Aquaforest.ExtendedOCR.Api Namespace

Classes

	Class	Description
	AdvancedDespeckle	The AdvancedDespeckle class provides advanced image noise reduction features by image despeckle filter.
	Binarization	The Binarization class offers image binarization features. Binarization is the first stage in image recognition. It converts the input image into a black and white image for faster processing.
	BlankPageRemoval	This class offers blank page detection and removal features.
	IHQCCompression	This class manages the intelligent High Quality Compression parameters for PDF(EnablePdfOutput=true) and XPS (EnableXpsOutput=true) output formats. **Note:** IHQCCompression cannot be used when OCRing PDF in Native mode with EnablePdfOutput set to true.
	Interpolation	This class offers image interpolation features.
	LicenseInfo	
	LineRemoval	LineRemoval class provides line removal capabilities. The class contains algorithms for removing vertical and horizontal lines in an image. The integrator can control various settings in line removal algorithm like: • minimum and maximum line length and thickness of horizontal lines • minimum and maximum line length and thickness of vertical lines • clean noisy pixels adjacent to the line • maximum gap between the lines
	Ocr	
	PDFImageCompression	This class manages the PDF image compression parameters
	PreProcessor	

Delegates

	Delegate	Description
	Ocr.StatusUpdateEventHandler	

Enumerations

	Enumeration	Description
	AsianOCREngine	The Asian OCR engine to use.
	BinarizationMode	The binarization algorithm to use.

⚙	ExtractImageMethod	Whether to convert the images in a PDF document to TIFF or not.
⚙	IHQCCompressionLevel	The IHQC compression level. **Note:** PDF 1.4 A-1a and PDF 1.4 A-1b only supports level 1.
⚙	IHQCQualityFactor	The IHQC quality factor.
⚙	InterpolationMode	Interpolation algorithms.
⚙	Jpeg2000CompressionMode	The JPEG 2000 compression mode.
⚙	Layout	The layout to use when generating non-PDF documents such as DOCX and RTF.
⚙	NonImagePDF	This allows control over the treatment of non image-only PDFs, i.e. PDFs that have some text in them as well as images.
⚙	PdfToImageBpp	The Bits Per Pixel to use for the rasterized PDF page. This only applies for documents that are not processed using Native mode.
⚙	PdfToImageCompression	The compression to use for images extracted or rasterized from each PDF page.
⚙	PdfToImageDpi	The DPI to use for the rasterized page.
⚙	PDFVersion	Set the PDF Version of the output file. **Note:**This will only affect documents that are not processed natively and documents that are converted to PDF/A.
⚙	SegmentationMode	IHQC segmentation modes.
⚙	SupportedLanguages	Extended OCR accepts up to 8 recognition languages at a time. This is helpful to process mixed documents but, because of the various character sets, not all combinations are allowed. For this reason, the multiple languages support is limited to a single alphabet. For example, Russian and French can't be mixed. **Note:** Asian languages can't be mixed. - Extended OCR SDK cannot load more than one Asian language at a time.
⚙	TextDirection	The text direction (for Asian languages only).
⚙	TextSpacing	The text spacings
⚙	TextType	The text types

AdvancedDespeckle Class

The AdvancedDespeckle class provides advanced image noise reduction features by image despeckle filter.

Inheritance Hierarchy

System.Object
  Aquaforest.ExtendedOCR.Api.AdvancedDespeckle

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public sealed class AdvancedDespeckle`

VB
`Public NotInheritable Class AdvancedDespeckle`

The **AdvancedDespeckle** type exposes the following members.

Constructors

	Name	Description
	AdvancedDespeckle	Initializes a new instance of the **AdvancedDespeckle** class

Properties

	Name	Description
	RemoveWhitePixels	By default, **AdvancedDespeckle** removes black pixels. If this setting is set to 'true', white pixels will be removed instead of black pixels.
	SpeckleSize	The size of speckles to remove

See Also

Aquaforest.ExtendedOCR.Api Namespace

AdvancedDespeckle Constructor

Initializes a new instance of the AdvancedDespeckle class

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public AdvancedDespeckle()

VB
Public Sub New

See Also

AdvancedDespeckle Class
Aquaforest.ExtendedOCR.Api Namespace

AdvancedDespeckle.AdvancedDespeckle Properties

The AdvancedDespeckle type exposes the following members.

Properties

	Name	Description
	RemoveWhitePixels	By default, AdvancedDespeckle removes black pixels. If this setting is set to 'true', white pixels will be removed instead of black pixels.
	SpeckleSize	The size of speckles to remove

See Also

AdvancedDespeckle Class

Aquaforest.ExtendedOCR.Api Namespace

AdvancedDespeckle.RemoveWhitePixels Property

By default, AdvancedDespeckle removes black pixels. If this setting is set to 'true', white pixels will be removed instead of black pixels.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> RemoveWhitePixels { get; set; }```

VB
```Public Property RemoveWhitePixels As Nullable(Of Boolean)   Get   Set```

Property Value
Type: Nullable(Boolean)

See Also

AdvancedDespeckle Class
Aquaforest.ExtendedOCR.Api Namespace

AdvancedDespeckle.SpeckleSize Property

The size of speckles to remove

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> SpeckleSize { get; set; }```

VB

```Public Property SpeckleSize As Nullable(Of Integer)```
```   Get```
```   Set```

Property Value
Type: Nullable(Int32)

See Also

AdvancedDespeckle Class
Aquaforest.ExtendedOCR.Api Namespace

AsianOCREngine Enumeration

The Asian OCR engine to use.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public enum AsianOCREngine

VB
Public Enumeration AsianOCREngine

Members

Member name	Value	Description
OcrAsian1	0	
OcrAsian2	1	

See Also

Aquaforest.ExtendedOCR.Api Namespace

Binarization Class

The Binarization class offers image binarization features. Binarization is the first stage in image recognition. It converts the input image into a black and white image for faster processing.

Inheritance Hierarchy

System.Object
  Aquaforest.ExtendedOCR.Api.Binarization

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public sealed class Binarization

VB
Public NotInheritable Class Binarization

The **Binarization** type exposes the following members.

Constructors

	Name	Description
⬤	Binarization	Initializes a new instance of the **Binarization** class

Properties

	Name	Description
	BinarizationMode	The binarization algorithm to use. Computes a black and white image with the input color/gray-scale image by using the smoothing, brightness, contrast levels and the threshold (for Threshold binarization only) stored in the instance of this class.
	Binarize	Whether or not to perform binarization on the document. The default value is 'false'.
	Brightness	The brightness threshold (higher values will darker the result). Enhances the image by increasing/decreasing the luminosity of the image. The default value is 0. **Note:** For this setting to work, Binarize needs to be set to 'true'.
	Contrast	The contrast threshold (lower values will darker the result).

		Enhances the image by increasing/decreasing the difference between image elements and their background. The default value is 40. **Note:** For this setting to work, Binarize needs to be set to 'true'.
	SmoothingLevel	Smoothing may be useful to binarize text with a colored background in order to avoid noisy pixels. (0 disables smoothing, higher values smooths more).
	Threshold	Sets the threshold for fixed threshold binarization (0 for automatic threshold computation). The default value is 128.

See Also

Aquaforest.ExtendedOCR.Api Namespace

Binarization Constructor

Initializes a new instance of the Binarization class

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public Binarization()`

VB
`Public Sub New`

See Also

Binarization Class
Aquaforest.ExtendedOCR.Api Namespace

Binarization.Binarization Properties

The Binarization type exposes the following members.

Properties

	Name	Description
📄	BinarizationMode	The binarization algorithm to use. Computes a black and white image with the input color/gray-scale image by using the smoothing, brightness, contrast levels and the threshold (for Threshold binarization only) stored in the instance of this class.
📄	Binarize	Whether or not to perform binarization on the document. The default value is 'false'.
📄	Brightness	The brightness threshold (higher values will darker the result). Enhances the image by increasing/decreasing the luminosity of the image. The default value is 0. **Note:** For this setting to work, Binarize needs to be set to 'true'.
📄	Contrast	The contrast threshold (lower values will darker the result). Enhances the image by increasing/decreasing the difference between image elements and their background. The default value is 40. **Note:** For this setting to work, Binarize needs to be set to 'true'.
📄	SmoothingLevel	Smoothing may be useful to binarize text with a colored background in order to avoid noisy pixels. (0 disables smoothing, higher values smooths more).
📄	Threshold	Sets the threshold for fixed threshold binarization (0 for automatic threshold computation). The default value is 128.

See Also

Binarization Class
Aquaforest.ExtendedOCR.Api Namespace

Binarization.BinarizationMode Property

The binarization algorithm to use.

 Computes a black and white image with the input color/gray-scale image by using the smoothing, brightness, contrast levels and the threshold (for Threshold binarization only) stored in the instance of this class.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<BinarizationMode> BinarizationMode { get; set; }```

VB

```Public Property BinarizationMode As Nullable(Of BinarizationMode)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(BinarizationMode)

See Also

Binarization Class
Aquaforest.ExtendedOCR.Api Namespace

Binarization.Binarize Property

Whether or not to perform binarization on the document. The default value is 'false'.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public bool Binarize { get; set; }```

VB

```Public Property Binarize As Boolean```
      ```Get```
      ```Set```

Property Value
Type: Boolean

See Also

Binarization Class
Aquaforest.ExtendedOCR.Api Namespace

Binarization.Brightness Property

The brightness threshold (higher values will darker the result).

Enhances the image by increasing/decreasing the luminosity of the image.

The default value is 0.

Note:

For this setting to work, Binarize needs to be set to 'true'.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> Brightness { get; set; }```

VB
```Public Property Brightness As Nullable(Of Integer)   Get   Set```

Property Value
Type: Nullable(Int32)

See Also

Binarization Class
Aquaforest.ExtendedOCR.Api Namespace

Binarization.Contrast Property

The contrast threshold (lower values will darker the result).

Enhances the image by increasing/decreasing the difference between image elements and their background.

The default value is 40.

Note:

For this setting to work, Binarize needs to be set to 'true'.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> Contrast { get; set; }```

VB

```Public Property Contrast As Nullable(Of Integer)```
```   Get```
```   Set```

Property Value
Type: Nullable(Int32)

See Also

Binarization Class
Aquaforest.ExtendedOCR.Api Namespace

Binarization.SmoothingLevel Property

Smoothing may be useful to binarize text with a colored background in order to avoid noisy pixels. (0 disables smoothing, higher values smooths more).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```csharp
public Nullable<int> SmoothingLevel { get; set; }
```

VB
```vb
Public Property SmoothingLevel As Nullable(Of Integer)
      Get
      Set
```

Property Value
Type: Nullable(Int32)

See Also

Binarization Class
Aquaforest.ExtendedOCR.Api Namespace

Binarization.Threshold Property

Sets the threshold for fixed threshold binarization (0 for automatic threshold computation). The default value is 128.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> Threshold { get; set; }```

VB

```Public Property Threshold As Nullable(Of Integer)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Int32)

See Also

Binarization Class
Aquaforest.ExtendedOCR.Api Namespace

BinarizationMode Enumeration

The binarization algorithm to use.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public enum BinarizationMode

VB
Public Enumeration BinarizationMode

Members

Member name	Value	Description
Threshold	0	Applies the fixed threshold for binarization. If the threshold provided is 0, the image is analyzed to automatically compute a threshold for the whole document.
Adaptive	1	Performs adaptive binarization. The adaptive binarization determines local thresholds.

See Also

Aquaforest.ExtendedOCR.Api Namespace

BlankPageRemoval Class

This class offers blank page detection and removal features.

Inheritance Hierarchy

System.Object
  Aquaforest.ExtendedOCR.Api.BlankPageRemoval

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public sealed class BlankPageRemoval`

VB
`Public NotInheritable Class BlankPageRemoval`

The **BlankPageRemoval** type exposes the following members.

Constructors

	Name	Description
	BlankPageRemoval	Initializes a new instance of the **BlankPageRemoval** class

Properties

	Name	Description
	DetectBlankPage	Detect blank pages in the source document. The default is 'true'. **Note:** This will not remove the blank pages detected. To remove blank pages set RemoveBlankPage to 'true' and ExtractImageMethod to ConvertToTiff.
	RemoveBlankPage	Remove blank pages from the output document. **Note:** You also need to set ExtractImageMethod to ConvertToTiff. The default value is 'false'.
	Sensitivity	The sensitivity (1 -100) for blank page detection. With high sensitivity, less blank pages are detected. Default value is 1.

See Also

Aquaforest.ExtendedOCR.Api Namespace

BlankPageRemoval Constructor

Initializes a new instance of the BlankPageRemoval class

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public BlankPageRemoval()

VB
Public Sub New

See Also

BlankPageRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

BlankPageRemoval.BlankPageRemoval Properties

The BlankPageRemoval type exposes the following members.

Properties

	Name	Description
	DetectBlankPage	Detect blank pages in the source document. The default is 'true'. **Note:** This will not remove the blank pages detected. To remove blank pages set RemoveBlankPage to 'true' and ExtractImageMethod to ConvertToTiff.
	RemoveBlankPage	Remove blank pages from the output document. **Note:** You also need to set ExtractImageMethod to ConvertToTiff. The default value is 'false'.
	Sensitivity	The sensitivity (1 -100) for blank page detection. With high sensitivity, less blank pages are detected. Default value is 1.

See Also

BlankPageRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

BlankPageRemoval.DetectBlankPage Property

Detect blank pages in the source document. The default is 'true'.

Note: This will not remove the blank pages detected. To remove blank pages set RemoveBlankPage to 'true' and ExtractImageMethod to ConvertToTiff.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public bool DetectBlankPage { get; set; }```

VB

```Public Property DetectBlankPage As Boolean```
```   Get```
```   Set```

Property Value
Type: Boolean

See Also

BlankPageRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

BlankPageRemoval.RemoveBlankPage Property

Remove blank pages from the output document. **Note:** You also need to set ExtractImageMethod to ConvertToTiff. The default value is 'false'.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public bool RemoveBlankPage { get; set; }```

VB

```Public Property RemoveBlankPage As Boolean```
 ```   Get```
 ```   Set```

Property Value
Type: Boolean

See Also

BlankPageRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

BlankPageRemoval.Sensitivity Property

The sensitivity (1 -100) for blank page detection. With high sensitivity, less blank pages are detected. Default value is 1.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public Nullable<int> Sensitivity { get; set; }

VB

Public Property Sensitivity As Nullable(Of Integer)
        Get
        Set

Property Value
Type: Nullable(Int32)

See Also

BlankPageRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

ExtractImageMethod Enumeration

Whether to convert the images in a PDF document to TIFF or not.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public enum ExtractImageMethod`

VB
`Public Enumeration ExtractImageMethod`

Members

Member name	Value	Description
Native	0	This method places the OCRed text directly into a copy of the original PDF rather than creating an entirely new PDF.
ConvertToTiff	1	The PDF is rasterized to create an image of each page in the PDF.

See Also

Aquaforest.ExtendedOCR.Api Namespace

IHQCCompression Class

This class manages the intelligent High Quality Compression parameters for PDF(EnablePdfOutput=true) and XPS (EnableXpsOutput=true) output formats.
Note: IHQCCompression cannot be used when OCRing PDF in Native mode with EnablePdfOutput set to true.

Inheritance Hierarchy

System.Object
  Aquaforest.ExtendedOCR.Api.IHQCCompression

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public sealed class IHQCCompression`

VB
`Public NotInheritable Class IHQCCompression`

The **IHQCCompression** type exposes the following members.

Constructors

	Name	Description
	IHQCCompression	Initializes a new instance of the **IHQCCompression** class

Properties

	Name	Description
	AdvancedBrightness	Contact technical support (support@aquaforest.com) for guidance on using this property.
	AdvancedContrast	Contact technical support (support@aquaforest.com) for guidance on using this property.
	AdvancedFactorPhoto	Contact technical support (support@aquaforest.com) for guidance on using this property.
	AdvancedSmooth	Contact technical support (support@aquaforest.com) for guidance on using this property.
	AutoAdjust	Contact technical support (support@aquaforest.com) for guidance on using this property.
	EnableIHQCCompression	Apply Intelligent High Quality Compression. The default value is 'false'.
	HighResolutionForeground	High resolution foreground feature enables a specific mode causing the use of high resolution for the foreground image.

		HighResolutionForeground enables a specific mode for working around a bug existing in the native readers of some platforms that does not read properly PDF (Mainly the native reader on MAC OS X and IOS). Those native readers do not follow some of the specification of the PDF format. In place of displaying the correct compressed image, they display a down-sampled version of it. This special high resolution foreground mode offers a workaround to customer who wants to use iHQC with these buggy applications. Most of the PDF readers, including Adobe Reader are able to read PDF with IHQC compression properly on any platform.
	IHQCCompressionLevel	The compression level to be used. Level 1 is the basic compression level. Level 3 is the most advanced intelligent High Quality Compression mode. **Note:** PDF 1.4 A-1a and PDF 1.4 A-1b only supports level 1.
	IHQCQualityFactor	Set the IHQC quality factor.
	IHQCThreshold	Contact technical support (support@aquaforest.com) for guidance on using this property.
	MinOutputResolution	Contact technical support (support@aquaforest.com) for guidance on using this property.
	SegmentationMode	Contact technical support (support@aquaforest.com) for guidance on using this property.
	Smoothing	Contact technical support (support@aquaforest.com) for guidance on using this property.

See Also
Aquaforest.ExtendedOCR.Api Namespace

IHQCCompression Constructor

Initializes a new instance of the IHQCCompression class

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public IHQCCompression()

VB
Public Sub New

See Also

IHQCCompression Class
Aquaforest.ExtendedOCR.Api Namespace

IHQCCompression.IHQCCompression Properties

The IHQCCompression type exposes the following members.

Properties

Name	Description
AdvancedBrightness	Contact technical support (support@aquaforest.com) for guidance on using this property.
AdvancedContrast	Contact technical support (support@aquaforest.com) for guidance on using this property.
AdvancedFactorPhoto	Contact technical support (support@aquaforest.com) for guidance on using this property.
AdvancedSmooth	Contact technical support (support@aquaforest.com) for guidance on using this property.
AutoAdjust	Contact technical support (support@aquaforest.com) for guidance on using this property.
EnableIHQCCompression	Apply Intelligent High Quality Compression. The default value is 'false'.
HighResolutionForeground	High resolution foreground feature enables a specific mode causing the use of high resolution for the foreground image. HighResolutionForeground enables a specific mode for working around a bug existing in the native readers of some platforms that does not read properly PDF (Mainly the native reader on MAC OS X and IOS). Those native readers do not follow some of the specification of the PDF format. In place of displaying the correct compressed image, they display a down-sampled version of it. This special high resolution foreground mode offers a workaround to customer who wants to use iHQC with these buggy applications. Most of the PDF readers, including Adobe Reader are able to read PDF with IHQC compression properly on any platform.
IHQCCompressionLevel	The compression level to be used. Level 1 is the basic compression level. Level 3 is the most advanced intelligent High Quality Compression mode. **Note:** PDF 1.4 A-1a and PDF 1.4 A-1b only supports level 1.
IHQCQualityFactor	Set the IHQC quality factor.
IHQCThreshold	Contact technical support (support@aquaforest.com) for guidance on using this property.
MinOutputResolution	Contact technical support (support@aquaforest.com) for guidance on using this property.
SegmentationMode	Contact technical support (support@aquaforest.com) for guidance on using this property.
Smoothing	Contact technical support (support@aquaforest.com) for guidance on using this property.

See Also
IHQCCompression Class
Aquaforest.ExtendedOCR.Api Namespace

IHQCCompression.AdvancedBrightness Property

Contact technical support (support@aquaforest.com) for guidance on using this property.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public Nullable<int> AdvancedBrightness { get; set; }`

VB

`Public Property AdvancedBrightness As Nullable(Of Integer)`
`        Get`
`        Set`

Property Value
Type: Nullable(Int32)

See Also

IHQCCompression Class
Aquaforest.ExtendedOCR.Api Namespace

IHQCCompression.AdvancedContrast Property

Contact technical support (support@aquaforest.com) for guidance on using this property.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> AdvancedContrast { get; set; }```

VB

```Public Property AdvancedContrast As Nullable(Of Integer)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Int32)

See Also

IHQCCompression Class
Aquaforest.ExtendedOCR.Api Namespace

IHQCCompression.AdvancedFactorPhoto Property

Contact technical support (support@aquaforest.com) for guidance on using this property.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> AdvancedFactorPhoto { get; set; }```

VB

```Public Property AdvancedFactorPhoto As Nullable(Of Integer)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Int32)

See Also

IHQCCompression Class
Aquaforest.ExtendedOCR.Api Namespace

IHQCCompression.AdvancedSmooth Property

Contact technical support (support@aquaforest.com) for guidance on using this property.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> AdvancedSmooth { get; set; }```

VB

```Public Property AdvancedSmooth As Nullable(Of Integer)```
 ``` Get```
 ``` Set```

Property Value
Type: Nullable(Int32)

See Also

IHQCCompression Class
Aquaforest.ExtendedOCR.Api Namespace

IHQCCompression.AutoAdjust Property

Contact technical support (support@aquaforest.com) for guidance on using this property.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public Nullable<bool> AutoAdjust { get; set; }`

VB

`Public Property AutoAdjust As Nullable(Of Boolean)`
`    Get`
`    Set`

Property Value
Type: Nullable(Boolean)

See Also

IHQCCompression Class
Aquaforest.ExtendedOCR.Api Namespace

IHQCCompression.EnableIHQCCompression Property

Apply Intelligent High Quality Compression. The default value is 'false'.

Namespace: [Aquaforest.ExtendedOCR.Api](#)
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public bool EnableIHQCCompression { get; set; }```

VB
```Public Property EnableIHQCCompression As Boolean   Get   Set```

Property Value
Type: [Boolean](#)

See Also

[IHQCCompression Class](#)
[Aquaforest.ExtendedOCR.Api Namespace](#)

IHQCCompression.HighResolutionForeground Property

High resolution foreground feature enables a specific mode causing the use of high resolution for the foreground image. HighResolutionForeground enables a specific mode for working around a bug existing in the native readers of some platforms that does not read properly PDF (Mainly the native reader on MAC OS X and IOS). Those native readers do not follow some of the specification of the PDF format. In place of displaying the correct compressed image, they display a down-sampled version of it. This special high resolution foreground mode offers a workaround to customer who wants to use iHQC with these buggy applications. Most of the PDF readers, including Adobe Reader are able to read PDF with IHQC compression properly on any platform.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```
public Nullable<bool> HighResolutionForeground { get; set; }
```

VB
```
Public Property HighResolutionForeground As Nullable(Of Boolean)
    Get
    Set
```

Property Value
Type: Nullable(Boolean)

See Also

IHQCCompression Class
Aquaforest.ExtendedOCR.Api Namespace

IHQCCompression.IHQCCompressionLevel Property

The compression level to be used. Level 1 is the basic compression level. Level 3 is the most advanced intelligent High Quality Compression mode.
Note: PDF 1.4 A-1a and PDF 1.4 A-1b only supports level 1.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<IHQCCompressionLevel> IHQCCompressionLevel { get; set; }```

VB

```Public Property IHQCCompressionLevel As Nullable(Of IHQCCompressionLevel)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(IHQCCompressionLevel)

See Also

IHQCCompression Class
Aquaforest.ExtendedOCR.Api Namespace

IHQCCompression.IHQCQualityFactor Property

Set the IHQC quality factor.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<IHQCQualityFactor> IHQCQualityFactor { get; set; }```

VB

```Public Property IHQCQualityFactor As Nullable(Of IHQCQualityFactor)```
```   Get```
```   Set```

Property Value
Type: Nullable(IHQCQualityFactor)

See Also

IHQCCompression Class
Aquaforest.ExtendedOCR.Api Namespace

IHQCCompression.IHQCThreshold Property

Contact technical support (support@aquaforest.com) for guidance on using this property.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<byte> IHQCThreshold { get; set; }```

VB

```Public Property IHQCThreshold As Nullable(Of Byte)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Byte)

See Also

IHQCCompression Class
Aquaforest.ExtendedOCR.Api Namespace

IHQCCompression.MinOutputResolution Property

Contact technical support (support@aquaforest.com) for guidance on using this property.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public Nullable<int> MinOutputResolution { get; set; }`

VB

`Public Property MinOutputResolution As Nullable(Of Integer)`
`        Get`
`        Set`

Property Value
Type: Nullable(Int32)

See Also

IHQCCompression Class
Aquaforest.ExtendedOCR.Api Namespace

IHQCCompression.SegmentationMode Property

Contact technical support (support@aquaforest.com) for guidance on using this property.

Namespace: [Aquaforest.ExtendedOCR.Api](#)
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<SegmentationMode> SegmentationMode { get; set; }```

VB
```Public Property SegmentationMode As Nullable(Of SegmentationMode)   Get   Set```

Property Value
Type: [Nullable](#)([SegmentationMode](#))

See Also

[IHQCCompression Class](#)
[Aquaforest.ExtendedOCR.Api Namespace](#)

IHQCCompression.Smoothing Property

Contact technical support (support@aquaforest.com) for guidance on using this property.

Namespace: [Aquaforest.ExtendedOCR.Api](#)
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> Smoothing { get; set; }```

VB

```Public Property Smoothing As Nullable(Of Boolean)```
 ```   Get```
 ```   Set```

Property Value
Type: [Nullable](#)([Boolean](#))

See Also

[IHQCCompression Class](#)
[Aquaforest.ExtendedOCR.Api Namespace](#)

IHQCCompressionLevel Enumeration

The IHQC compression level.
Note: PDF 1.4 A-1a and PDF 1.4 A-1b only supports level 1.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public enum IHQCCompressionLevel`

VB
`Public Enumeration IHQCCompressionLevel`

Members

Member name	Value	Description
Level1	1	Level I (Acrobat 5.0 and higher)
Level2a	2	Level II spec a
Level2b	3	Level II spec b
Level3	4	Level III (Acrobat 6.0 and higher)

See Also

Aquaforest.ExtendedOCR.Api Namespace

IHQCQualityFactor Enumeration

The IHQC quality factor.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public enum IHQCQualityFactor`

VB
`Public Enumeration IHQCQualityFactor`

Members

Member name	Value	Description
MinimalSize	1	Minimal size
VerySmallSize	2	Very small size
SmallSize	3	Small size
FavorSizeOverQuality	4	Favor size over quality
Medium	5	Medium
FavorQualityOverSize	6	Favor quality over size
HighQuality	7	High quality
VeryHighQuality	8	Very high quality
MaximalQuality	9	Maximal quality

See Also

Aquaforest.ExtendedOCR.Api Namespace

Interpolation Class

This class offers image interpolation features.

Inheritance Hierarchy

System.Object
  Aquaforest.ExtendedOCR.Api.Interpolation

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public sealed class Interpolation`

VB
`Public NotInheritable Class Interpolation`

The **Interpolation** type exposes the following members.

Constructors

	Name	Description
	Interpolation	Initializes a new instance of the **Interpolation** class

Properties

	Name	Description
	Interpolate	Whether or not to interpolate. Default is false.
	InterpolationMode	Set the interpolation mode. Default is 'Normal'.
	TargetResolution	Interpolates the source image to the given resolution. This value must be greater than the source image's resolution.

See Also

Aquaforest.ExtendedOCR.Api Namespace

Interpolation Constructor

Initializes a new instance of the [Interpolation](#) class

Namespace: [Aquaforest.ExtendedOCR.Api](#)
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public Interpolation()`

VB
`Public Sub New`

See Also

[Interpolation Class](#)
[Aquaforest.ExtendedOCR.Api Namespace](#)

Interpolation.Interpolation Properties

The Interpolation type exposes the following members.

Properties

Name	Description
Interpolate	Whether or not to interpolate. Default is false.
InterpolationMode	Set the interpolation mode. Default is 'Normal'.
TargetResolution	Interpolates the source image to the given resolution. This value must be greater than the source image's resolution.

See Also

Interpolation Class
Aquaforest.ExtendedOCR.Api Namespace

Interpolation.Interpolate Property

Whether or not to interpolate. Default is false.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public bool Interpolate { get; set; }`

VB

`Public Property Interpolate As Boolean`
`        Get`
`        Set`

Property Value
Type: Boolean

See Also

Interpolation Class
Aquaforest.ExtendedOCR.Api Namespace

Interpolation.InterpolationMode Property

Set the interpolation mode. Default is 'Normal'.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<InterpolationMode> InterpolationMode { get; set; }```

VB

```Public Property InterpolationMode As Nullable(Of InterpolationMode)```
```   Get```
```   Set```

Property Value
Type: Nullable(InterpolationMode)

See Also

Interpolation Class
Aquaforest.ExtendedOCR.Api Namespace

Interpolation.TargetResolution Property

Interpolates the source image to the given resolution. This value must be greater than the source image's resolution.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public Nullable<int> TargetResolution { get; set; }`

VB

`Public Property TargetResolution As Nullable(Of Integer)`
`    Get`
`    Set`

Property Value
Type: Nullable(Int32)

See Also

Interpolation Class
Aquaforest.ExtendedOCR.Api Namespace

InterpolationMode Enumeration

Interpolation algorithms.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public enum InterpolationMode```

VB
```Public Enumeration InterpolationMode```

Members

Member name	Value	Description
Normal	0	Normal interpolation mode.
Fast	1	Fast interpolation mode.

See Also

Aquaforest.ExtendedOCR.Api Namespace

Jpeg2000CompressionMode Enumeration

The JPEG 2000 compression mode.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public enum Jpeg2000CompressionMode`

VB
`Public Enumeration Jpeg2000CompressionMode`

Members

Member name	Value	Description
JPEG2000_QUALITY_FACTOR	0	Use quality factor.
JPEG2000_TARGET_FILE_SIZE	1	Use target file size.
JPEG2000_COMPRESSION_RATIO	2	Use compression ratio.

See Also

Aquaforest.ExtendedOCR.Api Namespace

Layout Enumeration

The layout to use when generating non-PDF documents such as DOCX and RTF.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public enum Layout`

VB
`Public Enumeration Layout`

Members

Member name	Value	Description
Standard	0	
Flow	1	

See Also

Aquaforest.ExtendedOCR.Api Namespace

LicenseInfo Class

Inheritance Hierarchy

System.Object
  Aquaforest.ExtendedOCR.Api.LicenseInfo

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public sealed class LicenseInfo

VB
Public NotInheritable Class LicenseInfo

The **LicenseInfo** type exposes the following members.

Properties

	Name	Description
	ArabicOCREnabled	Get whether Arabic OCR is enabled.
	AsianOCREnabled	Get whether Asian OCR is enabled.
	IHQCEnabled	Get whether Intelligent High Quality Compression (IHQC) is enabled.
	Instance	
	IsTrial	Get whether the license is trial.
	LicenseType	Get the license type: Basic, Standard, Advanced or Extended.
	SDKVersion	Get the version of the SDK that this license is for.
	Threads	Get the maximum number of threads allowed to run concurrently.

See Also

Aquaforest.ExtendedOCR.Api Namespace

LicenseInfo.LicenseInfo Properties

The [LicenseInfo](#) type exposes the following members.

Properties

	Name	Description
	ArabicOCREnabled	Get whether Arabic OCR is enabled.
	AsianOCREnabled	Get whether Asian OCR is enabled.
	IHQCEnabled	Get whether Intelligent High Quality Compression (IHQC) is enabled.
	Instance	
	IsTrial	Get whether the license is trial.
	LicenseType	Get the license type: Basic, Standard, Advanced or Extended.
	SDKVersion	Get the version of the SDK that this license is for.
	Threads	Get the maximum number of threads allowed to run concurrently.

See Also

LicenseInfo Class
Aquaforest.ExtendedOCR.Api Namespace

LicenseInfo.ArabicOCREnabled Property

Get whether Arabic OCR is enabled.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public bool ArabicOCREnabled { get; }

VB
Public ReadOnly Property ArabicOCREnabled As Boolean   Get

Property Value
Type: Boolean

See Also

LicenseInfo Class
Aquaforest.ExtendedOCR.Api Namespace

LicenseInfo.AsianOCREnabled Property

Get whether Asian OCR is enabled.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public bool AsianOCREnabled { get; }

VB
Public ReadOnly Property AsianOCREnabled As Boolean   Get

Property Value
Type: Boolean

See Also

LicenseInfo Class
Aquaforest.ExtendedOCR.Api Namespace

LicenseInfo.IHQCEnabled Property

Get whether Intelligent High Quality Compression (IHQC) is enabled.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public bool IHQCEnabled { get; }

VB
Public ReadOnly Property IHQCEnabled As Boolean   Get

Property Value
Type: Boolean

See Also

LicenseInfo Class
Aquaforest.ExtendedOCR.Api Namespace

LicenseInfo.Instance Property

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public static LicenseInfo Instance { get; }```

VB
```Public Shared ReadOnly Property Instance As LicenseInfo```   ```Get```

Property Value
Type: LicenseInfo

See Also

LicenseInfo Class
Aquaforest.ExtendedOCR.Api Namespace

LicenseInfo.IsTrial Property

Get whether the license is trial.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public bool IsTrial { get; }`

VB
`Public ReadOnly Property IsTrial As Boolean` `        Get`

Property Value
Type: Boolean

See Also

LicenseInfo Class
Aquaforest.ExtendedOCR.Api Namespace

LicenseInfo.LicenseType Property

Get the license type: Basic, Standard, Advanced or Extended.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public string LicenseType { get; }```

VB
```Public ReadOnly Property LicenseType As String```   ```   Get```

Property Value
Type: String

See Also

LicenseInfo Class
Aquaforest.ExtendedOCR.Api Namespace

LicenseInfo.SDKVersion Property

Get the version of the SDK that this license is for.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public string SDKVersion { get; }```

VB
```Public ReadOnly Property SDKVersion As String```   ```   Get```

Property Value
Type: String

See Also

LicenseInfo Class
Aquaforest.ExtendedOCR.Api Namespace

LicenseInfo.Threads Property

Get the maximum number of threads allowed to run concurrently.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public int Threads { get; }```

VB
```Public ReadOnly Property Threads As Integer``` ```   Get```

Property Value
Type: Int32

See Also

LicenseInfo Class
Aquaforest.ExtendedOCR.Api Namespace

LineRemoval Class

LineRemoval class provides line removal capabilities. The class contains algorithms for removing vertical and horizontal lines in an image. The integrator can control various settings in line removal algorithm like:

- minimum and maximum line length and thickness of horizontal lines

- minimum and maximum line length and thickness of vertical lines

- clean noisy pixels adjacent to the line

- maximum gap between the lines

Inheritance Hierarchy

System.Object
  Aquaforest.ExtendedOCR.Api.LineRemoval

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public sealed class LineRemoval`

VB
`Public NotInheritable Class LineRemoval`

The **LineRemoval** type exposes the following members.

Constructors

	Name	Description
	LineRemoval	Initializes a new instance of the **LineRemoval** class

Properties

	Name	Description
	HorizontalCleanX	Cleans noisy pixels attached to horizontal lines. The default value is 1.
	HorizontalCleanY	Cleans noisy pixels attached to horizontal lines. The default value is 1.
	HorizontalDilate	Helps with the detection of horizontal lines. The default value is 1.
	HorizontalMaxGap	The maximum horizontal line gap to close. It is useful to remove broken lines. The default value is 2.

	HorizontalMaxThickness	The maximum thickness of horizontal lines to remove. It is useful to keep vertical lines larger than this parameter. Can also be useful to keep vertical letter strokes. The default value is 16.
	HorizontalMinLength	The minimum length of horizontal lines to remove. The default value is 128.
	RemoveLines	Whether or not to remove lines from an image. The default value is 'false'. **Note:** The image must be black and white for this setting to work. If you want to use this setting on a color image then you have to set Binarize to 'true'.
	VerticalCleanX	Cleans noisy pixels attached to vertical lines. The default value is 1.
	VerticalCleanY	Cleans noisy pixels attached to vertical lines. The default value is 1.
	VerticalDilate	Helps with the detection of vertical lines. The default value is 1.
	VerticalMaxGap	The maximum vertical line gap to close. It is useful to remove broken lines. The default value is 2.
	VerticalMaxThickness	The maximum thickness of vertical lines to remove. It is useful to keep horizontal lines larger than this parameter. Can also be useful to keep horizontal letter strokes. The default value is 16.
	VerticalMinLength	The minimum length of vertical lines to remove. The default value is 128.

See Also
Aquaforest.ExtendedOCR.Api Namespace

LineRemoval Constructor

Initializes a new instance of the LineRemoval class

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public LineRemoval()`

VB
`Public Sub New`

See Also

LineRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

LineRemoval.LineRemoval Properties

The LineRemoval type exposes the following members.

Properties

	Name	Description
	HorizontalCleanX	Cleans noisy pixels attached to horizontal lines. The default value is 1.
	HorizontalCleanY	Cleans noisy pixels attached to horizontal lines. The default value is 1.
	HorizontalDilate	Helps with the detection of horizontal lines. The default value is 1.
	HorizontalMaxGap	The maximum horizontal line gap to close. It is useful to remove broken lines. The default value is 2.
	HorizontalMaxThickness	The maximum thickness of horizontal lines to remove. It is useful to keep vertical lines larger than this parameter. Can also be useful to keep vertical letter strokes. The default value is 16.
	HorizontalMinLength	The minimum length of horizontal lines to remove. The default value is 128.
	RemoveLines	Whether or not to remove lines from an image. The default value is 'false'. **Note:** The image must be black and white for this setting to work. If you want to use this setting on a color image then you have to set Binarize to 'true'.
	VerticalCleanX	Cleans noisy pixels attached to vertical lines. The default value is 1.
	VerticalCleanY	Cleans noisy pixels attached to vertical lines. The default value is 1.
	VerticalDilate	Helps with the detection of vertical lines. The default value is 1.
	VerticalMaxGap	The maximum vertical line gap to close. It is useful to remove broken lines. The default value is 2.
	VerticalMaxThickness	The maximum thickness of vertical lines to remove. It is useful to keep horizontal lines larger than this parameter. Can also be useful to keep horizontal letter strokes. The default value is 16.
	VerticalMinLength	The minimum length of vertical lines to remove. The default value is 128.

See Also

LineRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

LineRemoval.HorizontalCleanX Property

Cleans noisy pixels attached to horizontal lines. The default value is 1.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public Nullable<int> HorizontalCleanX { get; set; }`

VB

`Public Property HorizontalCleanX As Nullable(Of Integer)`
`        Get`
`        Set`

Property Value
Type: Nullable(Int32)

See Also

LineRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

LineRemoval.HorizontalCleanY Property

Cleans noisy pixels attached to horizontal lines. The default value is 1.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```csharp
public Nullable<int> HorizontalCleanY { get; set; }
```

VB
```vb
Public Property HorizontalCleanY As Nullable(Of Integer)
    Get
    Set
```

Property Value
Type: Nullable(Int32)

See Also

LineRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

LineRemoval.HorizontalDilate Property

Helps with the detection of horizontal lines. The default value is 1.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public Nullable<int> HorizontalDilate { get; set; }`

VB

`Public Property HorizontalDilate As Nullable(Of Integer)`
`        Get`
`        Set`

Property Value
Type: Nullable(Int32)

See Also

LineRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

LineRemoval.HorizontalMaxGap Property

The maximum horizontal line gap to close. It is useful to remove broken lines. The default value is 2.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> HorizontalMaxGap { get; set; }```

VB

```Public Property HorizontalMaxGap As Nullable(Of Integer)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Int32)

See Also

LineRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

LineRemoval.HorizontalMaxThickness Property

The maximum thickness of horizontal lines to remove. It is useful to keep vertical lines larger than this parameter. Can also be useful to keep vertical letter strokes. The default value is 16.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> HorizontalMaxThickness { get; set; }```

VB

```Public Property HorizontalMaxThickness As Nullable(Of Integer)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Int32)

See Also

LineRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

LineRemoval.HorizontalMinLength Property

The minimum length of horizontal lines to remove. The default value is 128.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public Nullable<int> HorizontalMinLength { get; set; }`

VB

`Public Property HorizontalMinLength As Nullable(Of Integer)`
`        Get`
`        Set`

Property Value
Type: Nullable(Int32)

See Also

LineRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

LineRemoval.RemoveLines Property

Whether or not to remove lines from an image. The default value is 'false'.

Note: The image must be black and white for this setting to work. If you want to use this setting on a color image then you have to set Binarize to 'true'.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public bool RemoveLines { get; set; }```

VB
```Public Property RemoveLines As Boolean   Get   Set```

Property Value
Type: Boolean

See Also

LineRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

LineRemoval.VerticalCleanX Property

Cleans noisy pixels attached to vertical lines. The default value is 1.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> VerticalCleanX { get; set; }```

VB

```Public Property VerticalCleanX As Nullable(Of Integer)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Int32)

See Also

LineRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

LineRemoval.VerticalCleanY Property

Cleans noisy pixels attached to vertical lines. The default value is 1.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> VerticalCleanY { get; set; }```

VB

```Public Property VerticalCleanY As Nullable(Of Integer)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Int32)

See Also

LineRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

LineRemoval.VerticalDilate Property

Helps with the detection of vertical lines. The default value is 1.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> VerticalDilate { get; set; }```

VB

```Public Property VerticalDilate As Nullable(Of Integer)```
```   Get```
```   Set```

Property Value
Type: Nullable(Int32)

See Also

LineRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

LineRemoval.VerticalMaxGap Property

The maximum vertical line gap to close. It is useful to remove broken lines. The default value is 2.

Namespace: [Aquaforest.ExtendedOCR.Api](#)
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> VerticalMaxGap { get; set; }```

VB
```Public Property VerticalMaxGap As Nullable(Of Integer)   Get   Set```

Property Value
Type: [Nullable](#)([Int32](#))

See Also

[LineRemoval Class](#)
[Aquaforest.ExtendedOCR.Api Namespace](#)

LineRemoval.VerticalMaxThickness Property

The maximum thickness of vertical lines to remove. It is useful to keep horizontal lines larger than this parameter. Can also be useful to keep horizontal letter strokes. The default value is 16.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> VerticalMaxThickness { get; set; }```

VB

```Public Property VerticalMaxThickness As Nullable(Of Integer)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Int32)

See Also

LineRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

LineRemoval.VerticalMinLength Property

The minimum length of vertical lines to remove. The default value is 128.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> VerticalMinLength { get; set; }```

VB
```Public Property VerticalMinLength As Nullable(Of Integer)   Get   Set```

Property Value
Type: Nullable(Int32)

See Also

LineRemoval Class
Aquaforest.ExtendedOCR.Api Namespace

NonImagePDF Enumeration

This allows control over the treatment of non image-only PDFs, i.e. PDFs that have some text in them as well as images.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public enum NonImagePDF`

VB
`Public Enumeration NonImagePDF`

Members

Member name	Value	Description
OCR	0	The document will be OCRed using the image extraction method defined by "ExtractImageMethod".
RaiseError	1	The task will terminate with an error.
Skip	2	The document will not be processed.
PassThrough	3	The file will not be processed, but a copy of the document will be made and named as if the processing had occurred.

See Also

Aquaforest.ExtendedOCR.Api Namespace

Ocr Class

Inheritance Hierarchy

System.Object
  Aquaforest.ExtendedOCR.Api.Ocr

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public sealed class Ocr : IDisposable`

VB
`Public NotInheritable Class Ocr` `      Implements IDisposable`

The **Ocr** type exposes the following members.

Constructors

	Name	Description
🔧	Ocr	Initializes a new instance of Ocr

Properties

	Name	Description
	AsianOCREngine	Set the Asian OCR engine to use when processing documents with Asian languages. Default is 'OcrAsian1'.
	Author	Set a custom Author in the output PDF document properties.
	CharactersPerInch	The number of characters per inch (cpi) (for fixed text spacing: Fixed).
	CreationDate	Set a custom creation date in the output PDF document properties. **Note:** If you are setting the date through the properties.xml, then the date string should be in the format 'yyyy-MM-dd HH:mm:ss'.
	EmbedFonts	Set the flag indicating whether to embed fonts in the output PDF document. **Note:**It is not recommended to embed fonts because it increases the file size. Unless you have some special needs, it is strongly advised to disable embedding fonts.

	EnableConsoleOutput	If enabled then progress messages will be sent to the console. Default is true.
	EnableCsvOutput	Enable or disable CSV file output.
	EnableDebugOutput	If enabled then debug messages will be written to the console output.
	EnableDocxOutput	Enable or disable DOCX file output.
	EnableEpubOutput	Enable or disable ePUB file output. EPUB (short for electronic publication) is an e-book standard by the International Digital Publishing Forum (IDPF). Files have the extension .epub.
	EnableExcelMlOutput	Enable or disable ExcelMlOutput file output.
	EnableHtmlOutput	Enable or disable HTML file output.
	EnableOpenDocumentTextOutput	Enable or disable Open document text file output.
	EnablePdfOutput	Enable or disable PDF file output.
	EnableRtfOutput	Enable or disable RTF file output.
	EnableTextOutput	Enable or disable text file output.
	EnableWordMlOutput	Enable or disable WordMl file output.
	EnableXlsxOutput	Enable or disable XLSX file output.
	EnableXpsOutput	Enable or disable XPS file output.
	EndPage	Sets the last page of the source file that the OCR process will be run to (for a multipage source).
	ExtractImageMethod	This allows control over the method used to extract images from PDF files for OCR processing. The default value is 'Native'.
	GetPdfTextTimeout	The maximum time (in milliseconds) to check if a PDF document contains text or not.
	HandleExceptionsInternally	When set to true the OCR object will catch any exceptions for method calls and simply return false from the method. The exceptions caught are stored in the LastException property overwriting any previous value.
	IHQCCompression	The IHQC Compression settings to use for PDF and XPS output formats. **Note:** You will need to have an IHQC license to use this option.
	Language	Select the language to use for OCR processing. This will determine the dictionary that is used. If no language is set by the user, English will be used.
	Languages	Extended OCR accepts up to 8 recognition languages at a time. This is helpful to process mixed documents but, because of the various character sets, not all combinations are allowed. For this reason, the multiple languages support is limited to a single alphabet. For example, Russian and French can't be mixed.

		Note: • Asian languages can't be mixed. • Extended OCR SDK cannot load more than one Asian language at a time.
	LastException	Stores last exception caught by the OCR object.
	Layout	Select the layout to be used for docx and rtf output. The default value is 'Standard'.
	License	Specifies the license key.
	LicenseInfo	Gets the current license information. **Note:** This needs to be called after setting **Ocr**.License.
	LimitCharsetCharacters	Use this property to limit the characters to recognize to only the value specified. Enter a Unicode string that contains the new character set.
	LogFilePath	Gets the path of the log file where output is written to if LogToFile property is set to true. **Note:** The value for this property must be accessed after calling Recognize(PreProcessor).
	LogToFile	Set this to true if you want to log output to a file. The file will be stored in the TempFolder specified.
	ModifiedDate	Set a custom modified date in the output PDF document properties. **Note:** If you are setting the date through the properties.xml, then the date string should be in the format 'yyyy-MM-dd HH:mm:ss'.
	NoOcr	Set this to true if you do not want to OCR the source document. This could be useful if you want to convert a TIFF to an image-only PDF.
	NumberPages	Returns the number pages in a document.
	PDFImageCompression	The compression parameters to use for color JPEG images in generated PDFs. **Note:** This won't be used if you set ExtractImageMethod to 'Native' because the 'Native' method does not make any changes to the PDF file.
	PdfToImageExtractionTimeout	The maximum time (in milliseconds) to take to get image(s) from a PDF page. Default value is 100,000 ms.
	PDFVersion	The PDF version of the resulting output document. The default value is 1.4. **Note:** This won't be used if you set ExtractImageMethod to

		'Native' because the 'Native' method does not make any changes to the PDF file.
	ProcessPageTimeout	The maximum time (in milliseconds) to take to process a page (i.e. apply pre-processing and OCR). Default value is 300,000 ms.
	ReadBmpViaFreeImage	Contact technical support (support@aquaforest.com) for guidance on using this property.
	RemoveExistingPDFText	This applies only when a PDF is being used as the source for OCR. When set to true this will not include any searchable text that already exists from the source document. Such functionality might be useful if the source document was created by OCR of an image only PDF or other image file and the quality of the text from the previous OCR is poor. **Note:** There is no way to distinguish text added as a result of OCR from text added by other means and as a result this option should be used with care.
	RetainBookmarks	Retains any bookmarks from the source file in the output when using ExtractImageMethod.ConvertToTiff.
	RetainCreationDate	Retains the creation date of the source file in the output PDF document properties.
	RetainMetadata	Retains any metadata from the source file in the output when using ExtractImageMethod.ConvertToTiff.
	RetainModifiedDate	Retains the modified date of the source file in the output PDF document properties.
	RetainViewerPreferences	Retains any PDF Viewer Preferences, Page Mode and Page Layout from source file in the output when using ExtractImageMethod.ConvertToTiff.
	StartPage	Sets the first page of the source file where OCR recognition will begin from (for a multi-page source).
	TempFolder	The temporary folder used to keep the intermediary processing files.
	TextDirection	Sets the text direction (for Asian languages only).
	TextSpacing	Sets the text spacing. Default is Unknown.
	TextType	Sets the text type. Default is Normal.
	ThreadCount	The number of worker threads that are used for performing page recognition. Please note that, in order to get the best performance, the number of worker threads need to be set accordingly to your system capabilities. Aquaforest Extended OCR provides an "automatic" mode for balancing the number of

		work threads based on the system where it is executed. In order to activate this mode, set this value to 0.
	UserLexicon	The path to the custom user lexicon file. Adds a user lexicon to the existing decision environment. This is particularly useful for texts having specific terms. A user lexicon is a list of words put into a Unicode text file. Each word is separated by either a space, a tab or a carriage return (CR).
	Version	Returns the version of the SDK being used.
	WorkDepth	This property can have values going from 0 to 255: 0 meaning "best speed" and 255 "best accuracy". Increasing the work depth allows the OCR engine to spend more time on low-quality documents. The "work depth" indicates how deep the engine is allowed to work in order to find a satisfactory result. For documents of good quality, the engine will find a satisfactory result at an early stage and there will not be much speed differences between "best speed" and "best accuracy". For bad quality documents, if allowed by "word depth", the engine will work deeper. The accuracy will be better but at the price of speed. Default value is 100.

Methods

Name	Description
Abort	Terminates processing.
DeleteTemporaryFiles	Delete temporary files used in the OCR processing
Dispose	Releases all resources used by the **Ocr**
GetPageImage	Retrieves the image for the specified page. **Note:** Make sure that you call this method before deleting temporary files.
ReadBMPSource(Stream)	Read a bitmap source from a Stream for OCR.
ReadBMPSource(String)	Read a bitmap source for OCR.
ReadDocumentString	Returns a string containing the words from all pages processed.
ReadImageSource	Reads an Image object checking the number of frames (pages).
ReadJPEGSource(Stream)	Read a JPEG source from a Stream for OCR.
ReadJPEGSource(String)	The source JPEG file to OCR.

≡●	[ReadPageString(Int32)](#)	This function will retrieve a string representing the words found for the page requested. **Note:** This should be called only after OCR has completed successfully for the page.
≡●	[ReadPageString(Int32, Rectangle)](#)	This function will retrieve a string representing the words found for the page requested and in the region specified. **Note:** This should be called only after OCR has completed successfully for the page.
≡●	[ReadPageWords(Int32)](#)	This function will retrieve a Words object representing the words found for the page requested. **Note:** This should be called only after OCR has completed successfully for the page.
≡●	[ReadPageWords(Int32, Rectangle)](#)	This function will retrieve a Words object representing the words found for the page requested and in the region specified. **Note:** This should be called only after OCR has completed successfully for the page.
≡●	[ReadPDFSource(Stream)](#)	Read a PDF source from a [Stream](#) for OCR.
≡●	[ReadPDFSource(String)](#)	The source PDF file to OCR.
≡●	[ReadPDFSource(Stream, String)](#)	Read a password protected PDF source from a [Stream](#) for OCR.
≡●	[ReadPDFSource(String, String)](#)	The source PDF file to OCR
≡●	[ReadTIFFSource(Stream)](#)	Read a TIFF source from a [Stream](#) for OCR.
≡●	[ReadTIFFSource(String)](#)	The source TIFF file to OCR.
≡●	[Recognize](#)	Call Recognize to perform the actual OCR once options have been set on the **Ocr** and [PreProcessor](#) objects.
≡●	[SaveCSVOutput(Stream)](#)	Save CSV output to a [Stream](#). **Note:** This function should only be called if [EnableCsvOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
≡●	[SaveCSVOutput(String, Boolean)](#)	Save CSV file to the specified output location. **Note:** This function should only be called if [EnableCsvOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
≡●	[SaveDOCXOutput(Stream)](#)	Save DOCX output to a [Stream](#). **Note:** This function should only be called if [EnableDocxOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
≡●	[SaveDOCXOutput(String, Boolean)](#)	Save DOCX file to the specified output location. **Note:** This function should only be called if

		EnableDocxOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveEPUBOutput(Stream)	Save EPUB output to a Stream. **Note:** This function should only be called if EnableEpubOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveEPUBOutput(String, Boolean)	Save EPUB file to the specified output location. **Note:** This function should only be called if EnableEpubOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveExcelMLOutput(Stream)	Save Excel Markup Language output to a Stream. **Note:** This function should only be called if EnableExcelMlOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveExcelMLOutput(String, Boolean)	Save Excel Markup Language file to the specified output location. **Note:** This function should only be called if EnableExcelMlOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveHTMLOutput(Stream)	Save HTML output to a Stream. **Note:** This function should only be called if EnableHtmlOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveHTMLOutput(String, Boolean)	Save HTML file to the specified location. **Note:** This function should only be called if EnableHtmlOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveOpenDocumentTextOutput(Stream)	Save ODT output to a Stream. **Note:** This function should only be called if EnableOpenDocumentTextOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveOpenDocumentTextOutput(String, Boolean)	Save Open Document Text file to the specified output location. **Note:** This function should only be called if EnableOpenDocumentTextOutput was set to true prior to the call to Recognize(PreProcessor).
	SavePDFOutput(Stream)	Save PDF output to a Stream. **Note:** This function should only be called if EnablePdfOutput was set to true prior to the call to Recognize(PreProcessor).
	SavePDFOutput(String, Boolean)	Save PDF file to the specified output location. **Note:** This function should only be called if EnablePdfOutput was set to true prior to the call to Recognize(PreProcessor).

▤◈	[SaveRTFOutput(Stream)](#)	Save RTF output to a [Stream](#). **Note:** This function should only be called if [EnableRtfOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
▤◈	[SaveRTFOutput(String, Boolean)](#)	Save RTF file to the specified output location. **Note:** This function should only be called if [EnableRtfOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
▤◈	[SaveTextOutput(Stream)](#)	Save TEXT output to a [Stream](#). **Note:** This function should only be called if [EnableTextOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
▤◈	[SaveTextOutput(String, Boolean)](#)	Save Text file to the specified output location. **Note:** This function should only be called if [EnableTextOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
▤◈	[SaveWordMLOutput(Stream)](#)	Save Word Markup Language output to a [Stream](#). **Note:** This function should only be called if [EnableWordMlOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
▤◈	[SaveWordMLOutput(String, Boolean)](#)	Save Word Markup Language file to the specified output location. **Note:** This function should only be called if [EnableWordMlOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
▤◈	[SaveXLSXOutput(Stream)](#)	Save XLSX output to a [Stream](#). **Note:** This function should only be called if [EnableXlsxOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
▤◈	[SaveXLSXOutput(String, Boolean)](#)	Save XLSX file to the specified output location. **Note:** This function should only be called if [EnableXlsxOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
▤◈	[SaveXPSOutput(Stream)](#)	Save XPS output to a [Stream](#). **Note:** This function should only be called if [EnableXpsOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
▤◈	[SaveXPSOutput(String, Boolean)](#)	Save XPS file to the specified output location. **Note:** This function should only be called if [EnableXpsOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).

Events

Name	Description
⚡ StatusUpdate	This event is raised when processing of a page is complete. The StatusUpdateEventArgs object provides access to information relating to the status of the page processed.

See Also

Aquaforest.ExtendedOCR.Api Namespace

Ocr Constructor

Initializes a new instance of Ocr

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public Ocr(
       string resourceFolder
)
```

VB

```vb
Public Sub New (
       resourceFolder As String
)
```

Parameters
resourceFolder
Type: System.String

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.Ocr Properties

The Ocr type exposes the following members.

Properties

	Name	Description
	AsianOCREngine	Set the Asian OCR engine to use when processing documents with Asian languages. Default is 'OcrAsian1'.
	Author	Set a custom Author in the output PDF document properties.
	CharactersPerInch	The number of characters per inch (cpi) (for fixed text spacing: Fixed).
	CreationDate	Set a custom creation date in the output PDF document properties. **Note:** If you are setting the date through the properties.xml, then the date string should be in the format 'yyyy-MM-dd HH:mm:ss'.
	EmbedFonts	Set the flag indicating whether to embed fonts in the output PDF document. **Note:**It is not recommended to embed fonts because it increases the file size. Unless you have some special needs, it is strongly advised to disable embedding fonts.
	EnableConsoleOutput	If enabled then progress messages will be sent to the console. Default is true.
	EnableCsvOutput	Enable or disable CSV file output.
	EnableDebugOutput	If enabled then debug messages will be written to the console output.
	EnableDocxOutput	Enable or disable DOCX file output.
	EnableEpubOutput	Enable or disable ePUB file output. EPUB (short for electronic publication) is an e-book standard by the International Digital Publishing Forum (IDPF). Files have the extension .epub.
	EnableExcelMlOutput	Enable or disable ExcelMlOutput file output.
	EnableHtmlOutput	Enable or disable HTML file output.
	EnableOpenDocumentTextOutput	Enable or disable Open document text file output.
	EnablePdfOutput	Enable or disable PDF file output.
	EnableRtfOutput	Enable or disable RTF file output.
	EnableTextOutput	Enable or disable text file output.
	EnableWordMlOutput	Enable or disable WordMl file output.
	EnableXlsxOutput	Enable or disable XLSX file output.
	EnableXpsOutput	Enable or disable XPS file output.

	EndPage	Sets the last page of the source file that the OCR process will be run to (for a multipage source).
	ExtractImageMethod	This allows control over the method used to extract images from PDF files for OCR processing. The default value is 'Native'.
	GetPdfTextTimeout	The maximum time (in milliseconds) to check if a PDF document contains text or not.
	HandleExceptionsInternally	When set to true the OCR object will catch any exceptions for method calls and simply return false from the method. The exceptions caught are stored in the LastException property overwriting any previous value.
	IHQCCompression	The IHQC Compression settings to use for PDF and XPS output formats. **Note:** You will need to have an IHQC license to use this option.
	Language	Select the language to use for OCR processing. This will determine the dictionary that is used. If no language is set by the user, English will be used.
	Languages	Extended OCR accepts up to 8 recognition languages at a time. This is helpful to process mixed documents but, because of the various character sets, not all combinations are allowed. For this reason, the multiple languages support is limited to a single alphabet. For example, Russian and French can't be mixed. **Note:** • Asian languages can't be mixed. • Extended OCR SDK cannot load more than one Asian language at a time.
	LastException	Stores last exception caught by the OCR object.
	Layout	Select the layout to be used for docx and rtf output. The default value is 'Standard'.
	License	Specifies the license key.
	LicenseInfo	Gets the current license information. **Note:** This needs to be called after setting Ocr.License.
	LimitCharsetCharacters	Use this property to limit the characters to recognize to only the value specified. Enter a Unicode string that contains the new character set.
	LogFilePath	Gets the path of the log file where output is written to if LogToFile property is set to true. **Note:** The value for this property must be accessed after calling Recognize(PreProcessor).

	LogToFile	Set this to true if you want to log output to a file. The file will be stored in the TempFolder specified.
	ModifiedDate	Set a custom modified date in the output PDF document properties. **Note:** If you are setting the date through the properties.xml, then the date string should be in the format 'yyyy-MM-dd HH:mm:ss'.
	NoOcr	Set this to true if you do not want to OCR the source document. This could be useful if you want to convert a TIFF to an image-only PDF.
	NumberPages	Returns the number pages in a document.
	PDFImageCompression	The compression parameters to use for color JPEG images in generated PDFs. **Note:** This won't be used if you set ExtractImageMethod to 'Native' because the 'Native' method does not make any changes to the PDF file.
	PdfToImageExtractionTimeout	The maximum time (in milliseconds) to take to get image(s) from a PDF page. Default value is 100,000 ms.
	PDFVersion	The PDF version of the resulting output document. The default value is 1.4. **Note:** This won't be used if you set ExtractImageMethod to 'Native' because the 'Native' method does not make any changes to the PDF file.
	ProcessPageTimeout	The maximum time (in milliseconds) to take to process a page (i.e. apply pre-processing and OCR). Default value is 300,000 ms.
	ReadBmpViaFreeImage	Contact technical support (support@aquaforest.com) for guidance on using this property.
	RemoveExistingPDFText	This applies only when a PDF is being used as the source for OCR. When set to true this will not include any searchable text that already exists from the source document. Such functionality might be useful if the source document was created by OCR of an image only PDF or other image file and the quality of the text from the previous OCR is poor. **Note:** There is no way to distinguish text added as a result of OCR from text added by other means and as a result this option should be used with care.
	RetainBookmarks	Retains any bookmarks from the source file in the output when using ExtractImageMethod.ConvertToTiff.
	RetainCreationDate	Retains the creation date of the source file in the output PDF document properties.
	RetainMetadata	Retains any metadata from the source file in the output when using ExtractImageMethod.ConvertToTiff.

	RetainModifiedDate	Retains the modified date of the source file in the output PDF document properties.
	RetainViewerPreferences	Retains any PDF Viewer Preferences, Page Mode and Page Layout from source file in the output when using ExtractImageMethod.ConvertToTiff.
	StartPage	Sets the first page of the source file where OCR recognition will begin from (for a multi-page source).
	TempFolder	The temporary folder used to keep the intermediary processing files.
	TextDirection	Sets the text direction (for Asian languages only).
	TextSpacing	Sets the text spacing. Default is Unknown.
	TextType	Sets the text type. Default is Normal.
	ThreadCount	The number of worker threads that are used for performing page recognition. Please note that, in order to get the best performance, the number of worker threads need to be set accordingly to your system capabilities. Aquaforest Extended OCR provides an "automatic" mode for balancing the number of work threads based on the system where it is executed. In order to activate this mode, set this value to 0.
	UserLexicon	The path to the custom user lexicon file. Adds a user lexicon to the existing decision environment. This is particularly useful for texts having specific terms. A user lexicon is a list of words put into a Unicode text file. Each word is separated by either a space, a tab or a carriage return (CR).
	Version	Returns the version of the SDK being used.
	WorkDepth	This property can have values going from 0 to 255: 0 meaning "best speed" and 255 "best accuracy". Increasing the work depth allows the OCR engine to spend more time on low-quality documents. The "work depth" indicates how deep the engine is allowed to work in order to find a satisfactory result. For documents of good quality, the engine will find a satisfactory result at an early stage and there will not be much speed differences between "best speed" and "best accuracy". For bad quality documents, if allowed by "word depth", the engine will work deeper. The accuracy will be better but at the

		price of speed. Default value is 100.

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.AsianOCREngine Property

Set the Asian OCR engine to use when processing documents with Asian languages. Default is 'OcrAsian1'.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<AsianOCREngine> AsianOCREngine { get; set; }```

VB

```Public Property AsianOCREngine As Nullable(Of AsianOCREngine)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(AsianOCREngine)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.Author Property

Set a custom Author in the output PDF document properties.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```csharp
public string Author { get; set; }
```

VB
```vb
Public Property Author As String
    Get
    Set
```

Property Value
Type: String

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.CharactersPerInch Property

The number of characters per inch (cpi) (for fixed text spacing: Fixed).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> CharactersPerInch { get; set; }```

VB

```Public Property CharactersPerInch As Nullable(Of Integer)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Int32)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.CreationDate Property

Set a custom creation date in the output PDF document properties.
Note: If you are setting the date through the properties.xml, then the date string should be in the format 'yyyy-MM-dd HH:mm:ss'.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```
public Nullable<DateTime> CreationDate { get; set; }
```

VB
```
Public Property CreationDate As Nullable(Of DateTime)
      Get
      Set
```

Property Value
Type: Nullable(DateTime)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.EmbedFonts Property

Set the flag indicating whether to embed fonts in the output PDF document.
Note: It is not recommended to embed fonts because it increases the file size. Unless you have some special needs, it is strongly advised to disable embedding fonts.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public Nullable<bool> EmbedFonts { get; set; }

VB

Public Property EmbedFonts As Nullable(Of Boolean)
        Get
        Set

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.EnableConsoleOutput Property

If enabled then progress messages will be sent to the console. Default is true.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> EnableConsoleOutput { get; set; }```

VB

```Public Property EnableConsoleOutput As Nullable(Of Boolean)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.EnableCsvOutput Property

Enable or disable CSV file output.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> EnableCsvOutput { get; set; }```

VB

```Public Property EnableCsvOutput As Nullable(Of Boolean)```
 ``` Get```
 ``` Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.EnableDebugOutput Property

If enabled then debug messages will be written to the console output.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> EnableDebugOutput { get; set; }```

VB

```Public Property EnableDebugOutput As Nullable(Of Boolean)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.EnableDocxOutput Property

Enable or disable DOCX file output.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> EnableDocxOutput { get; set; }```

VB

```Public Property EnableDocxOutput As Nullable(Of Boolean)```
```   Get```
```   Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.EnableEpubOutput Property

Enable or disable ePUB file output.
 EPUB (short for electronic publication) is an e-book standard by the International Digital Publishing Forum (IDPF). Files have the extension .epub.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> EnableEpubOutput { get; set; }```

VB

```Public Property EnableEpubOutput As Nullable(Of Boolean)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.EnableExcelMlOutput Property

Enable or disable ExcelMlOutput file output.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public Nullable<bool> EnableExcelMlOutput { get; set; }

VB

Public Property EnableExcelMlOutput As Nullable(Of Boolean)
        Get
        Set

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.EnableHtmlOutput Property

Enable or disable HTML file output.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public Nullable<bool> EnableHtmlOutput { get; set; }`

VB

`Public Property EnableHtmlOutput As Nullable(Of Boolean)`
`        Get`
`        Set`

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.EnableOpenDocumentTextOutput Property

Enable or disable Open document text file output.

Namespace: [Aquaforest.ExtendedOCR.Api](#)
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> EnableOpenDocumentTextOutput { get; set; }```

VB

```Public Property EnableOpenDocumentTextOutput As Nullable(Of Boolean)```
```   Get```
```   Set```

Property Value
Type: [Nullable](#)([Boolean](#))

See Also

[Ocr Class](#)
[Aquaforest.ExtendedOCR.Api Namespace](#)

Ocr.EnablePdfOutput Property

Enable or disable PDF file output.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public Nullable<bool> EnablePdfOutput { get; set; }`

VB

`Public Property EnablePdfOutput As Nullable(Of Boolean)`
`    Get`
`    Set`

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.EnableRtfOutput Property

Enable or disable RTF file output.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> EnableRtfOutput { get; set; }```

VB

```Public Property EnableRtfOutput As Nullable(Of Boolean)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.EnableTextOutput Property

Enable or disable text file output.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> EnableTextOutput { get; set; }```

VB

```Public Property EnableTextOutput As Nullable(Of Boolean)```
```   Get```
```   Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.EnableWordMlOutput Property

Enable or disable WordMl file output.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```
public Nullable<bool> EnableWordMlOutput { get; set; }
```

VB
```
Public Property EnableWordMlOutput As Nullable(Of Boolean)
      Get
      Set
```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.EnableXlsxOutput Property

Enable or disable XLSX file output.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public Nullable<bool> EnableXlsxOutput { get; set; }`

VB

`Public Property EnableXlsxOutput As Nullable(Of Boolean)`
`        Get`
`        Set`

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.EnableXpsOutput Property

Enable or disable XPS file output.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> EnableXpsOutput { get; set; }```

VB

```Public Property EnableXpsOutput As Nullable(Of Boolean)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.EndPage Property

Sets the last page of the source file that the OCR process will be run to (for a multipage source).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public int EndPage { get; set; }`

VB

`Public Property EndPage As Integer`
`    Get`
`    Set`

Property Value
Type: Int32

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ExtractImageMethod Property

This allows control over the method used to extract images from PDF files for OCR processing. The default value is 'Native'.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<ExtractImageMethod> ExtractImageMethod { get; set; }```

VB

```Public Property ExtractImageMethod As Nullable(Of ExtractImageMethod)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(ExtractImageMethod)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.GetPdfTextTimeout Property

The maximum time (in milliseconds) to check if a PDF document contains text or not.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public Nullable<int> GetPdfTextTimeout { get; set; }

VB

Public Property GetPdfTextTimeout As Nullable(Of Integer)
        Get
        Set

Property Value
Type: Nullable(Int32)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.HandleExceptionsInternally Property

When set to true the OCR object will catch any exceptions for method calls and simply return false from the method. The exceptions caught are stored in the [LastException](#) property overwriting any previous value.

Namespace: [Aquaforest.ExtendedOCR.Api](#)
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public bool HandleExceptionsInternally { get; set; }```

VB
```Public Property HandleExceptionsInternally As Boolean   Get   Set```

Property Value
Type: [Boolean](#)

See Also

[Ocr Class](#)
[Aquaforest.ExtendedOCR.Api Namespace](#)

Ocr.IHQCCompression Property

The IHQC Compression settings to use for PDF and XPS output formats.
Note: You will need to have an IHQC license to use this option.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public IHQCCompression IHQCCompression { get; set; }```

VB

```Public Property IHQCCompression As IHQCCompression```
 ```   Get```
 ```   Set```

Property Value
Type: IHQCCompression

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.Language Property

Select the language to use for OCR processing. This will determine the dictionary that is used. If no language is set by the user, English will be used.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<SupportedLanguages> Language { get; set; }```

VB

```Public Property Language As Nullable(Of SupportedLanguages)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(SupportedLanguages)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.Languages Property

Extended OCR accepts up to 8 recognition languages at a time. This is helpful to process mixed documents but, because of the various character sets, not all combinations are allowed. For this reason, the multiple languages support is limited to a single alphabet. For example, Russian and French can't be mixed.

Note:

- Asian languages can't be mixed.

- Extended OCR SDK cannot load more than one Asian language at a time.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public SupportedLanguages[] Languages { get; set; }```

VB
```Public Property Languages As SupportedLanguages()```

Property Value
Type: SupportedLanguages[]

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.LastException Property

Stores last exception caught by the OCR object.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Exception LastException { get; }```

VB
```Public ReadOnly Property LastException As Exception``` ```   Get```

Property Value
Type: Exception

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.Layout Property

Select the layout to be used for docx and rtf output. The default value is 'Standard'.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<Layout> Layout { get; set; }```

VB

```Public Property Layout As Nullable(Of Layout)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Layout)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.License Property

Specifies the license key.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public string License { set; }

VB
Public WriteOnly Property License As String   Set

Property Value
Type: String

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.LicenseInfo Property

Gets the current license information.
Note: This needs to be called after setting Ocr.License.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public LicenseInfo LicenseInfo { get; }```

VB
```Public ReadOnly Property LicenseInfo As LicenseInfo```   ```   Get```

Property Value
Type: LicenseInfo

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.LimitCharsetCharacters Property

Use this property to limit the characters to recognize to only the value specified.

 Enter a Unicode string that contains the new character set.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public string LimitCharsetCharacters { get; set; }```

VB
```Public Property LimitCharsetCharacters As String   Get   Set```

Property Value
Type: String

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.LogFilePath Property

Gets the path of the log file where output is written to if [LogToFile](#) property is set to true.
Note: The value for this property must be accessed after calling [Recognize(PreProcessor)](#).

Namespace: [Aquaforest.ExtendedOCR.Api](#)
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public string LogFilePath { get; }```

VB
```Public ReadOnly Property LogFilePath As String   Get```

Property Value
Type: [String](#)

See Also

[Ocr Class](#)
[Aquaforest.ExtendedOCR.Api Namespace](#)

Ocr.LogToFile Property

Set this to true if you want to log output to a file. The file will be stored in the TempFolder specified.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> LogToFile { get; set; }```

VB

```Public Property LogToFile As Nullable(Of Boolean)```
```   Get```
```   Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ModifiedDate Property

Set a custom modified date in the output PDF document properties.
Note: If you are setting the date through the properties.xml, then the date string should be in the format 'yyyy-MM-dd HH:mm:ss'.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<DateTime> ModifiedDate { get; set; }```

VB

```Public Property ModifiedDate As Nullable(Of DateTime)```
```   Get```
```   Set```

Property Value
Type: Nullable(DateTime)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.NoOcr Property

Set this to true if you do not want to OCR the source document. This could be useful if you want to convert a TIFF to an image-only PDF.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> NoOcr { get; set; }```

VB

```Public Property NoOcr As Nullable(Of Boolean)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.NumberPages Property

Returns the number pages in a document.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public int NumberPages { get; }`

VB
`Public ReadOnly Property NumberPages As Integer` `    Get`

Property Value
Type: Int32

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.PDFImageCompression Property

The compression parameters to use for color JPEG images in generated PDFs.
Note: This won't be used if you set ExtractImageMethod to 'Native' because the 'Native' method does not make any changes to the PDF file.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public PDFImageCompression PDFImageCompression { get; set; }```

VB

```Public Property PDFImageCompression As PDFImageCompression```
 ```   Get```
 ```   Set```

Property Value
Type: PDFImageCompression

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.PdfToImageExtractionTimeout Property

The maximum time (in milliseconds) to take to get image(s) from a PDF page. Default value is 100,000 ms.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> PdfToImageExtractionTimeout { get; set; }```

VB

```Public Property PdfToImageExtractionTimeout As Nullable(Of Integer)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Int32)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.PDFVersion Property

The PDF version of the resulting output document. The default value is 1.4.

Note: This won't be used if you set [ExtractImageMethod](#) to 'Native' because the 'Native' method does not make any changes to the PDF file.

Namespace: [Aquaforest.ExtendedOCR.Api](#)
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<PDFVersion> PDFVersion { get; set; }```

VB

```Public Property PDFVersion As Nullable(Of PDFVersion)```
 ```   Get```
 ```   Set```

Property Value
Type: [Nullable](#)([PDFVersion](#))

See Also

[Ocr Class](#)
[Aquaforest.ExtendedOCR.Api Namespace](#)

Ocr.ProcessPageTimeout Property

The maximum time (in milliseconds) to take to process a page (i.e. apply pre-processing and OCR). Default value is 300,000 ms.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> ProcessPageTimeout { get; set; }```

VB

```Public Property ProcessPageTimeout As Nullable(Of Integer)```
```   Get```
```   Set```

Property Value
Type: Nullable(Int32)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadBmpViaFreeImage Property

Contact technical support (support@aquaforest.com) for guidance on using this property.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> ReadBmpViaFreeImage { get; set; }```

VB

```Public Property ReadBmpViaFreeImage As Nullable(Of Boolean)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.RemoveExistingPDFText Property

This applies only when a PDF is being used as the source for OCR. When set to true this will not include any searchable text that already exists from the source document. Such functionality might be useful if the source document was created by OCR of an image only PDF or other image file and the quality of the text from the previous OCR is poor.

Note: There is no way to distinguish text added as a result of OCR from text added by other means and as a result this option should be used with care.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> RemoveExistingPDFText { get; set; }```

VB

```Public Property RemoveExistingPDFText As Nullable(Of Boolean)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.RetainBookmarks Property

Retains any bookmarks from the source file in the output when using
ExtractImageMethod.ConvertToTiff.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> RetainBookmarks { get; set; }```

VB

```Public Property RetainBookmarks As Nullable(Of Boolean)```
```   Get```
```   Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.RetainCreationDate Property

Retains the creation date of the source file in the output PDF document properties.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> RetainCreationDate { get; set; }```

VB

```Public Property RetainCreationDate As Nullable(Of Boolean)```
```   Get```
```   Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.RetainMetadata Property

Retains any metadata from the source file in the output when using ExtractImageMethod.ConvertToTiff.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public Nullable<bool> RetainMetadata { get; set; }`

VB

`Public Property RetainMetadata As Nullable(Of Boolean)`
`    Get`
`    Set`

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.RetainModifiedDate Property

Retains the modified date of the source file in the output PDF document properties.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> RetainModifiedDate { get; set; }```

VB

```Public Property RetainModifiedDate As Nullable(Of Boolean)```
```   Get```
```   Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.RetainViewerPreferences Property

Retains any PDF Viewer Preferences, Page Mode and Page Layout from source file in the output when using ExtractImageMethod.ConvertToTiff.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> RetainViewerPreferences { get; set; }```

VB

```Public Property RetainViewerPreferences As Nullable(Of Boolean)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Boolean)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.StartPage Property

Sets the first page of the source file where OCR recognition will begin from (for a multi-page source).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public int StartPage { get; set; }

VB

Public Property StartPage As Integer
    Get
    Set

Property Value
Type: Int32

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.TempFolder Property

The temporary folder used to keep the intermediary processing files.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public string TempFolder { get; set; }```

VB

```Public Property TempFolder As String```
 ```   Get```
 ```   Set```

Property Value
Type: String

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.TextDirection Property

Sets the text direction (for Asian languages only).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public Nullable<TextDirection> TextDirection { get; set; }

VB

Public Property TextDirection As Nullable(Of TextDirection)
        Get
        Set

Property Value
Type: Nullable(TextDirection)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.TextSpacing Property

Sets the text spacing.

 Default is Unknown.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<TextSpacing> TextSpacing { get; set; }```

VB

```Public Property TextSpacing As Nullable(Of TextSpacing)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(TextSpacing)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.TextType Property

Sets the text type.

 Default is Normal.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```csharp
public Nullable<TextType> TextType { get; set; }
```

VB
```vb
Public Property TextType As Nullable(Of TextType)
      Get
      Set
```

Property Value
Type: Nullable(TextType)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ThreadCount Property

The number of worker threads that are used for performing page recognition. Please note that, in order to get the best performance, the number of worker threads need to be set accordingly to your system capabilities. Aquaforest Extended OCR provides an "automatic" mode for balancing the number of work threads based on the system where it is executed. In order to activate this mode, set this value to 0.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> ThreadCount { get; set; }```

VB

```Public Property ThreadCount As Nullable(Of Integer)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Int32)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.UserLexicon Property

The path to the custom user lexicon file.

Adds a user lexicon to the existing decision environment. This is particularly useful for texts having specific terms.

A user lexicon is a list of words put into a Unicode text file. Each word is separated by either a space, a tab or a carriage return (CR).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public string UserLexicon { get; set; }`

VB

`Public Property UserLexicon As String`
`    Get`
`    Set`

Property Value
Type: String

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.Version Property

Returns the version of the SDK being used.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public string Version { get; }

VB
Public ReadOnly Property Version As String   Get

Property Value
Type: String

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.WorkDepth Property

This property can have values going from 0 to 255: 0 meaning "best speed" and 255 "best accuracy".
 Increasing the work depth allows the OCR engine to spend more time on low-quality documents.
 The "work depth" indicates how deep the engine is allowed to work in order to find a satisfactory result.
 For documents of good quality, the engine will find a satisfactory result at an early stage and there will not be much speed differences between "best speed" and "best accuracy".
 For bad quality documents, if allowed by "word depth", the engine will work deeper. The accuracy will be better but at the price of speed.
 Default value is 100.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<byte> WorkDepth { get; set; }```

VB

```Public Property WorkDepth As Nullable(Of Byte)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Byte)

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.Ocr Methods

The Ocr type exposes the following members.

Methods

Name	Description
Abort	Terminates processing.
DeleteTemporaryFiles	Delete temporary files used in the OCR processing
Dispose	
GetPageImage	Retrieves the image for the specified page. **Note:** Make sure that you call this method before deleting temporary files.
ReadBMPSource(Stream)	Read a bitmap source from a Stream for OCR.
ReadBMPSource(String)	Read a bitmap source for OCR.
ReadDocumentString	Returns a string containing the words from all pages processed.
ReadImageSource	Reads an Image object checking the number of frames (pages).
ReadJPEGSource(Stream)	Read a JPEG source from a Stream for OCR.
ReadJPEGSource(String)	The source JPEG file to OCR.
ReadPageString(Int32)	This function will retrieve a string representing the words found for the page requested. **Note:** This should be called only after OCR has completed successfully for the page.
ReadPageString(Int32, Rectangle)	This function will retrieve a string representing the words found for the page requested and in the region specified. **Note:** This should be called only after OCR has completed successfully for the page.
ReadPageWords(Int32)	This function will retrieve a Words object representing the words found for the page requested. **Note:** This should be called only after OCR has completed successfully for the page.
ReadPageWords(Int32, Rectangle)	This function will retrieve a Words object representing the words found for the page requested and in the region specified. **Note:** This should be called only after OCR has completed successfully for the page.
ReadPDFSource(Stream)	Read a PDF source from a Stream for OCR.
ReadPDFSource(String)	The source PDF file to OCR.

⬢	[ReadPDFSource(Stream, String)](#)	Read a password protected PDF source from a [Stream](#) for OCR.
⬢	[ReadPDFSource(String, String)](#)	The source PDF file to OCR
⬢	[ReadTIFFSource(Stream)](#)	Read a TIFF source from a [Stream](#) for OCR.
⬢	[ReadTIFFSource(String)](#)	The source TIFF file to OCR.
⬢	[Recognize](#)	Call Recognize to perform the actual OCR once options have been set on the [Ocr](#) and [PreProcessor](#) objects.
⬢	[SaveCSVOutput(Stream)](#)	Save CSV output to a [Stream](#). **Note:** This function should only be called if [EnableCsvOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
⬢	[SaveCSVOutput(String, Boolean)](#)	Save CSV file to the specified output location. **Note:** This function should only be called if [EnableCsvOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
⬢	[SaveDOCXOutput(Stream)](#)	Save DOCX output to a [Stream](#). **Note:** This function should only be called if [EnableDocxOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
⬢	[SaveDOCXOutput(String, Boolean)](#)	Save DOCX file to the specified output location. **Note:** This function should only be called if [EnableDocxOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
⬢	[SaveEPUBOutput(Stream)](#)	Save EPUB output to a [Stream](#). **Note:** This function should only be called if [EnableEpubOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
⬢	[SaveEPUBOutput(String, Boolean)](#)	Save EPUB file to the specified output location. **Note:** This function should only be called if [EnableEpubOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
⬢	[SaveExcelMLOutput(Stream)](#)	Save Excel Markup Language output to a [Stream](#). **Note:** This function should only be called if [EnableExcelMlOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
⬢	[SaveExcelMLOutput(String, Boolean)](#)	Save Excel Markup Language file to the specified output location. **Note:** This function should only be called if [EnableExcelMlOutput](#) was set to true prior to the call to [Recognize(PreProcessor)](#).
⬢	[SaveHTMLOutput(Stream)](#)	Save HTML output to a [Stream](#). **Note:** This function should only be called if

		EnableHtmlOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveHTMLOutput(String, Boolean)	Save HTML file to the specified location. **Note:** This function should only be called if EnableHtmlOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveOpenDocumentTextOutput(Stream)	Save ODT output to a Stream. **Note:** This function should only be called if EnableOpenDocumentTextOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveOpenDocumentTextOutput(String, Boolean)	Save Open Document Text file to the specified output location. **Note:** This function should only be called if EnableOpenDocumentTextOutput was set to true prior to the call to Recognize(PreProcessor).
	SavePDFOutput(Stream)	Save PDF output to a Stream. **Note:** This function should only be called if EnablePdfOutput was set to true prior to the call to Recognize(PreProcessor).
	SavePDFOutput(String, Boolean)	Save PDF file to the specified output location. **Note:** This function should only be called if EnablePdfOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveRTFOutput(Stream)	Save RTF output to a Stream. **Note:** This function should only be called if EnableRtfOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveRTFOutput(String, Boolean)	Save RTF file to the specified output location. **Note:** This function should only be called if EnableRtfOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveTextOutput(Stream)	Save TEXT output to a Stream. **Note:** This function should only be called if EnableTextOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveTextOutput(String, Boolean)	Save Text file to the specified output location. **Note:** This function should only be called if EnableTextOutput was set to true prior to the call to Recognize(PreProcessor).
	SaveWordMLOutput(Stream)	Save Word Markup Language output to a Stream. **Note:** This function should only be called if EnableWordMlOutput was set to true prior to the call to Recognize(PreProcessor).

SaveWordMLOutput(String, Boolean)	Save Word Markup Language file to the specified output location. **Note:** This function should only be called if EnableWordMlOutput was set to true prior to the call to Recognize(PreProcessor).
SaveXLSXOutput(Stream)	Save XLSX output to a Stream. **Note:** This function should only be called if EnableXlsxOutput was set to true prior to the call to Recognize(PreProcessor).
SaveXLSXOutput(String, Boolean)	Save XLSX file to the specified output location. **Note:** This function should only be called if EnableXlsxOutput was set to true prior to the call to Recognize(PreProcessor).
SaveXPSOutput(Stream)	Save XPS output to a Stream. **Note:** This function should only be called if EnableXpsOutput was set to true prior to the call to Recognize(PreProcessor).
SaveXPSOutput(String, Boolean)	Save XPS file to the specified output location. **Note:** This function should only be called if EnableXpsOutput was set to true prior to the call to Recognize(PreProcessor).

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.Abort Method

Terminates processing.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public void Abort()`

VB
`Public Sub Abort`

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.DeleteTemporaryFiles Method

Delete temporary files used in the OCR processing

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public void DeleteTemporaryFiles()`

VB
`Public Sub DeleteTemporaryFiles`

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.Dispose Method

Releases all resources used by the Ocr

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public void Dispose()```

VB
```Public Sub Dispose```

Implements
IDisposable.Dispose()

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.GetPageImage Method

Retrieves the image for the specified page.
Note: Make sure that you call this method before deleting temporary files.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```csharp
public Image GetPageImage(
	int pageNumber
)
``` |

VB
| ```vb
Public Function GetPageImage (
	pageNumber As Integer
) As Image
``` |

Parameters
pageNumber
Type: System.Int32
The page to retrieve the image from.

Return Value
Type: Image
The image for the specified page.

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadBMPSource Method

Overload List

	Name	Description
	ReadBMPSource(Stream)	Read a bitmap source from a Stream for OCR.
	ReadBMPSource(String)	Read a bitmap source for OCR.

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadBMPSource Method (Stream)

Read a bitmap source from a Stream for OCR.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public void ReadBMPSource(Stream source)```

VB
```Public Sub ReadBMPSource (source As Stream)```

Parameters
source
Type: System.IO.Stream
Source bitmap Stream

See Also

Ocr Class
ReadBMPSource Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadBMPSource Method (String)

Read a bitmap source for OCR.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public void ReadBMPSource(
      string sourcefile
)
``` |

VB
| ```
Public Sub ReadBMPSource (
      sourcefile As String
)
``` |

Parameters
sourcefile
Type: System.String
The full path of the source BMP file.

See Also

Ocr Class
ReadBMPSource Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadDocumentString Method

Returns a string containing the words from all pages processed.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public string ReadDocumentString()

VB
Public Function ReadDocumentString As String

Return Value
Type: String

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadImageSource Method

Reads an Image object checking the number of frames (pages).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```csharp
public void ReadImageSource(
      Image image
)
```

VB
```vb
Public Sub ReadImageSource (
      image As Image
)
```

Parameters
image
Type: System.Drawing.Image
The image object to OCR.

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadJPEGSource Method

Overload List

	Name	Description
	ReadJPEGSource(Stream)	Read a JPEG source from a Stream for OCR.
	ReadJPEGSource(String)	The source JPEG file to OCR.

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadJPEGSource Method (Stream)

Read a JPEG source from a Stream for OCR.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public void ReadJPEGSource(
      Stream source
)
``` |

VB
| ```
Public Sub ReadJPEGSource (
      source As Stream
)
``` |

Parameters
source
Type: System.IO.Stream
Source bitmap Stream

Return Value
Type:
Success

See Also

Ocr Class
ReadJPEGSource Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadJPEGSource Method (String)

The source JPEG file to OCR.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public void ReadJPEGSource(
      string sourcefile
)
``` |

VB
| ```
Public Sub ReadJPEGSource (
      sourcefile As String
)
``` |

Parameters
sourcefile
Type: System.String
The full path of the source JPEG file.

See Also

Ocr Class
ReadJPEGSource Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadPageString Method

Overload List

Name	Description
ReadPageString(Int32)	This function will retrieve a string representing the words found for the page requested. **Note:** This should be called only after OCR has completed successfully for the page.
ReadPageString(Int32, Rectangle)	This function will retrieve a string representing the words found for the page requested and in the region specified. **Note:** This should be called only after OCR has completed successfully for the page.

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadPageString Method (Int32)

This function will retrieve a string representing the words found for the page requested.
Note: This should be called only after OCR has completed successfully for the page.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```csharp
public string ReadPageString(
      int pageNumber
)
```

VB
```vb
Public Function ReadPageString (
      pageNumber As Integer
) As String
```

Parameters

pageNumber
Type: System.Int32
The page number of the page to get the string from.

Return Value

Type: String
The string containing the words from the specified page.

See Also

Ocr Class
ReadPageString Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadPageString Method (Int32, Rectangle)

This function will retrieve a string representing the words found for the page requested and in the region specified.

Note: This should be called only after OCR has completed successfully for the page.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public string ReadPageString(
        int pageNumber,
        Rectangle region
)
```

VB

```vb
Public Function ReadPageString (
        pageNumber As Integer,
        region As Rectangle
) As String
```

Parameters

pageNumber
Type: System.Int32
The page number of the page to get the string from.

region
Type: System.Drawing.Rectangle
Region to retrieve string from.

Return Value
Type: String
The string containing the words from the specified region and page.

See Also

Ocr Class
ReadPageString Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadPageWords Method

Overload List

Name	Description
ReadPageWords(Int32)	This function will retrieve a Words object representing the words found for the page requested. **Note:** This should be called only after OCR has completed successfully for the page.
ReadPageWords(Int32, Rectangle)	This function will retrieve a Words object representing the words found for the page requested and in the region specified. **Note:** This should be called only after OCR has completed successfully for the page.

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadPageWords Method (Int32)

This function will retrieve a Words object representing the words found for the page requested.
Note: This should be called only after OCR has completed successfully for the page.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public Words ReadPageWords(
    int pageNumber
)
``` |

VB
| ```
Public Function ReadPageWords (
    pageNumber As Integer
) As Words
``` |

Parameters

pageNumber
Type: System.Int32
The page number of the page to get the words from.

Return Value

Type: Words
Words object representing the words found for the page requested.

See Also

Ocr Class
ReadPageWords Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadPageWords Method (Int32, Rectangle)

This function will retrieve a Words object representing the words found for the page requested and in the region specified.

Note: This should be called only after OCR has completed successfully for the page.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public Words ReadPageWords(
      int pageNumber,
      Rectangle region
)
``` |

VB
| ```
Public Function ReadPageWords (
      pageNumber As Integer,
      region As Rectangle
) As Words
``` |

Parameters

pageNumber
Type: System.Int32
The page number of the page to get the words from.

region
Type: System.Drawing.Rectangle
Region to retrieve words from.

Return Value
Type: Words
Words object representing the words found for the specified region and page.

See Also

Ocr Class
ReadPageWords Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadPDFSource Method

Overload List

Name	Description
ReadPDFSource(Stream)	Read a PDF source from a Stream for OCR.
ReadPDFSource(String)	The source PDF file to OCR.
ReadPDFSource(Stream, String)	Read a password protected PDF source from a Stream for OCR.
ReadPDFSource(String, String)	The source PDF file to OCR

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadPDFSource Method (Stream)

Read a PDF source from a Stream for OCR.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public void ReadPDFSource(
      Stream source
)
```

VB

```vb
Public Sub ReadPDFSource (
      source As Stream
)
```

Parameters
source
Type: System.IO.Stream
Source PDF Stream

See Also

Ocr Class
ReadPDFSource Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadPDFSource Method (String)

The source PDF file to OCR.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public void ReadPDFSource(
        string sourcefile
)
```

VB

```vb
Public Sub ReadPDFSource (
        sourcefile As String
)
```

Parameters
sourcefile
Type: System.String
The full path of the source PDF file.

See Also

Ocr Class
ReadPDFSource Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadPDFSource Method (Stream, String)

Read a password protected PDF source from a Stream for OCR.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public void ReadPDFSource(
      Stream source,
      string password
)
``` |

VB
| ```
Public Sub ReadPDFSource (
      source As Stream,
      password As String
)
``` |

Parameters

source
Type: System.IO.Stream
Source PDF Stream

password
Type: System.String
The password to open the PDF.

See Also

Ocr Class
ReadPDFSource Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadPDFSource Method (String, String)

The source PDF file to OCR

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public void ReadPDFSource(
      string sourcefile,
      string password
)
```

VB

```vb
Public Sub ReadPDFSource (
      sourcefile As String,
      password As String
)
```

Parameters

sourcefile
Type: System.String
The full path of the source PDF file.

password
Type: System.String
The password to use to open the source PDF file.

See Also

Ocr Class
ReadPDFSource Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadTIFFSource Method

Overload List

	Name	Description
	ReadTIFFSource(Stream)	Read a TIFF source from a Stream for OCR.
	ReadTIFFSource(String)	The source TIFF file to OCR.

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadTIFFSource Method (Stream)

Read a TIFF source from a Stream for OCR.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public void ReadTIFFSource(
      Stream source
)
``` |

VB
| ```
Public Sub ReadTIFFSource (
      source As Stream
)
``` |

Parameters
source
Type: System.IO.Stream
Source TIFF Stream

See Also

Ocr Class
ReadTIFFSource Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.ReadTIFFSource Method (String)

The source TIFF file to OCR.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public void ReadTIFFSource(
       string sourcefile
)
``` |

VB
| ```
Public Sub ReadTIFFSource (
       sourcefile As String
)
``` |

Parameters
sourcefile
Type: System.String
The full path of the source TIFF file.

See Also

Ocr Class
ReadTIFFSource Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.Recognize Method

Call Recognize to perform the actual OCR once options have been set on the Ocr and PreProcessor objects.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```
public bool Recognize(
	PreProcessor preProcessor
)
```

VB

```
Public Function Recognize (
	preProcessor As PreProcessor
) As Boolean
```

Parameters

preProcessor
Type: Aquaforest.ExtendedOCR.Api.PreProcessor
The pre-processing settings to apply.

Return Value

Type: Boolean
Whether OCR has been successful or not.

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveCSVOutput Method

Overload List

Name	Description
SaveCSVOutput(Stream)	Save CSV output to a Stream. **Note:** This function should only be called if EnableCsvOutput was set to true prior to the call to Recognize(PreProcessor).
SaveCSVOutput(String, Boolean)	Save CSV file to the specified output location. **Note:** This function should only be called if EnableCsvOutput was set to true prior to the call to Recognize(PreProcessor).

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveCSVOutput Method (Stream)

Save CSV output to a Stream.

Note: This function should only be called if EnableCsvOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public bool SaveCSVOutput(
    out Stream outputStream
)
```

VB

```vb
Public Function SaveCSVOutput (
    <OutAttribute> ByRef outputStream As Stream
) As Boolean
```

Parameters

outputStream
Type: System.IO.Stream
The output Stream containing the results of the OCR in CSV format.

Return Value

Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveCSVOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveCSVOutput Method (String, Boolean)

Save CSV file to the specified output location.

Note: This function should only be called if EnableCsvOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public bool SaveCSVOutput(
      string outputfile,
      bool overwriteExisting
)
```

VB

```vb
Public Function SaveCSVOutput (
      outputfile As String,
      overwriteExisting As Boolean
) As Boolean
```

Parameters
outputfile
Type: System.String
The name of the file to save the output as.

overwriteExisting
Type: System.Boolean
Overwrite file with the same name.

Return Value
Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveCSVOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveDOCXOutput Method

Overload List

Name	Description
SaveDOCXOutput(Stream)	Save DOCX output to a Stream. **Note:** This function should only be called if EnableDocxOutput was set to true prior to the call to Recognize(PreProcessor).
SaveDOCXOutput(String, Boolean)	Save DOCX file to the specified output location. **Note:** This function should only be called if EnableDocxOutput was set to true prior to the call to Recognize(PreProcessor).

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveDOCXOutput Method (Stream)

Save DOCX output to a Stream.

Note: This function should only be called if EnableDocxOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public bool SaveDOCXOutput(
        out Stream outputStream
)
``` |

VB
| ```
Public Function SaveDOCXOutput (
        <OutAttribute> ByRef outputStream As Stream
) As Boolean
``` |

Parameters
outputStream
Type: System.IO.Stream
The output Stream containing the results of the OCR in DOCX format.

Return Value
Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveDOCXOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveDOCXOutput Method (String, Boolean)

Save DOCX file to the specified output location.

Note: This function should only be called if EnableDocxOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public bool SaveDOCXOutput(
      string outputfile,
      bool overwriteExisting
)
```

VB

```vb
Public Function SaveDOCXOutput (
      outputfile As String,
      overwriteExisting As Boolean
) As Boolean
```

Parameters

outputfile
Type: System.String
The name of the file to save the output as.

overwriteExisting
Type: System.Boolean
Overwrite file with the same name.

Return Value
Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveDOCXOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveEPUBOutput Method

Overload List

Name	Description
SaveEPUBOutput(Stream)	Save EPUB output to a Stream. **Note:** This function should only be called if EnableEpubOutput was set to true prior to the call to Recognize(PreProcessor).
SaveEPUBOutput(String, Boolean)	Save EPUB file to the specified output location. **Note:** This function should only be called if EnableEpubOutput was set to true prior to the call to Recognize(PreProcessor).

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveEPUBOutput Method (Stream)

Save EPUB output to a Stream.

Note: This function should only be called if EnableEpubOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public bool SaveEPUBOutput(
    out Stream outputStream
)
```

VB

```vb
Public Function SaveEPUBOutput (
    <OutAttribute> ByRef outputStream As Stream
) As Boolean
```

Parameters

outputStream
Type: System.IO.Stream
The output Stream containing the results of the OCR in EPUB format.

Return Value

Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveEPUBOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveEPUBOutput Method (String, Boolean)

Save EPUB file to the specified output location.

Note: This function should only be called if EnableEpubOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public bool SaveEPUBOutput(
    string outputfile,
    bool overwriteExisting
)
```

VB

```vb
Public Function SaveEPUBOutput (
    outputfile As String,
    overwriteExisting As Boolean
) As Boolean
```

Parameters
outputfile
Type: System.String
The name of the file to save the output as.

overwriteExisting
Type: System.Boolean
Overwrite file with the same name.

Return Value
Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveEPUBOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveExcelMLOutput Method

Overload List

Name	Description
SaveExcelMLOutput(Stream)	Save Excel Markup Language output to a Stream. **Note:** This function should only be called if EnableExcelMlOutput was set to true prior to the call to Recognize(PreProcessor).
SaveExcelMLOutput(String, Boolean)	Save Excel Markup Language file to the specified output location. **Note:** This function should only be called if EnableExcelMlOutput was set to true prior to the call to Recognize(PreProcessor).

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveExcelMLOutput Method (Stream)

Save Excel Markup Language output to a Stream.

Note: This function should only be called if EnableExcelMlOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public bool SaveExcelMLOutput(
      out Stream outputStream
)
``` |

VB
| ```
Public Function SaveExcelMLOutput (
      <OutAttribute> ByRef outputStream As Stream
) As Boolean
``` |

Parameters
outputStream
Type: System.IO.Stream
The output Stream containing the results of the OCR in Excel Markup Language format.

Return Value
Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveExcelMLOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveExcelMLOutput Method (String, Boolean)

Save Excel Markup Language file to the specified output location.

Note: This function should only be called if EnableExcelMlOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public bool SaveExcelMLOutput(
      string outputfile,
      bool overwriteExisting
)
```

VB

```vb
Public Function SaveExcelMLOutput (
      outputfile As String,
      overwriteExisting As Boolean
) As Boolean
```

Parameters
outputfile
Type: System.String
The name of the file to save the output as.

overwriteExisting
Type: System.Boolean
Overwrite file with the same name.

Return Value
Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveExcelMLOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveHTMLOutput Method

Overload List

Name	Description
SaveHTMLOutput(Stream)	Save HTML output to a Stream. **Note:** This function should only be called if EnableHtmlOutput was set to true prior to the call to Recognize(PreProcessor).
SaveHTMLOutput(String, Boolean)	Save HTML file to the specified location. **Note:** This function should only be called if EnableHtmlOutput was set to true prior to the call to Recognize(PreProcessor).

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveHTMLOutput Method (Stream)

Save HTML output to a Stream.

Note: This function should only be called if EnableHtmlOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public bool SaveHTMLOutput(
        out Stream outputStream
)
``` |

VB
| ```
Public Function SaveHTMLOutput (
        <OutAttribute> ByRef outputStream As Stream
) As Boolean
``` |

Parameters

outputStream
Type: System.IO.Stream
The output Stream containing the results of the OCR in HTML format.

Return Value

Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveHTMLOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveHTMLOutput Method (String, Boolean)

Save HTML file to the specified location.

Note: This function should only be called if EnableHtmlOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public bool SaveHTMLOutput(
      string outputfile,
      bool overwriteExisting
)
``` |

VB
| ```
Public Function SaveHTMLOutput (
      outputfile As String,
      overwriteExisting As Boolean
) As Boolean
``` |

Parameters
outputfile
Type: System.String
The name of the file to save the output as.

overwriteExisting
Type: System.Boolean
Overwrite file with the same name.

Return Value
Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveHTMLOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveOpenDocumentTextOutput Method

Overload List

Name	Description
SaveOpenDocumentTextOutput(Stream)	Save ODT output to a Stream. **Note:** This function should only be called if EnableOpenDocumentTextOutput was set to true prior to the call to Recognize(PreProcessor).
SaveOpenDocumentTextOutput(String, Boolean)	Save Open Document Text file to the specified output location. **Note:** This function should only be called if EnableOpenDocumentTextOutput was set to true prior to the call to Recognize(PreProcessor).

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveOpenDocumentTextOutput Method (Stream)

Save ODT output to a Stream.

Note: This function should only be called if EnableOpenDocumentTextOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public bool SaveOpenDocumentTextOutput(
    out Stream outputStream
)
```

VB

```vb
Public Function SaveOpenDocumentTextOutput (
    <OutAttribute> ByRef outputStream As Stream
) As Boolean
```

Parameters

outputStream
Type: System.IO.Stream
The output Stream containing the results of the OCR in ODT format.

Return Value

Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveOpenDocumentTextOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveOpenDocumentTextOutput Method (String, Boolean)

Save Open Document Text file to the specified output location.

Note: This function should only be called if EnableOpenDocumentTextOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```csharp
public bool SaveOpenDocumentTextOutput(
    string outputfile,
    bool overwriteExisting
)
```

VB
```vb
Public Function SaveOpenDocumentTextOutput (
    outputfile As String,
    overwriteExisting As Boolean
) As Boolean
```

Parameters
outputfile
Type: System.String
The name of the file to save the output as.

overwriteExisting
Type: System.Boolean
Overwrite file with the same name.

Return Value
Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveOpenDocumentTextOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SavePDFOutput Method

Overload List

Name	Description
SavePDFOutput(Stream)	Save PDF output to a Stream. **Note:** This function should only be called if EnablePdfOutput was set to true prior to the call to Recognize(PreProcessor).
SavePDFOutput(String, Boolean)	Save PDF file to the specified output location. **Note:** This function should only be called if EnablePdfOutput was set to true prior to the call to Recognize(PreProcessor).

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.SavePDFOutput Method (Stream)

Save PDF output to a Stream.

Note: This function should only be called if EnablePdfOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```csharp
public bool SavePDFOutput(
      out Stream outputStream
)
```

VB
```vb
Public Function SavePDFOutput (
      <OutAttribute> ByRef outputStream As Stream
) As Boolean
```

Parameters

outputStream
Type: System.IO.Stream
The output Stream containing the results of the OCR in PDF format.

Return Value

Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SavePDFOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SavePDFOutput Method (String, Boolean)

Save PDF file to the specified output location.

Note: This function should only be called if EnablePdfOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public bool SavePDFOutput(
      string outputfile,
      bool overwriteExisting
)
```

VB

```vb
Public Function SavePDFOutput (
      outputfile As String,
      overwriteExisting As Boolean
) As Boolean
```

Parameters
outputfile
Type: System.String
The name of the file to save the output as.

overwriteExisting
Type: System.Boolean
Overwrite file with the same name.

Return Value
Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SavePDFOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveRTFOutput Method

Overload List

Name	Description
SaveRTFOutput(Stream)	Save RTF output to a Stream. **Note:** This function should only be called if EnableRtfOutput was set to true prior to the call to Recognize(PreProcessor).
SaveRTFOutput(String, Boolean)	Save RTF file to the specified output location. **Note:** This function should only be called if EnableRtfOutput was set to true prior to the call to Recognize(PreProcessor).

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveRTFOutput Method (Stream)

Save RTF output to a Stream.

Note: This function should only be called if EnableRtfOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public bool SaveRTFOutput(
     out Stream outputStream
)
``` |

VB
| ```
Public Function SaveRTFOutput (
     <OutAttribute> ByRef outputStream As Stream
) As Boolean
``` |

Parameters

outputStream
Type: System.IO.Stream
The output Stream containing the results of the OCR in RTF format.

Return Value

Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveRTFOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveRTFOutput Method (String, Boolean)

Save RTF file to the specified output location.

Note: This function should only be called if EnableRtfOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public bool SaveRTFOutput(
      string outputfile,
      bool overwriteExisting
)
```

VB

```vb
Public Function SaveRTFOutput (
      outputfile As String,
      overwriteExisting As Boolean
) As Boolean
```

Parameters
outputfile
Type: System.String
The name of the file to save the output as.

overwriteExisting
Type: System.Boolean
Overwrite file with the same name.

Return Value
Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveRTFOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveTextOutput Method

Overload List

Name	Description
SaveTextOutput(Stream)	Save TEXT output to a Stream. **Note:** This function should only be called if EnableTextOutput was set to true prior to the call to Recognize(PreProcessor).
SaveTextOutput(String, Boolean)	Save Text file to the specified output location. **Note:** This function should only be called if EnableTextOutput was set to true prior to the call to Recognize(PreProcessor).

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveTextOutput Method (Stream)

Save TEXT output to a Stream.

Note: This function should only be called if EnableTextOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public bool SaveTextOutput(
    out Stream outputStream
)
```

VB

```vb
Public Function SaveTextOutput (
    <OutAttribute> ByRef outputStream As Stream
) As Boolean
```

Parameters

outputStream
Type: System.IO.Stream
The output Stream containing the results of the OCR in TEXT format.

Return Value

Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveTextOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveTextOutput Method (String, Boolean)

Save Text file to the specified output location.

Note: This function should only be called if EnableTextOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
``public bool SaveTextOutput(
string outputfile,
bool overwriteExisting
)```

VB
``Public Function SaveTextOutput (
outputfile As String,
overwriteExisting As Boolean
) As Boolean```

Parameters
outputfile
Type: System.String
The name of the file to save the output as.

overwriteExisting
Type: System.Boolean
Overwrite file with the same name.

Return Value
Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveTextOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveWordMLOutput Method

Overload List

Name	Description
SaveWordMLOutput(Stream)	Save Word Markup Language output to a Stream. **Note:** This function should only be called if EnableWordMlOutput was set to true prior to the call to Recognize(PreProcessor).
SaveWordMLOutput(String, Boolean)	Save Word Markup Language file to the specified output location. **Note:** This function should only be called if EnableWordMlOutput was set to true prior to the call to Recognize(PreProcessor).

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveWordMLOutput Method (Stream)

Save Word Markup Language output to a Stream.

Note: This function should only be called if EnableWordMlOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public bool SaveWordMLOutput(
      out Stream outputStream
)
``` |

VB
| ```
Public Function SaveWordMLOutput (
      <OutAttribute> ByRef outputStream As Stream
) As Boolean
``` |

Parameters
outputStream
Type: System.IO.Stream
The output Stream containing the results of the OCR in Word Markup Language format.

Return Value
Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveWordMLOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveWordMLOutput Method (String, Boolean)

Save Word Markup Language file to the specified output location.

Note: This function should only be called if EnableWordMlOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public bool SaveWordMLOutput(
      string outputfile,
      bool overwriteExisting
)
``` |

VB
| ```
Public Function SaveWordMLOutput (
      outputfile As String,
      overwriteExisting As Boolean
) As Boolean
``` |

Parameters
outputfile
Type: System.String
The name of the file to save the output as.

overwriteExisting
Type: System.Boolean
Overwrite file with the same name.

Return Value
Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveWordMLOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveXLSXOutput Method

Overload List

Name	Description
SaveXLSXOutput(Stream)	Save XLSX output to a Stream. **Note:** This function should only be called if EnableXlsxOutput was set to true prior to the call to Recognize(PreProcessor).
SaveXLSXOutput(String, Boolean)	Save XLSX file to the specified output location. **Note:** This function should only be called if EnableXlsxOutput was set to true prior to the call to Recognize(PreProcessor).

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveXLSXOutput Method (Stream)

Save XLSX output to a Stream.

Note: This function should only be called if EnableXlsxOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public bool SaveXLSXOutput(
      out Stream outputStream
)
``` |

VB
| ```
Public Function SaveXLSXOutput (
      <OutAttribute> ByRef outputStream As Stream
) As Boolean
``` |

Parameters
outputStream
Type: System.IO.Stream
The output Stream containing the results of the OCR in XLSX format.

Return Value
Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveXLSXOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveXLSXOutput Method (String, Boolean)

Save XLSX file to the specified output location.

Note: This function should only be called if EnableXlsxOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public bool SaveXLSXOutput(
       string outputfile,
       bool overwriteExisting
)
```

VB

```vb
Public Function SaveXLSXOutput (
       outputfile As String,
       overwriteExisting As Boolean
) As Boolean
```

Parameters
outputfile
Type: System.String
The name of the file to save the output as.

overwriteExisting
Type: System.Boolean
Overwrite file with the same name.

Return Value
Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveXLSXOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveXPSOutput Method

Overload List

Name	Description
⬦ SaveXPSOutput(Stream)	Save XPS output to a Stream. **Note:** This function should only be called if EnableXpsOutput was set to true prior to the call to Recognize(PreProcessor).
⬦ SaveXPSOutput(String, Boolean)	Save XPS file to the specified output location. **Note:** This function should only be called if EnableXpsOutput was set to true prior to the call to Recognize(PreProcessor).

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveXPSOutput Method (Stream)

Save XPS output to a Stream.

Note: This function should only be called if EnableXpsOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```csharp
public bool SaveXPSOutput(
     out Stream outputStream
)
```

VB
```vb
Public Function SaveXPSOutput (
     <OutAttribute> ByRef outputStream As Stream
) As Boolean
```

Parameters

outputStream
Type: System.IO.Stream
The output Stream containing the results of the OCR in XPS format.

Return Value

Type: Boolean
Whether save was successful or not.

See Also

Ocr Class
SaveXPSOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.SaveXPSOutput Method (String, Boolean)

Save XPS file to the specified output location.

Note: This function should only be called if EnableXpsOutput was set to true prior to the call to Recognize(PreProcessor).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#

```csharp
public bool SaveXPSOutput(
      string outputfile,
      bool overwriteExisting
)
```

VB

```vb
Public Function SaveXPSOutput (
      outputfile As String,
      overwriteExisting As Boolean
) As Boolean
```

Parameters

outputfile
Type: System.String
The name of the file to save the output as.

overwriteExisting
Type: System.Boolean
Overwrite file with the same name.

Return Value
Type: Boolean
Whether successful or not.

See Also

Ocr Class
SaveXPSOutput Overload
Aquaforest.ExtendedOCR.Api Namespace

Ocr.Ocr Events

The Ocr type exposes the following members.

Events

	Name	Description
⚡	StatusUpdate	This event is raised when processing of a page is complete. The StatusUpdateEventArgs object provides access to information relating to the status of the page processed.

See Also

Ocr Class

Aquaforest.ExtendedOCR.Api Namespace

Ocr.StatusUpdate Event

This event is raised when processing of a page is complete. The StatusUpdateEventArgs object provides access to information relating to the status of the page processed.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public event Ocr.StatusUpdateEventHandler StatusUpdate

VB
Public Event StatusUpdate As Ocr.StatusUpdateEventHandler

Value

Type: Aquaforest.ExtendedOCR.Api.Ocr.StatusUpdateEventHandler

See Also

Ocr Class
Aquaforest.ExtendedOCR.Api Namespace

Ocr.StatusUpdateEventHandler Delegate

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
| ```
public delegate void StatusUpdateEventHandler(
	Object sender,
	StatusUpdateEventArgs statusUpdateEventArgs
)
``` |

VB
| ```
Public Delegate Sub StatusUpdateEventHandler (
	sender As Object,
	statusUpdateEventArgs As StatusUpdateEventArgs
)
``` |

Parameters

sender
Type: System.Object
statusUpdateEventArgs
Type: Aquaforest.ExtendedOCR.Shared.StatusUpdateEventArgs

See Also

Aquaforest.ExtendedOCR.Api Namespace

PDFImageCompression Class

This class manages the PDF image compression parameters

Inheritance Hierarchy

System.Object
  Aquaforest.ExtendedOCR.Api.PDFImageCompression

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public sealed class PDFImageCompression`

VB
`Public NotInheritable Class PDFImageCompression`

The **PDFImageCompression** type exposes the following members.

Constructors

	Name	Description
	PDFImageCompression	Initializes a new instance of the **PDFImageCompression** class

Properties

	Name	Description
	EnableJpeg2000Compression	Use JPEG 2000 compression. If set to true, then the default compression mode is **JPEG2000_QUALITY_FACTOR**. The default value is false.
	Jpeg2000CompressionMode	The JPEG 2000 compression mode to use.
	Jpeg2000CompressionValue	The value to use for the selected compression mode.
	QualityFactor	Gets/Sets the image compression quality factor of color JPEG images in generated PDFs. The default value is 192. The quality factor determines the degree of loss caused by the compression process. Extended OCR accepts a value from 0 to 255: • 255 guarantees the highest image quality • 0 guarantees the best compression

See Also

Aquaforest.ExtendedOCR.Api Namespace

PDFImageCompression Constructor

Initializes a new instance of the PDFImageCompression class

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public PDFImageCompression()`

VB
`Public Sub New`

See Also

PDFImageCompression Class
Aquaforest.ExtendedOCR.Api Namespace

PDFImageCompression.PDFImageCompression Properties

The PDFImageCompression type exposes the following members.

Properties

	Name	Description
	EnableJpeg2000Compression	Use JPEG 2000 compression. If set to true, then the default compression mode is **JPEG2000_QUALITY_FACTOR**. The default value is false.
	Jpeg2000CompressionMode	The JPEG 2000 compression mode to use.
	Jpeg2000CompressionValue	The value to use for the selected compression mode.
	QualityFactor	Gets/Sets the image compression quality factor of color JPEG images in generated PDFs. The default value is 192. The quality factor determines the degree of loss caused by the compression process. Extended OCR accepts a value from 0 to 255: • 255 guarantees the highest image quality • 0 guarantees the best compression

See Also

PDFImageCompression Class
Aquaforest.ExtendedOCR.Api Namespace

PDFImageCompression.EnableJpeg2000Compression Property

Use JPEG 2000 compression. If set to true, then the default compression mode is
JPEG2000_QUALITY_FACTOR. The default value is false.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public bool EnableJpeg2000Compression { get; set; }```

VB

```Public Property EnableJpeg2000Compression As Boolean```
 ```   Get```
 ```   Set```

Property Value
Type: Boolean

See Also

PDFImageCompression Class
Aquaforest.ExtendedOCR.Api Namespace

PDFImageCompression.Jpeg2000CompressionMode Property

The JPEG 2000 compression mode to use.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public Nullable<Jpeg2000CompressionMode> Jpeg2000CompressionMode { get; set; }`

VB

`Public Property Jpeg2000CompressionMode As Nullable(Of Jpeg2000CompressionMode)`
`    Get`
`    Set`

Property Value

Type: Nullable(Jpeg2000CompressionMode)

See Also

PDFImageCompression Class
Aquaforest.ExtendedOCR.Api Namespace

PDFImageCompression.Jpeg2000CompressionValue Property

The value to use for the selected compression mode.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> Jpeg2000CompressionValue { get; set; }```

VB
```Public Property Jpeg2000CompressionValue As Nullable(Of Integer)   Get   Set```

Property Value
Type: Nullable(Int32)

See Also

PDFImageCompression Class
Aquaforest.ExtendedOCR.Api Namespace

PDFImageCompression.QualityFactor Property

Gets/Sets the image compression quality factor of color JPEG images in generated PDFs. The default value is 192.

The quality factor determines the degree of loss caused by the compression process. Extended OCR accepts a value from 0 to 255:

- 255 guarantees the highest image quality

- 0 guarantees the best compression

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```csharp
public Nullable<byte> QualityFactor { get; set; }
```

VB
```vb
Public Property QualityFactor As Nullable(Of Byte)
    Get
    Set
```

Property Value
Type: Nullable(Byte)

See Also

PDFImageCompression Class
Aquaforest.ExtendedOCR.Api Namespace

PdfToImageBpp Enumeration

The Bits Per Pixel to use for the rasterized PDF page. This only applies for documents that are not processed using Native mode.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public enum PdfToImageBpp`

VB
`Public Enumeration PdfToImageBpp`

Members

Member name	Value	Description
Bpp_1	1	
Bpp_24	24	

See Also

Aquaforest.ExtendedOCR.Api Namespace

PdfToImageCompression Enumeration

The compression to use for images extracted or rasterized from each PDF page.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public enum PdfToImageCompression`

VB
`Public Enumeration PdfToImageCompression`

Members

Member name	Value	Description
CCITT4	0	The CCITT4 compression schema is used.
LZW	1	The LZW compression schema is used.

See Also

Aquaforest.ExtendedOCR.Api Namespace

PdfToImageDpi Enumeration

The DPI to use for the rasterized page.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public enum PdfToImageDpi`

VB
`Public Enumeration PdfToImageDpi`

Members

Member name	Value	Description
DPI_72	72	72 DPI
DPI_100	100	100 DPI
DPI_150	150	150 DPI
DPI_200	200	200 DPI
DPI_300	300	300 DPI
DPI_400	400	400 DPI
DPI_500	500	500 DPI
DPI_600	600	600 DPI

See Also

Aquaforest.ExtendedOCR.Api Namespace

PDFVersion Enumeration

Set the PDF Version of the output file.

Note: This will only affect documents that are not processed natively and documents that are converted to PDF/A.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public enum PDFVersion`

VB
`Public Enumeration PDFVersion`

Members

Member name	Value	Description
PDF_1_4	0	PDF 1.4 - 2001 - Acrobat 5.0 + JBIG2; RC4 encryption key lengths greater than 40bits (128bits); Extensible Metadata Platform (XMP); Tagged PDF
PDF_A1b	1	PDF 1.4 A-1b. Requires ExtractImageMethod.ConvertToTiff.
PDF_A1a	2	PDF 1.4 A-1a. Requires ExtractImageMethod.ConvertToTiff.
PDF_1_5	3	PDF 1.5 - 2003 - Acrobat 6.0 + JPEG 2000
PDF_1_6	4	PDF 1.6 - 2005 - Acrobat 7.0 + AES encryption
PDF_1_7	5	PDF 1.7 - 2006 - Acrobat 8.0 -> 2008 - ISO 32000-1:2008
PDF_A2b	6	PDF 1.7 A-2b. Requires ExtractImageMethod.ConvertToTiff.
PDF_A2a	7	PDF 1.7 A-2a. Requires ExtractImageMethod.ConvertToTiff.
PDF_1_7_3	8	PDF 1.7 - 2008 - Adobe Extension Level 3 / Acrobat 9.0 + 256-bit AES encryption
PDF_1_7_5	9	PDF 1.7 - 2009 - Adobe Extension Level 5 / Acrobat 9.1
PDF_1_7_8	10	PDF 1.7 - 2011 - Adobe Extension Level 8 / Acrobat X
PDF_A3b	11	PDF A-3b. Requires ExtractImageMethod.ConvertToTiff.

See Also

Aquaforest.ExtendedOCR.Api Namespace

PreProcessor Class

Inheritance Hierarchy

System.Object
  Aquaforest.ExtendedOCR.Api.PreProcessor

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public sealed class PreProcessor`

VB
`Public NotInheritable Class PreProcessor`

The **PreProcessor** type exposes the following members.

Constructors

	Name	Description
	PreProcessor	Initializes a new instance of the **PreProcessor** class

Properties

	Name	Description
	AdvancedDespeckle	Set AdvancedDespeckle settings.
	Autorotate	When set to 'true', it will automatically rotate pages so that text flows from left to right. The default value is 'false'.
	Binarization	Set the various binarization settings.
	BlankPageRemoval	Set blank page removal settings.
	Bpp	**Obsolete.**
	Deskew	Used to straighten the image. The default value is 'false'.
	Despeckle	Removes specks below the specified pixel size from the image. The default value is 0 and the maximum value is 9.
	Dpi	**Obsolete.**
	ForceTableZones	Setting this property to true can improve CSV and XLSX output.
	Interpolation	Set the interpolation settings.
	KeepOriginalImage	Set this to true if you want to use the pre-processed image for OCR but keep the original image in the output document. Default value is 'true'.

	LineRemoval	Set the various line removal settings.
	NonImagePDF	This allows control over the treatment of non image-only PDFs, i.e. PDFs that have some text in them as well as images. The default value for this setting is 'OCR'.
	PdfToImageBpp	The Bits Per Pixel to use for the rasterized PDF page when using engine 1. This only applies for documents that are not processed using Native mode. The default value for this property is taken from the PDF page. **Note:** This should only be set with guidance from technical support.
	PdfToImageCompression	The compression to set to the images generated or rasterized from each page of the source PDF file. These images are then OCRed to create the searchable PDF. The default value for this property is taken from each page in the source PDF file.
	PdfToImageDpi	The DPI to set to the images rasterized from each page of the source PDF file. These images are then OCRed to create the searchable PDF. The default value for this property is taken from each page in the source PDF file.
	PdfToImageForceVectorCheck	This setting is useful when dealing with documents that contains vector objects (e.g. CAD drawings). By default, pages that contain only vector objects are rasterized. Pages that do not have any images but contain vector objects as well as electronic text are skipped from rasterization. However, sometimes there can be a page that contains vector objects (CAD drawings) but its title may be in electronic text. To force rasterizing pages like these, set this property to true.
	PdfToImageIncludeText	When set to False this will prevent the conversion of real text (i.e. electronically generated as opposed to text that is part of a scanned image) from being rendered in the page images extracted from the PDF. This is because the text is already searchable and so generally does not require OCR. The value can be set to True however if the OCR is required on this real text.
	PdfToImageMaxRes	The maximum resolution of the rasterized images. If the resolution retrieved from the PDF page is bigger than this value, it will be set to this value. The default value for this property is 600.
	PdfToImageMinRes	The minimum resolution of the rasterized images. If the resolution retrieved from the PDF page is lower than this value, it will be set to this value. The default value for this property is 72.

	RemoveDarkBorders	Removes the dark surrounding from bi-tonal, gray-scale or color images. The dark surrounding of the image is whitened. **Note:** The dark border should be touching the edge of the image/page for this to work.

See Also

Aquaforest.ExtendedOCR.Api Namespace

PreProcessor Constructor

Initializes a new instance of the [PreProcessor](#) class

Namespace: [Aquaforest.ExtendedOCR.Api](#)
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public PreProcessor()```

VB
```Public Sub New```

See Also

[PreProcessor Class](#)
[Aquaforest.ExtendedOCR.Api Namespace](#)

PreProcessor.PreProcessor Properties

The PreProcessor type exposes the following members.

Properties

Name	Description
AdvancedDespeckle	Set AdvancedDespeckle settings.
Autorotate	When set to 'true', it will automatically rotate pages so that text flows from left to right. The default value is 'false'.
Binarization	Set the various binarization settings.
BlankPageRemoval	Set blank page removal settings.
Bpp	**Obsolete.**
Deskew	Used to straighten the image. The default value is 'false'.
Despeckle	Removes specks below the specified pixel size from the image. The default value is 0 and the maximum value is 9.
Dpi	**Obsolete.**
ForceTableZones	Setting this property to true can improve CSV and XLSX output.
Interpolation	Set the interpolation settings.
KeepOriginalImage	Set this to true if you want to use the pre-processed image for OCR but keep the original image in the output document. Default value is 'true'.
LineRemoval	Set the various line removal settings.
NonImagePDF	This allows control over the treatment of non image-only PDFs, i.e. PDFs that have some text in them as well as images. The default value for this setting is 'OCR'.
PdfToImageBpp	The Bits Per Pixel to use for the rasterized PDF page when using engine 1. This only applies for documents that are not processed using Native mode. The default value for this property is taken from the PDF page. **Note:** This should only be set with guidance from technical support.
PdfToImageCompression	The compression to set to the images generated or rasterized from each page of the source PDF file. These images are then OCRed to create the searchable PDF. The default value for this property is taken from each page in the source PDF file.
PdfToImageDpi	The DPI to set to the images rasterized from each page of the source PDF file. These images are then OCRed to create the searchable PDF. The default value for this property is taken from each page in the source PDF file.

PdfToImageForceVectorCheck	This setting is useful when dealing with documents that contains vector objects (e.g. CAD drawings). By default, pages that contain only vector objects are rasterized. Pages that do not have any images but contain vector objects as well as electronic text are skipped from rasterization. However, sometimes there can be a page that contains vector objects (CAD drawings) but its title may be in electronic text. To force rasterizing pages like these, set this property to true.
PdfToImageIncludeText	When set to False this will prevent the conversion of real text (i.e. electronically generated as opposed to text that is part of a scanned image) from being rendered in the page images extracted from the PDF. This is because the text is already searchable and so generally does not require OCR. The value can be set to True however if the OCR is required on this real text.
PdfToImageMaxRes	The maximum resolution of the rasterized images. If the resolution retrieved from the PDF page is bigger than this value, it will be set to this value. The default value for this property is 600.
PdfToImageMinRes	The minimum resolution of the rasterized images. If the resolution retrieved from the PDF page is lower than this value, it will be set to this value. The default value for this property is 72.
RemoveDarkBorders	Removes the dark surrounding from bi-tonal, gray-scale or color images. The dark surrounding of the image is whitened. **Note:** The dark border should be touching the edge of the image/page for this to work.

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.AdvancedDespeckle Property

Set AdvancedDespeckle settings.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public AdvancedDespeckle AdvancedDespeckle { get; set; }`

VB

`Public Property AdvancedDespeckle As AdvancedDespeckle`
`    Get`
`    Set`

Property Value
Type: AdvancedDespeckle

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.Autorotate Property

When set to 'true', it will automatically rotate pages so that text flows from left to right. The default value is 'false'.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> Autorotate { get; set; }```

VB
```Public Property Autorotate As Nullable(Of Boolean)   Get   Set```

Property Value
Type: Nullable(Boolean)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.Binarization Property

Set the various binarization settings.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Binarization Binarization { get; set; }```

VB

```Public Property Binarization As Binarization```
```   Get```
```   Set```

Property Value
Type: Binarization

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.BlankPageRemoval Property

Set blank page removal settings.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public BlankPageRemoval BlankPageRemoval { get; set; }```

VB

```Public Property BlankPageRemoval As BlankPageRemoval```
`    Get`
`    Set`

Property Value
Type: BlankPageRemoval

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.Bpp Property

Note: This API is now obsolete.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
``[ObsoleteAttribute("Use PdfToImageBpp instead")]
public Nullable<PdfToImageBpp> Bpp { get; set; }```

VB
``<ObsoleteAttribute("Use PdfToImageBpp instead")>
Public Property Bpp As Nullable(Of PdfToImageBpp)
Get
Set```

Property Value
Type: Nullable(PdfToImageBpp)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.Deskew Property

Used to straighten the image. The default value is 'false'.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> Deskew { get; set; }```

VB

```Public Property Deskew As Nullable(Of Boolean)```
```   Get```
```   Set```

Property Value
Type: Nullable(Boolean)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.Despeckle Property

Removes specks below the specified pixel size from the image. The default value is 0 and the maximum value is 9.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> Despeckle { get; set; }```

VB

```Public Property Despeckle As Nullable(Of Integer)```
```   Get```
```   Set```

Property Value
Type: Nullable(Int32)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.Dpi Property

Note: This API is now obsolete.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```
[ObsoleteAttribute("Use PdfToImageDpi instead")]
public Nullable<PdfToImageDpi> Dpi { get; set; }
```

VB
```
<ObsoleteAttribute("Use PdfToImageDpi instead")>
Public Property Dpi As Nullable(Of PdfToImageDpi)
      Get
      Set
```

Property Value
Type: Nullable(PdfToImageDpi)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.ForceTableZones Property

Setting this property to true can improve CSV and XLSX output.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> ForceTableZones { get; set; }```

VB

```Public Property ForceTableZones As Nullable(Of Boolean)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Boolean)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.Interpolation Property

Set the interpolation settings.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Interpolation Interpolation { get; set; }```

VB

```Public Property Interpolation As Interpolation```
 ```   Get```
 ```   Set```

Property Value
Type: Interpolation

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.KeepOriginalImage Property

Set this to true if you want to use the pre-processed image for OCR but keep the original image in the output document. Default value is 'true'.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public Nullable<bool> KeepOriginalImage { get; set; }`

VB

`Public Property KeepOriginalImage As Nullable(Of Boolean)`
`    Get`
`    Set`

Property Value
Type: Nullable(Boolean)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.LineRemoval Property

Set the various line removal settings.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public LineRemoval LineRemoval { get; set; }`

VB

`Public Property LineRemoval As LineRemoval`
`    Get`
`    Set`

Property Value
Type: LineRemoval

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.NonImagePDF Property

This allows control over the treatment of non image-only PDFs, i.e. PDFs that have some text in them as well as images. The default value for this setting is 'OCR'.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<NonImagePDF> NonImagePDF { get; set; }```

VB

```Public Property NonImagePDF As Nullable(Of NonImagePDF)```
```   Get```
```   Set```

Property Value
Type: Nullable(NonImagePDF)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.PdfToImageBpp Property

The Bits Per Pixel to use for the rasterized PDF page when using engine 1. This only applies for documents that are not processed using Native mode.
 The default value for this property is taken from the PDF page.
Note: This should only be set with guidance from technical support.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<PdfToImageBpp> PdfToImageBpp { get; set; }```

VB

```Public Property PdfToImageBpp As Nullable(Of PdfToImageBpp)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(PdfToImageBpp)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.PdfToImageCompression Property

The compression to set to the images generated or rasterized from each page of the source PDF file. These images are then OCRed to create the searchable PDF.
 The default value for this property is taken from each page in the source PDF file.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<PdfToImageCompression> PdfToImageCompression { get; set; }```

VB

```Public Property PdfToImageCompression As Nullable(Of PdfToImageCompression)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(PdfToImageCompression)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.PdfToImageDpi Property

The DPI to set to the images rasterized from each page of the source PDF file. These images are then OCRed to create the searchable PDF.
 The default value for this property is taken from each page in the source PDF file.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```
public Nullable<PdfToImageDpi> PdfToImageDpi { get; set; }
```

VB
```
Public Property PdfToImageDpi As Nullable(Of PdfToImageDpi)
      Get
      Set
```

Property Value
Type: Nullable(PdfToImageDpi)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.PdfToImageForceVectorCheck Property

This setting is useful when dealing with documents that contains vector objects (e.g. CAD drawings).
By default, pages that contain only vector objects are rasterized. Pages that do not have any images but contain vector objects as well as electronic text are skipped from rasterization. However, sometimes there can be a page that contains vector objects (CAD drawings) but its title may be in electronic text. To force rasterizing pages like these, set this property to true.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> PdfToImageForceVectorCheck { get; set; }```

VB

```Public Property PdfToImageForceVectorCheck As Nullable(Of Boolean)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Boolean)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.PdfToImageIncludeText Property

When set to False this will prevent the conversion of real text (i.e. electronically generated as opposed to text that is part of a scanned image) from being rendered in the page images extracted from the PDF. This is because the text is already searchable and so generally does not require OCR. The value can be set to True however if the OCR is required on this real text.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<bool> PdfToImageIncludeText { get; set; }```

VB

```Public Property PdfToImageIncludeText As Nullable(Of Boolean)```
```   Get```
```   Set```

Property Value
Type: Nullable(Boolean)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.PdfToImageMaxRes Property

The maximum resolution of the rasterized images. If the resolution retrieved from the PDF page is bigger than this value, it will be set to this value. The default value for this property is 600.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```csharp
public Nullable<int> PdfToImageMaxRes { get; set; }
```

VB
```vb
Public Property PdfToImageMaxRes As Nullable(Of Integer)
    Get
    Set
```

Property Value
Type: Nullable(Int32)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.PdfToImageMinRes Property

The minimum resolution of the rasterized images. If the resolution retrieved from the PDF page is lower than this value, it will be set to this value. The default value for this property is 72.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
```public Nullable<int> PdfToImageMinRes { get; set; }```

VB

```Public Property PdfToImageMinRes As Nullable(Of Integer)```
 ```   Get```
 ```   Set```

Property Value
Type: Nullable(Int32)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

PreProcessor.RemoveDarkBorders Property

Removes the dark surrounding from bi-tonal, gray-scale or color images. The dark surrounding of the image is whitened.

Note: The dark border should be touching the edge of the image/page for this to work.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
public Nullable<bool> RemoveDarkBorders { get; set; }

VB

Public Property RemoveDarkBorders As Nullable(Of Boolean)
        Get
        Set

Property Value
Type: Nullable(Boolean)

See Also

PreProcessor Class
Aquaforest.ExtendedOCR.Api Namespace

SegmentationMode Enumeration

IHQC segmentation modes.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public enum SegmentationMode`

VB
`Public Enumeration SegmentationMode`

Members

Member name	Value	Description
Color	0	IHQC segmentation only with input color image.
BWhite_and_color	1	IHQC segmentation with dual stream images.
Dualstream	2	IHQC segmentation with input color image and input black and white image.

See Also

Aquaforest.ExtendedOCR.Api Namespace

SupportedLanguages Enumeration

Extended OCR accepts up to 8 recognition languages at a time. This is helpful to process mixed documents but, because of the various character sets, not all combinations are allowed. For this reason, the multiple languages support is limited to a single alphabet. For example, Russian and French can't be mixed.

- **Note:**
  Asian languages can't be mixed.

- Extended OCR SDK cannot load more than one Asian language at a time.

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public enum SupportedLanguages`

VB
`Public Enumeration SupportedLanguages`

Members

Member name	Value	Description
English	0	English (American)
German	1	
French	2	
Spanish	3	
Italian	4	
British	5	
Swedish	6	
Danish	7	
Norwegian	8	
Dutch	9	
Portuguese	10	
Brazilian	11	
Galician	12	
Icelandic	13	

Greek	14	
Czech	15	
Hungarian	16	
Polish	17	
Romanian	18	
Slovak	19	
Croatian	20	
Serbian	21	
Slovenian	22	
Luxemb	23	
Finnish	24	
Turkish	25	
Russian	26	
Byelorussian	27	
Ukrainian	28	
Macedonian	29	
Bulgarian	30	
Estonian	31	
Lithuanian	32	
Afrikaans	33	
Albanian	34	
Catalan	35	
Irish_Gaelic	36	
Scottish_Gaelic	37	
Basque	38	
Breton	39	
Corsican	40	
Frisian	41	
Nynorsk	42	
Indonesian	43	
Malay	44	
Swahili	45	
Tagalog	46	
Japanese	47	You will need to have the Asian OCR license to use this language.
Korean	48	You will need to have the Asian OCR license to use this language.

Schinese	49	You will need to have the Asian OCR license to use this language.
Tchinese	50	You will need to have the Asian OCR license to use this language.
Quecha	51	
Aymara	52	
Faroese	53	
Friulian	54	
Greenlandic	55	
Haitian_Creole	56	
Rhaeto_Roman	57	
Sardinian	58	
Kurdish	59	
Cebuano	60	
Bemba	61	
Chamorro	62	
Fijan	63	
Ganda	64	
Hani	65	
Ido	66	
Interlingua	67	
Kicongo	68	
Kinyarwanda	69	
Malagasy	70	
Maori	71	
Mayan	72	
Minangkabau	73	
Nahuatl	74	
Nyanja	75	
Rundi	76	
Samoan	77	
Shona	78	
Somali	79	
Sotho	80	
Sundanese	81	
Tahitian	82	
Tonga	83	

Tswana	84	
Wolof	85	
Xhosa	86	
Zapotec	87	
Javanese	88	
Pidgin_Nigeria	89	
Occitan	90	
Manx	91	
Tok_Pisin	92	
Bislama	93	
Hiligaynon	94	
Kapampangan	95	
Balinese	96	
Bikol	97	
Ilocano	98	
Madurese	99	
Waray	100	
None	101	No language, Latin alphabet
Serbian_Latin	102	
Latin	103	
Latvian	104	
Numeric	114	This language limits recognition to numeric characters.
Esperanto	115	
Maltese	116	
Zulu	117	
Afaan	118	
Asturian	119	
AzeriLatin	120	
Luba	121	
Papamiento	122	
Tatar	123	
Turkmen	124	
Welsh	125	
Arabic	126	You will need to have the Arabic OCR license to use this language.

		Note: Works only for Arabic texts with embedded English words. The result for a zone with only English will be empty.
Farsi	127	You will need to have the Arabic OCR license to use this language.
Mexican	128	
BosnianLatin	129	Bosnian (Latin). CharsetCategory.E
BosnianCyrillic	130	Bosnian (Cyrillic). CharsetCategory.D
Moldovan	131	Moldovan. CharsetCategory.E
SwissGerman	132	German (Switzerland). CharsetCategory.C
Tetum	133	Tetum. CharsetCategory.C
Kazakh	134	Kazakh (Cyrillic). CharsetCategory.D
MongolianCyrillic	135	Mongolian (Cyrillic). CharsetCategory.D
UzbekLatin	136	Uzbek (Latin). CharsetCategory.C

See Also
Aquaforest.ExtendedOCR.Api Namespace

TextDirection Enumeration

The text direction (for Asian languages only).

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public enum TextDirection`

VB
`Public Enumeration TextDirection`

Members

Member name	Value	Description
Unknown	0	Unknown text direction.
LeftToRight	1	The text runs from left to right.
TopToBottom	2	Text goes top to bottom.
RightToLeft	3	Text goes right to left.

See Also

Aquaforest.ExtendedOCR.Api Namespace

TextSpacing Enumeration

The text spacings

Namespace: Aquaforest.ExtendedOCR.Api
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public enum TextSpacing`

VB
`Public Enumeration TextSpacing`

Members

Member name	Value	Description
Fixed	0	Fixed text spacing. You need to set a value for CharactersPerInch when using fixed text spacing.
Proportional	1	Proportional text spacing.
Unknown	2	Unknown text spacing.

See Also

Aquaforest.ExtendedOCR.Api Namespace

TextType Enumeration

The text types

Namespace: [Aquaforest.ExtendedOCR.Api](#)
Assembly: Aquaforest.ExtendedOCR.Api (in Aquaforest.ExtendedOCR.Api.dll) Version: 2.20

Syntax

C#
`public enum TextType`

VB
`Public Enumeration TextType`

Members

Member name	Value	Description
Normal	0	Normal text.
DotMatrix	1	Dot matrix text.

See Also

[Aquaforest.ExtendedOCR.Api Namespace](#)

Aquaforest.ExtendedOCR.Shared Namespace

Classes

Class	Description
CharacterData	Contains information describing a single character extracted from the Extended OCR engine.
StatusUpdateEventArgs	
WordData	
Words	This class contains a collection of WordData objects which are available on a page by page basis.

CharacterData Class

Contains information describing a single character extracted from the Extended OCR engine.

Inheritance Hierarchy

System.Object
  Aquaforest.ExtendedOCR.Shared.CharacterData

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
`public sealed class CharacterData`

VB
`Public NotInheritable Class CharacterData`

The **CharacterData** type exposes the following members.

Constructors

	Name	Description
⬖	CharacterData	Initializes a new instance of the **CharacterData** class

Properties

	Name	Description
🔧	Baseline	Gets the Y-coordinate of the bottom edge of the character in pixels.
🔧	Character	Gets the string representation of the character.
🔧	Height	Gets the height of the character in pixels.
🔧	Width	Gets the width of the character in pixels.
🔧	[X](X)	Gets the X-coordinate of the left edge of the character in pixels.
🔧	[Y](Y)	Gets the Y-coordinate of the top edge of the character in pixels.

See Also

Aquaforest.ExtendedOCR.Shared Namespace

CharacterData Constructor

Initializes a new instance of the [CharacterData](#) class

Namespace: [Aquaforest.ExtendedOCR.Shared](#)
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#

```csharp
public CharacterData(
      string character,
      int x,
      int y,
      int width,
      int height,
      int baseline
)
```

VB

```vb
Public Sub New (
      character As String,
      x As Integer,
      y As Integer,
      width As Integer,
      height As Integer,
      baseline As Integer
)
```

Parameters
character
Type: [System.String](#)
x
Type: [System.Int32](#)
y
Type: [System.Int32](#)
width
Type: [System.Int32](#)
height
Type: [System.Int32](#)
baseline
Type: [System.Int32](#)

See Also

[CharacterData Class](#)
[Aquaforest.ExtendedOCR.Shared Namespace](#)

CharacterData.CharacterData Properties

The CharacterData type exposes the following members.

Properties

	Name	Description
	Baseline	Gets the Y-coordinate of the bottom edge of the character in pixels.
	Character	Gets the string representation of the character.
	Height	Gets the height of the character in pixels.
	Width	Gets the width of the character in pixels.
	X	Gets the X-coordinate of the left edge of the character in pixels.
	Y	Gets the Y-coordinate of the top edge of the character in pixels.

See Also

CharacterData Class
Aquaforest.ExtendedOCR.Shared Namespace

CharacterData.Baseline Property

Gets the Y-coordinate of the bottom edge of the character in pixels.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public int Baseline { get; }```

VB
```Public ReadOnly Property Baseline As Integer```   ```   Get```

Property Value
Type: Int32

See Also

CharacterData Class
Aquaforest.ExtendedOCR.Shared Namespace

CharacterData.Character Property

Gets the string representation of the character.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public string Character { get; }```

VB
```Public ReadOnly Property Character As String``` ```   Get```

Property Value
Type: String

See Also

CharacterData Class
Aquaforest.ExtendedOCR.Shared Namespace

CharacterData.Height Property

Gets the height of the character in pixels.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
`public int Height { get; }`

VB
`Public ReadOnly Property Height As Integer` `    Get`

Property Value
Type: Int32

See Also

CharacterData Class
Aquaforest.ExtendedOCR.Shared Namespace

CharacterData.Width Property

Gets the width of the character in pixels.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public int Width { get; }```

VB
```Public ReadOnly Property Width As Integer``` ```   Get```

Property Value
Type: Int32

See Also

CharacterData Class
Aquaforest.ExtendedOCR.Shared Namespace

CharacterData.X Property

Gets the X-coordinate of the left edge of the character in pixels.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public int X { get; }```

VB
```Public ReadOnly Property X As Integer``` ```   Get```

Property Value
Type: Int32

See Also

CharacterData Class
Aquaforest.ExtendedOCR.Shared Namespace

CharacterData.Y Property

Gets the Y-coordinate of the top edge of the character in pixels.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public int Y { get; }```

VB
```Public ReadOnly Property Y As Integer```   ```   Get```

Property Value
Type: Int32

See Also

CharacterData Class
Aquaforest.ExtendedOCR.Shared Namespace

StatusUpdateEventArgs Class

Inheritance Hierarchy

System.Object
  System.EventArgs
    Aquaforest.ExtendedOCR.Shared.StatusUpdateEventArgs

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public sealed class StatusUpdateEventArgs : EventArgs```

VB
```Public NotInheritable Class StatusUpdateEventArgs Inherits EventArgs```

The **StatusUpdateEventArgs** type exposes the following members.

Constructors

Name	Description
StatusUpdateEventArgs()	Initializes a new instance of the **StatusUpdateEventArgs** class
StatusUpdateEventArgs(StatusUpdateEventArgs)	Initializes a new instance of the **StatusUpdateEventArgs** class

Properties

Name	Description
BlankPage	Indicates whether the page was detected as blank.
ImageAvailable	Indicates whether an image was successfully extracted (after applying all the appropriate pre-processing settings).
PageNumber	Returns page for which the object relates to.
Rotation	The rotation in Degrees (°) of the current page. If Autorotate is set to false this will always be 0.
TextAvailable	Indicates whether text was extracted for the page.

See Also

Aquaforest.ExtendedOCR.Shared Namespace

StatusUpdateEventArgs Constructor

Overload List

	Name	Description
▤◆	StatusUpdateEventArgs()	Initializes a new instance of the StatusUpdateEventArgs class
▤◆	StatusUpdateEventArgs(StatusUpdateEventArgs)	Initializes a new instance of the StatusUpdateEventArgs class

See Also

StatusUpdateEventArgs Class
Aquaforest.ExtendedOCR.Shared Namespace

StatusUpdateEventArgs Constructor

Initializes a new instance of the StatusUpdateEventArgs class

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
public StatusUpdateEventArgs()

VB
Public Sub New

See Also

StatusUpdateEventArgs Class
StatusUpdateEventArgs Overload
Aquaforest.ExtendedOCR.Shared Namespace

StatusUpdateEventArgs Constructor (StatusUpdateEventArgs)

Initializes a new instance of the StatusUpdateEventArgs class

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
| ```
public StatusUpdateEventArgs(
      StatusUpdateEventArgs original
)
``` |

VB
| ```
Public Sub New (
      original As StatusUpdateEventArgs
)
``` |

Parameters
original
Type: Aquaforest.ExtendedOCR.Shared.StatusUpdateEventArgs

See Also

StatusUpdateEventArgs Class
StatusUpdateEventArgs Overload
Aquaforest.ExtendedOCR.Shared Namespace

StatusUpdateEventArgs.StatusUpdateEventArgs Properties

The StatusUpdateEventArgs type exposes the following members.

Properties

	Name	Description
	BlankPage	Indicates whether the page was detected as blank.
	ImageAvailable	Indicates whether an image was successfully extracted (after applying all the appropriate pre-processing settings).
	PageNumber	Returns page for which the object relates to.
	Rotation	The rotation in Degrees (°) of the current page. If Autorotate is set to false this will always be 0.
	TextAvailable	Indicates whether text was extracted for the page.

See Also

StatusUpdateEventArgs Class
Aquaforest.ExtendedOCR.Shared Namespace

StatusUpdateEventArgs.BlankPage Property

Indicates whether the page was detected as blank.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
`public bool BlankPage { get; }`

VB
`Public ReadOnly Property BlankPage As Boolean` `    Get`

Property Value
Type: Boolean

See Also

StatusUpdateEventArgs Class
Aquaforest.ExtendedOCR.Shared Namespace

StatusUpdateEventArgs.ImageAvailable Property

Indicates whether an image was successfully extracted (after applying all the appropriate pre-processing settings).

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
`public bool ImageAvailable { get; }`

VB
`Public ReadOnly Property ImageAvailable As Boolean` `    Get`

Property Value
Type: Boolean

See Also

StatusUpdateEventArgs Class
Aquaforest.ExtendedOCR.Shared Namespace

StatusUpdateEventArgs.PageNumber Property

Returns page for which the object relates to.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public int PageNumber { get; }```

VB
```Public ReadOnly Property PageNumber As Integer```   ```   Get```

Property Value

Type: Int32

See Also

StatusUpdateEventArgs Class

Aquaforest.ExtendedOCR.Shared Namespace

StatusUpdateEventArgs.Rotation Property

The rotation in Degrees (°) of the current page. If Autorotate is set to false this will always be 0.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
public int Rotation { get; }

VB
Public ReadOnly Property Rotation As Integer   Get

Property Value
Type: Int32

See Also

StatusUpdateEventArgs Class
Aquaforest.ExtendedOCR.Shared Namespace

StatusUpdateEventArgs.TextAvailable Property

Indicates whether text was extracted for the page.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public bool TextAvailable { get; }```

VB
```Public ReadOnly Property TextAvailable As Boolean``` ```   Get```

Property Value
Type: Boolean

See Also

StatusUpdateEventArgs Class
Aquaforest.ExtendedOCR.Shared Namespace

WordData Class

Inheritance Hierarchy

System.Object
  Aquaforest.ExtendedOCR.Shared.WordData

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
`public sealed class WordData`

VB
`Public NotInheritable Class WordData`

The **WordData** type exposes the following members.

Constructors

	Name	Description
	WordData	Initializes a new instance of the **WordData** class

Properties

	Name	Description
	Bottom	Gets the Y-coordinate of the bottom edge of the word in pixels.
	CharacterList	Gets the list of characters in the word.
	Height	Gets the height of the word in pixels.
	Left	Gets the X-coordinate of the left edge of the word in pixels.
	Top	Gets the Y-coordinate of the top edge of the word in pixels.
	Width	Gets the width of the word in pixels.
	Word	Gets the string representation of the word.

Methods

	Name	Description
	Add	Add the specified character data.

See Also

Aquaforest.ExtendedOCR.Shared Namespace

WordData Constructor

Initializes a new instance of the WordData class

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
`public WordData()`

VB
`Public Sub New`

See Also

WordData Class
Aquaforest.ExtendedOCR.Shared Namespace

WordData.WordData Properties

The WordData type exposes the following members.

Properties

	Name	Description
	Bottom	Gets the Y-coordinate of the bottom edge of the word in pixels.
	CharacterList	Gets the list of characters in the word.
	Height	Gets the height of the word in pixels.
	Left	Gets the X-coordinate of the left edge of the word in pixels.
	Top	Gets the Y-coordinate of the top edge of the word in pixels.
	Width	Gets the width of the word in pixels.
	Word	Gets the string representation of the word.

See Also

WordData Class
Aquaforest.ExtendedOCR.Shared Namespace

WordData.Bottom Property

Gets the Y-coordinate of the bottom edge of the word in pixels.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
`public int Bottom { get; }`

VB
`Public ReadOnly Property Bottom As Integer` `    Get`

Property Value
Type: Int32

See Also

WordData Class
Aquaforest.ExtendedOCR.Shared Namespace

WordData.CharacterList Property

Gets the list of characters in the word.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
`public List<CharacterData> CharacterList { get; }`

VB
`Public ReadOnly Property CharacterList As List(Of CharacterData)` `        Get`

Property Value
Type: List(CharacterData)

See Also

WordData Class
Aquaforest.ExtendedOCR.Shared Namespace

WordData.Height Property

Gets the height of the word in pixels.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public int Height { get; }```

VB
```Public ReadOnly Property Height As Integer```   ```   Get```

Property Value
Type: Int32

See Also

WordData Class
Aquaforest.ExtendedOCR.Shared Namespace

WordData.Left Property

Gets the X-coordinate of the left edge of the word in pixels.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public int Left { get; }```

VB
```Public ReadOnly Property Left As Integer``` ```   Get```

Property Value
Type: Int32

See Also

WordData Class
Aquaforest.ExtendedOCR.Shared Namespace

WordData.Top Property

Gets the Y-coordinate of the top edge of the word in pixels.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
`public int Top { get; }`

VB
`Public ReadOnly Property Top As Integer` `        Get`

Property Value
Type: Int32

See Also

WordData Class
Aquaforest.ExtendedOCR.Shared Namespace

WordData.Width Property

Gets the width of the word in pixels.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public int Width { get; }```

VB
```Public ReadOnly Property Width As Integer```   ```   Get```

Property Value
Type: Int32

See Also

WordData Class
Aquaforest.ExtendedOCR.Shared Namespace

WordData.Word Property

Gets the string representation of the word.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public string Word { get; }```

VB
```Public ReadOnly Property Word As String``` ```   Get```

Property Value
Type: String

See Also

WordData Class
Aquaforest.ExtendedOCR.Shared Namespace

WordData.WordData Methods

The [WordData](#) type exposes the following members.

Methods

	Name	Description
=◆	[Add](#)	Add the specified character data.

See Also

[WordData Class](#)
[Aquaforest.ExtendedOCR.Shared Namespace](#)

WordData.Add Method

Add the specified character data.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
| ```
public void Add(
        CharacterData characterData
)
``` |

VB
| ```
Public Sub Add (
        characterData As CharacterData
)
``` |

Parameters

characterData

Type: Aquaforest.ExtendedOCR.Shared.CharacterData
The character data to add.

See Also

WordData Class
Aquaforest.ExtendedOCR.Shared Namespace

Words Class

This class contains a collection of WordData objects which are available on a page by page basis.

Inheritance Hierarchy

System.Object
  Aquaforest.ExtendedOCR.Shared.Words

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
``public sealed class Words : IList<WordData>,
ICollection<WordData>, IEnumerable<WordData>, IEnumerable```

VB
``Public NotInheritable Class Words
Implements IList(Of WordData), ICollection(Of WordData),
IEnumerable(Of WordData), IEnumerable```

The **Words** type exposes the following members.

Constructors

	Name	Description
	Words	Initializes a new instance of the **Words** class

Properties

	Name	Description
	Count	Gets the number of items contained in the IList.
	Height	Gets the height of the character at the current index.
	IsReadOnly	Gets a value indicating whether the IList is read-only.
	Item	Gets or sets the element at the specified index.
	Width	Gets the width of the character at the current index.

Methods

	Name	Description
	Add	Adds an item to the IList.
	Clear	Removes all items from the IList.
	Contains	Determines whether the IList contains the specified item.
	CopyTo	Copies the elements of the IList to an Array, starting at a particular Array index.

	GetEnumerator	Returns an enumerator that iterates through the collection.
	GetFirst	Gets the first item in the IList.
	GetHeight	Gets the height of the item at the specified index.
	GetNext	Gets the next item in the IList.
	GetWidth	Gets the width of the item at the specified index.
	IndexOf	Determines the index of the specified item in the IList.
	Insert	Inserts an item to the IList at the specified index.
	Remove	Removes the first occurrence of a specific item from the IList.
	RemoveAt	Removes the IList item at the specified index.

See Also
Aquaforest.ExtendedOCR.Shared Namespace

Words Constructor

Initializes a new instance of the Words class

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
`public Words()`

VB
`Public Sub New`

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.Words Properties

The Words type exposes the following members.

Properties

	Name	Description
	Count	Gets the number of items contained in the IList.
	Height	Gets the height of the character at the current index.
	IsReadOnly	Gets a value indicating whether the IList is read-only.
	Item	Gets or sets the element at the specified index.
	Width	Gets the width of the character at the current index.

See Also

Words Class

Aquaforest.ExtendedOCR.Shared Namespace

Words.Count Property

Gets the number of items contained in the IList.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public int Count { get; }```

VB
```Public ReadOnly Property Count As Integer```   ```   Get```

Property Value
Type: Int32

Implements
ICollection(T).Count

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.Height Property

Gets the height of the character at the current index.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
`public int Height { get; }`

VB
`Public ReadOnly Property Height As Integer` `    Get`

Property Value
Type: Int32

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.IsReadOnly Property

Gets a value indicating whether the IList is read-only.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
`public bool IsReadOnly { get; }`

VB
`Public ReadOnly Property IsReadOnly As Boolean` `    Get`

Property Value
Type: Boolean

Implements
ICollection(T).IsReadOnly

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.Item Property

Gets or sets the element at the specified index.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
| ```
public WordData this[
      int index
] { get; set; }
``` |

VB
| ```
Public Default Property Item (
      index As Integer
) As WordData
      Get
      Set
``` |

Parameters
index
Type: System.Int32
The zero-based index of the element to get or set.

Return Value
Type: WordData
The element at the specified index.

Implements
IList(T).Item(Int32)

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.Width Property

Gets the width of the character at the current index.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public int Width { get; }```

VB
```Public ReadOnly Property Width As Integer```   ```   Get```

Property Value
Type: Int32

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.Words Methods

The Words type exposes the following members.

Methods

	Name	Description
	Add	Adds an item to the IList.
	Clear	Removes all items from the IList.
	Contains	Determines whether the IList contains the specified item.
	CopyTo	Copies the elements of the IList to an Array, starting at a particular Array index.
	GetEnumerator	Returns an enumerator that iterates through the collection.
	GetFirst	Gets the first item in the IList.
	GetHeight	Gets the height of the item at the specified index.
	GetNext	Gets the next item in the IList.
	GetWidth	Gets the width of the item at the specified index.
	IndexOf	Determines the index of the specified item in the IList.
	Insert	Inserts an item to the IList at the specified index.
	Remove	Removes the first occurrence of a specific item from the IList.
	RemoveAt	Removes the IList item at the specified index.

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.Add Method

Adds an item to the IList.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
| ```
public void Add(
      WordData item
)
``` |

VB
| ```
Public Sub Add (
      item As WordData
)
``` |

Parameters

item
Type: Aquaforest.ExtendedOCR.Shared.WordData
The WordData to add to the IList.

Implements
ICollection(T).Add(T)

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.Clear Method

Removes all items from the IList.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public void Clear()```

VB
```Public Sub Clear```

Implements
ICollection(T).Clear()

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.Contains Method

Determines whether the IList contains the specified item.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
| ```
public bool Contains(
      WordData item
)
``` |

VB
| ```
Public Function Contains (
      item As WordData
) As Boolean
``` |

Parameters

item

Type: Aquaforest.ExtendedOCR.Shared.WordData

The item to locate in the IList.

Return Value

Type: Boolean

true if the item is found; otherwise, false.

Implements

ICollection(T).Contains(T)

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.CopyTo Method

Copies the elements of the IList to an Array, starting at a particular Array index.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
| ```
public void CopyTo(
      WordData[] array,
      int arrayIndex
)
``` |

VB
| ```
Public Sub CopyTo (
      array As WordData(),
      arrayIndex As Integer
)
``` |

Parameters

array
Type: Aquaforest.ExtendedOCR.Shared.WordData[]
The one-dimensional Array that is the destination of the elements copied from IList. The Array must have zero-based indexing.

arrayIndex
Type: System.Int32
The zero-based index in array at which copying begins.

Implements
ICollection(T).CopyTo(T[], Int32)

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.GetEnumerator Method

Returns an enumerator that iterates through the collection.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public IEnumerator<WordData> GetEnumerator()```

VB
```Public Function GetEnumerator As IEnumerator(Of WordData)```

Return Value
Type: IEnumerator(WordData)
An IEnumerator object that can be used to iterate through the collection of WordData.

Implements
IEnumerable(T).GetEnumerator()

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.GetFirst Method

Gets the first item in the IList.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
```public WordData GetFirst()```

VB
```Public Function GetFirst As WordData```

Return Value
Type: WordData

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.GetHeight Method

Gets the height of the item at the specified index.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#

```csharp
public int GetHeight(
    int index
)
```

VB

```vb
Public Function GetHeight (
    index As Integer
) As Integer
```

Parameters
index
Type: System.Int32
The zero-based index of the item.

Return Value
Type: Int32
The height of the item at the specified index if found in the list; otherwise, 0.

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.GetNext Method

Gets the next item in the IList.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
`public WordData GetNext()`

VB
`Public Function GetNext As WordData`

Return Value
Type: WordData

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.GetWidth Method

Gets the width of the item at the specified index.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
| ```
public int GetWidth(
    int index
)
``` |

VB
| ```
Public Function GetWidth (
    index As Integer
) As Integer
``` |

Parameters
index
Type: System.Int32
The zero-based index of the item.

Return Value
Type: Int32
The width of the item at the specified index if found in the list; otherwise, 0.

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.IndexOf Method

Determines the index of the specified item in the IList.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#

```public int IndexOf(```
 ```   WordData item```
 ```)```

VB

```Public Function IndexOf (```
 ```   item As WordData```
 ```) As Integer```

Parameters
item
Type: Aquaforest.ExtendedOCR.Shared.WordData
The item to locate in the IList.

Return Value
Type: Int32
The index of value if found in the list; otherwise, -1.

Implements
IList(T).IndexOf(T)

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.Insert Method

Inserts an item to the IList at the specified index.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
| ```
public void Insert(
    int index,
    WordData item
)
``` |

VB
| ```
Public Sub Insert (
    index As Integer,
    item As WordData
)
``` |

Parameters
index
Type: System.Int32
The zero-based index at which value should be inserted.

item
Type: Aquaforest.ExtendedOCR.Shared.WordData
The item to insert into the IList.

Implements
IList(T).Insert(Int32, T)

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.Remove Method

Removes the first occurrence of a specific item from the IList.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

```C#
public bool Remove(
    WordData item
)
```

```VB
Public Function Remove (
    item As WordData
) As Boolean
```

Parameters
item
Type: Aquaforest.ExtendedOCR.Shared.WordData
The item to remove from the IList.

Return Value
Type: Boolean

Implements
ICollection(T).Remove(T)

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace

Words.RemoveAt Method

Removes the IList item at the specified index.

Namespace: Aquaforest.ExtendedOCR.Shared
Assembly: Aquaforest.ExtendedOCR.Shared (in Aquaforest.ExtendedOCR.Shared.dll) Version: 2.20

Syntax

C#
| ```
public void RemoveAt(
    int index
)
``` |

VB
| ```
Public Sub RemoveAt (
    index As Integer
)
``` |

Parameters
index
Type: System.Int32
The zero-based index of the item to remove.

Implements
IList(T).RemoveAt(Int32)

See Also

Words Class
Aquaforest.ExtendedOCR.Shared Namespace