

OCR with best dataset with PDF extractor SDK in C# with ByteScout Data Extraction Suite

Learn OCR with best dataset with PDF extractor SDK in C#

An easy to understand guide to learn how to OCR with best dataset with PDF extractor SDK in C#. OCR with best dataset with PDF extractor SDK in C# can be applied with ByteScout Data Extraction Suite. 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.

C# code snippet like this for ByteScout Data Extraction Suite works best when you need to quickly implement OCR with best dataset with PDF extractor SDK in your C# application. To use OCR with best dataset with PDF extractor SDK in your C# project or application just copy & paste the code and then run your app! Updated and detailed documentation and tutorials are available along with installed ByteScout Data Extraction Suite if you'd like to learn more about the topic and the details of the API.

On our website you may get trial version of ByteScout Data Extraction Suite for free. Source code samples are included to help you with your C# application.

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

Source Code Files:

```
using System.Diagnostics;
using 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'.

namespace OCRExample
{
    class Program
    {
        static void Main(string[] args)
        {
            // Create Bytescout.PDFExtractor.TextExtractor instance
            TextExtractor extractor = new TextExtractor();
            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";

            // Set OCR language
            extractor.OCRLanguage = "eng"; // "eng" for english, "deu" for German, "fre" for French
            // Find more language files at https://github.com/bytescout/ocrdata/tree/master

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

// Save extracted text to file
extractor.SaveTextToFile("output.txt");

// Cleanup
extractor.Dispose();

// Open result document in default associated application (for demo purpos
ProcessStartInfo processStartInfo = new ProcessStartInfo("output.txt");
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 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](http://www.ByteScout.com)

[visit www.PDF.co](http://www.PDF.co)

www.bytescout.com