How to read text from noisy image with pdf extractor sdk in VB.NET and ByteScout Data Extraction Suite

This code in VB.NET shows how to read text from noisy image with pdf extractor sdk with this how to tutorial

Sample source code below will display you how to manage a complex task like read text from noisy image with pdf extractor sdk in VB.NET. 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 and you can use it to read text from noisy image with pdf extractor sdk with VB.NET.

Want to quickly learn? This fast application programming interfaces of ByteScout Data Extraction Suite for VB.NET plus the guidelines and the code below will help you quickly learn how to read text from noisy image with pdf extractor sdk. This VB.NET sample code is all you need for your app. Just copy and paste the code, add references (if needs to) and you are all set! Applying VB.NET application mostly includes various stages of the software development so even if the functionality works please test it with your data and the production environment.

ByteScout Data Extraction Suite free trial version is available on our website. VB.NET and other programming languages are supported.

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

Source Code Files:

Program.vb

```vb
Imports Bytescout.PDFExtractor

Module Program

    Sub Main()

        Try

            Using extractor As New TextExtractor()

                ' Load noisy image document
                extractor.LoadDocumentFromFile("sample.png")

                ' Set the font repairing OCR mode
                extractor.OCRMode = OCRMode.TextFromImagesAndVectorsAndRepairedFonts

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

                ' 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

                ' You can also apply various preprocessing filters
                ' to improve the recognition on low-quality scans.

                Console.WriteLine("Please wait while PDF Extractor SDK is processing n

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

                ' Remove vertical Or horizontal lines (sometimes helps to avoid OCR eng
                ' 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 OCRAnalyser class to find an optimal set of image
                ' filters for your specific document.
```

```vbnet
                ' See "OCR Analyser" example.

                ' Read all text
                Dim allText = extractor.GetText()

                Console.Clear()
                Console.WriteLine("Extracted Text: ")
                Console.WriteLine(allText)

            End Using

        Catch ex As Exception
            Console.Clear()
            Console.WriteLine("Exception: " + ex.Message)
        End Try

        Console.WriteLine("Press any key to exit...")
        Console.ReadLine()

    End Sub

End Module
```

VIDEO

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