

How to recognize text from document in PowerShell using ByteScout Text Recognition SDK

Tutorial on how to recognize text from document in PowerShell

Learn how to recognize text from document in PowerShell with this source code sample. ByteScout Text Recognition SDK: the text recognition SDK to help with extraction of text using OCR from scanned images and documents. Supports English and non-Latin languages, can take PDF as input. It can recognize text from document in PowerShell.

You will save a lot of time on writing and testing code as you may just take the PowerShell code from ByteScout Text Recognition SDK for recognize text from document below and use it in your application. In order to implement the functionality, you should copy and paste this code for PowerShell below into your code editor with your app, compile and run your application. Implementing PowerShell application typically includes multiple stages of the software development so even if the functionality works please test it with your data and the production environment.

Our website provides trial version of ByteScout Text Recognition SDK for free. It also includes documentation and source code samples.

PowerShell - GeneralExample.ps1

```
# Add reference to ByteScout.TextRecognition.dll assembly
Add-Type -Path "c:\Program Files\ByteScout Text Recognition
SDK\net40\ByteScout.TextRecognition.dll"

$InputDocument = "..\..\invoice-sample.png"
$OutputDocument = ".\result.txt"

# Create and activate TextRecognizer instance
$textRecognizer = New-Object ByteScout.TextRecognition.TextRecognizer
$textRecognizer.RegistrationName = "demo"
$textRecognizer.RegistrationKey = "demo"

try {
    # Load document (image or PDF)
    $textRecognizer.LoadDocument($InputDocument)

    # Set the location of OCR language data files
    $textRecognizer.OCRLanguageDataFolder = "c:\Program Files\ByteScout Text
Recognition SDK\ocrdata_best\"

    # Set OCR language.
    # "eng" for english, "deu" for German, "fra" for French, "spa" for Spanish, etc.
    - according to files in "ocrdata" folder
```

```
# Find more language files at https://github.com/bytescout/ocrdata
$textRecognizer.OCRLanguage = "eng"

# Recognize text from all pages and save it to file
$textRecognizer.SaveText($OutputDocument)

# Open the result file in default associated application (for demo purposes)
& $OutputDocument
}
catch {
    # Display exception
    Write-Host $_.Exception.Message
}

$textRecognizer.Dispose()
```

PowerShell - run.bat

```
@echo off

powershell -NoProfile -ExecutionPolicy Bypass -Command "& .\GeneralExample.ps1"
echo Script finished with errorlevel=%errorlevel%

pause
```

FOR MORE INFORMATION AND FREE TRIAL:

[Download Free Trial SDK \(on-premise version\)](#)

[Read more about ByteScout Text Recognition SDK](#)

[Explore documentation](#)

[Visit www.ByteScout.com](#)

or

[Get Your Free API Key for www.PDF.co Web API](#)