Research Journal of Engineering Sciences ISSN 2278 – 9472
Vol. **7(7),** 5-8, July **(2018)** *Res. J. Engineering Sci.*

Review Paper

Character identification using document image analysis

Antara Mukherjee and KGS Sharma*
Department of ETE, Bhilai Institute of Technology, Raipur, CSVTU, Bhilai, CG, India
ganpatishrinivas01@bitraipur.ac.in

Available online at: www.isca.in, www.isca.me
Received 6th April 2018, revised 17th July 2018, accepted 24th July 2018

Abstract

The interdisciplinary field of computer vision combined with the powerful algorithms of machine learning can be used to build intelligent electronic systems for various domains of data analysis. Document layout analysis deals with the identification and categorization of the geometric and logical characteristics of text elements from the scanned documents. An OCR engine is used to convert images of typed or handwritten text into machine code. This research uses a Raspberry Pi3 processor module along with suitable peripherals and the entire system is supported by Python programming.

Keywords: OCR, pattern recognition, character recognition, raspberry Pi3 board, Camera Pi module, layout-analysis, GUI.

Introduction

Although the technology of document imaging or layout analysis has been in existence since the past couple of decades, the limit of its application domains has been pushed farther. The rapid development of embedded systems, artificial intelligence and computing techniques has redefined the possibilities of encoding-decoding, script reading and translations, handwriting study etc. In recent times, it has also made a few successful attempts at developing high level security solutions as well.

The issue of lack of reference material for decoding ancient scripts can be resolved by using powerful pattern and character recognition algorithms. Programmable electronic systems can be built to capture the data from these scripts, evaluate them and convert them into the desired output.

Advanced applications of image processing, edge detection techniques[1], computerized mathematical tools, geometrical and graphical analysis of each of the scanned characters is used to create a reference data set. This data set is fed into the internal memory of the Raspberry Pi controlled system. Using Python programming, various characteristics of each alphabet such as number of straight lines, curves, dots, pixel spaces etc. are studied and matched for each input element from the scanned document. The output of this system can be either audio signal, digital signal, LCD output, HDMI display etc[2].

Text image capturing

Optical character recognition: The process of identifying and evaluating the printed characters by using photoelectric devices and computer software is the principle of 'Optical Character Recognition' (OCR). It converts images of typed, handwritten or printed text into machine encoded text from scanned document or from subtitle text superimposed on an image[3]. It is a widely used method of digitizing printed texts so that they can be edited, searched and more compactly stored on most of the electronic devices. This electronic text should also be suitably displayed on-line, and used in machine processes such as cognitive computing, machine translation, text-to-speech, key data and text mining. OCR is a deep field of research in pattern recognition, artificial intelligence and computer vision.

Character classification: Each character of scanned textual document holds a unique set of graphical and layout features. Hence the first step in recognizing characters of the text is to distinguish them on the basis of these features. This process is called 'feature extraction'. For this purpose, a set of features are extracted for each class that helps distinguish it from other classes, while remaining invariant to characteristic differences within the class[4]. In[4], various feature extraction methods are classified in three major groups: i. Global transformation and series expansion, ii. Statistical features, iii. Geometrical and topological features.

Pattern recognition: After classifying the characters, this data is further processed by studying the features at the binary level. The pixels are given digital values (0 or 1) based on thickness of line, height and width of the character etc. This level data processing is done by pattern recognition algorithms, which are designed so as to identify and match the features of the sample data and the scanned image data.

Hardware

Raspberry Pi 3 model B: Raspberry Pi consists of a Broadcom system on a chip (SoC) including ARM compatible CPU and an on chip graphics processing unit GPU. Python is considered to be the most widely used programming language on this

processor. In this research, Raspberry Pi is [5]programmed to take input of scanned documents, identify their features and produce the desired output. Since it consists of four USB ports, an audio output and an HDMI output port, a variety of output peripherals can be used, depending on the type of application.

Camera Pi Module: This project uses an RPI NOIR camera board. The camera plugs directly into the CSI connector on the Raspberry Pi. It is able to deliver clear 5MP resolution image, or 1080p HD video recording at 30fps. The module attaches to Raspberry Pi, by way of a 15 pin Ribbon Cable, to the dedicated 15 pin MIPI Camera Serial Interface (CSI), which was designed especially for interfacing cameras. The CSI bus is capable of extremely high data rates and it exclusively carries pixel data to the BCM2835 processor[6]. This camera does not consist of any infrared filter making it suitable to take photographs even in dim light.

Output Interfaces: The output of the document analysis system can be either an audio signal or a display on an LCD screen or a monitor. In case of text to audio conversion, the output audio jack of the Raspberry Pi reads the textual document in the form of a speech signal to the user. Hence the audio jack acts as the interface while a speaker phone acts as an output peripheral.

In applications which require the conversion of printed text into electronically readable format, visual display peripherals are used at the output. An LCD screen is used for basic display while a computer monitor equipped with suitable software can be enabled to read as well as edit the document.

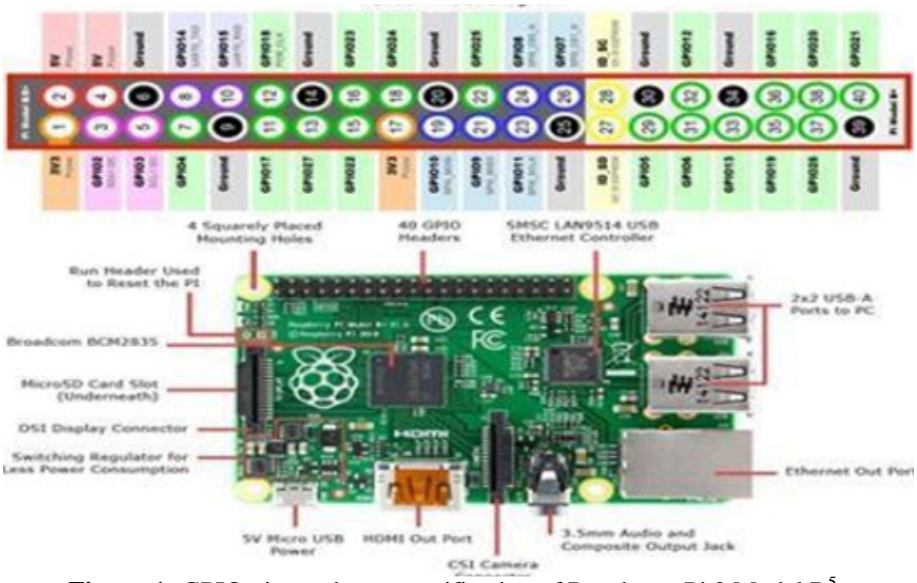

Figure-1: GPIO pins and port specification of Raspberry Pi 3 Model B[5].

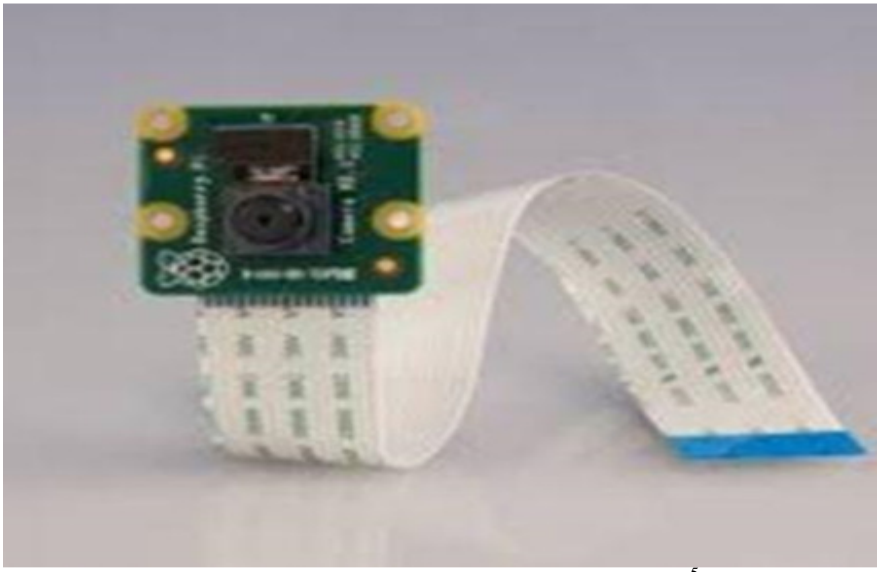

Figure-2: Raspberry Pi 5MP Camera Board v1.3[5].

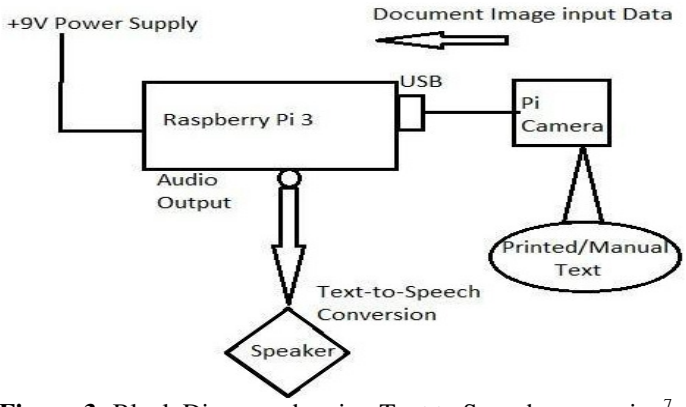

Figure-3: Block Diagram showing Text to Speech conversion[7].

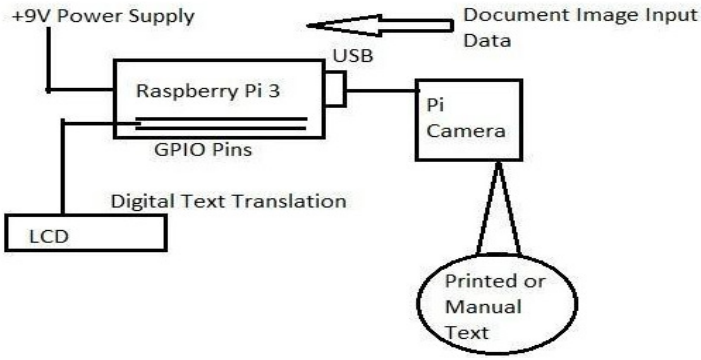

Figure-4: Block Diagram showing Printed to Digital Text Conversion[7].

Results

As a part of the proposed system, numerical digit identification can also be taken into consideration as well. The results below show the process of identifying the digit '4' as an example. This is executed in the python development environment (IDLE Python 3.6) using the K-Nearest Neighbor machine learning algorithm.

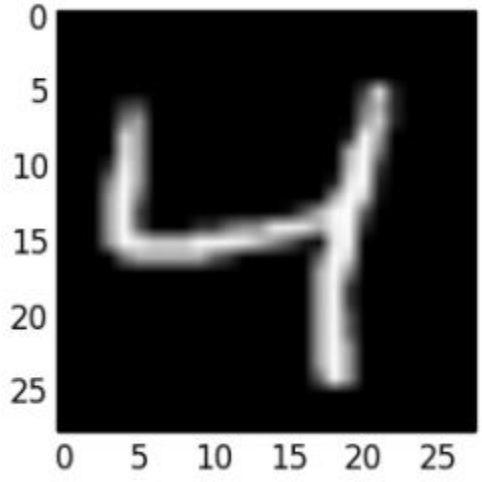

Figure-5: Figure showing input taken from the camera or handwritten input from user. Pixel characteristics are classified on a 25 X 25 scale[8].

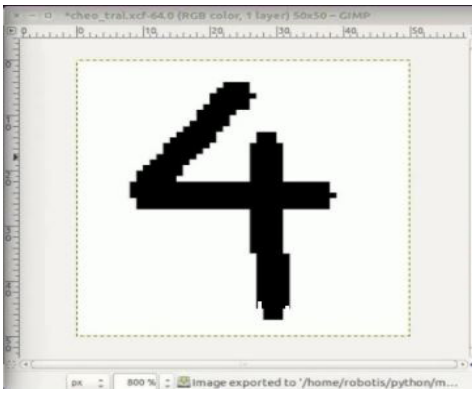

Figure-6: Figure considering the closest matching character from the dataset available. The dataset includes 150 fonts to compare each digit[8].

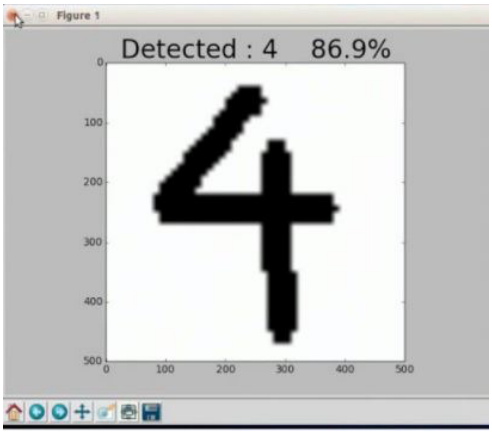

Figure-7: The final identified digit after comparison shown on a 500 X 500 scale, where extent of matching character identification is shown in percentage[8].

Future scope: A number of prototypes have already been developed and further research is still being carried out to discover new areas of applications. Currently, in the age of digital revolution, it is very important to preserve the authenticity of data. While data needs to be secured, it should also prove beneficial for humans as well.

Recently, a new domain of 'Forensic Document Analysis' has come into the picture. These security features have been embedded together with the existing printing processes to identify the authenticity of the documents. This is especially useful for professionals working with government security agencies or national defense services. This technique combined with powerful cryptographic solutions can also be used to carry out fingerprint analysis[9], face recognition and store confidential document information as well.

Visually impaired people report numerous difficulties with accessing printed text using existing technology, including problems with alignment, focus, accuracy, mobility and efficiency[10]. The concept of converting the scanned images of textual documents into audio signal output can be particularly

helpful for people who are visually challenged. Since they do not possess the normal sight, it becomes difficult for them to understand visual or textual information. The proposal of a method of developing a Raspberry Pi controlled system is made which is capable to convert the input text into audio output. Hence, a document analysis system can be used as an automatic book reader for the blind.

Conclusion

In this research, we have described the working methodology of using a document image analysis system. In order to enable a machine to read a printed or handwritten textual information, it should be able to study and store orientation, edge features, geometrical characteristics etc of each pixelated image. Using GUI application, image processing and machine learning algorithms, intelligence can be developed into the computing system.

Although there can be a wide range of application specific domain, this research studies and discusses the possibilities of creating high end security systems and the ones beneficial for the social aspects of human interactions with machines.

References

1. Maini R. and Aggarwal H. (2009). Study and comparison of various image edge detection techniques. *International journal of image processing (IJIP)*, 3(1), 1-11. ISSN 1985-2304.

2. Balamurugan E., Sangeetha K. and Sengottuvelan P. (2011). Document Image Analysis -A Review. *International journal of Computer application*, 1(1). ISSN-2250-1797.

3. Aaron James S., Sanjana S. and Monisha M. (2015). OCR based automatic book reader for the visually impaired using Raspberry PI. *International Journal of Innovative Research in Computer and Communication Engineering*, 4(7), 1111-1118. ISSN-2320-9801.

4. Dongre V.J. and Mankar V.H. (2010). A Review of Research on Devnagari Character Recognition. *International Journal of Computer Applications*, 12(2), 2. ISSN NO 0975- 8887.

5. Element 14 Community (2015). Raspberry pi 3 model GPIO 40 pin block Pinout. https://www.element14.com/community/docs/DOC-73950/l/raspberry-pi-3-model-b-gpio-40-pin-block-pinout raspberrypi.org/forums 2015.

6. Senthilkumar G., Gopalakrishnan K. and Kumar V.S. (2014). Embedded image capturing system using raspberry pi system. *International Journal of Emerging Trends & Technology in Computer Science*, 3(2), 213-215. ISSN 2278-6856.

7. Engelsma J.J., Cao K. and Jain A.K. (2017). Raspi Reader: Open Source Fingerprint Reader. arXiv preprint arXiv:1712.09392.

8. Gurav M.D., Salimath S.S., Hatti S.B., Byakod V.I. and Kanade S. (2017). B-LIGHT: A Reading aid for the Blind People using OCR and Open CV. *International Journal of Scientific Research Engineering & Technology (IJSRET)*, 6(5). 546-548. ISSN 2278 0882.

9. Jabeen F.A., Ramamurthy B. and Latha N.A. (2017). Development and implementation using Arduino and Raspberry Pi based Ignition control system. *Advances in Computational Sciences and Technology*, 10(7), 989-2004. ISSN 0973-6107

10. Bukhari S.S., Shafait F. and Breuel T.M. (2012). Layout analysis of Arabic script documents. In Guide to OCR for Arabic scripts, Springer, London, 35-53. ISBN978-1-4471-4072-6.

11. Jones J.D., Witek K., Verweij W., Jupe F., Cooke D., Dorling S. and Foster S. (2014). Elevating crop disease resistance with cloned genes. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 369(1639), 20130087.