www.bytescout.com

pdf ocr (optical character recognition) with pdf extractor sdk in VBScript using ByteScout PDF Suite

Learn to code in VBScript to make pdf ocr (optical character recognition) with pdf extractor sdk with this simple How-To tutorial

Every ByteScout tool includes sampleVBScript source codes that you can find here or in the folder with installed ByteScout product. Pdf ocr (optical character recognition) with pdf extractor sdk in VBScript can be applied with ByteScout PDF Suite. ByteScout PDF Suite is the set that includes 6 SDK products to work with PDF from generating rich PDF reports to extracting data from PDF documents and converting them to HTML. This bundle includes PDF (Generator) SDK, PDF Renderer SDK, PDF Extractor SDK, PDF to HTML SDK, PDF Viewer SDK and PDF Generator SDK for Javascript.

This rich and prolific sample source code in VBScript for ByteScout PDF Suite contains various functions and options you should do calling the API to implement pdf ocr (optical character recognition) with pdf extractor sdk. Follow the steps-by-step instructions from the scratch to work and copy and paste code for VBScript into your editor. VBScript application implementation mostly involves various stages of the software development so even if the functionality works please check it with your data and the production environment.

Our website gives free trial version of ByteScout PDF Suite. It includes all these source code samples with the purpose to assist you with your VBScript application implementation.

FOR MORE INFORMATION AND FREE TRIAL:

Download Free Trial SDK (on-premise version)

Read more about ByteScout PDF Suite

Explore API Documentation

Get Free Training for ByteScout PDF Suite

Get Free API key for Web API

visit www.ByteScout.com

Source Code Files:

ExtractTextFromImageOrScannedPdfUsingOcr.vbs

```
' This example demonstrates the use of Optical Character Recognition (OCR) to extract t
' from scanned PDF documents and raster images.


' Create TextExtractor object
Set extractor = CreateObject("Bytescout.PDFExtractor.TextExtractor")
extractor.RegistrationName = "demo"
extractor.RegistrationKey = "demo"

' Load sample PDF document
extractor.LoadDocumentFromFile("..\..\sample_ocr.pdf")


' Enable Optical Character Recognition (OCR)
extractor.OCRMode = 1 ' OCRMode.Auto = 1

' Set the location of OCR language data files
extractor.OCRLanguageDataFolder = "c:\Program Files\Bytescout PDF Extractor SDK\ocrdata

' Set OCR language
' "eng" for english, "deu" for German, "fra" for French, "spa" for Spanish etc - accord
extractor.OCRLanguage = "eng"
' Find more language files at https://github.com/bytescout/ocrdata

' Set PDF document rendering resolution
extractor.OCRResolution = 300


' You can also apply various preprocessing filters to improve the recognition on low-qu
' But they significantly hit the performance, so do not enable them by default.

' Automatically deskew skewed scans
'extractor.OCRImagePreprocessingFilters.AddDeskew()

' Remove vertical or horizontal lines (sometimes helps to avoid OCR engine's page segme
'extractor.OCRImagePreprocessingFilters.AddVerticalLinesRemover()
'extractor.OCRImagePreprocessingFilters.AddHorizontalLinesRemover()

' Repair broken letters
'extractor.OCRImagePreprocessingFilters.AddDilate()

' Remove noise
'extractor.OCRImagePreprocessingFilters.AddMedian()

' Apply Gamma Correction
'extractor.OCRImagePreprocessingFilters.AddGammaCorrection()

' Add Contrast
' extractor.OCRImagePreprocessingFilters.AddContrast(20)


' (!) You can use new OCRAnalyzer class to find an optimal set of image preprocessing
```

```
' filters for your specific document.
' See "OCR Analyser" example.


' Save extracted text to file
extractor.SaveTextToFile("output.txt")

WScript.Echo "Extracted text saved as 'output.txt'."

Set extractor = Nothing
```

VIDEO

https://www.youtube.com/watch?v=NEwNs2b9YN8

ON-PREMISE OFFLINE SDK

60 Day Free Trial or Visit ByteScout PDF Suite Home Page
Explore ByteScout PDF Suite Documentation
Explore Samples
Sign Up for ByteScout PDF Suite Online Training

ON-DEMAND REST WEB API

Get Your API Key
Explore Web API Docs
Explore Web API Samples

visit www.ByteScout.com

visit www.PDF.co