

# Detection and Recognition of Hypertexts in Imagery using Text Recognition Technique

Rajath C S<sup>1</sup>, Rakesh C S<sup>3</sup>, Sumathi M<sup>2</sup>

<sup>1</sup>Student, Dept. of ISE, DBIT, Bangalore, Karnataka, India

<sup>2</sup>Student, Dept. of ISE, SE-Jain University, Bangalore, Karnataka, India

<sup>3</sup>Assistant professor, Dept. of ISE, DBIT, Bangalore, Karnataka, India

\*\*\*

**Abstract** - Image analytics is ever growing vast field with variety of applications. The ability of an electronic device to read textual characters from an image is the process of text recognition. This paper proposes a technique to analyze the hyperlink type texts present in an image, received or captured, on a hand held device type. Recognizing a particular type of pattern in the given text and processing it to predict that certain pattern is of URL, email or others hyperlink type is the main goal of this paper. We also discuss the methods in which these hyperlink patterns on detection and selection can be sent to a browser. Establishing an interaction between computing devices and humans epitomizes the greater levels of computer science, thus combining the features of image recognition, text and pattern recognition by establishing co - relativity to the field of natural language processing to build a system development program.

**Key Words:** Image processing, text recognition, hyperlinks, URL, browser, pattern recognition.

## 1. INTRODUCTION

Visual media is the frontier in much of today's communication. From image tagging to emojis to selfies, images are everywhere and have become greater part of the modern lexicon – especially on social media. Till date, major portion of social media analysis has been on textual content rather than images.

Image capturing today is the most sort out activity by mobile phone users\*. There are multiple reasons for capturing image, the main being to snap one self and capture surrounding. The mobile camera comes with many in built salient features; most of them are AI driven features. In fact today we need not have an external application for scanning the Bar codes and QR codes. Ensuring cognitive communication through human – machine interaction is a step ahead in the world of Artificial Intelligence. Image processing is a vital process in scanning of any type of images. Discovering hyperlinks stored in typical images, such as images captured in outer doors, product images, blogs, newspaper articles, academic papers, etc. is the focal point of this paper. With advances in image analytics there arises the need to enhance and automate best of image handling techniques for better usability of end users. Recognizing or identifying hyperlinks/URLs are of great importance as they enhance camera features and media functionality. The compiler processes the text at every line and recognizes the image for its pattern type to understand and parse it to the console for its respective output. The text being scanned may be an email id, a website link, a relative path to a host etc.

The paper is organized as follows: in Section 2, we discuss the related work done in field of image to text recognition. Section 3, overview of text recognition system. Section 4, we discuss about proposed system. Section 5, discusses about the system being proposed. In Section 6, we discuss the future work and the applications. Section 7, finally, conclusion is given.

\* - The word mobile phone users also refers to all other kinds of hand held devices using which image capturing and viewing is possible.

## 2. LITERATURE REVIEW

As discussed earlier text recognition from images is still an active research in the field of pattern recognition. To address the issues related to text recognition many researchers have proposed different methodologies, each approach or technology tries to address the issues in different ways. In forthcoming section we present a detailed survey of approaches proposed to handle the issues related to text recognition.

Bolan Su et al.[1] has proposed a novel adaptive character extraction method for the images present on web. This approach processes images faster, so that it is more suitable for real-time processing and applicable for mobile devices. They evaluated this adaptive method on complex scene images.

Sankaran et al. [2] has proposed a novel recognition approach that result in a 15% decrease in word error rate on heavily degraded Indian language document images.

Gur et al. [3] has discussed some problems in text recognition and retrieval. Automated optical character recognition (OCR) tools do not supply a complete solution and in most cases human inspection is required. They suggest a novel text recognition algorithm based on usage of fuzzy logic rulesrelying on statistical data of the analyzed font.

Antonio Moreno et al. [5] has studied relativity of text mining to artificial intelligence and have proposed a methodology on how huge set of textual data can be used to produce intuitive results. They have also provided step by step analysis of how exactly can processing of text can be performed and at which point we produce a predictive model.

Jawahar et al. [6] has proposed a recognition scheme for the Indian script of devanagari. The used approach does not require word to character segmentation, which is one of the most common reasons for high word error rate. The have reported a reduction of more than 20% in word error rate and over 9% reduction in character error rate while comparing with the best available OCR system.

### 3. SYSTEM DESIGN

Figure I briefly describe the flow for building the system for immediate capturing of hyperlink type of texts present in an image and developing a feature that provides an option to select the utility of the recognized text.

The overall data processing and system result has been broadly classified into two categories:

- a. High Level Operations.
- b. System Level Operations. High Level operation takes place in the initial stage and end stage, whereas system operations are an intermediate processes. Initially, high level operates in the sequential order accordingly to below steps:
  - a. User input in the form of image capture or in phone image is received.
  - b. The hand held device screen recognizes that the current screen space is being used by an image type.
  - c. If the current screen status is contained by image type then th1e software developed gets activated.

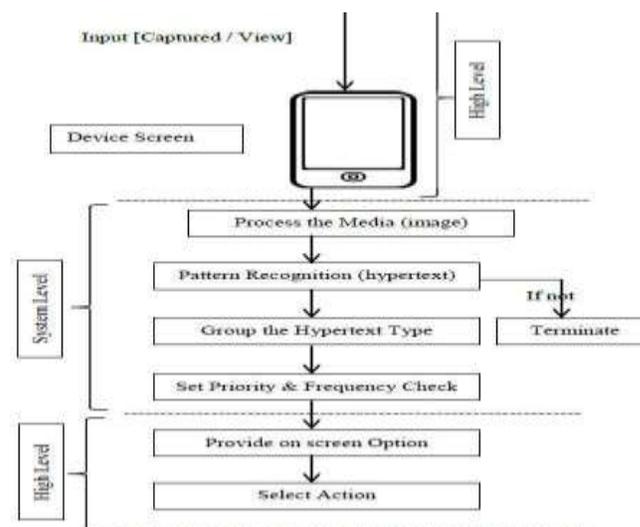


Figure 1 – System Design

Further processing takes place at system level.

- d. Based on this, the in-built system software internally scans / processes the image.
- e. The very next step is to search for required pattern type.
- f. If there are hypertext types then it is read and copied into a buffer. Buffer is used for multiple purposes, like: in case there are multiple patterns of similar type to group them. If the software fails to detect any of the required type of pattern the background processing thread is terminated.
- g. If there are multiple hypertexts then we need to order them based on the need of access hence we need to assign priority tags and if there are duplicates it is important to have a count of them and eliminate it. If there is multiple hypertext type then each of it will be stored in a buffer stack in the form of a list. Based on the priority set and user choice the required pattern typed would be fetched and redirected to its respective application for its utility.

The control now is transferred back to high level where post processing happens and the user gets to choose the further necessary actions. The user now selects his choice of hypertext and based on selection the redirection takes place. For example, if it is a URL selected then automatically the respective text is redirected to browser, if a mail address then to a default mailing application.

#### 4. EXISTING SYSTEM

Google's inbuilt feature on Google Photos named Google lens is closely related product. With Google lens or other systems there is question on data privacy, since these all work on their own cloud the data extracted per image of a user may or may not be considered for further analysis which would result is predicting user behavior. Having an in built camera feature enhances over all device specification.

Assume there is an image of addidas shirt on your phone and at the very corner with very small size text there is Addidas website printed on it. The current system does not detect this type of text. There are convolutional neural networks models that perform image textual mining which are limited to visited cards, block texts; identity boards etc. whereas they need to be extended to a complete document with end to end textual analysis.

Optical Character Recognition technique is one of the oldest and highly used for text extraction but it is not complete and efficient more over it is an external method. You'd require third party applications or any other type of OCR devices.

#### 5. PROPOSED SYSTEM

With the abundance of internet usage there comes a situation where almost everything everywhere is referenced to the web, knowing this it's important to simplify the process to access the internet or have a go to feature in all aspects and one among them being considered in this paper is multimedia. We propose to build system software that scans every active image\* and provide the person viewing the image with a choice for further actions. This process eliminates the cause for containing an additional step, wherein one has to send the image to a scanner application for identification of the required type of URLs. This proposal technique can be extended to both electronic and hand written type of text recognition.

For digital images, that is filled with 80% of text, a first stage is OCR processing. This OCR application operates as a background thread. Even though OCR is used as an end to end functionality here it does not make presence throughout application. The decision on whether the proposed technique of image processing shall occur automatically or based on user (manual) selection can be taken post initial experimental analysis.

Referring to the example in previous section, we insist that these type of URLs, the one's present on products, can directly lead to products page thus it greatly eases users work where one need not manually with difficulty search for the exact product instead find it with a click on the image. This process improves sequentially image mining, text recognition, pattern type identification, internal browsing power and search matching. This provides a greater boost for much of commercial daily product seller industry those majorly use images for marketing.

For better computation and faster examination of images, as users continually swipe to and fro for images the prediction rate should be higher, parallel processing of images can be performed across multiple cores. On the same side it is important to ensure data security since the concept is a direct implementation on end users choices, we need to ensure that none of user

personal data is being captured or the image processing does not store the content or have a backup of pixels, hence we propose a security layer implementation that performs an one way encryption to tackle handle data privacy and theft.

\* - image currently present on the device screen

## 6. CONCLUSIONS AND FUTURE WORK

Image data mining is at a rudimentary stage. The existing technology is not robust to retrieve efficient and reliable results; hence it can be enhanced by considering demand, time and research. In this paper, we proposed and discussed the method of text recognition by suggesting that advanced OCR is a wide area for researchers in pattern recognition. We introduced image mining a method for data mining of images which is based on processing of various feature components of an image. We also emphasized on the topic to establish human – machine interaction by introducing newer techniques that enhances cognitive features of devices. A lot of research work is being done on OCR and advance OCR for solving challenging problems in the field of image processing. Each image feature as its own significance and should be designed perfectly for better results.

The fundamental step ahead would be to build a model that outputs a highly efficient and accurate prediction rate for recognizing a specific pattern. We will be implementing the proposed methodology on a large set of images by developing and training the software such that any image type input should analyze for hyperlinks.

The amount of next generation media contents will be so large and filled with rich features that the ability to use it should be smooth and ease hence applications would be on the daily basis capturing images on large scale. The technologies around text analytics are currently being applied in several industries, for instance, sentiment and opinion analysis in media, finance, healthcare, marketing, branding or consumer markets. Insights are extracted not only from the traditional enterprise data sources, but also from online and social media, since more general public has turned out to be the large generator of text content. Social media is a great application; the current system on social media is such that only those hyperlinks that are manually attached are recognized whereas daily there are tons of images floating around, 30% of which are said to contain URLs. The same concept can be extended to videos. Infographics today are all over and contain wide category of text that are meant to serve the user hence this high featured text on infograph type images can be mined and hyperlinks in it can be grouped so that people viewing the infograph can gain immediate and detailed information about it.

## REFERENCES

- [1] Bolan Su, Shijian Lu, Trung Quy Phan, and Chew Lim Tan. "Character Extraction in Web Image for Text Recognition" in 21st International Conference on Pattern Recognition (ICPR 2012), Tsukuba, Japan, November 11-15, 2012.
- [2] Shrey Dutta, Naveen Shankran, PramodSankar K., C.V. Jawahar, "Robost recognition of Degraded Documents Using Character N-Grams". IEEE, 2012.
- [3] Koen Deschacht, Marie-Francine Moens. "Text Analysis for Automatic Image Annotation". The 45th Annual Meeting of the Association for Computational Linguistics, Prague, June 2007.
- [4] Amarjot Singh, Ketan Bacchuwar, and Akshay Bhasin. "A Survey of OCR Applications" International Journal of Machine Learning and Computing, Vol. 2, No. 3, June 2012.
- [5] Erran Carmel, Stephen Crawford. and Hsinchun Chen. "Browsing in Hypertext: A Cognitive Study". IEEE Transactions on Systems, Man and Cybernetics, Vol. 22. No. 5, September – October, 1992.
- [6] Naveen Sankaran and C.V Jawahar, "Recognition of Printed Devanagari Text Using BLSTM Neural Network," IEEE, 2012.s
- [7] Antonio Moreno, Teófilo Redondo. "Text Analytics: the convergence of Big Data and Artificial Intelligence". International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 3, N°6.
- [8] Dat T. Huynh, Tru H. Cao, Phuong H.T. Pham, Toan N. Hoang. "Using Hyperlink Texts to Improve Quality of Identifying Document Topics Based on Wikipedia". International Conference on Knowledge and Systems Engineering, 2009.
- [9] Frank Vanhoenshoven, Gonzalo N'apoles, Rafael Falcon, Koen Vanhoof and Mario K'oppen. "Detecting Malicious URLs using Machine Learning Techniques". Computational Intelligence (SSCI), IEEE Symposium Series on Dec 6-9, 2016.

- [10] Pratik Madhukar Manwatkar, H Shashank Yadav. "Text Recognition from Images". IEEE Sponsored 2nd International Conference on Innovations in Information, Embedded and Communication systems (ICIIECS), 2015.
- [11] N.Nathiya, K.Pradeepa. "Optical Character Recognition for Scene Text Detection, Mining and Recognition". IEEE International Conference on Computational Intelligence and Computing Research, 2013.
- [12] Sen Wang, Li Cai, Huanqing Cui, Chaowen Feng, Xiaokuo Yang "Three-dimensional Quantum Cellular Neural Network and Its Application to Image Processing". IEEE ICIS 2017, May 24-26, 2017, May 24-26, 2017, Wuhan, China.
- [13] Zhang Weigong, Di Suran, Zhang Yongxiang and Li Liming. Image Capture and Display Based on Embedded Linux". International Conference on Solid State Devices and Materials Science, 2012.
- [14] Monika Sahu, madhup Shrivastava. "Image Mining: A New Approach For Data Mining Based On Texture". Third International Conference on Computer and Communication Technology, 2012.
- [15] Qixiang Ye. "Text Detection and Recognition in Imagery: A Survey". IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 37, NO. 7, JULY 2015.