

filter watermark text in VB.NET with ByteScout PDF Extractor SDK

ByteScout PDF Extractor SDK is the SDK is designed to help developers with pdf tables and pdf data extraction from unstructured documents like pdf, tiff, scans, images, scanned and electronic forms. The library is powered by OCR, computer vision and AI to provide unique functionality like table detection, automatic table structure extraction, data restoration, data restructuring and reconstruction. Supports PDF, TIFF, PNG, JPG images as input and can output CSV, XML, JSON formatted data. Includes full set of utilities like pdf splitter, pdf merger, searchable pdf maker and other utilities.

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:

ExtractText.VS2005.vbproj

```
<Project DefaultTargets="Build" xmlns="http://schemas.microsoft.com/developer/msbuild/2003">
  <PropertyGroup>
    <Configuration Condition="'$(Configuration)' == ''">Debug</Configuration>
    <Platform Condition="'$(Platform)' == ''">AnyCPU</Platform>
    <ProductVersion>8.0.50727</ProductVersion>
    <SchemaVersion>2.0</SchemaVersion>
    <ProjectGuid>{7B47AE2F-F68A-456C-A7E3-6FFA0FCF8AED}</ProjectGuid>
    <OutputType>Exe</OutputType>
    <AppDesignerFolder>Properties</AppDesignerFolder>
    <RootNamespace>ExtractText</RootNamespace>
    <AssemblyName>ExtractText</AssemblyName>
  </PropertyGroup>
</Project>
```

```

</PropertyGroup>
<PropertyGroup Condition=" '$(Configuration)|$(Platform)' == 'Debug|AnyCPU' ">
  <DebugSymbols>>true</DebugSymbols>
  <DebugType>full</DebugType>
  <Optimize>>false</Optimize>
  <OutputPath>bin\Debug\</OutputPath>
  <DefineConstants>DEBUG,TRACE</DefineConstants>
  <ErrorReport>prompt</ErrorReport>
  <WarningLevel>4</WarningLevel>
</PropertyGroup>
<PropertyGroup Condition=" '$(Configuration)|$(Platform)' == 'Release|AnyCPU' ">
  <DebugType>pdbonly</DebugType>
  <Optimize>>true</Optimize>
  <OutputPath>bin\Release\</OutputPath>
  <DefineConstants>TRACE</DefineConstants>
  <ErrorReport>prompt</ErrorReport>
  <WarningLevel>4</WarningLevel>
</PropertyGroup>
<Import Project="$(MSBuildBinPath)\Microsoft.VisualBasic.Targets" />
<ItemGroup>
  <Import Include="Microsoft.VisualBasic" />
  <Import Include="System" />
  <Reference Include="Bytescout.PDFExtractor, Version=1.0.0.12, Culture=neutral, processorArchitecture=MSIL">
    <SpecificVersion>False</SpecificVersion>
  </Reference>
  <Reference Include="System" />
  <Reference Include="System.Data" />
  <Reference Include="System.Xml" />
  <Compile Include="Program.vb" />
  <Compile Include="Properties\AssemblyInfo.vb" />
  <Content Include="..\..\DocumentWithWatermark.pdf">
    <CopyToOutputDirectory>Always</CopyToOutputDirectory>
    <Link>DocumentWithWatermark.pdf</Link>
  </Content>
</ItemGroup>
</ItemGroup>
</ItemGroup>
</Project>

```

ExtractText.VS2008.vbproj

```

<Project DefaultTargets="Build" xmlns="http://schemas.microsoft.com/developer/msbuild/2003" ToolsVersion="3.5">
  <PropertyGroup>
    <Configuration Condition=" '$(Configuration)' == '' ">Debug</Configuration>
    <Platform Condition=" '$(Platform)' == '' ">AnyCPU</Platform>
    <ProductVersion>9.0.21022</ProductVersion>
    <SchemaVersion>2.0</SchemaVersion>
    <ProjectGuid>{7B47AE2F-F68A-456C-A7E3-6FFA0FCF8AED}</ProjectGuid>
    <OutputType>Exe</OutputType>
    <AppDesignerFolder>Properties</AppDesignerFolder>
    <RootNamespace>ExtractText</RootNamespace>
    <AssemblyName>ExtractText</AssemblyName>
    <OldToolsVersion>2.0</OldToolsVersion>
    <TargetFrameworkVersion>v3.5</TargetFrameworkVersion>
  </PropertyGroup>
  <PropertyGroup Condition=" '$(Configuration)|$(Platform)' == 'Debug|AnyCPU' ">
    <DebugSymbols>>true</DebugSymbols>
    <DebugType>full</DebugType>
    <Optimize>>false</Optimize>
    <OutputPath>bin\Debug\</OutputPath>
    <DefineConstants>DEBUG,TRACE</DefineConstants>
    <ErrorReport>prompt</ErrorReport>
    <WarningLevel>4</WarningLevel>
  </PropertyGroup>
  <PropertyGroup Condition=" '$(Configuration)|$(Platform)' == 'Release|AnyCPU' ">
    <DebugType>pdbonly</DebugType>

```

```

<Optimize>true</Optimize>
<OutputPath>bin\Release\</OutputPath>
<DefineConstants>TRACE</DefineConstants>
<ErrorReport>prompt</ErrorReport>
<WarningLevel>4</WarningLevel>
</PropertyGroup>
<Import Project="$(MSBuildToolsPath)\Microsoft.VisualBasic.Targets" />
<ItemGroup>
  <Import Include="Microsoft.VisualBasic" />
  <Import Include="System" />
  <Reference Include="Bytescout.PDFExtractor, Version=1.0.0.12, Culture=neutral, processorArchitecture=MSIL">
    <SpecificVersion>False</SpecificVersion>
  </Reference>
  <Reference Include="System" />
  <Reference Include="System.Data" />
  <Reference Include="System.Xml" />
  <Compile Include="Program.vb" />
  <Compile Include="Properties\AssemblyInfo.vb" />
  <Content Include="..\..\DocumentWithWatermark.pdf">
    <CopyToOutputDirectory>Always</CopyToOutputDirectory>
    <Link>DocumentWithWatermark.pdf</Link>
  </Content>
</ItemGroup>
<ItemGroup>
</ItemGroup>
</Project>

```

ExtractText.VS2010.vbproj

```

<Project DefaultTargets="Build" xmlns="http://schemas.microsoft.com/developer/msbuild/2003" ToolsVersion="4.0">
  <PropertyGroup>
    <Configuration Condition="'$(Configuration)' == ''">Debug</Configuration>
    <Platform Condition="'$(Platform)' == ''">AnyCPU</Platform>
    <ProductVersion>
    </ProductVersion>
    <SchemaVersion>2.0</SchemaVersion>
    <ProjectGuid>{7B47AE2F-F68A-456C-A7E3-6FFA0FCF8AED}</ProjectGuid>
    <OutputType>Exe</OutputType>
    <AppDesignerFolder>Properties</AppDesignerFolder>
    <RootNamespace>ExtractText</RootNamespace>
    <AssemblyName>ExtractText</AssemblyName>
    <OldToolsVersion>3.5</OldToolsVersion>
    <TargetFrameworkVersion>v4.0</TargetFrameworkVersion>
  </PropertyGroup>
  <PropertyGroup Condition="'$(Configuration)|$(Platform)' == 'Debug|AnyCPU' ">
    <DebugSymbols>true</DebugSymbols>
    <DebugType>full</DebugType>
    <Optimize>>false</Optimize>
    <OutputPath>bin\Debug\</OutputPath>
    <DefineConstants>DEBUG,TRACE</DefineConstants>
    <ErrorReport>prompt</ErrorReport>
    <WarningLevel>4</WarningLevel>
  </PropertyGroup>
  <PropertyGroup Condition="'$(Configuration)|$(Platform)' == 'Release|AnyCPU' ">
    <DebugType>pdbonly</DebugType>
    <Optimize>true</Optimize>
    <OutputPath>bin\Release\</OutputPath>
    <DefineConstants>TRACE</DefineConstants>
    <ErrorReport>prompt</ErrorReport>
    <WarningLevel>4</WarningLevel>
  </PropertyGroup>
  <Import Project="$(MSBuildToolsPath)\Microsoft.VisualBasic.Targets" />
  <ItemGroup>
    <Import Include="Microsoft.VisualBasic" />
    <Import Include="System" />
    <Reference Include="Bytescout.PDFExtractor, Version=1.0.0.12, Culture=neutral, processorArchitecture=MSIL">
      <SpecificVersion>False</SpecificVersion>

```

```

</Reference>
<Reference Include="System" />
<Reference Include="System.Data" />
<Reference Include="System.Xml" />
<Compile Include="Program.vb" />
<Compile Include="Properties\AssemblyInfo.vb" />
<Content Include="..\..\DocumentWithWatermark.pdf">
  <CopyToOutputDirectory>Always</CopyToOutputDirectory>
  <Link>DocumentWithWatermark.pdf</Link>
</Content>
</ItemGroup>
<ItemGroup>
</ItemGroup>
</Project>

```

Program.vb

```

Imports Bytescout.PDFExtractor

Class Program
  Friend Shared Sub Main(args As String())

    ' Create Bytescout.PDFExtractor.TextExtractor instance
    Dim extractor As New TextExtractor()
    extractor.RegistrationName = "demo"
    extractor.RegistrationKey = "demo"

    ' Load sample PDF document
    extractor.LoadDocumentFromFile("..\DocumentWithWatermark.pdf")

    ' Filter text using text filter
    extractor.AddFilter("^COPY{code}quot;; caseSensitive:=True, useRegex:=True)

    ' Filter text using appearance filter
    ' extractor.AddFilter("Arial", fontSize:=203, exclude:=True)

    ' Save extracted text to file
    extractor.SaveTextToFile("..\result.txt")

    ' Cleanup
    extractor.Dispose()

    ' Open result file in default associated application (for demo purposes)
    System.Diagnostics.Process.Start("..\result.txt")
  End Sub
End Class

```

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)

www.bytescout.com