

check if OCR is required for PDF in VB.NET using ByteScout PDF Extractor SDK

Tutorial: how to do check if OCR is required for PDF in VB.NET

Today you are going to learn how to check if OCR is required for PDF in VB.NET. ByteScout PDF Extractor SDK was made to help with check if OCR is required for PDF in VB.NET. ByteScout PDF Extractor SDK is the SDK that helps developers to extract data from unstructured documents, pdf, images, scanned and electronic forms. Includes AI functions like automatic table detection, automatic table extraction and restructuring, text recognition and text restoration from pdf and scanned documents. Includes PDF to CSV, PDF to XML, PDF to JSON, PDF to searchable PDF functions as well as methods for low level data extraction.

You will save a lot of time on writing and testing code as you may just take the code below and use it in your application. Follow the instruction from the scratch to work and copy and paste code for VB.NET into your editor. Further enhancement of the code will make it more vigorous.

ByteScout PDF Extractor SDK is available as free trial. You may get it from our website along with all other source code samples for VB.NET applications.

FOR MORE INFORMATION AND FREE TRIAL:

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

[Read more about ByteScout PDF Extractor SDK](#)

[Explore API Documentation](#)

[Get Free Training for ByteScout PDF Extractor SDK](#)

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

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

Source Code Files:

```
Imports Bytescout.PDFExtractor
```

```
Module Program
```

```
    Sub Main()
```

```
        Try
```

```
            ' Loop through all files in directory and check whether OCR operation is recommended
            For Each filePath As String In System.IO.Directory.GetFiles("InputFiles")
                _CheckOCRRequired(filePath)
            Next
```

```
        Catch ex As Exception
```

```
            Console.WriteLine("Error: " + ex.Message)
```

```
        End Try
```

```
        Console.WriteLine("Press enter key to exit...")
```

```
        Console.ReadLine()
```

```
    End Sub
```

```
''' <summary>
```

```
''' Check whether OCR Operation is required
```

```
''' </summary>
```

```
''' <param name="filePath"></param>
```

```
Private Sub _CheckOCRRequired(ByVal filePath As String)
```

```
    ' Read all file content...
```

```
    Using extractor As TextExtractor = New TextExtractor()
```

```
        extractor.RegistrationKey = "demo"
```

```
        extractor.RegistrationName = "demo"
```

```
    ' Load document
```

```
    extractor.LoadDocumentFromFile(filePath)
```

```
    Console.WriteLine("{1}*****{1}{1}FilePath: {0}", filePath, vbCrLf)
```

```
    Dim pageIndex As Int32 = 0
```

```
    ' Identify OCR operation is recommended for page
```

```
    If (extractor.IsOCRRecommendedForPage(pageIndex)) Then
```

```
        Console.WriteLine("{0}OCR Recommended: True", vbCrLf)
```

```
        ' 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 Extrac
```

```
        ' Set OCR language
```

```
        extractor.OCRLanguage = "eng" ' "eng" for english, "deu" for German, "
```

```
        ' Find more language files at https://github.com/bytescout/ocrdata
```

```
        ' Set PDF document rendering resolution
        extractor.OCRResolution = 300

    Else
        Console.WriteLine("{0}OCR Recommended: False", vbCrLf)
    End If

    ' Read all text
    Dim allExtractedText = extractor.GetText()
    Console.WriteLine("{1}Extracted Text:{1}{0}{1}{1}", allExtractedText, vbCrLf)

End Using

End Sub

End Module
```

VIDEO

https://www.youtube.com/watch?v=s28W3_KMraU

ON-PREMISE OFFLINE SDK

[60 Day Free Trial](#) or [Visit ByteScout PDF Extractor SDK Home Page](#)
[Explore ByteScout PDF Extractor SDK Documentation](#)
[Explore Samples](#)
[Sign Up for ByteScout PDF Extractor SDK 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)

