A SURVEY ON IMAGE RETRIEVAL USING MACHINE LEARNING

Nikhil Dalvi[1], Manish Khapre[2], Ankur Repal[3], Shubham Kadam[4], Prof. Yogesh Thorat[5]

[1,2,3,4]Students, Department of Computer Engineering, Dr. D.Y. Patil School of Engineering, Savitribai Phule Pune University, Pune
[5]Professor of Department of Computer Engineering, Dr. D.Y. Patil School of Engineering, Savitribai Phule Pune University, Pune

---***---

Abstract - *Currently there is vast and wide use of digitization technology. Text Recognition is likewise called as OCR, it comprises of PC framework intended to decipher pictures of typewritten content (most part filtered by scanner) into machine editable text or to convert images of text character into standard encoding system representing them. In machine learning OCR has become important field of research and also in computational vision aspects. Text Recognition is used in various official tasks in which the large type of data from universities, multinational companies, government offices etc., from where we can collect information from an image containing text. People usually scan the data and have the text of that data available in a .txt or .docx format and also wish to translate the data in their own specified Indian language. **Optical Character Recognition (OCR)** is procedure by which text image is given as input to a proposed system thus by providing the system with an image for manipulation purpose. The proposed framework utilizes an **OCR Engine-** a program with the specific procedure of speculating which letter (conspicuous to a computer) an image (recognizable to a human) represents. OCR is a **Machine-Learning** technique used to transform images that contain text into actual text content*

Keywords: Optical Character Recognition (OCR), image retrieval, OCR engine, Recognition (text recognition), Tesseract, Machine Learning and computational vision.

1. INTRODUCTION

Human beings travelling to various locations for distinct motives, at that time they discover the difficulty in verbal exchange with those nearby people due to the fact the visitor person does not know the accurate language which is used in that location. They do not understand the words that are written on board. So, at that point they wanted a man or woman to recognize and translate character from that exclusive picture. Our task is to identify the hassle in the image captured in digital camera by resizing, filtering and making it understandable to the individual over the picture after which it translates to human understandable language. So, the people cannot create some other problems for conversation or examine some other region statistics.

Text recognition utilized in project wherein the large information sources like places of work, banks, schools and so on., in actual existing applications where we want to collect some statistics from text written photograph. Human beings wish to test a document and feature the text of that document available in a .txt or .docx format.

The goal of proposed system is to know the common apprehended structure of information and suit those facts into fashions that may be understood and used by human beings.

Even though gadget mastering is an area inside computer science, it differs from traditional computational tactics. In traditional computing, algorithms are sets of explicitly programmed commands used by computers to calculate or clear up the problem. System getting to know algorithms rather than to allow the computer systems to teach on records inputs and use statistical analysis as a way to output values that fall within a specific variety. Due to this, machine gaining knowledge of computer systems in constructing models from sample facts as a way to automate selection-making tactics based on information inputs.

Any user has benefitted from gadget getting to know about the facial popularity era permits social media platforms to help users tag and get to know the number of people in an image**. Optical Character Recognition (OCR)** technology involves photo scanning of the text character-by-character, analysis of the scanned-in image, and then translation of the character image into character codes. Advance engines, powered by means of system mastering, propose what movies or television indicates to watch subsequent based totally on consumer choices. Self-riding cars that depend upon machine mastering to navigate may quickly be available to purchasers.

2. LITERATURE REVIEW

A.Kannan, Dr. V. Mohan, Dr. N. Anbazhagan, "Image Clustering and Retrieval using Image Mining Techniques" **[1]** Picture recovery is the essential necessity undertaking in the present situation. Content Based Image Retrieval is the famous picture recovery framework by which the objective picture to be recovered dependent on the valuable highlights of the given picture. In opposite end, picture mining is the emerging idea which can be utilized to remove potential data from the general accumulation of pictures. Target or close Images can be recovered in somewhat quick on the off chance that it is bunched in a correct way. In this paper, the ideas of CBIR and Image mining have been joined and another bunching strategy has been acquainted all together with increment the speed of the picture recovery framework.

Raniah A. Alghamdi, Mounira Taileb, Mohammad Ameen, "A New Multimodal Fusion Method Based on Association Rules Mining for Image Retrieval". **[2]** The

recovering strategy proposed in this paper uses the combination of the pictures' multimodal data (printed and visual) which is an ongoing pattern in picture recovery inquiries about. It joins two unique information mining methods to recover semantically related pictures: grouping and affiliation rules mining calculation. The semantic affiliation rules mining is built at the disconnected stage where the affiliation rules are found between the writings semantic groups and the visual bunches of the pictures to utilize it later at the online stage. The analysis was directed on in excess of 54,500 pictures of Image CLEF 2011 Wikipedia gathering. It was contrasted with an online picture recovering framework called MM Retrieval and to the proposed framework yet without utilizing affiliation rules. The acquired outcomes demonstrate that the proposed strategy accomplished the best exactness score among various question classifications.

Xin Zhou, Adrien Depeursinge, Henning Muller, "Information Fusion for Combining Visual and Textual Image Retrieval". **[3]** In this paper, established methodologies, for example, most extreme mixes (combMAX), entirety mixes (combSUM) and the result of the greatest and a non– zero number (combMNZ) were utilized and the trade– off between two combination impacts (chorale and dim steed impacts) was contemplated dependent on the total of n maximums. Different standardization methodologies were gone for. The combination calculations are assessed utilizing the best four visual and printed keeps running of the ImageCLEF therapeutic picture recovery assignment 2008 and 2009. The outcomes demonstrate that melded runs beat the best unique runs and multi– methodology combination factually outflanks single methodology combination. The logarithmic position punishment shows to be the steadiest standardization. The dim steed impact is in rivalry with the chorale impact and every one of them can create best combination execution relying upon the idea of the info information.

Pradeep K. Atrey, M. Anwar Hossain, "Multimodal fusion for multimedia analysis: a survey". **[4]** The combination strategies are depicted from the viewpoint of the fundamental idea, points of interest, shortcomings, and their use in different examination undertakings as detailed in the writing. Besides, a few particular issues that impact a multimodal combination process, for example, the utilization of relationship and autonomy, certainty level, relevant data, synchronization between various modalities, and the ideal methodology choice are additionally featured. At last, we present the open issues for further research in the region of multimodal combination.

Herbert Bay, Tinne Tuytelaars, and Luc Van Gool, "SURF: Speeded Up Robust Features". **[5]** In this paper, we present a novel scale-and turn invariant intrigue point finder and descriptor, instituted SURF (Speeded Up Robust Features). It approximates or even beats recently proposed plans concerning repeatability, peculiarity, and strength, yet can be processed and looked at a lot quicker. This is accomplished by depending on vital pictures for picture convolutions; by expanding on the qualities of the main existing indicators and descriptors (on the off chance that, utilizing a Hessian framework-based measure for the identifier, and a dispersion-based descriptor); and by rearranging these strategies to the fundamental. This prompts a blend of novel recognition, depiction, and coordinating advances. The paper presents trial results on a standard assessment set, and in addition on symbolism acquired with regards to a genuine question acknowledgment application. Both demonstrate SURF's solid execution.

3. SYSTEM ARCHITECTURE OVERVIEW

An approach is proposed for retrieval based on combination of colour, texture and edge features of image. Performance assessment of considered image retrieval techniques and proposed method is evaluated using parameters like Sensitivity, Specificity, Retrieval score and Accuracy. Experimental results of performance evaluation demonstrate that proposed technique outperforms other techniques. The purpose of proposed system is that the content based for the image text retrieval system that is the people understandable format various person cannot understand the advertised image text that is why we convert that image text to the retrieval text using Tesseract.

System can take a stored image and then process that image starting from the text detection. After the image conversion we can process this character for character Detection. Character Detection is the technique which is detects the character over the images.

Optical Character Recognition (OCR) is a process through which text images are input to a system by providing that system with an image. The computer system uses an OCR Engine- which basically is a computer program having the specific function based on guessing which letter an image represents.

OCR is a **Machine-Learning** technique which is used to transform images that contains text into actual text content. OCR uses different libraries to digitize and preserve their holdings. It is also used to process cheques and for sorting the mails. On daily basis huge number of magazines, letters are sorted by OCR machine which leads in speeding up mail delivery.

Character Segmentation is the method which is used to categorize the character in multiple forms.
Tesseract is an OCR engine that which is assisted by Unicode and has the capability to recognize and identify **many languages**. It contains trained data to recognize other languages.

Tesseract library is shipped with a handy command line tool called **tesseract**. We can use this tool to perform OCR on images and the output is stored in a text file. If we want to integrate Tesseract in the code, we will use Tesseract's API.
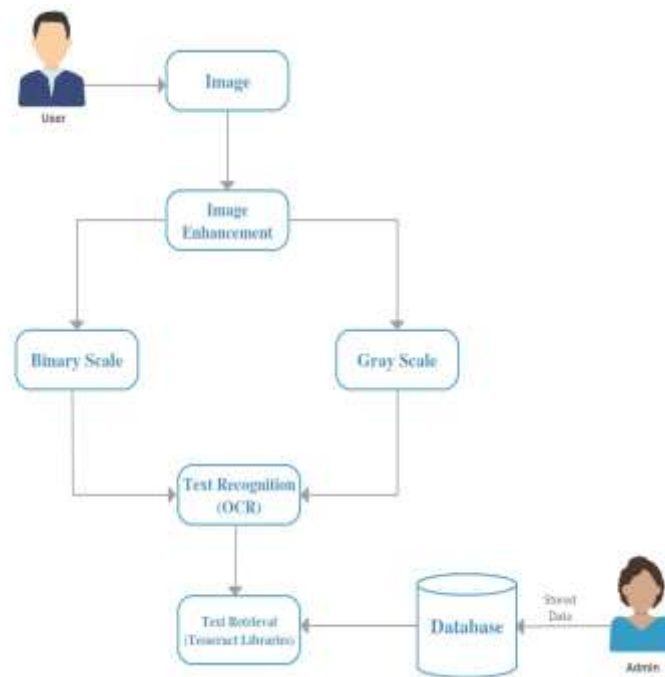
Fig.1 (System Architecture)

4. COMPARISON

Sr. No.	Paper Title	Authors	Description	Advantages	Disadvantages
01	A Novel Technique for Content Based Image Retrieval Using Color, Texture and Edge Features.	ManpreetKaur, NeelofarSohi	In this examination, recovery systems dependent on highlights like HSV, Color Moment, HSV and Color Moment, Gabor Wavelet and Wavelet Transform, Edge Gradient are contemplated and executed	A methodology is proposed for recovery dependent on mix of shading, surface and edge highlights of picture.	With expanding volume of computerized information, inquiry and recovery of applicable pictures from huge datasets in exact and productive way is a testing issue
02	Image Clustering and Retrieval using Image Mining Techniques.	A. Kannan, Dr. V. Mohan Dr. N. Anbazhagan	In this paper pictures are grouped dependent on RGB Components, Texture esteems and Fuzzy C mean calculation.	Convert Content Based Image Retrieval framework. To group the restorative pictures so as to analyze the correct ailment confirmed before.	To anticipate the picture just in restorative examination or malady expectation.
03	A New Multimodal Fusion Method Based on Association Rules Mining for Image Retrieval.	Raniah A. Alghamdi Mounira Taileb Mohammad Ameen.	In this paper semantic affiliation rules mining are developed at the disconnected stage where the affiliation rules are found between the content semantic bunches and the visual groups of the pictures to utilize it later at the online stage.	The picture recovering strategy is combination of the pictures' multimodal data (literary and visual) which is an ongoing pattern in picture recovery examines.	Some unmistakable issues (e.g. connection, certainty) that impact the combination procedure are likewise explained in more prominent detail.

Sr. No.	Paper Title	Authors	Description	Advantages	Disadvantages
04	SURF: Speeded Up Robust Features.	Herbert Bay, Tinne Tuytelaars, and Luc Van Gool	In this paper, creator introduces a novel scale-and pivot invariant intrigue point indicator and descriptor, instituted SURF (Speeded Up Robust Features).	The paper presents exploratory outcomes on a standard assessment set, and in addition on symbolism got with regards to a genuine protest acknowledgment application. Both demonstrate SURF's solid execution. Well as on imagery obtained in the context of a real-life object recognition application. Both show SURF's strong performance.	Which in contrast with the best in class are quicker to figure, while not relinquishing execution
05	Overview of the Wikipedia Retrieval Task at Image CLEF 2010	Adrian Popescu and Theodora Tsikrika and Jana Kludas	It abridges the recovery approaches utilized by the partaking gatherings, and gives an examination of the fundamental assessment results.	The point is to research recovery approaches with regards to an expansive and heterogeneous gathering of pictures that are looked for by clients with assorted data needs.	In this framework can't foresee the specific point that will be researched (i.e. points are not known to the framework ahead of time).

5. CONCLUSION

Although device learning is an area inside computer technology, it differs from traditional computational tactics. In conventional computing, algorithms are units of explicitly programmed commands utilized by computers to calculate or resolve the trouble. System studying algorithms instead permit computers to educate on statistics inputs and use statistical analysis so one can get output values that fall inside a particular range. Because of this, device studying helps computers in building models from pattern information so that you can automate selection-making tactics based totally on records inputs. Any era person nowadays has benefitted from device gaining knowledge of. Facial recognition generation lets in social media systems to assist customers tag and proportion pix of pals. Optical character recognition (OCR) generation converts photographs of text into movable type. Advance engines, powered by means of device mastering, advocate what movies or television shows suggests looking at subsequent based on person options. Self-using vehicles that rely upon device gaining knowledge of to navigate may soon be available to purchasers.

ACKNOWLEDGEMENT

REFERENCES

1. A. Kannan, Dr. V. Mohan, Dr. N. Anbazhagan, "Image Clustering and Retrieval using Image Mining Techniques" 2010 IEEE International Conference on Computational Intelligence and Computing Research.

2. Raniah A. Alghamdi, Mounira Taileb, Mohammad Ameen, "A New Multimodal Fusion Method Based on Association Rules Mining for Image Retrieval". 17th IEEE Mediterranean Electro Technical Conference, Beirut, Lebanon, 13-16 April 2014.

3. Xin Zhou, Adrien Depeursinge, Henning Muller, "Information Fusion for Combining Visual and Textual Image Retrieval". 2010 International Conference on Pattern Recognition.

4. Pradeep K. Atrey, M. Anwar Hossain, "Multimodal fusion for multimedia analysis: a survey". Multimedia Systems (2010) 16:345–379

5. Herbert Bay, Tinne Tuytelaars, and Luc Van Gool, "SURF: Speeded Up Robust Features". The ACM SIGACCESS International Conference on Computers and Accessibility (ASSETS'10), 2010.