

www.bytescout.com

OCR analyser with PDF extractor SDK in C# and ByteScout Premium Suite

How to apply ByteScout Premium Suite for OCR analyser with PDF extractor SDK in C#

This page helps you to learn from code samples for programming in C#. ByteScout Premium Suite was made to help with OCR analyser with PDF extractor SDK in C#. ByteScout Premium Suite is the set that includes 12 SDK products from ByteScout including tools and components for PDF, barcodes, spreadsheets, screen video recording.

If you want to quickly learn then these fast application programming interfaces of ByteScout Premium Suite for C# plus the guideline and the C# code below will help you quickly learn OCR analyser with PDF extractor SDK. If you want to know how it works, then this C# sample code should be copied and pasted into your application's code editor. Then just compile and run it. Use of ByteScout Premium Suite in C# is also described in the documentation included along with the product.

Trial version along with the source code samples for C# can be downloaded from our website

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 \[www.ByteScout.com\]\(http://www.ByteScout.com\)](#)

Source Code Files:

Program.cs

```
using System;
using System.Drawing;
using System.Diagnostics;
using Bytescout.PDFExtractor;

// This example demonstrates the use of OCR Analyser - a tooling class for analysis of
// in PDF or raster image formats to find best parameters for Optical Character Recognition
// provide highest recognition quality.

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

namespace OCRAnalyser
{
    class Program
    {
        static void Main(string[] args)
        {
            // Input document
            string inputDocument = @".\sample_ocr.pdf";

            // Document page index
            int pageIndex = 0;

            // Area of the document page to perform the analysis (optional).
            // RectangleF.Empty means the full page.
            RectangleF rectangle = RectangleF.Empty; // new RectangleF(100, 50, 350, 250);

            // Location of language data files
            string ocrLanguageDataFolder = @"c:\Program Files\Bytescout PDF Extractor\Languages\en-US\";
            string ocrLanguage = "eng"; // "eng" for english, "deu" for German, "fra" for French
            // Find more language files at https://github.com/bytescout/ocrdata

            // Create OCRAnalyzer instance and activate it with your registration info
            using (OCRAnalyzer ocrAnalyzer = new OCRAnalyzer("demo", "demo"))
            {
                // Display analysis progress
                ocrAnalyzer.ProgressChanged += (object sender, string message, double progress)
                {
                    Console.WriteLine(message);
                };

                // Load document to OCRAnalyzer
                ocrAnalyzer.LoadDocumentFromFile(inputDocument);

                // Setup OCRAnalyzer
                ocrAnalyzer.OCRLanguage = ocrLanguage;
                ocrAnalyzer.OCRLanguageDataFolder = ocrLanguageDataFolder;

                // Set page area for analysis (optional)
                ocrAnalyzer.SetExtractionArea(rectangle);

                // Perform analysis and get results
                OCRAnalysisResults analysisResults = ocrAnalyzer.AnalyzeByOCRConfidence();
            }
        }
    }
}
```

```
// Now extract the text using detected OCR parameters

string outputDocument = @".\result.txt";

// Create TextExtractor instance
using (TextExtractor textExtractor = new TextExtractor("demo", "demo"))
{
    // Load document to TextExtractor
    textExtractor.LoadDocumentFromFile(inputDocument);

    // Setup TextExtractor
    textExtractor.OCRMode = OCRMode.Auto;
    textExtractor.OCRLanguageDataFolder = ocrLanguageDataFolder;
    textExtractor.OCRLanguage = ocrLanguage;

    // Apply analysis results to TextExtractor instance
    ocrAnalyzer.ApplyResults(analysisResults, textExtractor);

    // Set extraction area (optional)
    textExtractor.SetExtractionArea(rectangle);

    // Save extracted text to file
    textExtractor.SaveTextToFile(outputDocument);

    // Open result document in default associated application (for demo)
    ProcessStartInfo processStartInfo = new ProcessStartInfo(outputDocument);
    processStartInfo.UseShellExecute = true;
    Process.Start(processStartInfo);
}
}
}
}
```

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](#)

[visit www.PDF.co](#)

[www.bytescout.com](#)