www.bytescout.com

How to read text from noisy image with PDF extractor SDK in C# using ByteScout Data Extraction Suite

Learn to code in C# to read text from noisy image with PDF extractor SDK with this step-by-step tutorial

The documentation is designed for a specific purpose to help you to apply the features on your side. 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. It can be applied to read text from noisy image with PDF extractor SDK using C#.

Want to quickly learn? This fast application programming interfaces of ByteScout Data Extraction Suite for C# plus the guidelines and the code below will help you quickly learn how to read text from noisy image with PDF extractor SDK. Just copy and paste the code into your C# application's code and follow the instructions. If you want to use these C# sample examples in one or many applications then they can be used easily.

You can download free trial version of ByteScout Data Extraction Suite from our website with this and other source code samples for C#.

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.cs

```
using Bytescout.PDFExtractor;
using System;

namespace ReadTextFromNoisyImage
{
    class Program
    {
        static void Main(string[] args)
        {
            try
            {
                //Read all text from noisy image file
                using (TextExtractor extractor = new TextExtractor())
                {
                    // Load noisy image document
                    extractor.LoadDocumentFromFile("sample.png");

                    // Set the font repairing OCR mode
                    extractor.OCRMode = OCRMode.TextFromImagesAndVectorsAndRepairedFont

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

                    // Set OCR language
                    extractor.OCRLanguage = "eng"; // "eng" for english, "deu" for Ger
                                                   // Find more language files at https

                    // Set 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 Extra

                    // Automatically deskew skewed scans
                    extractor.OCRImagePreprocessingFilters.AddDeskew();

                    // Remove vertical or horizontal lines (sometimes helps to avoid O
                    //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
```

```
                    // filters for your specific document.
                    // See "OCR Analyser" example.

                    //Read all text
                    string allText = extractor.GetText();

                    Console.Clear();
                    Console.WriteLine("Extracted Text: \n\n" + allText);
                }
            }
            catch (Exception ex)
            {
                Console.Clear();
                Console.WriteLine("Exception: " + ex.Message);
            }

            Console.ReadLine();
        }
    }
}
```

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

visit www.PDF.co

www.bytescout.com