

PART NUMBER, BUILD NUMBER

Part#	1041/74
Build#	11.1.14.63

NEW FEATURES AND IMPROVEMENTS

WINDOWS 10 SUPPORT

This release officially supports Microsoft Windows 10.

LARGE DOCUMENT CONVERSION TO SEARCHABLE PDFS

Use new API for converting PDF documents over 500 pages in order to reduce processing time up to 4 times. Please check “Exporting Large Documents” section in the Help file.

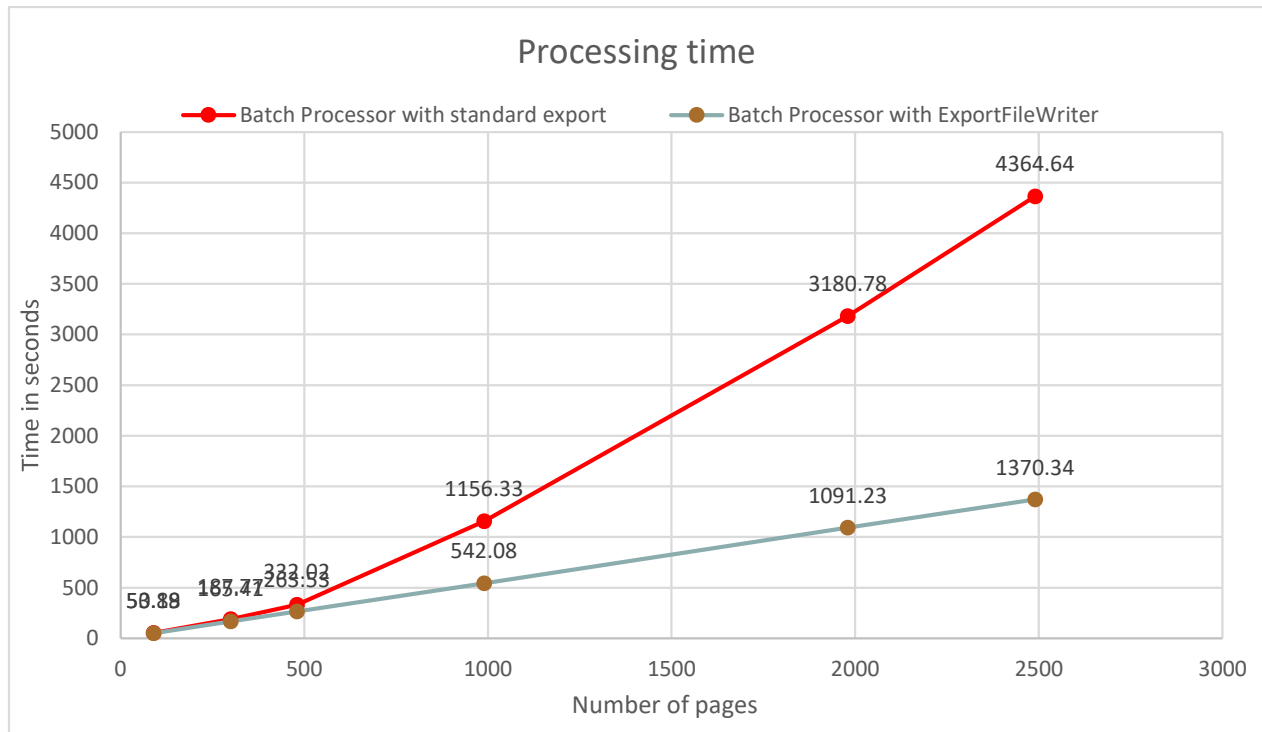
The main advantages of the new scenario:

- Dramatic speed increase when exporting large documents
- Less RAM resources required
- Convenient error handling without losing all the exported data. In case of simultaneous export a single error may cause failure of the whole export stage. If you are processing a large number of pages, the restarting of export can take significant time. Using the new functionality you do not have to stop processing. The processing errors can be caught and handled inside subsets of pages and export will be finished faster even if some errors occur during the processing.

Recommendations

- Using the new export mode is reasonable for documents of 50 pages or more. For 500 pages and more using ExportFileWriter is highly recommended.
- For the best possible speed choose Batch Processor (see “Processing using Batch Processor” in the help file).
- Export a fixed number of pages at a time. You may need to do experiments to pick the best number of pages for your documents. During internal ABBYY tests the portion size of 30 pages was found best for generic documents.

The results of speed testing are presented on the diagram.



Testing machine: Intel® Core™ i5-3450 (3.10 GHz, 4 physical cores), 8 GB of RAM, the number of simultaneously run processes is 4. Documents were saved by 30 pages at a time.

JAPANESE OCR IMPROVEMENTS

FineReader Engine 11 R6 contains the following improvements in Japanese OCR, resulting in 5% reduction of recognition errors:

- JapaneseModern is new OCR language with excluded outdated Kanji characters and it improves the accuracy of OCR for modern documents.
- Several frequent modern Japanese characters were trained and added into the recognition set.
- There are new models are added, which help with similar characters and mixed language recognition:
 - o Date: 2014 年 10 月 19 日 (日) .
 - o Address: 東京都港区赤坂 2 丁目 1 7 - 2 2 .
- Updated dictionary

ABILITY TO PROCESS SINGLE-PAGE DOCUMENTS FROM MEMORY IN BATCH MODE

IBatchProcessor interface allowed to process single-page documents in batch mode, i.e. using multi-processing on all available CPU cores. In previous releases of FineReader Engine it worked with image files only.

Starting from this release it is possible to process single-page documents stored in memory in batch mode using new API:

- `IImageSourceEx`. Comparing to `IImageSource` interface the new one allows implementing callbacks for image files and images stored in memory.
- `IBatchProcessor::StartEx()`. Comparing to `Start()` method it accepts new `IImageSourceEx` interface instead of outdated `IImageSource` one.

ARABIC FATHATAN (َ) DIACRITICS SUPPORT

FineReader Engine 11 R6 supports [Arabic Fathatan diacritics](#), which increases Arabic OCR accuracy.

THAI OCR ACCURACY IMPROVEMENT

Improved support of the most popular Thai fonts (Tahoma, TH SarabunPSK, TH Niramit AS, TH Kodchasa, TH Baijam, TH Fah Kwang) resulted in 11%-13% better Thai recognition accuracy overall, according to internal tests.

EXTENDED INFORMATION ON CHARACTER POSITION IN A WORD

New `IsWordLeftmost` and `IsWordFirst` allows capturing the accurate location of word's characters.

Previously the `IsWordStart` property supposed to define if a character starts a word. For right-to-left writing, this parameter was ambiguous: logically a word starts from the right end whereas geometrically it starts from the left. To solve the ambiguity we added these two new properties to API and XML export:

- `IsWordLeftmost`. Specifies whether the character is the leftmost character in a word.
- `IsWordFirst`. Specifies whether the character is the first character in a word.

The old property `IsWordStart` is declared deprecated and will be removed from API and XSD in FRE12.

SMALLER PDF WITH PAGES OF MIXED COLORITY

In previous FRE 11 releases, we added PDF exporting API for setting up an individual compression type for each page. Now we introduce additional API for setting up individual resolution:

- `BwPictureResolutionParams`
- `ColorPictureResolutionParams`
- `GrayPictureResolutionParams`

Note: New API does not work in MRC mode. MRC and individual resolution/compression setting are two different approaches to the same goal – visual quality and small size.

'BARCODERECOGNITION_ACCURACY' BECAME MORE ACCURATE

FineReader Engine 11 R6 includes the following improvements in barcode detection:

- QRCode: 5.6% less errors
- EAN 13 with supplemental: 20.5% less errors
- EAN 13: 15.38% less errors
- UPC-E with supplemental: 10.5% less errors
- EAN 8 with supplemental: 5.9% less errors

The accuracy improvement brings also a slowdown, so it was applied to accuracy profile only. The slowdown varies from 5% to 70% depending on image complexity, detection type, and test-bed content (length and specifics).

CALLBACK FROM PARALLEL PROCESSING

While using a parallel processing, one of Engine processors may 'hang up' with an image taking more time to process than it is allowed by default. There are several reasons for that:

- Limited hardware resources
- Specific image (large, complex layout, etc.)
- Logical error in FineReader Engine

Previously that led to process breaking with an error, that is not always a desirable behavior. In order to unlock new scenarios FRE 11 R6 presents new API:

- `IParallelProcessingCallback` to implement desired behavior on a client side
- `IEngine::SetParallelProcessingCallback()` to activate the handling.

WARNINGS ARE CONNECTED TO PARTICULAR PAGES

FineReader Engine 11 R6 introduces an updated `IFRDocument::IFRDocumentEventsEx` interface with extended `OnWarning()` callback returning `PageIndex` property.

UPDATED ADMIN'S GUIDE

We updated the 'Admin's Guide' describing various installation scenarios of the product. It covers installation and de-installation processes for the following main distribution scenarios:

- Standalone mode using `setup.exe` with silent command line SN activation
- Standalone mode w/o `setup.exe` with SN activation using License Manager GUI
- Network mode using `setup.exe`, Licensing Service set up with command line SN activation

- Network mode w/o setup.exe, Licensing Service set up SN activation using GUI

'CLASSIFICATION' SAMPLE BECAME FASTER

We speeded up the 'Classification' sample by adding multi-processing to the most calculation intensive step – OCR.

CODEMETER DRIVERS UPDATE

The release includes version 5.21 (5.21.1478.500) of runtime and driver for CodeMeter protection.

'IS REDUNDANCY' OPTION IS LISTED IN LM

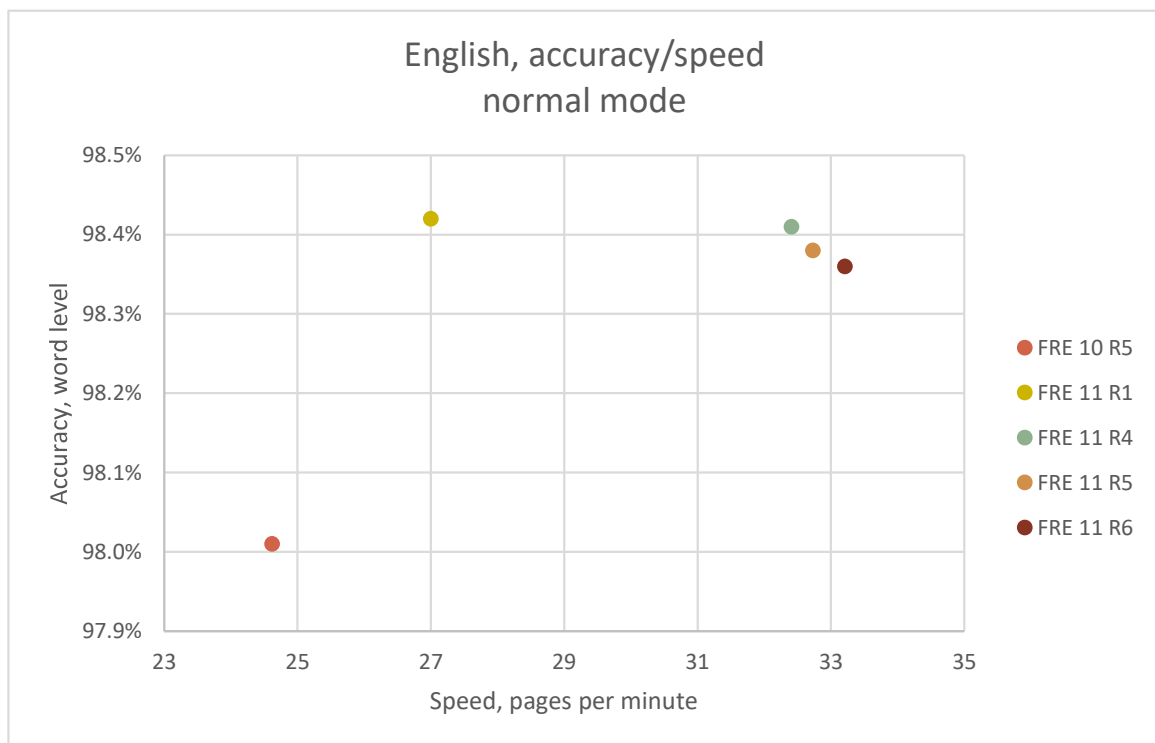
License Manager is showing now 'Is Redundancy' licensing option introduced in FineReader Engine 11 R5.

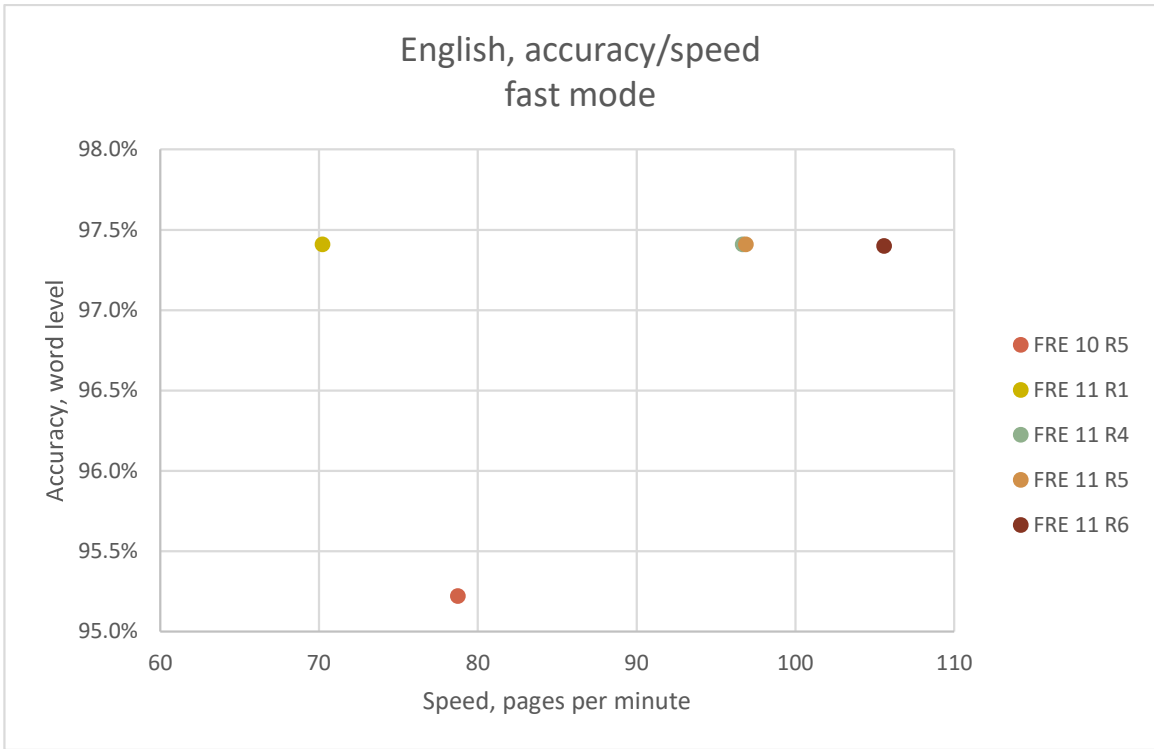
PERFORMANCE RESULTS

This section compares the performance of FRE 11 R6 to previous releases and the latest version of FRE 10.5. The test machine's processor was an Intel® Core™ i5-4690 CPU (3.50GHz, 4 physical cores) with 8GB of RAM. Test batches for different languages may include significantly different documents, making it impossible to accurately compare documents featuring different languages.

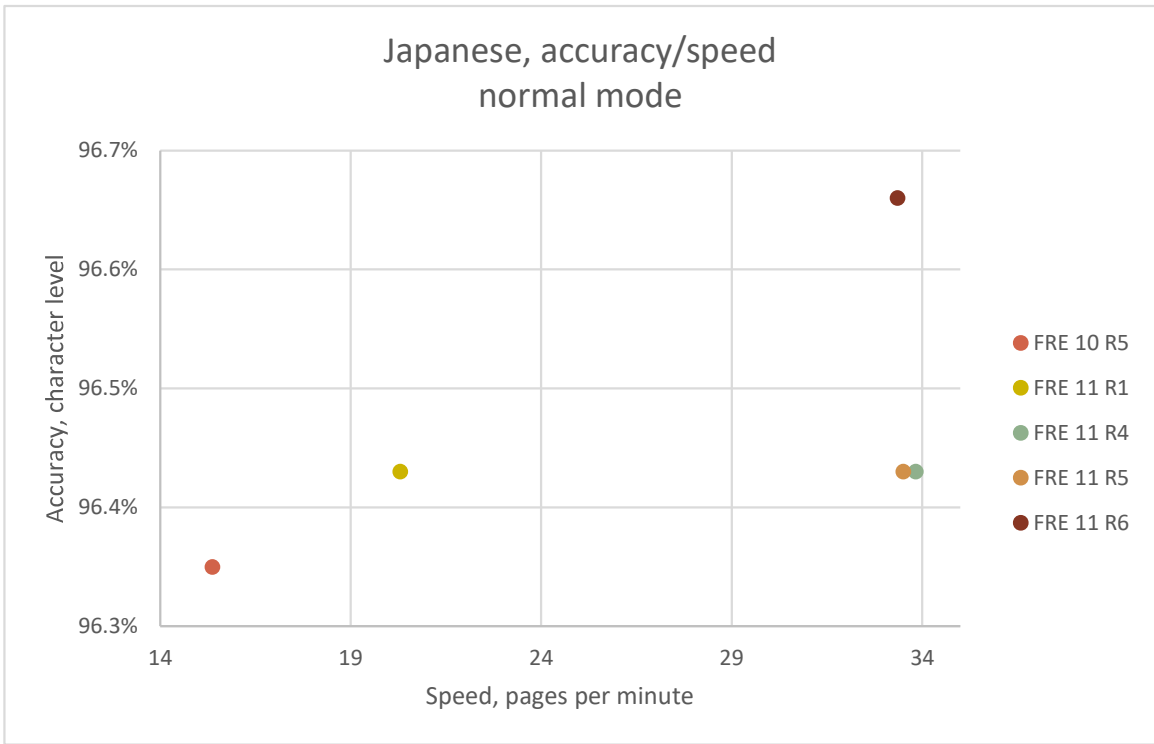
Here are the results of our comparisons:

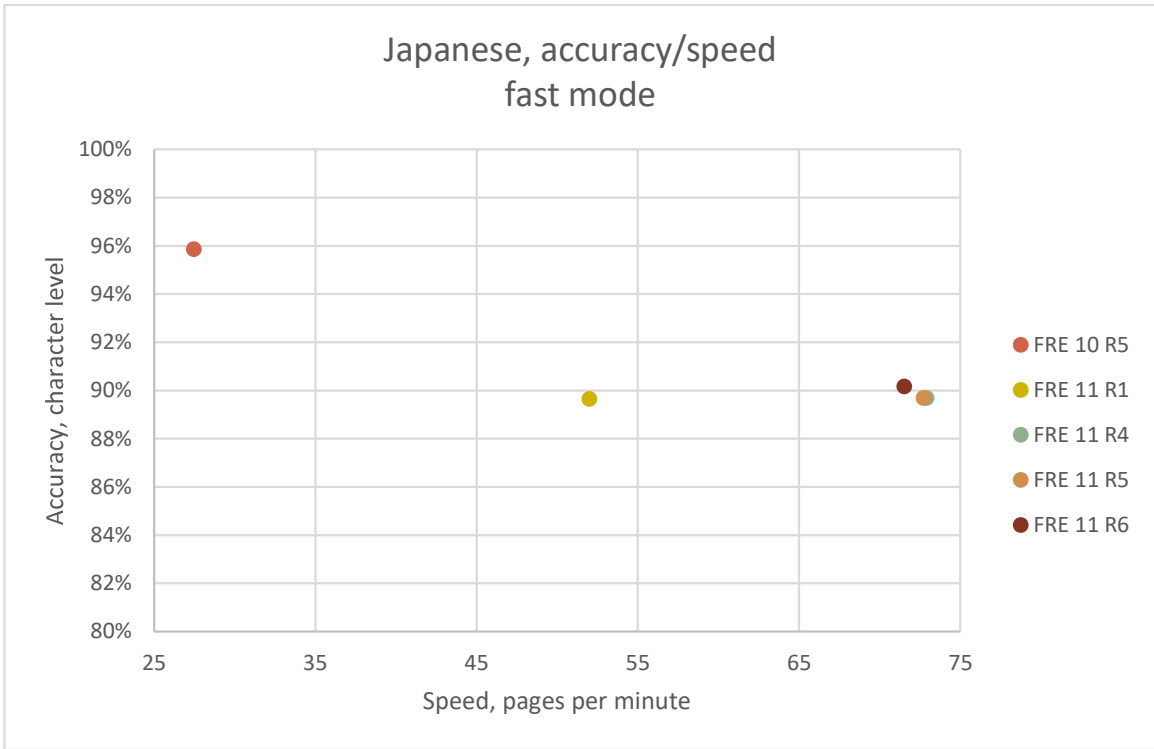
ENGLISH



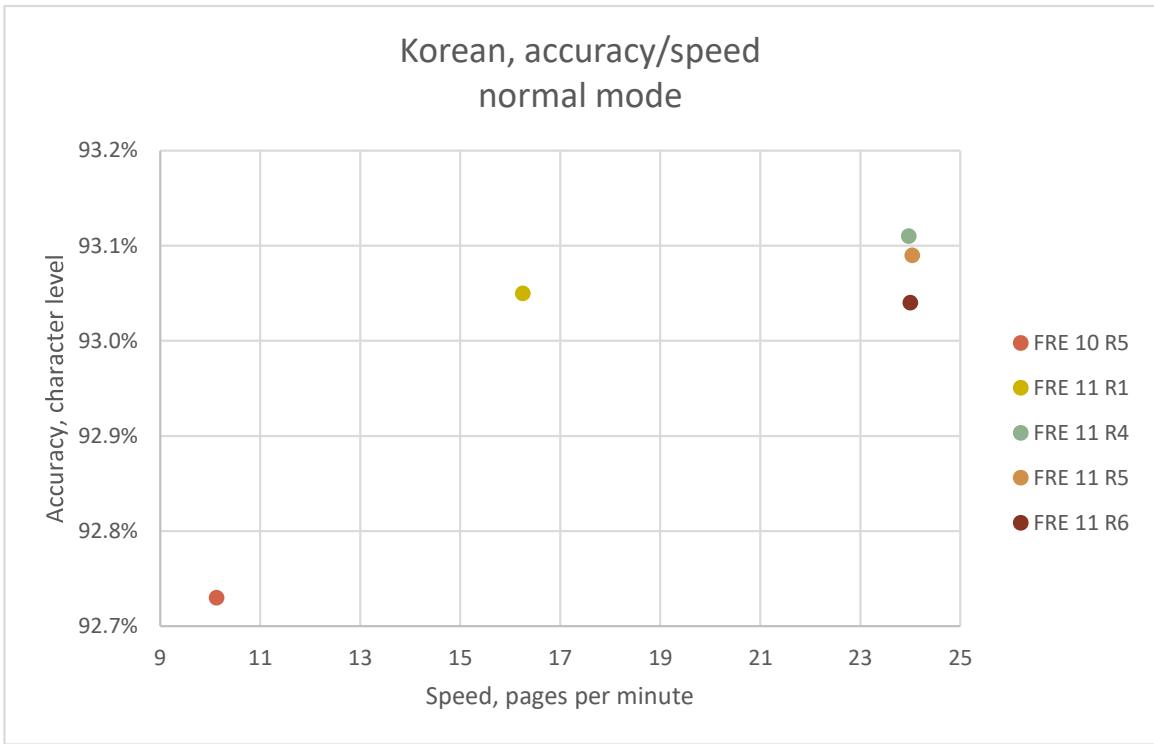


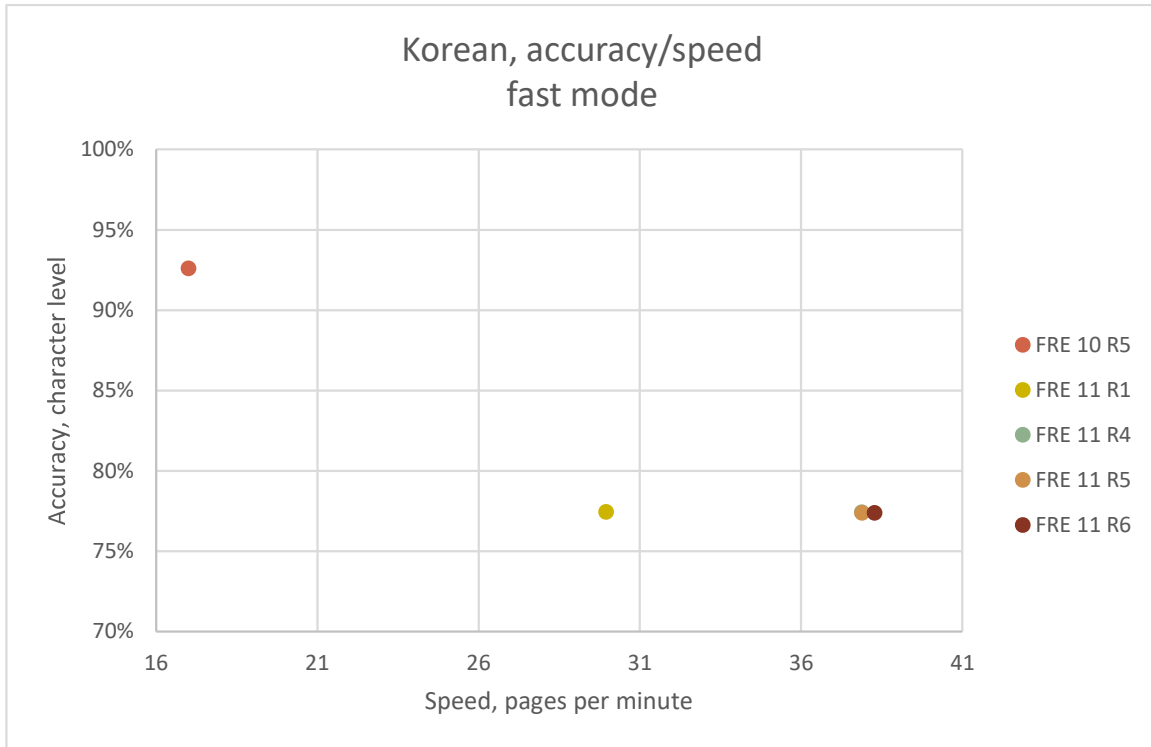
JAPANESE



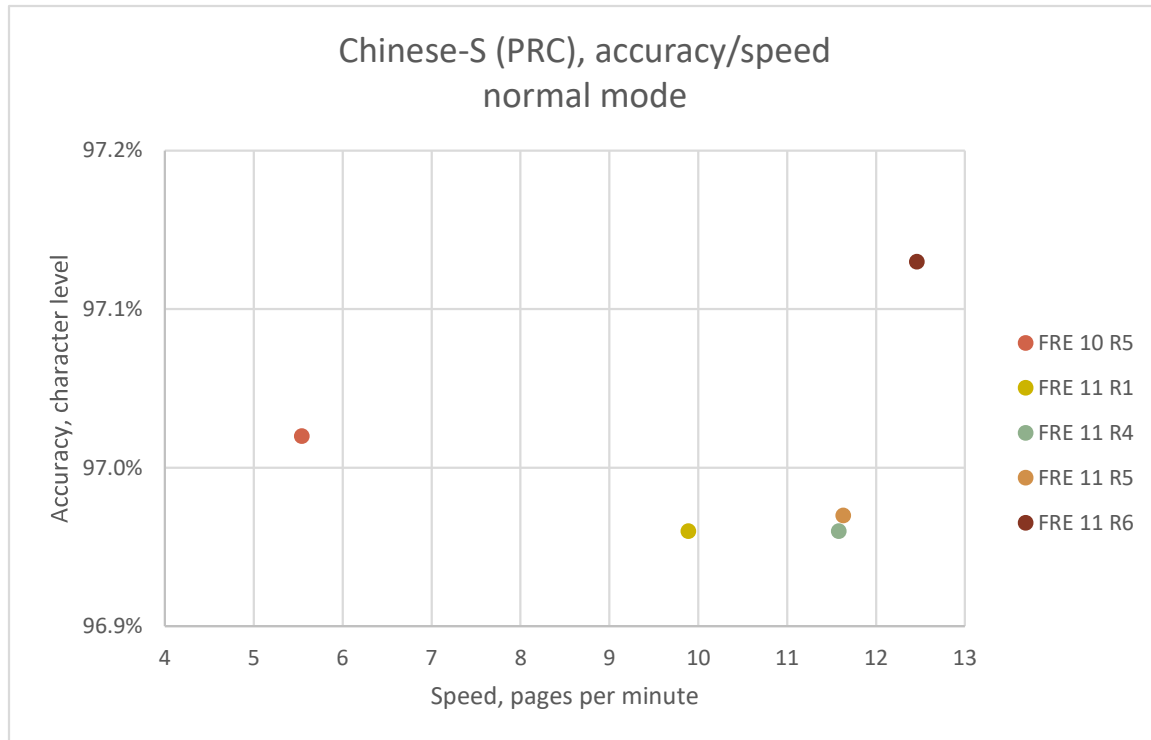


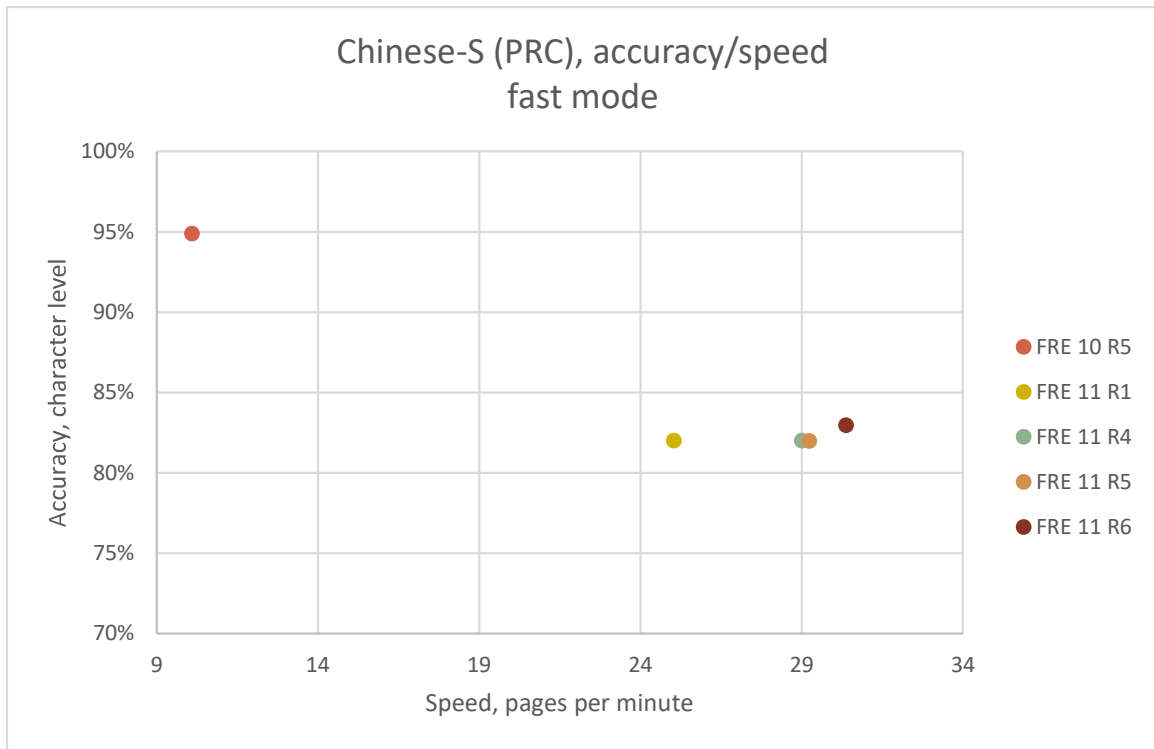
KOREAN



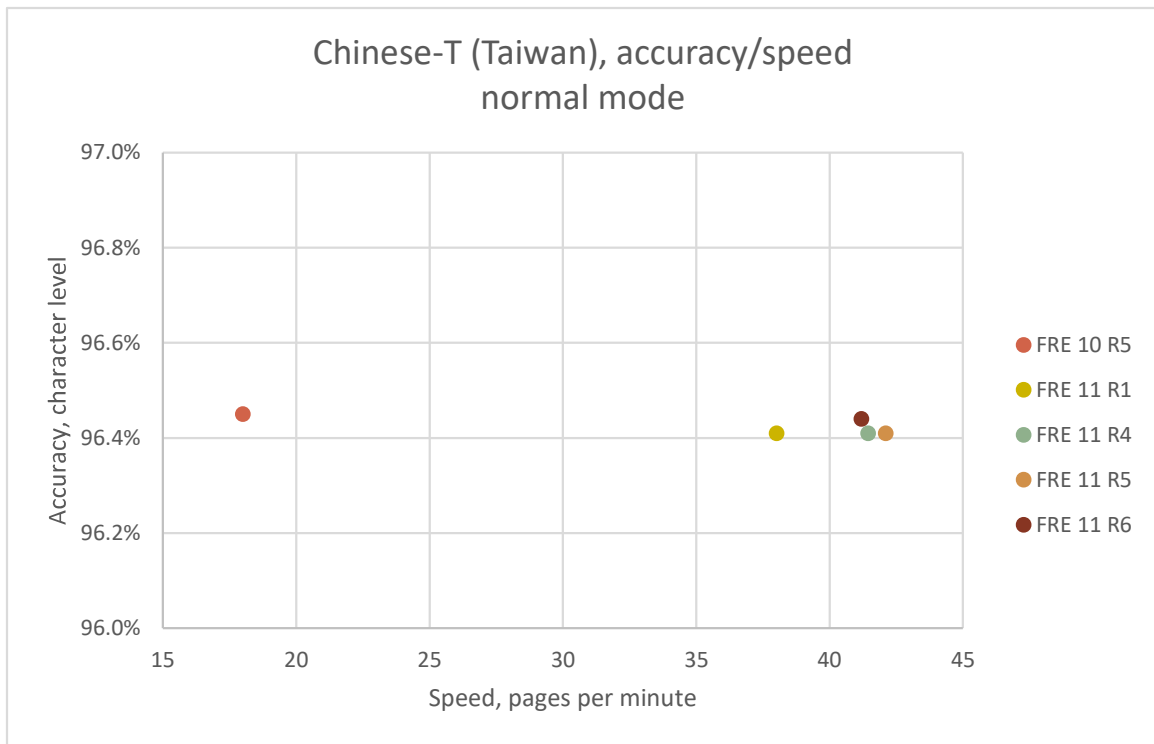


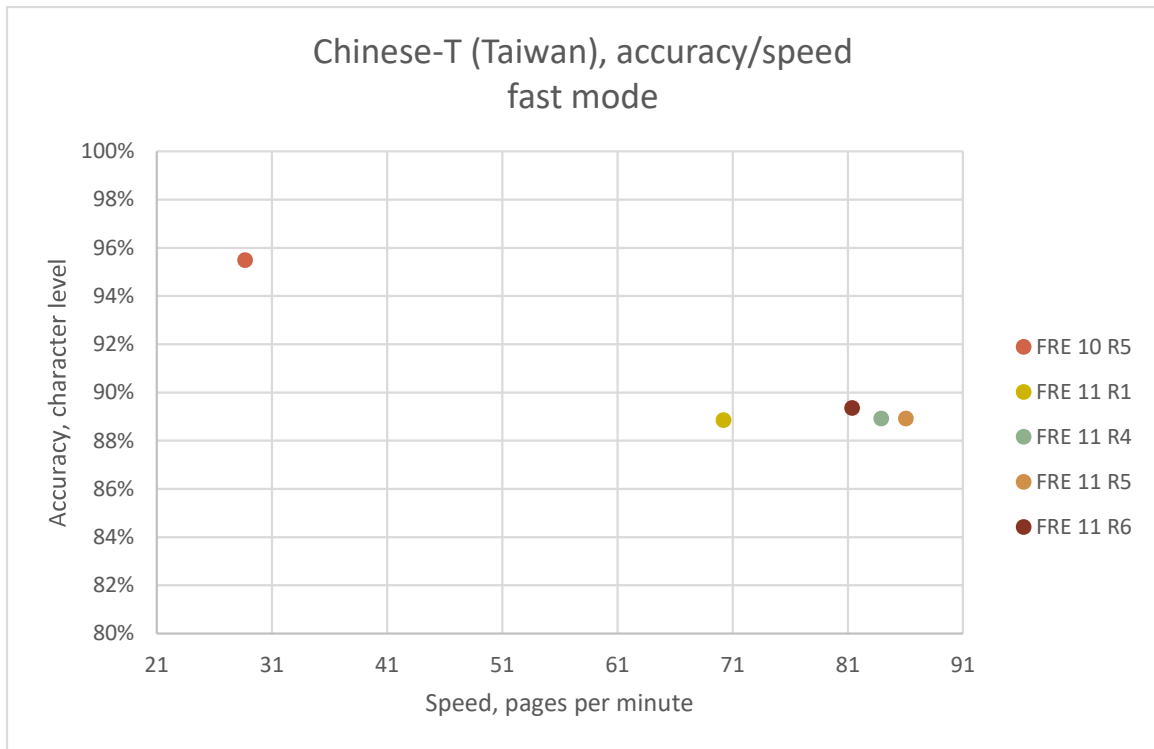
CHINESE PRC



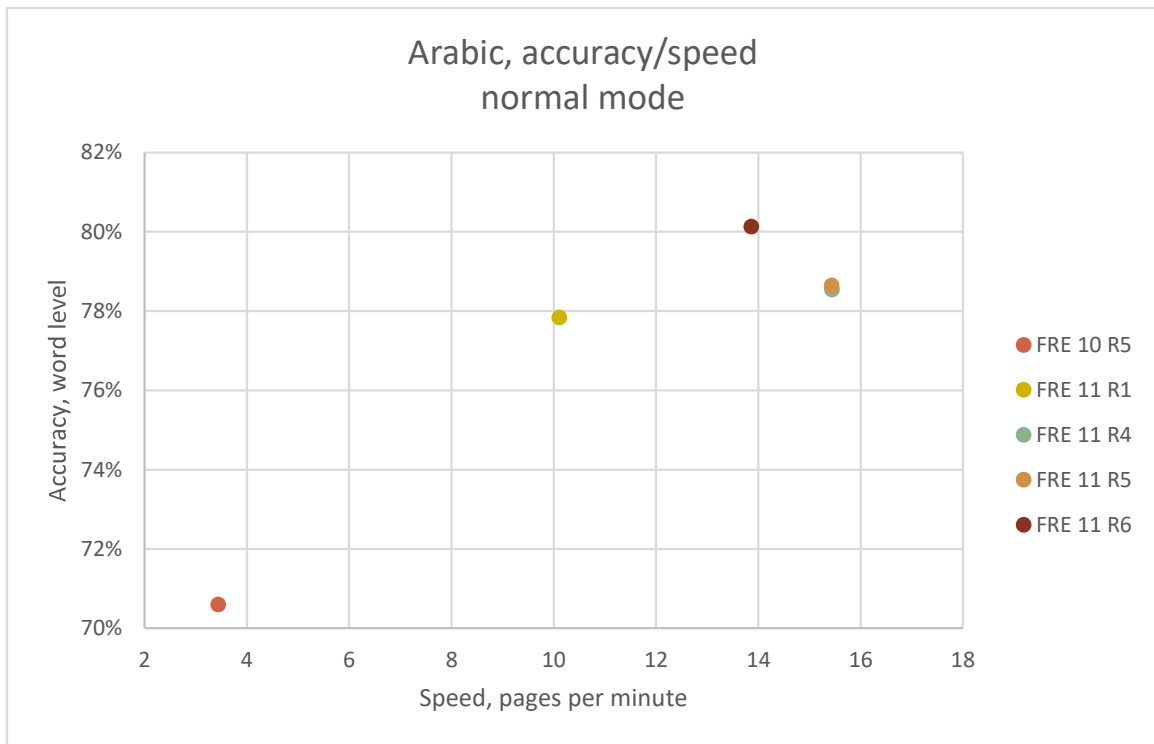


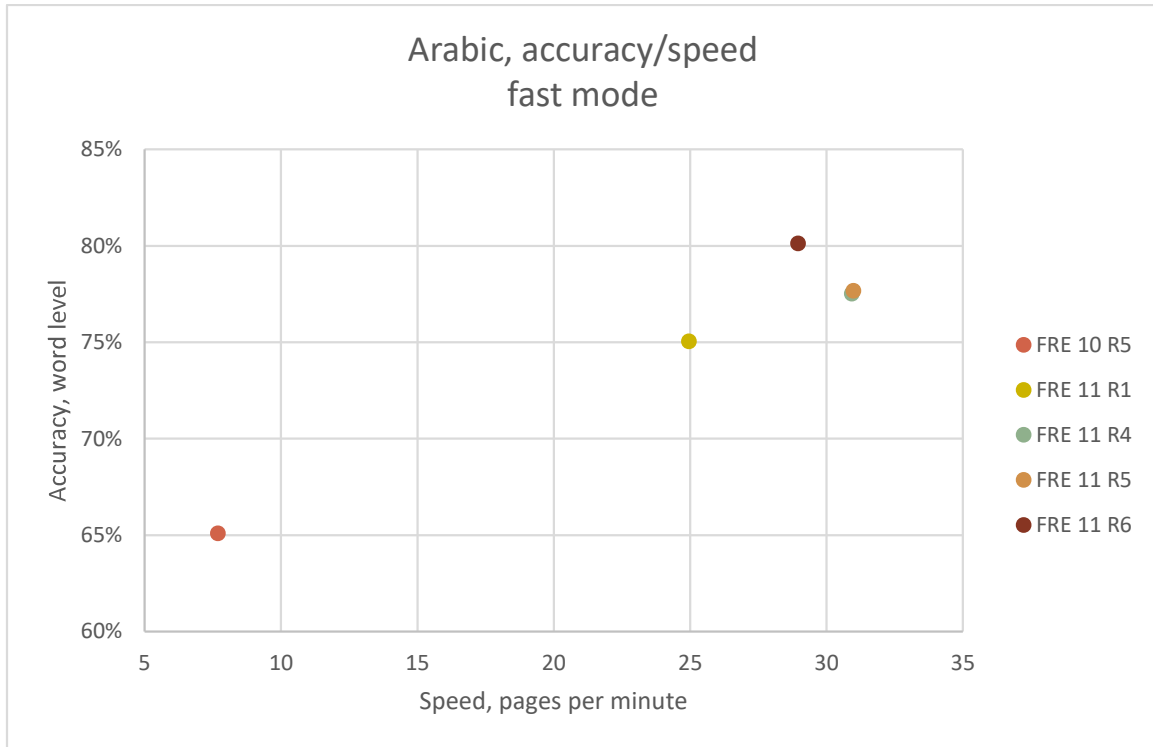
CHINESE TAIWAN





ARABIC





FIXED BUGS

This section contains a list of bugs that have been fixed.

This four-point scale will help you evaluate the severity of each issue, enabling you to make informed decisions on how important updates are for your system.

Critical	A bug that causes crashes or hangings of software. Critical bugs can include access violations, internal program errors, stack overflow, out of memory or other exceptions that can lead to program failure.
Major	A bug that does not cause program failure but affects major functionality of a feature or impairs the system's performance. Major bugs can include disparity between feature functionality and internal specifications, memory leaks or data corruption.
Minor	A bug that leads to feature malfunctions or affects minor functionality of the software. Minor bugs can include recognition errors, missing or lost objects, incorrect color detection, incorrect document analysis, license counter errors, etc.
Trivial	A cosmetic issue that does not affect the functionality of the product but which can cause inconveniences. Trivial bugs can include Help file errors, log errors, incomplete information in error messages, etc.

The following table contains bugs fixed in this release, sorted in descending order of severity. If the bugs have workarounds, root causes or side effects, they will be mentioned in the Description section.

Severity	Description	Technology Subsystem
Critical	Unhandled exception on a customer's image with PageAnalysisParams::DetectBarcodes=true property set.	Document Analysis
Major	One of simultaneously running processes executing conversion to PDF in multi-processing mode crashes soon after start.	Export
Major	In 'PDF with MRC' scenario overridden image resolution is not applied. Root cause: FRE initializes not all internal parameters for MRC passed to Export subsystem.	API
Minor	Conversion to RTF 'Editable Copy'. The output contains more pages than original document even with the PageAnalysisParams::PaperSizeDetectionMode = PSDM_CloseToImageSize parameter set. Root cause: DA makes a mistake in text block detection, Synthesis is not able to correct the DA mistake.	Synthesis
Minor	Incorrect table cell text indent on an image.	Document Analysis
Minor	Text frame has incorrect parameters on a customer's image.	Synthesis
Minor	FRE Java interop does not install itself properly being renamed.	Install
Minor	' ǀ ' is recognized as ' ǁ ' on a customer's image. Workaround: add undesired character to the list of prohibited characters.	Recognizer
Minor	Output contains no tabs between text line parts (words) as they are on a customer's image.	Synthesis
Minor	In case of Hebrew or any other right-to-left writing languages the property IsWordStart = true for the leftmost character instead of the rightmost. Root cause: the property was designed for left-to-right writing and does not work properly for reversed writing. Workaround: check for a word language and interpret the property value correspondingly.	Synthesis
Minor	Incorrectly recognized words in English in a document with a text mixture of English and Japanese words.	Recognizer
Minor	Incorrectly recognized words in English in a document with a text mixture of English and Chinese words.	Recognizer
Minor	Incorrect business cards splitting on customer images.	Document Analysis
Minor	In 'source PDF content reuse' scenario 'Franklin Gothic Heavy' fonts is taken for output instead of 'Arial' on a file provided by a customer.	Synthesis
Trivial	Incomplete article 'Using Text Type Autodetection'.	Help
Trivial	The 'Business Card Recognition' sample uses incorrect image preprocessing sequence that may lead to skew correction malfunctioning. Workaround: do frDocument->Preprocess with pageSplittingParams->SplitType = PST_BusinessCardSplit instead of frDocument->SplitPages.	Samples
Trivial	Java sample code for Engine Pool creation uses string variables where enums are more appropriate.	Samples
Trivial	The product documentation attributes NrmlPart.ssc and NrmlPart.slp files as required for CJK OCR only, whereas they are required also for Arabic OCR.	Help

KNOWN ISSUES AND WORKAROUNDS

SOME API IS NOT IMPLEMENTED

The following API is not implemented in FRE 10 and FRE 11:

- `IFootnoteSeries::HasSeparator`. Always returns “true”.
- `ITextPicture::ColumnNumber`. Always returns “0”.
- `ICharParams::IsWordStart`. Always returns “false”. It is true only for character parameters got through `IWordRecognitionVariants` interface.
- `IIncut::TextWrapping`. Always returns “TW_Undefined”.
- `IRunningTitlesSeriesText::HasSeparator`. Always returns “false”.

The implementation is not planned.

FARSI LANGUAGE IS BASED ON ARABIC LANGUAGE

Farsi OCR output indicates Language ID as Arabic (`CharParams::LanguageId = LI_ArabicSaudiArabia`). This is because Farsi technology preview is based on Arabic OCR language.

The fix will be available in FRE 12.

IPDFMRCPARAMS::MONOCHROMETEXT DOESN'T WORK CORRECTLY

Different algorithms of compressions during export to PDF are used with `IPDFMRCPParams::MonochromeText` set to TRUE.

'USER PATTERN TRAINING UTILITY' SAMPLE CRASHES FOR 'FIELDLEVELRECOGNITION' PROFILE

Run the sample, select 'FieldLevelRecognition' processing profile, recognize an image. The Text Editor will show “The page has been recognized, but no page view has been created. Click Read to complete the process.” message. A subsequent recognition command will result into IPE: “...\\Interfaces\\Implementations\\inc\\NotifyImpl.h, 62”.

The fix will be available in the next release.

INFINITE LOOP IN JAVA CRASHES

In 32-bit JDK 8 Update 45, JDK 8 Update 51 applications infinite loop for 'Hello' sample processing crashes after 5-9 cycles.

The fix will be available in the FRE 11 R6 patch.

PASSWORD PROTECTED PDF FILES COULDN'T BE OPENED IN PARALLEL PROCESSING MODE

PDF file may be password protected and to open a file for processing a corresponding password must be provided. API has two possibilities:

- `AddImageFileWithPassword`, accepts a password as a parameter.
- `AddImageFileWithPasswordCallback`, calls for a password as soon as it is required.

We found that in parallel processing mode `AddImageFileWithPasswordCallback` scenario does not work properly.

The fix will be available in the FRE 11 R6 patch.

JAPANESEMODERN OCR LANGUAGE IS NOT AVAILABLE FOR BCR

'JapaneseModern' OCR language is not available for BCR.

The fix will be available in the FRE 11 R6 patch.

LANGUAGE AUTO DETECTION USES CJK RESOURCES THOUGH NONE IS SELECTED FOR OCR

'Cjk.*' resource files are used while `RecognizerParams::LanguageDetectionMode = TSPV_Yes` even though recognition languages set does not include any of CJK languages.

The fix will be available in the FRE 11 R7.

R6 GM PATCH –

PART NUMBER, BUILD NUMBER

Part#	1041/76
Build#	11.1.14.64

NEW FEATURES AND IMPROVEMENTS

There are no new features.

PERFORMANCE RESULTS

The same as in FRE 11 R6.

FIXED BUGS

This section contains a list of bugs that have been fixed.

This four-point scale will help you evaluate the severity of each issue, enabling you to make informed decisions on how important updates are for your system.

Critical	A bug that causes crashes or hangings of software. Critical bugs can include access violations, internal program errors, stack overflow, out of memory or other exceptions that can lead to program failure.
Major	A bug that does not cause program failure but affects major functionality of a feature or impairs the system's performance. Major bugs can include disparity between feature functionality and internal specifications, memory leaks or data corruption.
Minor	A bug that leads to feature malfunctions or affects minor functionality of the software. Minor bugs can include recognition errors, missing or lost objects, incorrect color detection, incorrect document analysis, license counter errors, etc.
Trivial	A cosmetic issue that does not affect the functionality of the product but which can cause inconveniences. Trivial bugs can include Help file errors, log errors, incomplete information in error messages, etc.

The following table contains bugs fixed in this release, sorted in descending order of severity. If the bugs have workarounds, root causes or side effects, they will be mentioned in the Description section.

Severity	Description	Technology Subsystem	HD #	Office
Critical	In 32-bit JDK 8 Update 45, JDK 8 Update 51 applications infinite loop for 'Hello' sample processing crashes after 5-9 cycles.	API	510416	US

Major	In parallel processing mode AddImageFileWithPasswordCallback (password protected PDF files opening) scenario does not work properly.	API	n/a	n/a
Major	'JapaneseModern' OCR language is not available for BCR.	API	n/a	n/a
Minor	IEngine::AvailablePredefinedLanguages returns no JapaneseModern OCR language.	API	n/a	n/a

KNOWN ISSUES AND WORKAROUNDS

SOME API IS NOT IMPLEMENTED

The following API is not implemented in FRE 10 and FRE 11:

- IFootnoteSeries::HasSeparator. Always returns "true".
- ITextPicture::ColumnNumber. Always returns "0".
- ICharParams::IsWordStart. Always returns "false". It is true only for character parameters got through IWordRecognitionVariants interface.
- IIncut::TextWrapping. Always returns "TW_Undefined".
- IRunningTitlesSeriesText::HasSeparator. Always returns "false".

The implementation is not planned.

FARSI LANGUAGE IS BASED ON ARABIC LANGUAGE

Farsi OCR output indicates Language ID as Arabic (CharParams::LanguageId = LI_ArabicSaudiArabia). This is because Farsi technology preview is based on Arabic OCR language.

The fix will be available in FRE 12.

IPDFMRCPARAMS::MONOCHROMETEXT DOESN'T WORK CORRECTLY

Different algorithms of compressions during export to PDF are used with IPDFMRCPParams::MonochromeText set to TRUE.

'USER PATTERN TRAINING UTILITY' SAMPLE CRASHES FOR 'FIELDLEVELRECOGNITION' PROFILE

Run the sample, select 'FieldLevelRecognition' processing profile, recognize an image. The Text Editor will show "The page has been recognized, but no page view has been created. Click Read to complete the process." message. A subsequent recognition command will result into IPE: "...\\Interfaces\\Implementations\\inc\\NotifyImpl.h, 62".

The fix will be available in the next release.

LANGUAGE AUTO DETECTION USES CJK RESOURCES THOUGH NONE IS SELECTED FOR OCR

'Cjk.*' resource files are used while RecognizerParams::LanguageDetectionMode = TSPV_Yes even though recognition languages set does not include any of CJK languages.

The fix will be available in the FRE 11 R7.

PDF/A VALIDATION REPORT

The following issues are detected by Adobe Acrobat 11.0.9 (Preflight 11.0.9) for PDF/A files produced by this release:

1. In SourceContentReuseMode = CRM_ContentOnly the Engine sometimes (5% of the test base in Korean language) generates PDF/A-2a and PDF/A-3a files which Adobe Acrobat Preflight validates with the following message "Text is mapped to Unicode Private Use Area but no ActualText entry is present".