

## ocr with mean dataset with pdf extractor sdk in VBScript with ByteScout PDF Suite

ocr with mean dataset with pdf extractor sdk in VBScript

Easy to understand coding instructions are written to assist you to try-out the features without the requirement to write your own code. Ocr with mean dataset 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.

Want to speed up the application development? Then this VBScript, code samples for VBScript, developers help to speed up the application development and writing a code when using ByteScout PDF Suite. To use ocr with mean dataset with pdf extractor sdk in your VBScript project or application just copy & paste the code and then run your app! Use of ByteScout PDF Suite in VBScript is also described in the documentation included along with the product.

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

Source Code Files:

## ExtractTextFromImageOrScannedPdfUsingMeanOcrDataset.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 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](http://www.ByteScout.com)

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

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