International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 07 | July 2020 www.irjet.net p-ISSN: 2395-0072

AUTOMATION OF MANUAL DATA ENTRY BY DATA EXTRACTION THROUGH IMAGE PROCESSING

Lohit Marla[1], P. Vijaya Manikanta Reddy [2], Bhargava Sai Gopal Kolli

[1,2,3]Student, Department of Computer Science and Engineering, GITAM deemed to be University, Visakhapatnam, Andhra Pradesh, India.

Abstract - In today's digitalized world, everything is being online where the users can get the information right from the websites and also a lot of information that is being generated and published in internet sources is being entered from some sort of documents, lists or by other means. The problem here is that we are manually entering the data from the written documents into the database where a lot of time that is being invested is wasted and also the effort of the data entry teams. So, we in this project collaboratively worked on to find a solution by choosing a use case of an online food delivery unit where the customers of it, order food online platform through the web based applications or android applications on watching the food menus of different hotels and restaurants which were manually entered by various data entry teams. The solution that we came up with is to group all menu cards that were sent by the customers together and extract the data through Google cloud vision[5](Data Extraction) by pre-processing from the uploaded menu cards of online food delivery units and display them in an online website. There will be substantial decrement of the cost of paying salaries and the additional effort of a data entry team that is being kept and makes the entire process much more easier and efficient by automation.

***Key Words*: Google Cloud Vision, Automation, Pre-Processing and Data Extraction.**

1. INTRODUCTION

From the past to the present, a lot of written information is resided in the books or in other written formats. Apart from them a lot of documents like land, house registration documents in government were there in written format where they all had to make or convert them in the digital format to store them in a database. So they can be used for extracting hidden insights along with the information. There are certain tools to extract the data from various written formats such documents, number plates etc,.

Here we used the Google cloud vision API for extracting the data from menu cards which we have taken one use case of online display of the food items along with their prices in the web apps, online sites where they have been manually entered in them. Google cloud vision API[1] identifies different languages and converts them into a text format.

The purpose of this project is to get the image from the user using a web site, this image is sent to the OCR. Finally the extracted text is sorted into Item and Price, this sorted Item and Price are displayed to the user. Users could upload one image at a time.

1.1 PROBLEM DEFINITION

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

1.2 PURPOSE

The purpose of this document is to identify characters from the written documents or in printed forms. Ability to sort the recognized handwritten character is a challenging area.

1.3 SCOPE

Based on our assessment of the applications in the field today, a majority of applications are used in the field of character recognition. The scope of this design document is to achieve the features of 3 the system such as pre-process the images, feature extraction, segmentation and display the text present in the image.

1.4 Technology Used

- Google Cloud Vision
- Python
- Flack

2. PROBLEM IDENTIFICATION

The progression of advanced technology has constantly intrigued us by leaving routine work as boring. The problem arises when there is a substantial loss of time and money on doing repetitive work daily where it can be incorporated with technology. The time that is wasted can be utilized to the development of advanced technology. In our daily life, taking from government offices to hospitals,

hotels to hostels most of the time there is a lot of time wasted by entering the forms into the database manually. Along with time, money is also wasted on investing them on the person who was manually entering the data.

If we take a scenario of online food order delivery. The content that we usually see in the application is manually entered. The vendor which receives the items list from his customers has to manually enter the details into the database where the web team will take further action on them to make visible to the consumers. The problem here to the vendor is that they are investing some amount of time to manually enter the data and also money to do that which in turn takes much long time to complete their customers request. This problem can be reduced by automating the process of manually entering the data into the database.

3. Objectives

The main objectives[6] that we have worked for to include in our solution are as follows:

- **Easy to use**

The software is easy to use with less knowledge on the background details and the process where there is not much complex structure is involved to complicate the issue and things can be done easily.

- **Less training cost**

There will be no much training involved for the vendors as the process is easy to use as the process is to upload the images from the customers side and run the program.

- **Better customer experience**

Better customer experience can be provided to the customers as their job is just to upload the images from their side and see the reflected data in their portals that is the e-commerce sites.

- **Provide accurate details**

The results obtained from the process are accurate as we have incorporated Google rest API which gives us more than 95% accuracy and the noisy data is cleaned from the images uploaded.

- **Reduce costs**

The main advantage is that we can reduce the cost of paying wages to the workers from the vendor side where they usually work on to manually enter the data from the uploaded images of the customer and also the respective costs can be reduced to the vendor on working with this automation.

- **Less expensive**

The whole setup requires less space to implement and less hardware which reduces the extra expenses to the vendor. It requires an open source development programming language python to be installed on their system and Google rest API which requires less amount of hardware related things like space, ram and also the internet connection.

- **Reduce time**

The other major advantage with this system is that we can reduce the time that is being invested on manually entering the data from the images those where uploaded by the customers.

- **Automating the process**

The entire process can be automated for better experience.

- **Less human interference**

We can reduce the human resources that are working on manually entering the data from the images and can reduce the costs related in turn.

4. SYSTEM METHODOLOGY

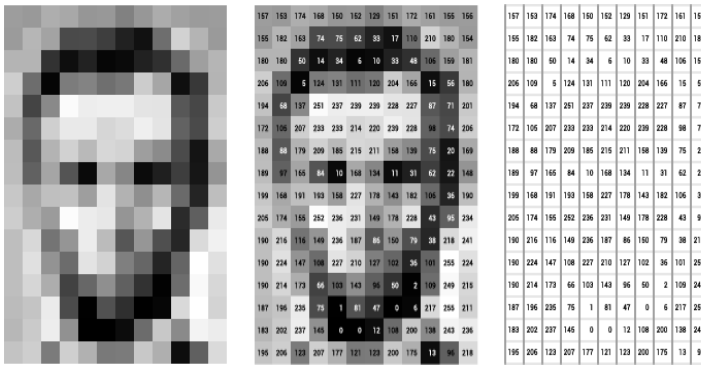
4.1 DATA EXTRACTION

To automate the process, first the data has to be extracted from the images through the Google rest API in the form of a string[9]. In the process of data extraction, the image has to be selected from the folder that we have saved all the images. The selected image has to be passed to the Google API call which works on OCR as backdrop and extracts the data in the form a string. In this process the entire data will be extracted including all the noisy data. Among all the languages we are primarily focusing on the English language.

Technology Used in OCR

Convolutional Neural Network

In Neural Network, CNN[3] is mainly used for image recognition, image classification etc. Computers see an image in a different way as compared to humans, the computer world consists of numbers and they are represented in the form of a 2-D array known as pixels.



Convolution of an image with different filters helps to improve edge detection, blur detection.

The Google Cloud Vision has developed the OCR which will pass the image through different filters. So that we will be having a better output while using Google Cloud Vision

OCR works in the following way

1. Partition the image into overlapping image tiles.
2. Feed each image tile to a small Neural Network.
3. Stores the result in a new array.
4. Downsampling.
5. Make Predictions.

4.2 DATA PRE-PROCESSING

The second step is data pre-processing where the noisy data has to be cleaned up and the data will be stored as a dictionary data structure in python for simplification of the process. The data that is extracted from an image through the Api call will be stored in the form of a string which includes noisy data along with the actual data that we are looking for. The noisy data will be cleaned by splitting the data based on the spaces and converted into a list by removing all the unwanted characters from the string on iterating the list and items as keys and prices as values were identified and stored in the dictionary format.

4.3 DISPLAYING DATA

The data dictionary elements which are items as keys and values as prices has to be displayed to the customers in the web pages and are done through the flask where we will upload the data to the customers in their portal.

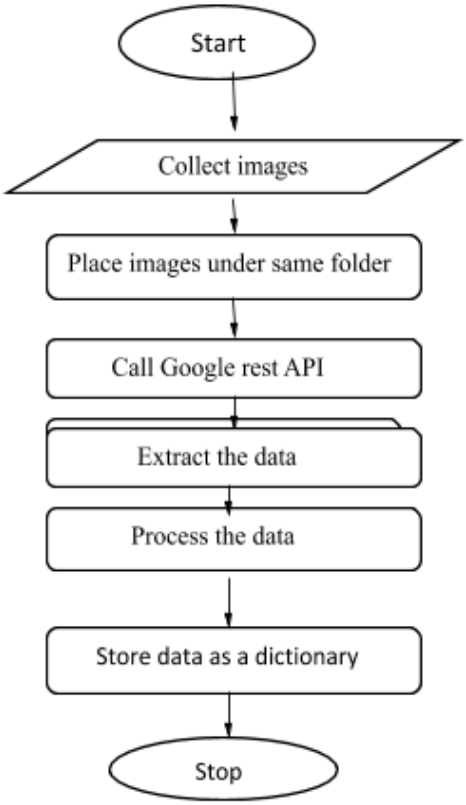
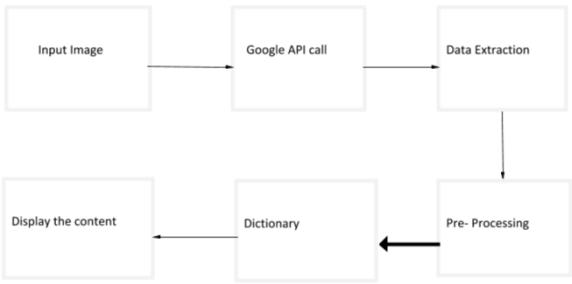


Chart -1: Flow Chart of Data Extraction and Data Pre-processing.

Above flow chart describes the process of images collection from the customers into a folder to the passing image to the Google rest API, from data extraction to the data pre-processing, from the data structure formation to the display of the content in the web based application with the food items as keys and prices as values.

5. Architecture of the system



First stage is the input stage, where the images where collected from the customers according to the use case that is chosen we collect menu cards from the customers and next stage is passing images to the Google API call and the data is collected from image in the form of a string and the string is pre-processed into a data dictionary which is in the form of a key-value pair and displays the content to the customers on the web based applications.

6. Time and space Complexity

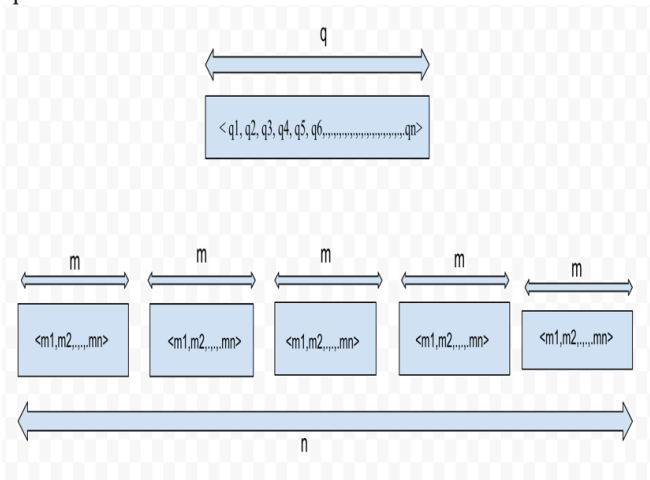
The main part that influences the time of the system is to remove the noisy data from the obtained string and convert the extracted data in the form a key-value pair. The noisy data includes the special characters(!@#$%^&*) and the data which is irrelevant.

Method 1

First, the string has to be converted into individual tokens based on the common part(spaces). Later, each token has to be selected and checked for the noisy data from each token. Finally, items and price values has to identified and grouped them as a data dictionary.

Let the length of the string be **q**,
the number of tokens be **n**,
the length of each token be **m**.
To remove the noisy data we iterate through each token which is n times and to each token whose length m for checking noisy data.

The total time complexity be O(n*m).
For checking a noisy data it requires a q operations as we check for all the q characters in the string and makes the equation as



T(n)= O(n*m)+q
- ⇨ O(n*n)+q. #m=n
- ⇨ O(n^2)+q.
- ⇨ O(n^2). #additional q operations can be negligible as it is a constant.

As in some cases the length of the token can be equal to the number of tokens and even more than that. So, m can be made n for making our calculation easy.

From the above calculation, it is clear that the time complexity is around O(n^2) in both average and worst case.

Best case
Let number of tokens be **n,**
the length of each token be **m,**
for each iteration through the list of tokens take O(n) and for checking noisy data it takes **q** operations and the equation is as follows

T(n) = O(n*m)+q
- ⇨ O(n*1)+q. # m = 1
- ⇨ O(n)+q
- ⇨ O(n)

Let the length of each token as 1, that is m = 1.
As the q operations is constant we can ignore it. So, time complexity in the best case would be O(n).

Space Complexity

Let the space utilized by the list of tokens be **n**,
Space for each token be **m,**
Variables utilized for data pre-processing is **6**.

S(n) = O(n*m)+6
- ⇨ O(n*n)+6
- ⇨ O(n^2)+6. # 6 as constant can be ignored
- ⇨ O(n^2)

From the above analysis it is clear that the space complexity in average and worst case would be O(n^2).

In case of best case,
The space for each token be **1**, that is **m =1**.
Space complexity equation is as follows,

S(n) = O(n*m)+6
- ⇨ O(n*1)+6
- ⇨ O(n)+6. # 6 as constant can be ignored
- ⇨ O(n)

Method 2

First, each character in a string will be validated for its relevancy and items and price values will be decided with the space separating each word in the string[19].

Let the length of the string be **n,**
As we iterate through the length of the string it takes O(n),
There will n operations to identify the noisy data.
So, time complexity is O(n)+n.
But the n operations can be ignored.

< n_1, n_2, n_3, n_4, n_5, n_6,.,.,.,.,.,.,.,.,.,.,.,.,.,.,.n_n>

Total time complexity is O(n) in all the cases(Best, Average, Worst).

Space complexity

Let the space utilized by the string be **n,**
Variables utilized those for data pre-processing is **6**.

S(n) = O(n)+6
⇨ O(n)+6 # 6 as constant can be ignored
⇨ O(n)

From the above analysis it is clear that the space complexity in all the cases(best, average and worst) would be O(n).

Performance Analysis			
Method 1	Best	Average	Worst
Time Complexity	O(n)	$O(n^2)$	$O(n^2)$
Space Complexity	O(n)	$O(n^2)$	$O(n^2)$
Method 2			
Time Complexity	O(n)	O(n)	O(n)
Space Complexity	O(n)	O(n)	O(n)

7. Results

For testing we are going to take some images interactively and we have placed them in a single folder. Before that, the server will be set to running state. From the web interface we can see an option to choose a file from the static file location that we have implicitly given and clicks on the upload button option to upload the image. Images in this context will be of menu cards that were collected from various restaurants.

Fig -1: Menu_1

We can see some noisy data that is irrelevant to the customer on the image that is uploaded. The image also includes some special characters like full stops and also some designs(printed format designs on the image). The uploaded images can have any format of text on it along with any kind of image.

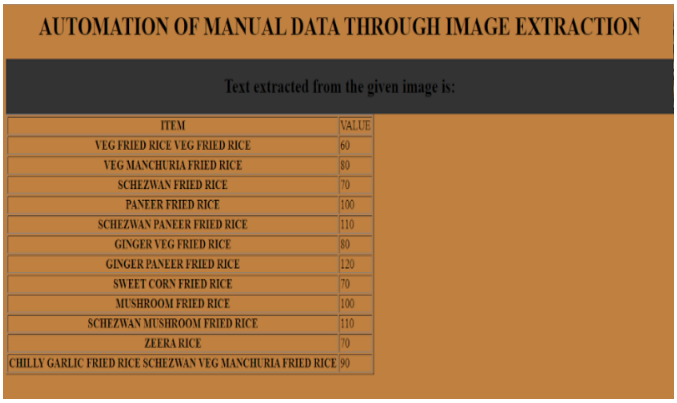


Fig -2: Output_1

8. Accuracy

Input consists of the data which includes both the items along with their prices and also the noisy data(irrelevant data). The key part of the solution to the problem is to remove the noisy data in the data pre-processing stage of the system. From the input Menu_1, we can see total 13 items along with their respective prices were passed to the system and where total 12 items along with prices obtained as output giving us more than 90 percent accuracy. There is a direct proportion of the accuracy to the result that we obtained from the Google Cloud Vision.

9. Other Use Cases

One of the major problems faced by many government organizations is the management of assets.

There are many issues every day regarding false claims of assets as the transactions involved are not properly tracked. This was due to the transformation from traditional methods to Digitalization. Earlier all the transactions are recorded on papers and are stored, which now become a problem while referring them. Our model can be used to efficiently convert all the old documents to the digitalized versions which can be stored in the blockchain. This makes asset verification easier and avoids false claims.

10. CONCLUSIONS

The process of data extraction and data pre-processing was implemented and tested successfully with accurate results. Some of our key findings are:

- Elimination of manual data entry.
- Reduction of time to the users.
- Reduction of costs, on the data entry team.
- Improves Confidentiality and Privacy
- Time complexity is o(n).
- Space complexity is o(n)+6.
- Reduces Human Errors.

11. FUTURE WORK

Our model has a wider scope in the current market and also in the near future. Various Domains had the necessity of automating the manual data entry, as in many work environments such as schools, hospitals, government offices, malls and also certain E-commerce companies such as Zomato, Swiggy, Uber eats, Flipkart, Amazon, SnapDeal etc, where data is being entered from the lists, menu cards and documents. There has been a lot of data which were manually entered by the workers into the database which is cost-intensive, time-consuming and prone to human errors.

Our future work focuses on designing a cost-effective model with high ease of use, containing web-based automation framework that automatically takes images in different formats from the customer and transfers them through a secure channel to the vendors where they can extract the data, pre-process it and store in multiple formats. Customers will also have the flexibility to choose the type of framework they require which will be linked to a compactable design on the backend. For example, if a customer chooses our framework to manage assets a blockchain-based[10] microservice deployed in the docker will be linked to the model However speed, accuracy and security will be considered as the highest priority and will be maintained throughout the product development life cycle.

Eliminating manual data entry from common paper-based workflows brings some obvious advantages. Chief among them is that recognition software is ultimately less expensive than human labour. Also, recognition software is more accurate, doesn't get tired, take a vacation or need an HR department to manage it. Automated processing solution also brings a higher level of security by eliminating privacy breaches caused by humans handling sensitive data. Additionally, if third-party processing is being utilized, organizations can be limited as to where the documents can be processed

REFERENCES

[1] Evelyn Chan, "Optimal Character Recognition With Google Cloud Vision API" jun, 2018

[2] Daphne Cornelisse, "An intuitive guide to Convolutional Neural Networks" Apr, 2018

[3] Adit Deshpande, "Beginners Guide to Understanding Convolutional Neural Networks" Engineer at Forward, UCLA CS'19

[4] Ujjwalkarn, "An Intuitive Explanation of Convolutional Neural Networks" Aug, 2016.

[5] Yunmai, "Use of Optical Character in our daily life: 9 usage of OCR", Agu, 2016. Software

[6] Nisha Gopinath Menon, "Everything You Need to know about Google Cloud Vision API", jun, 2017. cognitive Clouds.

[7] Adrian Frutiger, "Univers and OCR-B" Dec, 2015. multimediaman

[8] Krunal, "Python String Replace Example | Python Replace()" Dec, 2019. AppDivend

[9] Aishwarys Singh, "3 Beginner-Friendly Techniques to Extract Features from Image Data using Python" Aug 2019. Analytics Vidhya

[10] Melanie Swan " Blockchain: Blueprint for a New Economy" - 2015

[11] Xiwei Xu, Ingo Weber, Mark Staples "Architecture for Blockchain Applications"- 2019

[12] James Turnbull " The Docker Book: Containerization Is the New Virtualization "-2014

[13] Edwin M Sarmiento "The SQL Server DBA's Guide to Docker Containers" – 2020

[14] Adrian Mouat "Using Docker: Developing and Deploying Software with Containers" – 2015

[15] Binildas Christudas " Practical Microservices Architectural Patterns " – 2019

[16] Raymond Pairan "The Road to Microservice Architecture" – 2019

[17] Kasun Indrasiri, Prabath Siriwardena " Microservices for the Enterprise " – 2018

[18] John C. Cherniavsky, J.A. Storer "An Introduction to Data Structures and Algorithms" – 2012

[19] Dr. Basant Agarwal Benjamin Baka "Hands-On Data Structures and Algorithms with Python" -2018

BIOGRAPHIES

M. Lohit,
B.Tech Final year,
Dept. of Computer Science and Engineering, GITAM deemed to be University

P. Vijaya Manikanta Reddy,
B.Tech Final year,
Dept. of Computer Science and Engineering, GITAM deemed to be University

Bhargava Sai Gopal Kolli,
B.Tech Final year,
Dept. of Computer Science and Engineering, GITAM deemed to be University