of life and advancement in photography gadgetry has made exploitation of digital image easy with the help of image editing software. Therefore, it is crucial to detect such image forgery operations in the images. The image forgery detection can be done based on object removal, object addition, unusual size modifications in the image. Images are one of the powerful media for communication. In this project we have used algorithms such as Copy move, Canny Edge Defection, Structure Similarity index, Hierarchical Agglomerate Clustering and Neural Network algorithms with improved accuracy rate.

INTRODUCTION

Introduction to project

Forgery is an illegal means of manipulating images or documents without prior access. Images are tampered for different reasons either to create false evidence or to earn money in an illegal way. An pictorial representation of image conveys much better idea than the words of human. Due to the progression in digital technology, images are processed using several tools like Adobe Photoshop, GIMP and Corel Paint Shop and they ended up with a threat for the authenticity of digital images. Generally, image manipulations are of two types

a) Allowed manipulation

b) Malignant manipulation.

Allowed or incidental manipulations are the ones which never alters the semantic sense of information and are acceptable by any authentication system. The edits made should be very minor and subtle. Manipulation of images is generally allowed when correcting the color, tuning the brightness and contrast of the photo, fitting a layout using cropping a frame, reducing the noise like dust, dirt or scratches in the photo. Combining certain parts of whole image or leaving out certain parts of an image is acceptable unless they are mentioned and differentiated preferable by using boxes that portray the different parts of the image.

Semantic sense is really changed in malignant manipulation and this fashion should not be repeated. Moreover it never performs adding, moving or removing objects within the frame, changing the color other than to restore

what the picture actually looks like, to alter its interpretation, cropping a frame in order, flopping an image either left or right reversal, and lastly painting a photograph in other than its true orientation. Image forgery has two flavours namely Active and passive based approach.

Digital watermarking and digital signatures are some of the examples. Watermarking involves injecting a watermark which is used for the authenticity of the digital image which is indivisible from the image. On the other hand, Passive methods are the non- intrusive/blind methods and it never needs any prior information to include in the digital image. A digital image can be tampered by different attacks like resizing, addition of noise, blurring, rotation, scaling compressing, image splicing, copy-move and many more.

Introduction to Domain

MACHINE LEARNING

In the statistical context, Machine Learning is defined as an application of artificial intelligence where available information is used through algorithms to process or assist the processing of statistical data. While Machine Learning involves concepts of automation, it requires human guidance. Machine Learning involves a high level of generalization in order to get a system that performs well on yet unseen data instances. Machine learning is a relatively new discipline within Computer Science that provides a collection of data analysis techniques. Some of these techniques are based on well established statistical methods (e.g. logistic regression and principal component analysis) while many others are not. Most statistical techniques follow the paradigm of determining a particular probabilistic model that best describes observed data among a class of related models. Similarly, most machine learning techniques are designed to find models that best fit data (i.e. they solve certain optimization problems), except that these machine learning models are no longer restricted to probabilistic ones Therefore, an advantage of machine learningtechniques over statistical ones is that the latter require underlying probabilistic models while the former do not. Even though some machine learning techniques use probabilistic models, the classical statistical techniques are most often too stringent for the on coming Big Dataera, because data sources are increasingly complex and multi-faceted.

DIGITAL IMAGE PROCESSING

Two principal research paths evolve under the name of Digital Image Processing. The first one includes methods that attempt at answering question, By performing some kind of ballistic analysis to identify the device that captured the image or at least to determine which devices did not capture it. The history of a digital image can be represented as a composition of several steps, collected into three main phases: acquisition, coding, and editing. These methods will be collected in the following under the common name of image source device identification techniques.

Digital image processing allows one to enhance image features of interest

while attenuating detail irrelevant to a given application, and then extract useful information about the scene from the enhanced image. This introduction is a practical guide to the challenges, and the hardware and algorithms used to meet them. Images are produced by a variety of physical devices, including still and video cameras, xray devices, electron microscopes, radar, and ultrasound, and used for a variety of purposes, including entertainment, medical, business, industrial, military, civil, security, and scientific. The goal in each case is for an observer, human or machine, to extract useful information about the scene being imaged. An example of an

challenges, and the hardware and algorithms used to meet them. Images are produced by a variety of physical devices, including still and video cameras, x-ray devices, electron microscopes, radar, and ultrasound, and used for a variety of purposes, including entertainment, medical, business (e.g. documents), industrial, military, civil (e.g. traffic), security, and scientific.

IMAGE ANALYSIS

Image enhancement processing by an observer to extract information is called image analysis. Enhancement and analysis are distinguished by their output, images Vs scene information, and by the challenges faced and methods employed. Image enhancement has been done by chemical, optical, and electronic means, while analysis has been done mostly by humans and electronically. Digital image processing is a subset of the electronic domain wherein the image is converted to an array of small integers called pixels, representing a physical quantity such as scene radiance stored in a digital memory, and processed by computer or other digital hardware. Digital image processing, either as enhancement for human observers or performing autonomous analysis, offers advantages in cost, speed, and flexibility, and with the rapidly falling price and rising performance of personal computers it has become the dominant method in use.

AIM

The main aim of Image Forgery Detection is using copy move technique to identify certain features over canny edge detection , structural similarity index , hierarchical agglomerative clustering with improved accuracy. The project depicts the whether the original image is altered or not and point outthe features if image is altered.

OBJECTIVE

From the early days an image has generally been accepted as a proof of occurrence of the depicted event. Computer becoming more prevalent in business and other field, accepting digital image as official document has become a common practice. The availability of lowcost hardware and software tools, makes it easy to create, alter, and manipulated digital images with no obvious traces of having been subjected to any of these operations. As result we are rapidly reaching a situation where one can no longer take the integrity and authenticity of digital images for granted. Detecting forgery in digital images is an emerging research field with important implications for ensuring the credibility of digital images.

In the recent past large amount of digital image manipulation could be seen in tabloid magazine, fashion Industry, Scientific Journals, Court rooms, main media outlet and photo hoaxes we receive in our email. Digital image forgery detection techniques are classified into active and passive approach. In active approach, the digital image requires some preprocessing such as watermark embedding or signature generation at the time of creating the image, which would limit their application in practice. Moreover, there are millions of digital images in internet without digital signature or watermark. In such scenario active approach could not be used to find the authentication of the image.

SCOPE OF THE PROJECT

In today's time because of less-cost and more-resolution digital cameras, there is sample amount of digital images across globe. Digital images havea crucial presence in specific domains like in insurance process, forensic lab work, monitoring systems, services of intelligence, medical imaging and journalism. The most needed requirement is the images we see should be authentic. With the availability of effective image processing software'slike Adobe Photoshop the possibility is very high to modify an artificial picture. Copymove forgery is a very regular category of the digital fraud.

There are basically two techniques for identifying copy-move fraud whicha re Block based method and Key point based method

Python Imaging Library (PIL) is the de facto image processing package for Python language. It incorporates lightweight image processing tools that aids in editing, creating

and saving images. Support for Python Imaging Library got discontinued in 2011, but a project named pillow forked the original PIL project and added Python3.x support to it.

Pillow was announced as a replacement for PIL for future usage. Pillow supports a large number of image file formats including BMP, PNG, JPEG, and TIFF. The library

encourages adding support for newer formats in the library by creating new file decoders.

METHODOLOGY

Digital Photo images are everywhere, on the covers of magazines, in newspapers, in courtrooms, and all over the Internet. We are exposed to them throughout the day and most of the time. Ease with which images can be manipulated; we need to be aware that seeing does not always imply believing. We propose methodologies to identify such unbelievable photo images and succeeded to identify forged region by given only the forged image. Formats are additive tag for every file system and contents are relatively expressed with extension based on most popular digital camera uses JPEG and Other image formats like png, bmp etc. We have designed algorithm running behind with the concept of abnormal anomalies and identify the forgery regions.

COMPARATIVE STUDY OF EXISTING AND PROPOSED SYSTEM

As discussed in existing we will be using the concept of only Canny Edge Defection, Structure Similarity index, Hierarchical Agglomerate Clustering and Neural Network algorithms for classification method which will lead to less accuracy in predicting the image forgery and also noise ratio is too much high. In our proposed system we will be using the Copy move block key point technique. From this we are getting less noise ratio and good accuracy so we can state that our proposed system works better than the existing system.

OUTPUT SCREENS

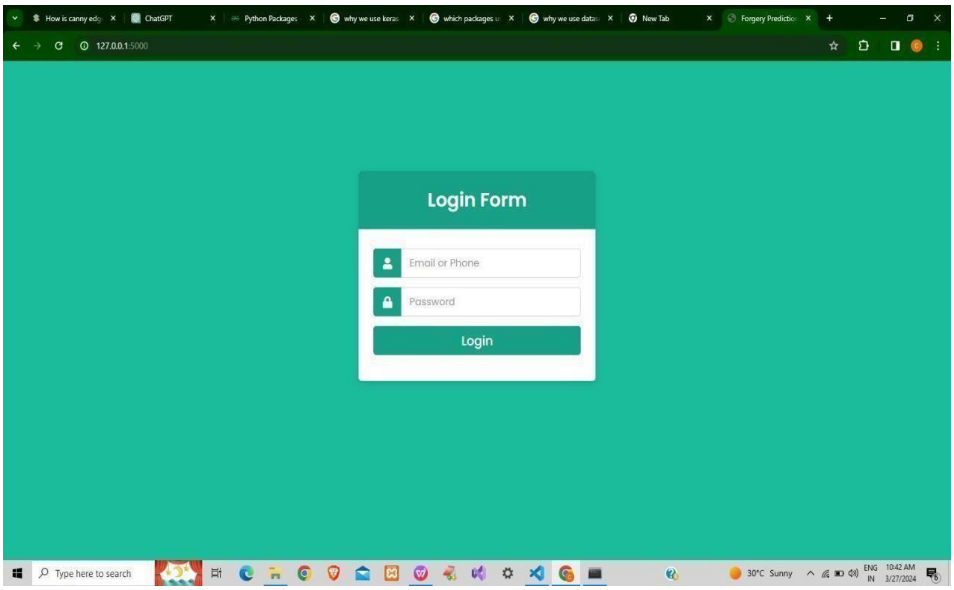

Fig 5.1 Login Page

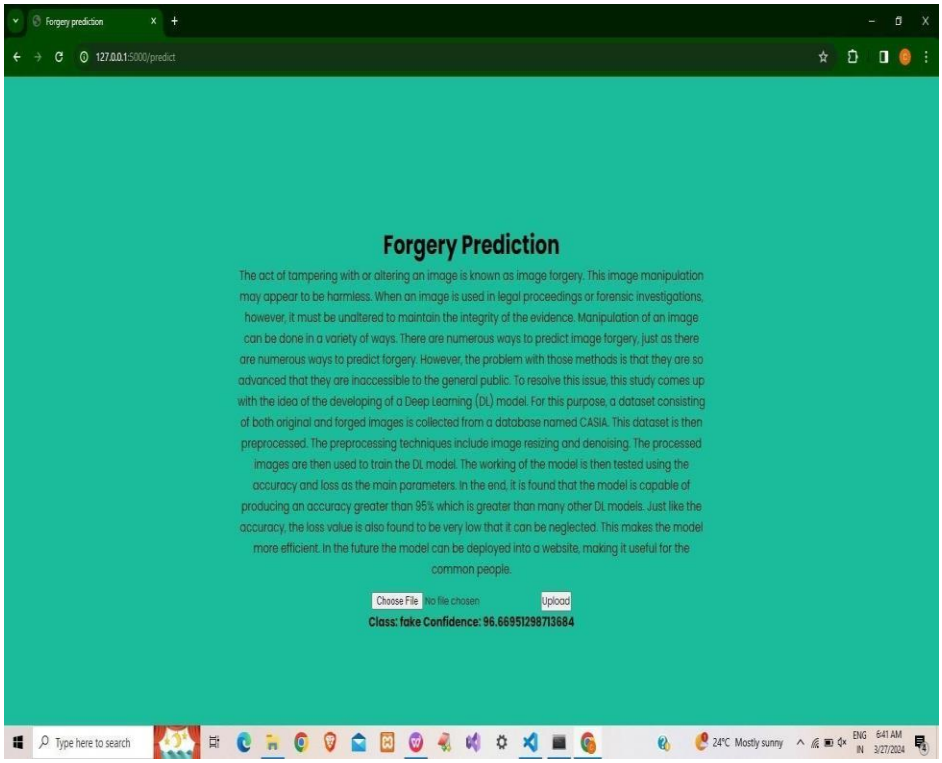

Fig 5.3 output of Image Detection

Integration Testing

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

Functional Testing

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted. Invalid Input : identified classes of invalid input must be rejected.
Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised. Systems/Procedures: interfacing systems or procedures must beinvoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests

CONCLUSION

The proposed scheme for the detection of image forgery uses feature point of extraction and morphological operation. The algorithm used in the proposed experiment can achieve good performance under various challenging conditions such as geometrical transform and JPEG compression.Hence the system is providing an accurate result in detecting copy move forgery without the help of any pre existing data set for the forged image.

FUTURE ENHANCEMENTS

The future work may focus on increasing the accuracy rate of the proposed algorithm in images as well as in video forgery detection. Another future direction in the proposed system can be of using a variable size of overlapping blocks which are used for the morphological operations. The usage of this system is generally limited to the forensics, in future this system can also be implemented to filter out the content on the social media to eliminate fake news and malicious content.

BIBLOGRAPHY

[1] A.C. Popscu, and H.Farid, "Statistical Tools for Digital Forensics" ,in Proc.The 6th international workshop on information hiding ,Toronto, Canada 2019.

[2] Shivani Thakur, RamanpreetKaur, Dr. Raman Chadha,JasmeetKaur, "AReview

Paper on Image Forgery Detection In Image Processing", IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278- 0661,p-ISSN: 2278- 8727, Volume 18, Issue
4,Ver. I (Jul.-Aug. 2019), PP 86-89.

[3] DevanshiChauhan, DipaliKasat, SanjeevJain, VilasThakare, "Survery on KeyPoint based Copymove Forgery Detection Methods onImages", sciencedirect volume 85, 2019.

[4] Ali Qureshi, M., and M. Deriche. "A review on copy move image forgerydetection techniques." IEEE, 2019.

[5] Qazi, Tanzeela. "Survey on blind image forgery detection."IET, 2019.

[6] M. Qiao,Sung, Q. Liu and B. Ribeiro, "A novel approach for detection of copy- move forgery," Fifth International Conference on ADVCOMP (Advanced Engineering Computing and Applications Sciences, 2019.

[7] Gajanan K. Birajdar, Vijay H. Mankar, "Digital image forgery detection using passive techniques: A survey," Digital Investigation, vol. 10, no.3, 2019, pp. 226- 245.

[8] GagandeepKaur, Manoj Kumar, "Study Of Various Copy Move Forgery Attack Detection In Digital Images", International Journal Of Research In Computer Applications And Robotics, Vol.3 Issue 9, Pg.: 30-34 September 2019.

[9]R.C. Gonzalez, R.E. Woods, "Digital Image Processing", 2nd edition, Addison- Wesley, 2019.

[10]L.Kang, X.-P. Cheng, "Copy-Move Forgery Detection in Digital Image", International Congress on Image and Signal Processing, IEEE Computer Society, 2019, pp. 2419-21.[2]

[11] ManpreetKaur, Richa Sharma, "Optimization of Copy- Move Forgery Detection Technique", International Journal of advanced Research in Computer Science and software Engineering, Volume 3, Issue 4, April 2019.[1]

[12] J.Fridrich, "Methods for "Methods for Tamper Detection in Digital Images", Proc. ACMWorkshop on Multimedia and Security, Orlando, FL, October 3031, 2019, pp. 1923.

[13] A.C. Popescu and H. Farid, "Exposing Digital Forgeries by Detecting Traces of Re-sampling", IEEE Transactions on Signal Processing, vol. 53, no. 2, pp. 758-767, 2005.