

## ocr with fast dataset with pdf extractor sdk in VBScript and ByteScout Data Extraction Suite

How to apply ByteScout Data Extraction Suite for ocr with fast dataset with pdf extractor sdk in VBScript

Every ByteScout tool includes sample VBScript source codes that you can find here or in the folder with installed ByteScout product. Ocr with fast dataset with pdf extractor sdk in VBScript can be applied with ByteScout Data Extraction Suite. ByteScout Data Extraction Suite is the set that includes 3 SDK products for data extraction from PDF, scans, images and from spreadsheets: PDF Extractor SDK, Data Extraction SDK, Barcode Reader SDK.

This rich and prolific sample source code in VBScript for ByteScout Data Extraction Suite contains various functions and options you should do calling the API to implement ocr with fast dataset with pdf extractor sdk. To use ocr with fast dataset with pdf extractor sdk in your VBScript project or application just copy & paste the code and then run your app! Check these VBScript sample code examples to see if they acknowledge to your needs and requirements for the project.

ByteScout Data Extraction Suite free trial version is available for download from our website. Free trial also includes programming tutorials along with source code samples.

FOR MORE INFORMATION AND FREE TRIAL:

[Download Free Trial SDK \(on-premise version\)](#)

[Read more about ByteScout Data Extraction Suite](#)

[Explore API Documentation](#)

[Get Free Training for ByteScout Data Extraction Suite](#)

[Get Free API key for Web API](#)

[visit www.Bytescout.com](http://www.Bytescout.com)

Source Code Files:

## ExtractTextFromImageOrScannedPdfUsingFastOcrDataset.vbs

```
' This example demonstrates the use of Optical Character Recognition (OCR) to extract text
' 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 - according to
extractor.OCRLanguage = "eng"
' Find more language files at https://github.com/bytescout/ocrdata/tree/master/ocrdata

' Set PDF document rendering resolution
extractor.OCRResolution = 300

' You can also apply various preprocessing filters to improve the recognition on low-quality scans
' 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 segmentation errors)
'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 Data Extraction Suite Home Page](#)  
[Explore ByteScout Data Extraction Suite Documentation](#)  
[Explore Samples](#)  
[Sign Up for ByteScout Data Extraction Suite Online Training](#)

## ON-DEMAND REST WEB API

[Get Your API Key](#)  
[Explore Web API Docs](#)  
[Explore Web API Samples](#)

[visit www.ByteScout.com](http://www.ByteScout.com)

[visit www.PDF.co](http://www.PDF.co)

[www.bytescout.com](http://www.bytescout.com)