How to convert scanned pdf to json with pdf extractor sdk in VB.NET using ByteScout Premium Suite

Learn to code in VB.NET to convert scanned pdf to json with pdf extractor sdk with this step-by-step tutorial

Sample source code below will display you how to manage a complex task like convert scanned pdf to json with pdf extractor sdk in VB.NET. Want to convert scanned pdf to json with pdf extractor sdk in your VB.NET app? ByteScout Premium Suite is designed for it. ByteScout Premium Suite is the set that includes 12 SDK products from ByteScout including tools and components for PDF, barcodes, spreadsheets, screen video recording.

Want to save time? You will save a lot of time on writing and testing code as you may just take the VB.NET code from ByteScout Premium Suite for convert scanned pdf to json with pdf extractor sdk below and use it in your application. IF you want to implement the functionality, just copy and paste this code for VB.NET below into your code editor with your app, compile and run your application. Further improvement of the code will make it more robust.

ByteScout provides the free trial version of ByteScout Premium Suite along with the documentation and source code samples.

FOR MORE INFORMATION AND FREE TRIAL:

Download Free Trial SDK (on-premise version)

Read more about ByteScout Premium Suite

Explore API Documentation

Get Free Training for ByteScout Premium Suite

Get Free API key for Web API

visit www.ByteScout.com

Source Code Files:

Program.vb

```vb
Imports Bytescout.PDFExtractor

' This example demonstrates the use of Optical Character Recognition (OCR) to extract
' from scanned PDF documents and raster images.

' To make OCR work you should add the following references to your project:
' "Bytescout.PDFExtractor.dll", "Bytescout.PDFExtractor.OCRExtension.dll".

Class Program

    Friend Shared Sub Main(args As String())

        ' Create Bytescout.PDFExtractor.JSONExtractor instance
        Dim extractor As New JSONExtractor()
        extractor.RegistrationName = "demo"
        extractor.RegistrationKey = "demo"

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

        ' Enable Optical Character Recognition (OCR)
        ' in .Auto mode (SDK automatically checks if needs to use OCR or not)
        extractor.OCRMode = OCRMode.Auto

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

        ' Set OCR language
        extractor.OCRLanguage = "eng"  ' "eng" for english, "deu" for German, "fra" for
        ' 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-quality scans.

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

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

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

        ' Remove noise
        'extractor.OCRImagePreprocessingFilters.AddMedian()

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

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


        ' (!) You can use new OCRAnalyzer class to find an optimal set of image prepro
        ' filters for your specific document.
        ' See "OCR Analyser" example.


        ' Save extracted text to file
        extractor.SaveJSONToFile("output.json")

        ' Cleanup
        extractor.Dispose()

        ' Open output file in default associated application
        System.Diagnostics.Process.Start("output.json")

    End Sub

End Class
```

VIDEO

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

ON-PREMISE OFFLINE SDK

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

ON-DEMAND REST WEB API

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

visit www.ByteScout.com