|  |
| --- |
|  |
| Case Study on OCR with UiPath |
| Experimenting various option with different OCR engine to solve RPA Problem |

|  |
| --- |
| Satish Prasad  1-20-2020 |

Table of Contents

[Step by Step Guide to use OCR Engines with UiPath 2](#_Toc30580252)

[Different Types of OCR 2](#_Toc30580253)

[Problem Statement 2](#_Toc30580254)

[Input Data 2](#_Toc30580255)

[Test Cases 2](#_Toc30580256)

[Output Data 3](#_Toc30580257)

[Prepare Environment 3](#_Toc30580258)

[Approach to Solve Problem 3](#_Toc30580259)

[Using Tesseract OCR 3](#_Toc30580260)

[Using Google Cloud OCR 4](#_Toc30580261)

[What is the difference between Google Cloud OCR and Google OCR? 4](#_Toc30580262)

[Using Microsoft Cloud OCR – 4](#_Toc30580263)

[Using Abbyy Cloud OCR 5](#_Toc30580264)

[Key Points you need to remember 7](#_Toc30580265)

[Conclusions 7](#_Toc30580266)

# Step by Step Guide to use OCR Engines with UiPath

## Different Types of OCR

* Tesseract /Google OCR - This actually uses the open source Tesseract OCR Engine, so it is free to use. Also, this processing is done on the local machine where UiPath is running.
* Google Cloud OCR - This requires a Google Cloud API Key, which has a free trial.
* Microsoft OCR - This uses the MODI OCR Engine, which is also free to use, and the processing is done locally like Google OCR.
* Microsoft Cloud OCR - This uses the Microsoft Computer Vision API, which is also free to sign up for.
* Abbyy OCR - This requires you to install Abbyy FineReader on your local machine and purchase a license.
* Abbyy Cloud OCR - This requires a subscription.

[There are few other option available but we selected top 6 to try and test …feel free to add your comments in case you done same with other OCR ]

## Problem Statement

The problem statement is generic in nature – What we are intended to do is to decide and see what works best with which OCR Engine…

We will build matrix by end of exercise to give idea about when to use which OCR…

On High-level -

1. Which OCR Engine works best with UiPath
2. Differences in terms of processing when using Paid vs Free OCR
3. Which one is recommended for Handwritten Materials (like meals receipts, hotels invoices, taxi fares, parking receipts)
4. Which engines read low-quality scans perfectly
5. What to consider before starting the OCR Project

## Input Data

(Set of 10 Pdf Files)

1. Few sample invoice
2. Few sample handwritten invoice
3. Few structured sample invoice

## Test Cases

## Output Data

Extracted data in (key, value) pair or excel sheet

## Prepare Environment

First thing first – what all dependencies need to be added into project?

This requires multiple things such as installing uipath package getting some external package installed, sign-up on external website and create keys for integration etc…

So we have broken the procedure step by step

# Approach to Solve Problem

# Using Tesseract OCR

* The Tesseract OCR engine used in UiPath is updated now to version 4.0. That contains an OCR engine - libtesseract and a command line program - tesseract.
* Tesseract 4 adds a new neural net (LSTM) based OCR engine which is focused on line recognition, but also still supports the legacy Tesseract OCR engine of Tesseract 3 which works by recognizing character patterns.

UiPath.Core.Activities.GoogleOCR

# Using Google Cloud OCR

### What is the difference between Google Cloud OCR and Google OCR?

* The difference is the engine as Google Could OCR is using the Google Cloud OCR engine and Google OCR is using Tesseract OCR Engine.
* Also, Google OCR is using the Tesseract engine which is deployed locally (comes with UiPath Studio) and the image processing and text extraction is done locally, on your computer. While Google Cloud OCR is uploading the image to be processed to Google server (cloud) and you get back the resulting text. So all the processing is done remotely on Google servers and you just get the result.
* Google OCR is free while you need to pay for Google Cloud OCR (free trial is available with limitation on Usages)

# Using Microsoft Cloud OCR –

Step to Start –

* Signup for free account on Azure or Login using your pay-as-you-go account
* Sign in into Azure portal and add Computer Vision
* Check how to embed Computer Vision with quickstarts and documentation.

Read more about - <https://azure.microsoft.com/en-us/services/cognitive-services/computer-vision/>

Extracting the text from images using Computer Vision API to extract printed and handwritten text from images/pdfs into machine-readable character stream is super easy all you need to know is the

1. Endpoints of vision api ;
2. Keys to connect those services

Bit caution here as azure provide two variants of Computer Vision API

1. Read API
2. OCR API

**Azure Computer Vision OCR API** recognizes printed text and supports a large variety of languages.

**Azure Computer Vision Read API** recognizes the handwritten and printed text, but temporary is available only in English.

The major difference among these two is that Read API uses the model that support only English language as of now while OCR supports more than 25 languages with auto detection and rotation of recognized text from Image.

Image Requirements –

* The image must be presented in JPEG, PNG, GIF, or BMP format
* The file size of the image must be less than 4 megabytes (MB)
* The dimensions of the image must be greater than 50 x 50 pixels
  + For the Read API, the dimensions of the image must be between 50 x 50 and 10000 x 10000 pixels.

Distill actionable information from images ( 5,000 transactions, 20 per minute.)

**Endpoints**

https://<<location>>.api.cognitive.microsoft.com/vision

Key 1: e1c530c745d7425eae1c056edcf4479c

Key 2: e6e8040014474c78839d221d3290a0b5

# Using Abbyy Cloud OCR

* Abbyy Cloud OCR SDK supports the recognition of printed text in more than 200 languages, including most Asian languages: Chinese, Japanese, Korean, Arabic, Farsi, Vietnamese, Thai and others using industry leading FineReader OCR technology.
* Abbyy Cloud OCR SDK recognizes both printed and hand-printed text within specific fields (zonal OCR).
* Its Cloud OCR recognition features are used for reading invoices, receipts, bills, business cards and many other document category. Not Only this it also support handwritten or manually filed forms extraction as well.
* Convert image/PDF to searchable PDF, PDF/A
* Convert image/PDF to Microsoft Word, Excel, PowerPoint

Must Read - Comparison Cloud OCR SDK vs. FineReader Engine SDK <https://abbyy.technology/en:features:comparisons:comp_onlinesdk-fre>

# Key Points you need to remember

1. You should note that in many cases, in order to get better OCR results, you'll need to **improve the quality of the image** you are giving to **OCR engine**.

# Conclusions

* Unsurprisingly, the paid OCR engines performed the best, especially with scanned documents. None of the engines read low-quality scans perfectly, but the cloud options were closest.
* If OCR is a key part of your project, I recommend trying all of your available options for the specific document types you're working with to find the best option that works within your project budget.

References –

1. Various OCR Activities - <https://docs.uipath.com/studio/docs/ocr-activities>
2. <https://docs.uipath.com/activities/docs/microsoft-azure-computer-vision-ocr>