Tungsten

OmniPage Capture SDK™
for Windows

Product Summary

TUNGSTEN
AUTOMATION

Tungsten OmniPage Capture SDK for Windows

The Market-Leading OCR and Imaging SDK

Integrate High-Tech Imaging, OCR Recognition and PDF Capabilities into Your Most Critical Applications

Established as the core technology behind all Tungsten imaging products, OmniPage Capture SDK is widely recognized as the top-tier imaging and OCR toolkit on the market today.

The updated OmniPage CSDK 22 provides everything you need to add robust imaging, OCR recognition and PDF capabilities into your most critical applications. In addition to outstanding speed and accuracy with numerous product performance options and best-in-class PDF management, OmniPage CSDK delivers powerful form-recognition and document classification technologies and unique tools to create reliable OCR and document processing applications with minimal programming requirement.

Used by commercial vendors who are serious about high OCR accuracy and quality document imaging in their applications, the OmniPage CSDK provides scalable recognition, extraordinary PDF support and a simple API that lets you create high-value, competitive products.

OmniPage Capture SDK Benefits

The MOST accurate OCR available

OmniPage CSDK provides a scalable voting interface and significant throughput management capabilities. Combined with highly accurate machine-print OCR (OCR, OCR-A, OCR-B and MICR), handprint (ICR), checkmark (OMR) and barcode (1D and 2D) recognition engines, it delivers unmatched flexibility and accuracy.

Simple to use APIs

The OmniPage CSDK offers easy-to-use, yet versatile APIs that enable you to productively build your application and control the characteristics of the recognition and conversion processes. It supports programming in C/C++, .NET, C#, VB.NET, Java, and Python and is available in 32-bit and 64-bit versions.

Enhanced Table Algorithms

OmniPage CSDK 22 includes greater sophisticated parsing logic that can better handle situations such as tables with complex headers, merged cells, financial tables, overrun columns, sparse tables with little data, and tables with faint grid lines that would get dropped during binarization, providing 22% fewer errors than the previous version.

Document Classifier

Documents exist in different types and layouts and it's a common task to identify and categorize them based on their diverse contents. The new Document Classifier module enables users to separate different documents and sort them based on layout and/or textual similarities. This intelligent and trainable tool can be used in a wide range of business processes, such as email forwarding, mail-room automation, spam filtering, form processing and data extraction.

PDF supercompression with new high level MRC compression

Save valuable storage space and consume less bandwidth with PDF super compression. MRC (mixed raster content) PDF compression reduces file sizes down to 2% of their original size. This technology separates PDF images into multiple layers and applies optimized compression algorithms to each layer leading to dramatic compression ratios. OmniPage CSDK 22 now provides a level of compression which triples the resolution of text layer and smooths out character contours, providing the highest quality output of page reconstruction.

Powerful forms processing capabilities

OmniPage CSDK offers various tools and capabilities required to create complete form processing applications. It enables users to easily define form templates for all the document types they wish to process and contains a convenient API to execute form matching and data extraction.

Reliable, highly available applications

The OmniPage CSDK includes new, innovative capabilities that enable developers to create reliable, long running document processing applications that can automatically scale up and take advantage of all or the designated CPU power of a multicore machine. All these can be achieved with low programming effort.

Broad range of output formats

The OmniPage CSDK supports a wide range of image and application formats for outputting the conversion results, including BMP, GIF, TIF, PDF, HTML, Microsoft Office formats (.DOCX, .XLSX, .PPTX), TXT, XML, ePub and more. New in OmniPage CSDK 22 is support for both hOCR and ALTO output standards.

Comprehensive language support and updated Eastern language support

Tungsten's advanced OCR technology supports more than 125 languages of the world based on the Latin, Greek, Cyrillic alphabets as well as Arabic, Hebrew, Chinese (Traditional and Simplified), Vietnamese, Japanese, Korean and Thai.

Multilingual content is constantly improved upon, the latest version providing Asian multi-threading for Chinese, Japanese, Korean and Arabic. Arabic and Hebrew OCR improvements now lead to significantly more accurate results with embedded English characters.

Integrated PDF toolkit with Added PDF/A-4 support

Extensive PDF capabilities—including unique PDF overlay matching that achieves up to 100% accuracy in PDF conversion—allow you to significantly reduce development costs and time-to-market. OmniPage CSDK supports output to all PDF/A-4 levels (the highest format available), ideal for long term document archival and can generate PDF-MRC (mixed raster content) files optimized for file size and quality.

Easy and flexible licensing

OmniPage CSDK's licensing options were created with the users' needs in mind. It does not require the use of dongles. Choose your preferred type of license:

- Developer licenses to install OmniPage CSDK and use it to develop custom solutions. It also includes a number of test runtime licenses to try out the built application in the target environment

- Runtime licenses to distribute and install the developed applications together with the required OmniPage CSDK components. Runtime license options are available on the page count bases or for permanent use

- Licenses for add-ons of advanced document imaging technologies are available both as developer's and runtime licenses to expand the capabilities of the application as the business requires.

Key Benefits

- The world's most accurate OCR solution increases productivity, lowers costs and maximizes ROI

- Delivers everything you need for scanning, OCR, ICR, OMR, barcode, PDF and document conversion

- Enables developers to provide higher value to customers with updated and enhanced functionality

- Provides an easy-to-use API to shorten development cycles and accelerates time to market

- Helps your organization differentiate itself from the competition with the most advanced scanning, OCR and PDF technologies

Product Configurations

The OmniPage Capture SDK is available in one base kit with optional add-ons:

The Base Western OCR Kit

- C/C++ Libraries for Western OCR engines

- Access to individual OCR engines for application optimization

- OCR-A, OCR-B, E-13B (MICR)

- PDF Technology: Support for PDF 2.0, PDF/A, export to PDF normal, PDF image-only, PDF image on text formats and high-compression rate PDF-MRC
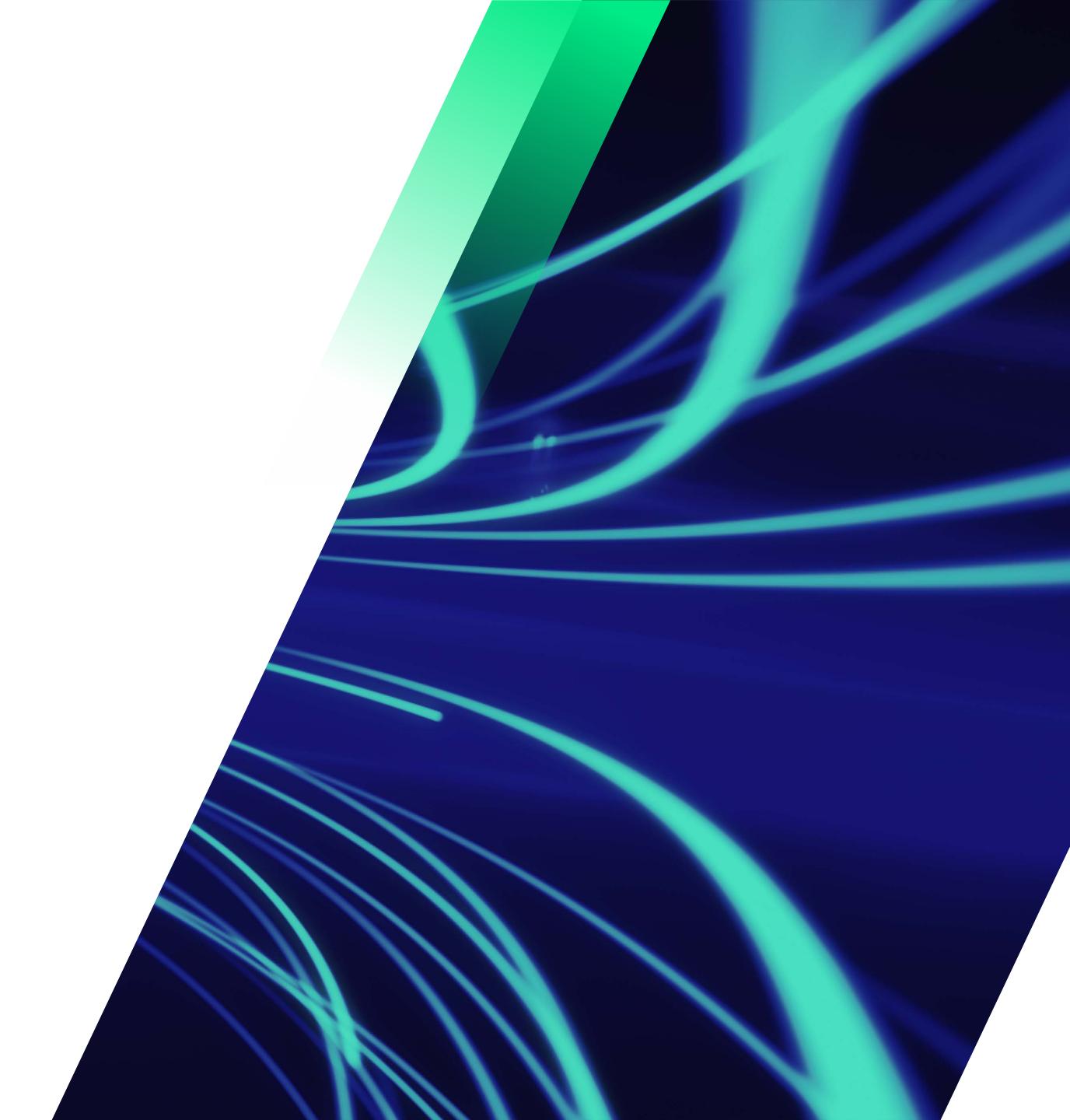
Add-on options

- Asian Plus Add-on: OCR engines for Chinese (Simplified and Traditional), Vietnamese, Japanese, Korean, Hebrew, Arabic and Thai

- Recognition Add-on: Barcode, ICR, OMR, Form Template Editor, Document Classifier Assistant

System Requirements

Development/Runtime environment

- Windows 10, 11

- Windows Server, 2012 R2, 2016, 2019, and 2022 Windows Server Core in Docker containers

- Microarchitecture Feature Level x86-64-v2 compatible processor required

- Microarchitecture Feature Level x86-64-v3 compatible processor recommended

  Note:
  - Microarchitecture Feature Levels defined by AMD, Intel, Red Hat, and SUSE
  - Examples for x86-64-v2: minimum 1st gen. Intel Core i5/i7, AMD K10
  - Examples for x86-64-v3: minimum 4th gen. Intel Core, AMD Excavator

- 4 GB RAM (6 GB recommended)

- 4 GB free disk space - Development 450 MB free disk space (less if not all recognition modules are distributed) - Runtime Microsoft Visual C++ Runtime 2015-2019 or later Microsoft .Net Framework 4.6.2 or later (for .Net applications) Microsoft .Net Core 2.1 or later (for .Net Core applications) Java 1.2 or later (for Java applications) Python 3.7, 3.8, or 3.9 (for Python applications)

To discover more about Tungsten OmniPage Capture SDK, please click here.

About Tungsten Automation

Tungsten Automation, formerly Kofax, is the global leader in intelligent automation solutions with a trusted legacy of nearly 40 years, with a team of 2,000+ employees in 40 countries, serving 25,000+ global customers. Our dedication to innovation and customer success has earned us industry recognition, including being named a leader in Intelligent Automation (IA), Intelligent Document Processing (IDP) and Process Orchestration by top analysts. We are trusted to help businesses gain unprecedented efficiencies and reduce costs through AI-powered workflow automations that propel their businesses into the future.

To learn more, visit **www.TungstenAutomation.com**

www.TungstenAutomation.com

TUNGSTEN AUTOMATION