



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



---

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume: 5      Issue: III      Month of publication: March 2017**

**DOI: <http://doi.org/10.22214/ijraset.2017.3223>**

**[www.ijraset.com](http://www.ijraset.com)**

**Call:  08813907089**

**E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)**

# **Intelligent Forms Processing System**

Tharani B<sup>1</sup>, Ramalakshmi. R<sup>2</sup>, Pavithra. S<sup>3</sup>, Reka. V. S<sup>4</sup>, Sivaranjani. J<sup>5</sup>

<sup>1</sup>Assistant Professor, <sup>2,3,4,5</sup>UG Students, Dept. of ECE

Sri Shakthi Institute of Engg and Technology, Coimbatore, TN, India.

**Abstract:** *The reading part of words is one of the most complex tasks in automated forms processing. The project describes an integrated real time system to read names and addresses on forms. The Name and Address Block Reader (NABR) system accepts both machine printed and hand printed address block images as input. The data is then fed to an RDBMS for further processing. The application software has two major steps: document analysis and document recognition. The functional architecture, software design, system architecture and hardware implementation are described. Useful application evaluation on machine printed and handwritten addresses are presented.*

**Keywords:** *automated forms processing, Name and Address Block reader, document analysis.*

## **I. INTRODUCTION**

Nowadays entering the details of the person and storing the data had become a tedious job in all the organizations. Our proposed system has made easy for data entering people, helping them to fill the machine printed or hand written paper forms automatically. An IFPS can be valued if there is a need to process hundreds or thousands of images every day. This system reduces the risk of a person in entering the information's and reduces time. It can be used in the large organization. We made brief explanation about each part of our project.

## **II. OBJECTIVE**

The main objective of this project is to capture the paper form as an image and scans the content of the image. The extracted data images are interpreted by an OCR process that reads typical mono space fonts. The information is filled accordingly in the webpage with the help of IOT. The Intelligent form processing system (IFPS) also provides capabilities for efficiently storing form images. Thus intelligent form processing scans letters from the form, captured by the camera, and the details of the form is stored.

## **III. INTELLIGENT FORMS PROCESSING SYSTEM**

The project "intelligent form processing system" is implemented to read the details of the person from the form and store it to the webpage. It previously reads the contents of the form only the machine printed format and stores it in the spreadsheet. So, we have found the solution through our project. The first part of this project is camera, it captures the pages of the document and it sends the captured image file to raspberry pi 2B which contains the Tesseract OCR (optical character recognition). The Tesseract software which turns the image or printed document into machine encoded language. It involves character code that had used in data processing.

## **IV. LITERATURE SURVEY**

In the literature survey, the project related survey's like electronic intelligent forms processing system, scan the data that capture by using the camera module and print the data on the website it can be viewed only when the now the username and password, it can also be downloaded and kept in a excel sheet are explained briefly. With the increasing importance of information processing in industries arises the need for efficient data logging systems that are compatible with existing measuring devices such as LCD and LED meters. This paper presents the framework for an Internet of Things (IoT) device as an automated industrial meter reader that uploads the collected numeral data to a cloud storage for centralized data processing. The implementation of the device is done using Raspberry Pi as the platform. The device follows a four-step process- Image Acquisition using Raspberry Pi camera module, Optical Character Recognition using feature extraction technique, Internet Upload Mechanism using Google Forms and Online Data Processing using Google Spreadsheet. With the increasing importance of information processing in industries arises the need for efficient data logging systems that are compatible with existing measuring devices such as LCD and LED meters. This paper presents the framework for an Internet of Things (IoT) device as an automated industrial meter reader that uploads the collected numeral data to a cloud storage for centralized data processing. The implementation of the device is done using Raspberry Pi as the

## International Journal for Research in Applied Science & Engineering Technology (IJRASET)

platform.

The device follows a four-step process- Image Acquisition using Raspberry Pi camera module, Optical Character Recognition using feature extraction technique, Internet Upload Mechanism using Google Forms and Online Data Processing using Google Spreadsheet. Up-gradation of all existing meters is won't be economically viable for most enterprises. Also, it would be desirable if a device can be built, which can be modified flexibly to read any type of digital industrial meter. Further, centralized storage and processing of data collected from different locations would increase the convenience of data management. Outsourcing of data for knowledge processing will also be easier if it is collected centrally. This is the motivation behind the design of an Internet-enabled automated industrial meter reader.

Dr.Radha Shankarmani designed the Digitization and Paperless Processing through the use of mobile imaging Technology. Even today, a large number of organizations collect data using paper forms. However, it can be difficult to aggregate, and analyze the data collected using paper forms. Better management and processing of forms and applications is indispensable to improving customer experience. But, typing the form data into a spreadsheet is time-consuming, mundane and may result in errors. Various attempts have been made to automate the process but the solutions require the use of expensive specialized hardware.

### V. PROPOSED SYSTEM

Nowadays processing a thousands of forms is a tedious and complicated job. Instead of wasting time by doing it in a manual manner, we have proposed a system called an intelligent forms processing system, for entering the details of the forms automatically using raspberry pi. An intelligent forms processing system (IFPS) which provides capabilities for automatically indexing form documents for storage/retrieval to/from a document library and for capturing information from scanned form images using intelligent character recognition. the primary features used in this step are the pattern of lines defining data areas on the form. The last phase is to transfer the data to the webpage which has separate fields and the data are stored in the respective fields accordingly.

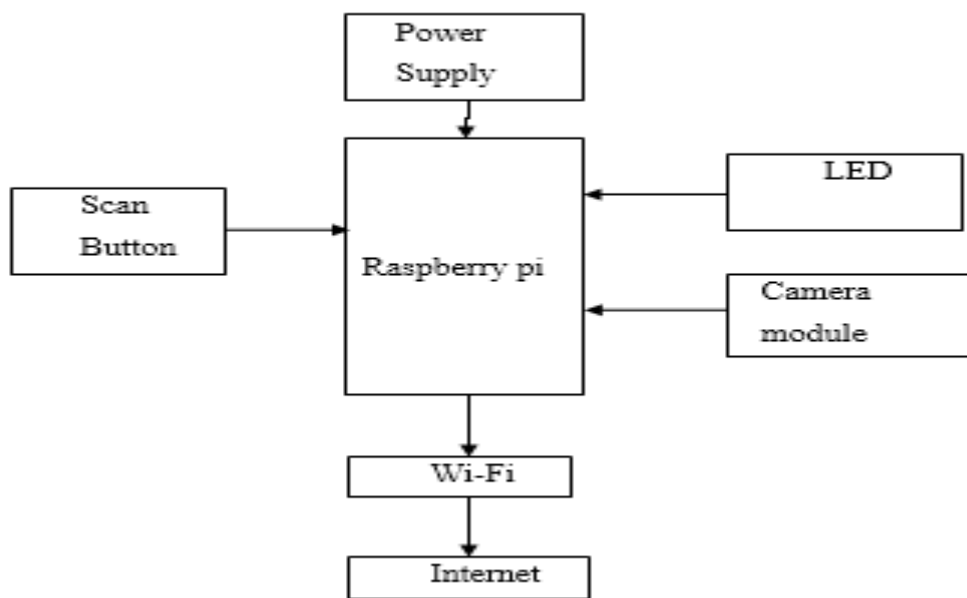


Fig 1 Proposed Block Diagram

When the scan button is pressed, the form is captured by the camera module which consists of 13 MB, connected to the raspberry pi. The camera module is capable of taking full HD photo and can be controlled programmatically. The image is saved in the format as .jpeg. The captured image is sent to the raspberry pi where it has functions such as optical character recognition. OCR which converts the captured image into text file. It is system that provides full alphanumeric recognition of printed or hand written characters at electronic speed by simply scanning the form. It includes form definition, scanning, image preprocessing and recognition capabilities. Here open cv and Tesseract software is used for this function. It recognize the characters and extracts the hand written or printed letters and converts this letters into machine encoded text. This is saved as text file. The text is segregated

## International Journal for Research in Applied Science & Engineering Technology (IJRASET)

using text segregation process. Then the details in the forms is stored in the local database. Then it is upload to the collecting website using a login id. In this webpage the download option is given at the bottom. If we press the download button the details stored in the webpage stored in to XL sheet. The LED is used for indication purposes. Here we use red and green LED's for indication. The green LED is glow if the camera captures the form properly and uploads in the web page. If the camera cannot recognize the form image properly it indicates by glowing red LED.

A camera section is an picture sensor integrated with control electronics and an interface like CSI, Ethernet or plain raw low-voltage differential signaling. A charge-coupled device (CCD) is a device for the movement of electrical charge, usually from within the device to an area where the charge can be manipulated, for example conversion into a digital value.



Fig 2 Web Camera

In recent years CCD has become a major technology for digital imaging. In a CCD image sensor, pixels are represented by p-doped MOS capacitors. the CCD is then used to read out these charges. CCD image sensors are widely used in professional, medical, and scientific applications where high-quality image data are required. In applications with less exacting quality demands, such as consumer and professional digital cameras, active pixel sensors (CMOS) are generally used; the large quality advantage CCDs enjoyed early on has narrowed over time.

### . VI. ALGORITHM

- Step 1: Wait for to press the scan button.
- Step 2: Keep the form, press the scan button.
- Step 3: Take an image using the digital camera.
- Step 4: Performs image processing algorithms.
- Step 5: Text extraction.
- Step 6: Text segregation.
- Step 7: Save in local data base.
- Step 8: Upload to the collecting website.
- Step 9: Show the finishing indication.

### VII. RESULTS AND DISCUSSIONS



Fig 3 Snapshot of hardware setup

# International Journal for Research in Applied Science & Engineering Technology (IJRASET)



Fig 4 Snapshot of Raspberry pi

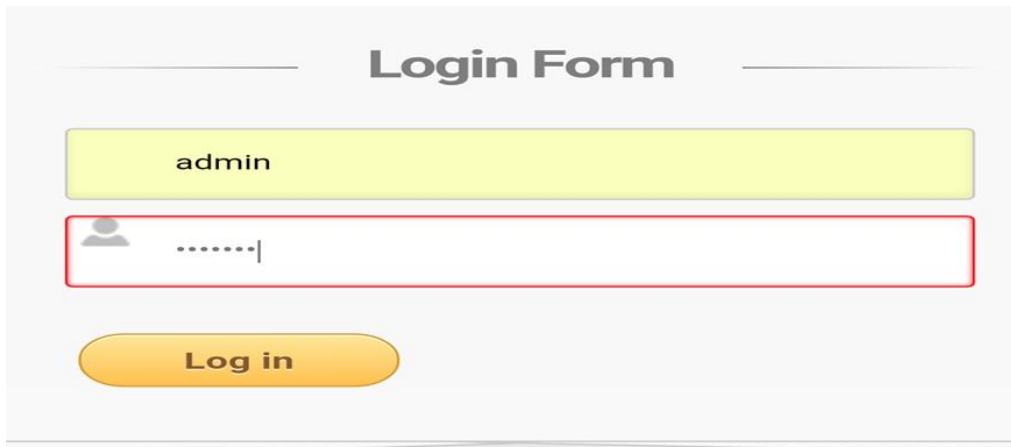
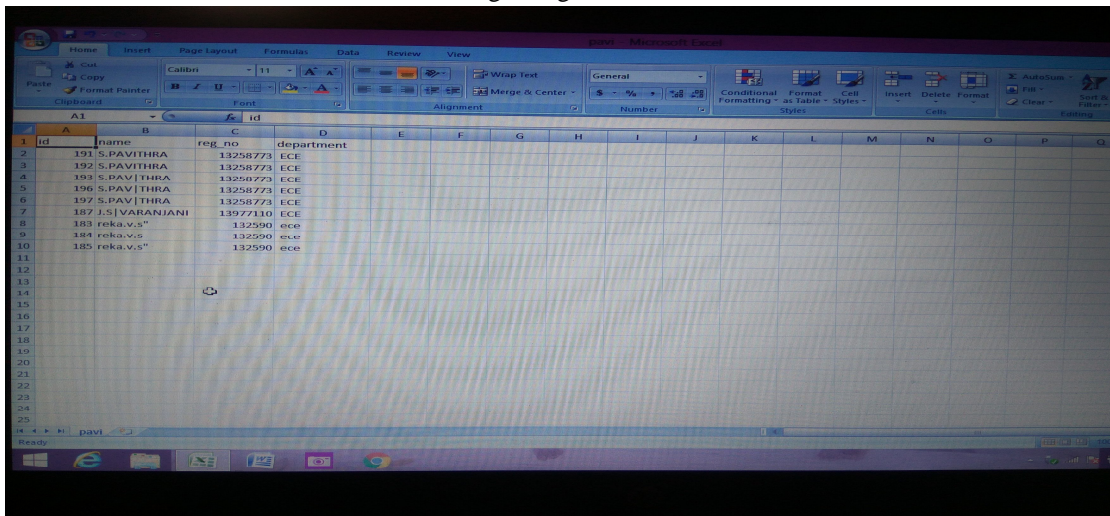


Fig 5 Log in form



id	name	reg_no	department
191	S.PAVITHRA	13258773	ECE
192	S.PAVITHRA	13258773	ECE
193	S.PAVITHRA	13258773	ECE
196	S.PAVITHRA	13258773	ECE
197	S.PAVITHRA	13258773	ECE
187	J.S VARANJANI	13977110	ECE
183	reka.v.s	132590	ece
184	reka.v.s	132590	ece
185	reka.v.s	132590	ece

Fig 6 Snapshot of output in excel format

# International Journal for Research in Applied Science & Engineering Technology (IJRASET)

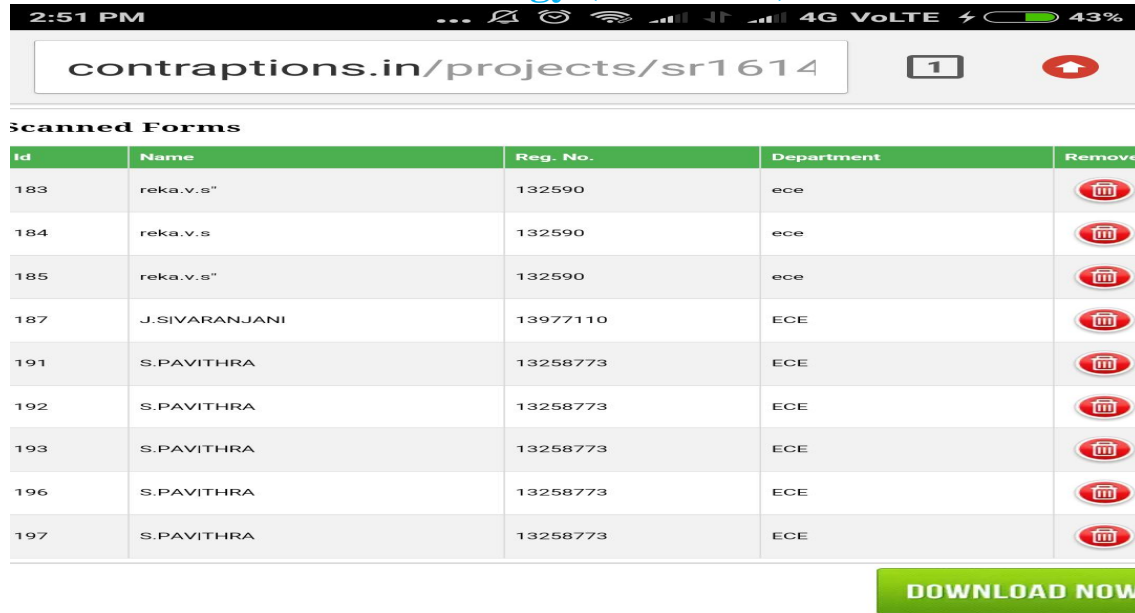


Fig 7 Snapshot of final output

## VIII. CONCLUSIONS

Nowadays, there are various organizations still using paper forms mainly because they are simple to distribute and collect the information. Converting this information from hard copy to soft copy requires manual power. So we have presented a device using Raspberry pi that scans and captures the image of the printed and handwritten forms. The IFPS device converts the image into text format and uploads it in the webpage, once the user manually presses the button. Once the press is recognized by the device, the webcam starts scanning and the process continues to load the data in the webpage. The system is simple and efficient to use. The total time required by the hardware for the entire process from scanning to uploading is 30 seconds.

## REFERENCES

- [1] K. Jung, K.I. Kim, A.K. Jain, "Text information extraction in images and video: a survey", *Pattern Recognition*, 37 977-997S, 200
- [2] Lead Tools Form Recognition Implementation Strategies for LargeEnterprises.<https://www.leadtools.com/whitepapers/2013/formsrecognition-implementation-strategies-for-large-enterprises.pdf>.
- [3] Y.-L. Boureau, F. Bach, Y. LeCun, and J. Ponce. Learning mid-level features for recognition. In *Computer Vision and Pattern Recognition (CVPR)*, 2010 IEEE Conference on, pages 2559–2566. IEEE, 2010.
- [4] wikipedia data base resource.



10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)