An Experimental Performance Analysis on Robotics Process Automation (RPA) With Open Source OCR Engines: Microsoft Ocr And Google Tesseract OCR

Malathi T¹, Selvamuthukumaran D², Diwaan Chandar C S³, Niranjan V⁴, Swashthika A K⁵

¹Assistant Professor, Department of Computer Science & Engineering
²Assistant Professor, Department of Mechanical Engineering
³Second Year, Department of Information Science & Engineering
⁴,⁵Second Year, Department of Computer Technology

Bannari Amman Institute of Technology, Sathyamangalam, Erode, Tamilnadu - 638401

Email ID: tmalathime@gmail.com

Abstract. Robotic Process Automation (RPA) plays an important role in the solution for automating human digital interaction in various software domains like IT, Business Strategy and so on. Optical Character Recognition (OCR) superimposes subtitled characters on an image. Here we use two Open source OCR engines, Google Tesseract OCR - It literally makes use of the open source Tesseract OCR Engine [1]. Microsoft OCR - This is another open source OCR engine accessible in the Robotics Process Automation tool, UiPath [1]. The technique of optical character recognition (OCR) has been used to transform printed text into editable text. In various applications, OCR is a very useful and common technique. OCR precision can be dependent on algorithms for text preprocessing and segmentation. Often because of different scale, design, orientation, complex picture context, etc., it is difficult to retrieve text from the image. We begin this paper with an introduction to the Optical Character Recognition (OCR) process, in this paper, we propose a research analysis to estimate the accuracy of two open source OCR engines with Robotics Process Automation (RPA). These RPA processes are created using the UiPath tool and our results compare the OCR engines performance during the execution with the workflow and save the results to our local storage or on cloud storage if cloud support enabled and here we used local storage for purpose, and the result analysis and error analysis was evaluated by manual process which also included excel formulas to calculate the accuracy of the extracted string with the original string.

Keywords: Robotic process automation, optical Character recognition

1 Introduction

In this 21st century, we are in the new automation era. Great advances in automation, RPA and Machine Learning (ML) algorithms are used to free the workers from mundane, boring repetitive tasks and they focus on tasks which involve creativity and human decisions. Automation is a process of replacing a manual step with one that happens automatically. It’s not actually a human replacement or any cost saving play, It can save time, reduce errors, increase consistency and a way to centralise solutions and mistakes making them easier to fix to make organizations reinvent faster methods to explore their services to the society. RPA is a key way to achieve this. RPA is a tool used to imitate human digital interaction with the help of software bots.

RPA is not actually a humanoid robot which is built to resemble the human body. UiPath is helpful for beginners and marketed as “nocoding” solution. Learning a computer programming language is like learning something else. It will be easier for some people, and more difficult for others. But in RPA there
is no such problem. **RPA neither requires** the development of code, nor a direct access to the code [4] instead we just need to drag and drop the activities in the workflow.

### 2 OCR engines in RPA

1. Abbyy cloud OCR
2. Abbyy OCR
3. Google Cloud vision OCR
4. Microsoft Azure Computer Vision OCR
5. Microsoft OCR
6. Microsoft Project Oxford Online OCR
7. Tesseract OCR

#### 2.1 Microsoft OCR:

It uses built in Microsoft OCR if windows 10 is installed, else it will extract using Microsoft Office Document Imaging (MODI) OCR which is used in Microsoft Office (Up to MS office 2007)

#### 2.2 Google Tesseract OCR:

It obtains the string information from an indicated UI element or image using Tesseract OCR Engine.

| OCR Engines/Properties | Microsoft OCR | Tesseract OCR |
|------------------------|---------------|---------------|
| Language               | Microsoft OCR uses only “English” to take out the characters from the image file. It must be given as ”english” in the Engine properties | For the Google OCR engine, this field needs to contain the language file prefix, such as “en” for English, “ita” for Italian. |
| Profile                | To select a type for the image to recognize characters even more better. The options available in profile are: None - Recognize without a profile Screen - It is suitable for applications Scan - It is better for scanned documents. Legacy - It is the default profile in both the OCR engines | |
| Invert                 | NA            | The colors of the UI element are inverted before scraping. This is more familiar with dark background and light text. |
| Denied Characters      | NA            | The OCR engine extracts the given string without taking into account the characters specified here. |
| Extract words info     | To extract the position of the text, check the ExtractedWords in the properties window | NA |
| Extra Options          | NA            | AllowedCharacters - The OCR engine returns a given string according to the characters specified in this option |
| Scale                  | Scaling property is used to scale the image. It is used to resize the image for better performance. It is most recommended for smaller images |
3 Data Scraping using RPA OCR Engines

Google Tesseract and Microsoft OCR were the two open source OCR engines available in UiPath. There were several differences and similarities between these OCRs. Actually Both OCR uses a common architecture, and a scalable workflow to scrape data from the image variable and recognition capabilities, this experimental research defines how different aspects such as velocity, accuracy, time are more comfortable in each engines, where each engines has been tested with set of different sources, and implemented to run on certain Machines with specifications below.

### System Specifications

#### Minimal Hardware requirements
Central Processing Unit: Pentium 200 MHz
Random Access Memory: 32 MegaBytes
Hard Disk Space: 4 GigaBytes

#### Performed Specifications

**Tesseract OCR**
- Operating system: Windows 10
- Central processing unit: Core i5 i5-8265U
- Random access memory - 8 GigaBytes
- Processor: Intel core 15-8265 4 x 1.6-3.9GHz Whiskey Lake-U

**Microsoft OCR**
- Operating system: Windows 10
- Central processing unit: AMD Ryzen 3
- Random access memory - 4 GigaBytes
- Processor: AMD Pro A4-4350B R4, compute cores 2C+3G 2.50GH

### 4 Architecture & Process workflow

UiPath is one among the top tools available in the market of Robotic Process Automation[5]. It is used to automate Business processes and IT systems. Repetitive tasks may be boring for users. By the way UiPath is used to automate such boring tasks. Moreover UiPath is a code free environment. So, users with minimal or without coding background can automate the process using drag and drop functionality for all activities available in UiPath.

**4.1 UiPath Architecture**
The UiPath tool consists of three main components
- UiPath Studio
- UiPath Robot
- UiPath Orchestrator

**4.1.1 UiPath Studio**
UiPath is a designer tool using which a user can build workflows for automation tasks. It has several predefined activities and libraries. The users can drag and drop activities to model the workflow for the automation process. In simple terms, UiPath Studio is a tool which is used to model the automation workflow using predefined activities and libraries to automate repetitive processes.
4.1.2 **UiPath Robots**
UiPath Robots is a program that hosts the process that is built in UiPath studio, that lets us execute our projects on any machine with or without human monitoring (or) supervision.

4.1.3 **UiPath Orchestrator**
UiPath Orchestrator is a web platform that helps us to monitor our robots (i.e OCR workflows). It is a web application that lets us to manage the creation, monitoring, and deployment of our resources in our environment or machine. It allows us to launch and schedule our bots in a desktop or virtual machines, control status of the bot and analyse the results of their work.

4.2 **Workflow Design**

Workflows can be created in UiPath Studio and hence monitored in UiPath Orchestrator to save results to cloud storage or we can run the workflow in our local machine to save results in our Local Machine such as excel files and other word files.

Images used and Output:
https://github.com/Niranjan849/AN-EXPERIMENTAL-PERFORMANCE-ANALYSIS-ON-ROBOTICS-PROCESS-AUTOMATION-RPA-WITH-OPEN-SOURCE-OCR-ENGINEERING

Proposed approach:
5 Result analysis
Tabulation:
a.) Accuracy - The similarity between the original text and the extracted text is calculated and the average value is calculated among them
b.) Time taken - The time taken to extract 100 images is calculated and tabulated
c.) Average time taken - The average time taken to extract each image is calculated from the total time taken and is calculated and tabulated.

| Characteristics/ Engine | Microsoft OCR Engine | Tesseract OCR Engine |
|-------------------------|----------------------|----------------------|
| Average accuracy(%)    | 76.18835979          | 43.05                |
| Time taken(hrs)        | 00:25:00.00          | 00:20:32.00          |
| Average time taken(s)  | 00:00:15.00          | 00:00:12.32          |

6 Error analysis
Image properties
All the images were in .jpg or .JPEG format

When images with background were scraped, many additional text and special characters which were not in the images were extracted
Original Text
YOU DEFINE YOU.

Output Obtained
```
"= 1
fl',
., r - .w l.
~ YOU DEFINE YOU. 3
J
w, a;
`", 7", _", !`.4`
'··· -a", !!!.`
```

Data with more Accuracy:(one data for sample included)

**Good font for the OCR**

**Difficult font for the OCR**

Too small font for OCR

**Good font size for OCR**

Original Text
Good font for the OCR
Difficult font for the OCR
Too small font for the OCR
Good font size for the OCR

Output obtained
Good font for the OCR
Difficult font for the OCR
Too small font for OCR
Good font size for OCR
Images with Dark background and light text are not recognized

![Image of Dreams don't work unless you do]

*Figure 6.3*

**Original Text**
Dreams
Don’t work
unless
you do

**Extracted Text**
a: 1’... {-1
£9 unless’ij
/

*Vertical images are not recognized*

![Image of Vertical Label]

*Figure 6.4*

**Original Text**
Vertical Label

**Extracted Text**
·O
2
©
ad
S
2
~
dn
ov
>`
Similarity is calculated manually based on the count of recognized words which may contain many errors

Different set of images are used for Google tesseract and Microsoft OCR

| Tesseract OCR: | Original | Extracted Text |
|---------------|----------|----------------|
| < 20%         | Family WHERE LIFE BEGINS & LOVE NEVER ENDS | : WHERE Ll'lI BUSINS 5L l.()iT. NEVER [X |
| < 40%         | BE YOURSELF but be your best self | BE YOURS ELF M49101" MM |
| < 60%         | find your own way | [IND OWN 0 WAY $4 |
| < 80%         | I attend film school ny, background is acting and directing. But the acting does best, I like to work on the stage | I attend film school. my background is ucmlg and directing. But I'lce acting docs best. I like to work |
| 100%          | I'm such a hypocrite I tell people to stay strong when I am the weakest person in the world. | I'm such a hypocrite I tell people to stay strong when I am the weakest person in the world |

| Microsoft OCR: | original | Extracted Text |
|---------------|----------|----------------|
| < 25%         | life was is will be great just belive | LIFTGREAT. |
| < 50%         | The road to success is always under construction | TO success is uwaus umer construcTion. |
| < 75%         | Don't let yesterday take up too much of today | DON'T LET YESTER my TIAKE UP - roo MUCU90f |
| 100%          | Hello World! | Hello xWorld! |
7 Future works
There are several ways to extend this experiment. There were a lot of errors made during the experiment. 100 different images were used for Tesseract OCR and Microsoft OCR, That mistake would be rectified in the next experiment. Systems with different system requirements were used for the experiment and will use the same machine for extracting the images in the upcoming experiments. [7] The biggest problem made is that the similarity is checked manually which might not be much efficient and contains errors, so that could be rectified by using Levenshtein Distance Algorithm-String Comparison which is better in calculating the similarity between two strings and returns the similarity in percentage.

8 Conclusion
Bringing to an end of the experiment by having a squint at various frameworks such as accuracy, time taken. The results of microsoft OCR are more accurate when compared with the results of Tesseract OCR in many instances. But also in some cases, Tesseract OCR provides better results. We have used different types of images such as images with white background, dark background, different fonts, colored fonts, grey scaled, etc.[9] This provides different values of precision in Tesseract OCR and Microsoft OCR. But while considering the time taken to recognize the characters in the images, Tesseract OCR has provided a better result compared to Microsoft OCR. Also we had made some errors in calculating values for these parameters which we would rectify in our future research.

9 Acknowledgement
The authors thank Ms. Malathi T, Assistant Professor, Department of Computer Science & Engineering, Bannari Amman Institute of Technology for all the help she made to accomplish this research.

10 References
[1] https://support.bp-3.com/hc/en-us/articles/360017969373-Comparing-the-OCR-Engines-Available-in UiPath
[2] https://www.guru99.com/robotic-process-automation-tutorial.html
[3] R. Smith, “An Overview of the Tesseract OCR Engine.” In Proceedings of Document analysis and Recognition, ICDAR 2007, IEEE Ninth International Conference.
[4] The Tesseract open source OCR engine, http://code.google.com/p/tesseract-ocr, Tesseract Vs Gocr A Comparative Study 83Published By: Blue Eyes Intelligence Engineering & Sciences Publication, Retrieval Number: D0788092413 /2013©BEIESP
[5] https://www.edureka.co/blog/uipath-rpa-architecture/
[6] R.D Lins and N.F Alves, “A New Technique for Accessing the Performance of OCRs.” IADIS International Conference on Applied Computing, 2005.
[7] C.A.B Mello and R.D Lins, “A Comparative Study on OCR Tools.” Vision Interface '99, Trois-Rivières, Canada, 19-21 May.
[8] R. Mithe, S. Indalkar and N. Divekar, 2013, “Optical Character Recognition.” International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, 2(1).
[9] C. Patel, A. Patel and D. Patel, 2012, “Optical Character Recognition by Open Source OCR Tool Tesseract: A Case Study.” International Journal of Computer Applications (0975 – 8887) 55–100.