

R7 GM –

PART NUMBER, BUILD NUMBER

Part#	1041/89
Build#	11.1.16.94

NEW FEATURES AND IMPROVEMENTS

CONVERSION DOCUMENTS TO SEARCHABLE PDFS AT FULL THROTTLE REQUIRES NO PAGE NUMBERS SETUP ANY MORE

It is possible to convert documents containing undefined number of pages to searchable PDF using ExportFileWriter. The PagesCount parameter must be set to -1 in this case (this parameter is deprecated and will be removed in future versions).

We have introduced this feature in FRE 11 in the R6 release. And from now our customers have no necessity to know an amount of pages in the document at the moment of creating a session of recognition.

It can be useful for effective work with scanners when one has to process a lot of pages and doesn't know the amount of them in the document until the end of scanning, but needs the processing to start before the scanning will be finished.

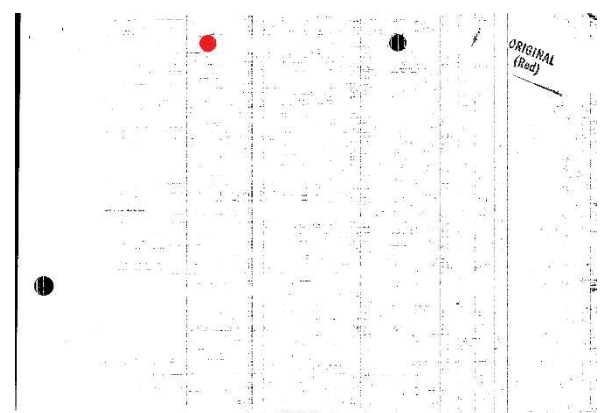
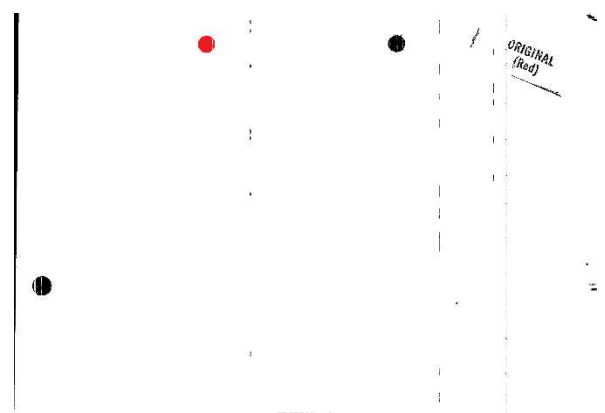
SIMULTANEOUS NETWORK AND STANDALONE USAGE OF ONE FRE INSTALLATION

Now it is possible to define network and standalone licenses in one LicensingSettings.xml file.

This possibility was important for one of our customers when he decided to put his application into a network folder and let local (to a host of the folder) and remote (other PC in the same network) end-users run the same copy of the application. It is reasonable that 'local' end-users should use a standalone FRE license, whereas 'remote' end-users should use a network FRE license. To do so we specified two licensing configurations (standalone and network) in one LicensingSettings.xml file.

GARBAGE REMOVAL FROM COLOR IMAGES

New method ImageDocument::RemoveGarbageEx works with both color and white-and-black images. From now it is possible to remove garbage from color images using this method.

Input image	Output image
<p>ABBYY FineReader Engine 10 Introduction</p> <p>Document Analysis</p> <p>Basic document analysis features</p> <p>Document Analysis is a set of functions for automatic detection of the following objects on a page:</p> <ul style="list-style-type: none"> • Text blocks • Pictures • Tables and table cells • Barcodes • Separators <p>Additionally, document analysis provides some special features to prepare image for OCR:</p> <ul style="list-style-type: none"> • process detection of page orientation – 90, 180, and 270 degrees • split double pages • process vertical text detection in table cells 	<p>ABBYY FineReader Engine 10 Introduction</p> <p>Document Analysis</p> <p>Basic document analysis features</p> <p>Document Analysis is a set of functions for automatic detection of the following objects on a page:</p> <ul style="list-style-type: none"> • Text blocks • Pictures • Tables and table cells • Barcodes • Separators <p>Additionally, document analysis provides some special features to prepare image for OCR:</p> <ul style="list-style-type: none"> • process detection of page orientation – 90, 180, and 270 degrees • split double pages • process vertical text detection in table cells
<p>упоминалось. Лишь в разделе УП "Законодатель 1, прокурорском надзоре, арбитраже и охране на глава "Судоустройство. Суды Союза ССР". ассификаторе отраслей законодательства" в ра арственного строя" не включена судебная сис 3-м месте, к концу длинного списка отраслей включен п. "Суд. Юстиция. Прокуратура". ни в какой мере не соответствует действитель ту и роли суда в государственном устройстве титуции Российской Федерации устанавливает,</p>	<p>упоминалось. Лишь в разделе УП "Законодатель 1, прокурорском надзоре, арбитраже и охране на глава "Судоустройство. Суды Союза ССР". ассификаторе отраслей законодательства" в ра арственного строя" не включена судебная сис 3-м месте, к концу длинного списка отраслей включен п. "Суд. Юстиция. Прокуратура". ни в какой мере не соответствует действитель ту и роли суда в государственном устройстве титуции Российской Федерации устанавливает,</p>
	

Dealing with black-and-white images the method works the same way as ImageDocument::RemoveGarbage which will be removed in future versions of FRE.

POSSIBILITY TO INJECT TEXT LAYER TO SELECTED PAGES OF PDF DOCUMENTS

A new method `IEngine::InjectTextLayerEx` allows to process specific pages in the "image only" or "image on text" PDF documents. It creates a searchable PDF file which contains the same page images and the invisible text layer created from the recognized text of the document.

This new method works the same way as `IEngine::InjectTextLayer` and has new arguments:

- `PageIndices`. This parameter refers to the `IIntsCollection` object which specifies the indices of the document pages, to which the text will be injected. This parameter is optional and may be 0, in which case the text will be injected to all the pages of the document.
- `ProcessingEvents`. Refers to the interface of the user-implemented object that is used for reporting events to the listeners. This parameter may be 0, in which case no callback will be attached.

The `IEngine::InjectTextLayer` method is now deprecated and will be removed in future versions.

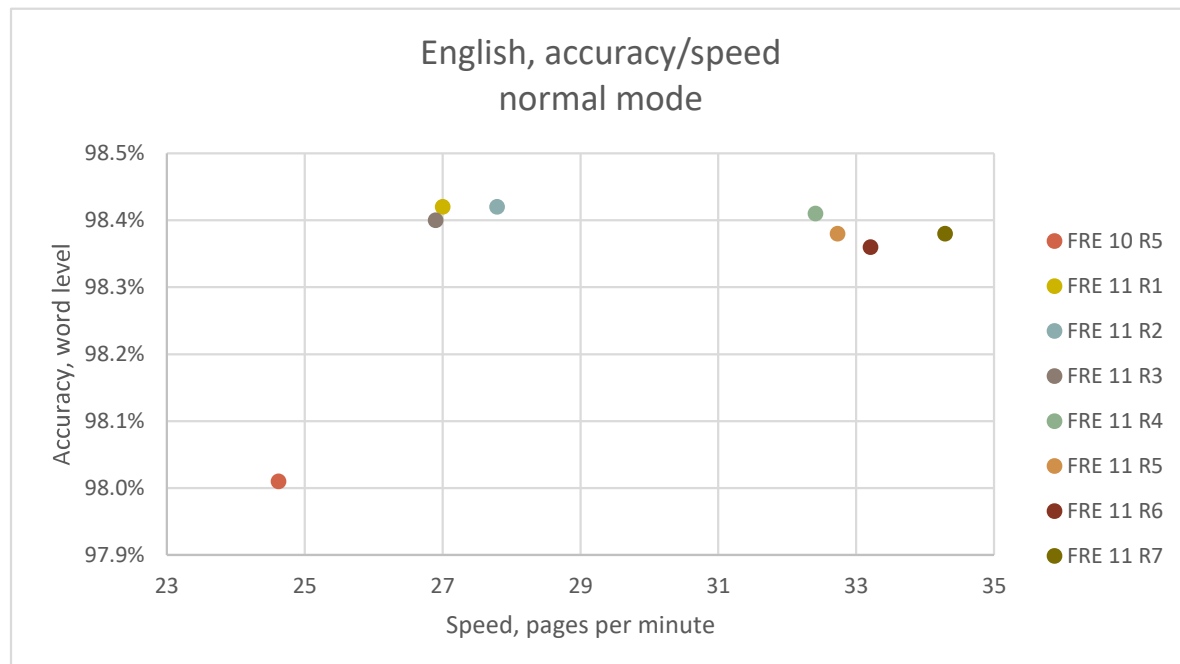
A NEW ARTICLE ABOUT WORKING WITH SCREENSHOTS IN DOCUMENTATION

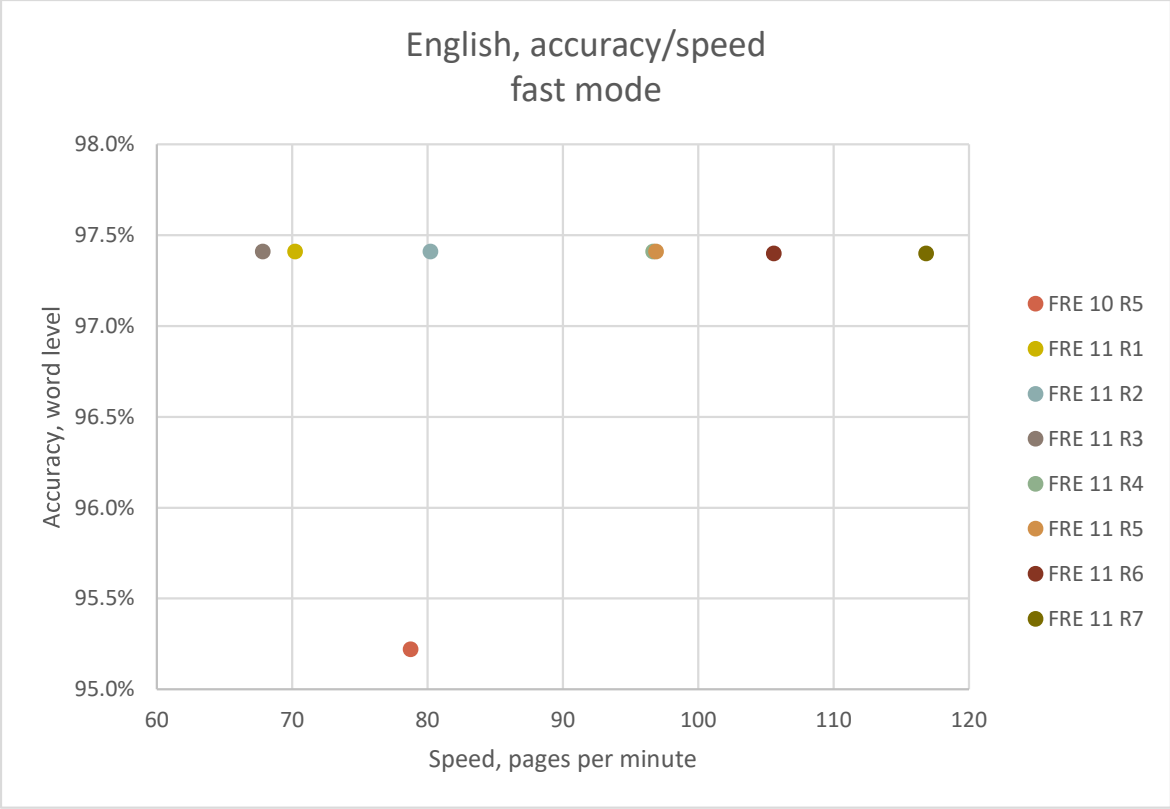
An article 'I am working with the screenshot image. Are there any special recommendations for screenshot processing?' about suitable settings of FRE for screenshot processing was added. It is in the FAQ section.

PERFORMANCE RESULTS

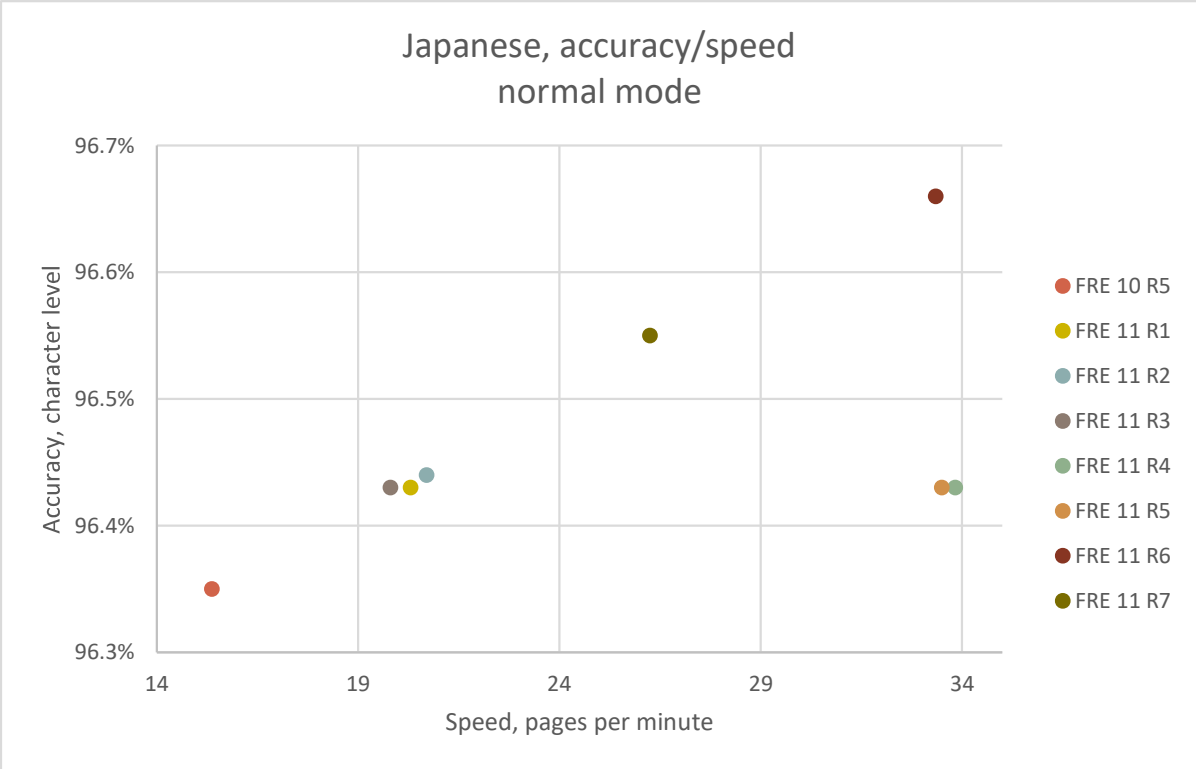
This section contains performance results of the release comparing to the previous releases. The processor of the testing machine is Intel® Core™ i5-4690 CPU (3.50GHz, 4 physical cores) with 8GB of RAM.

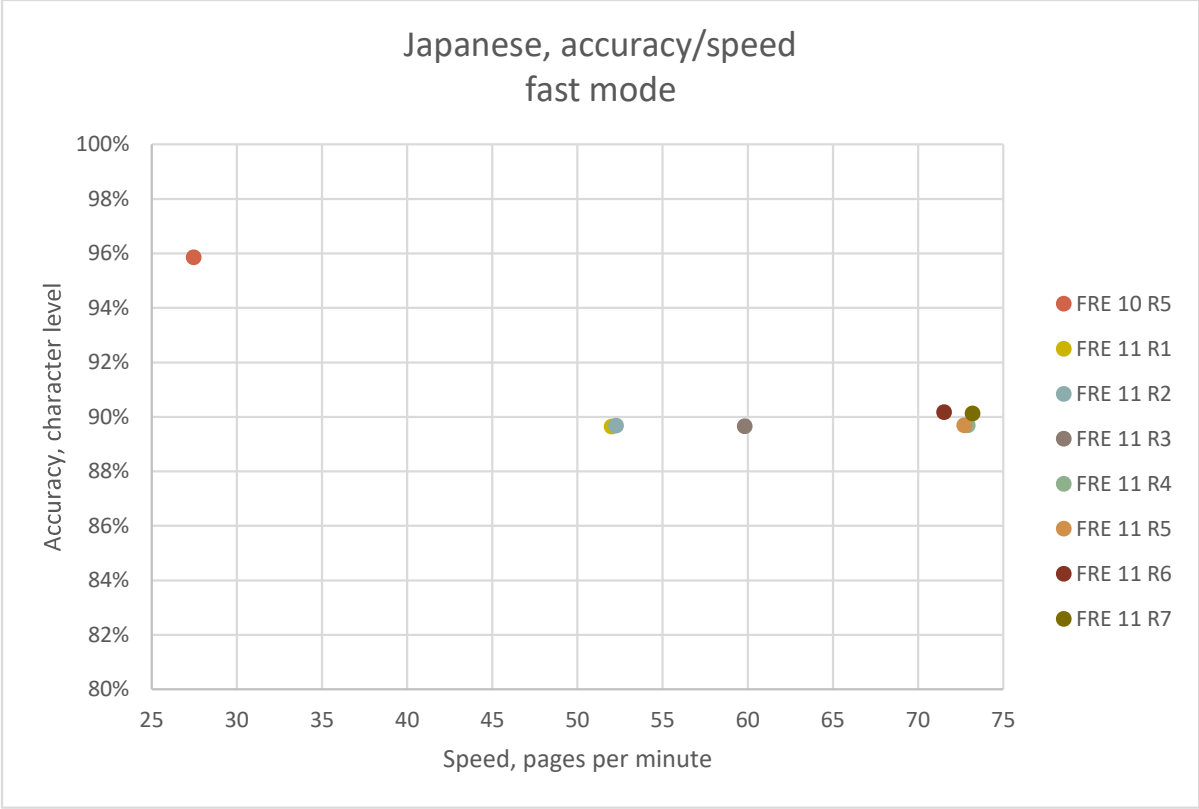
ENGLISH



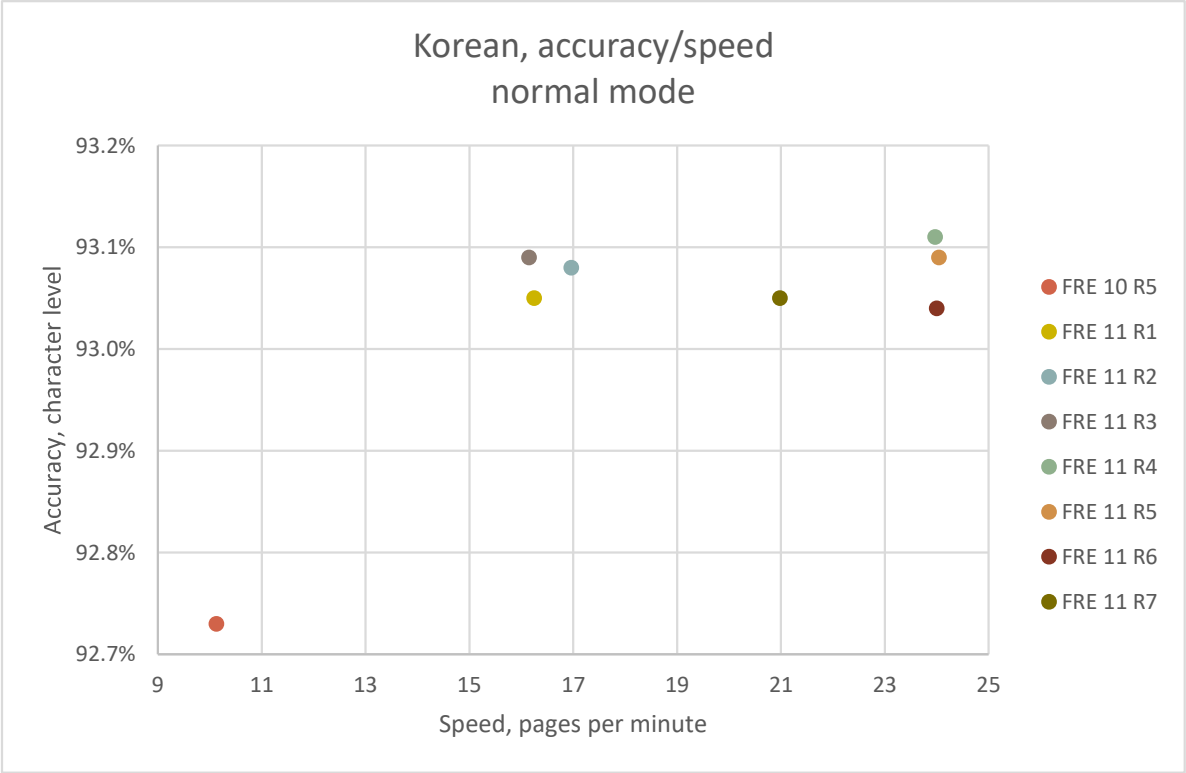


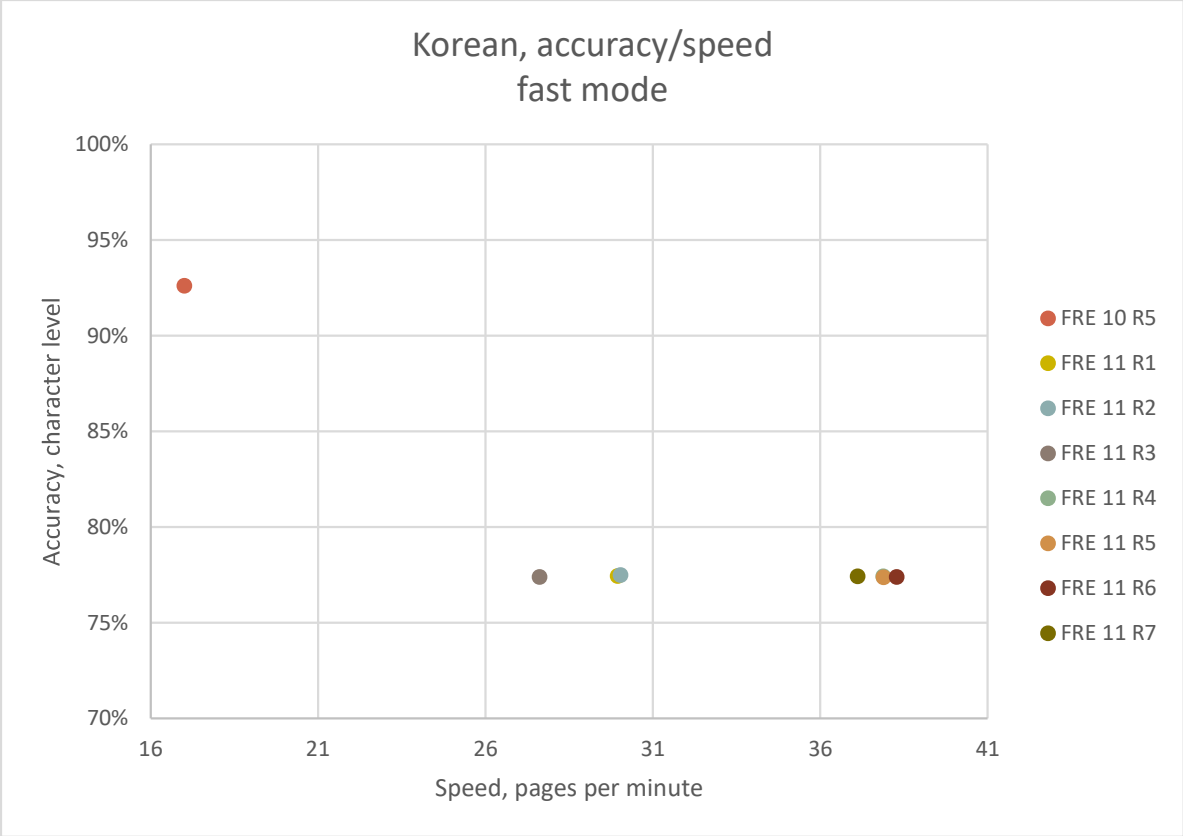
JAPANESE



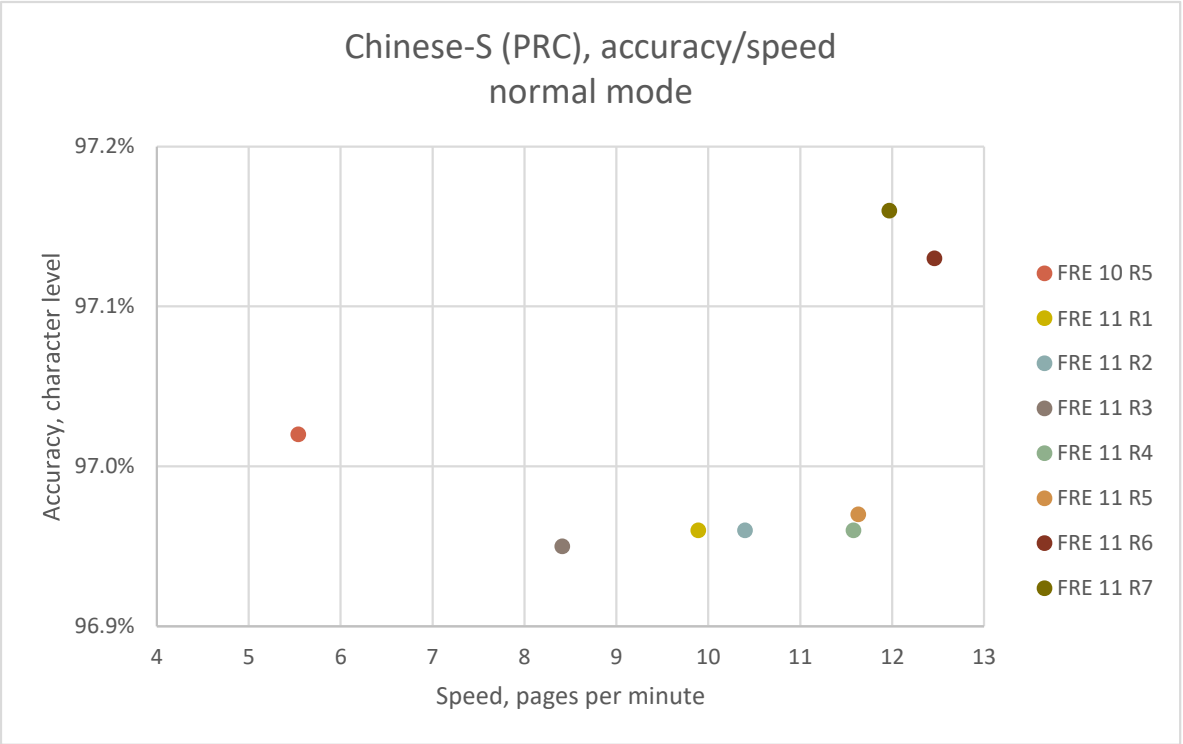


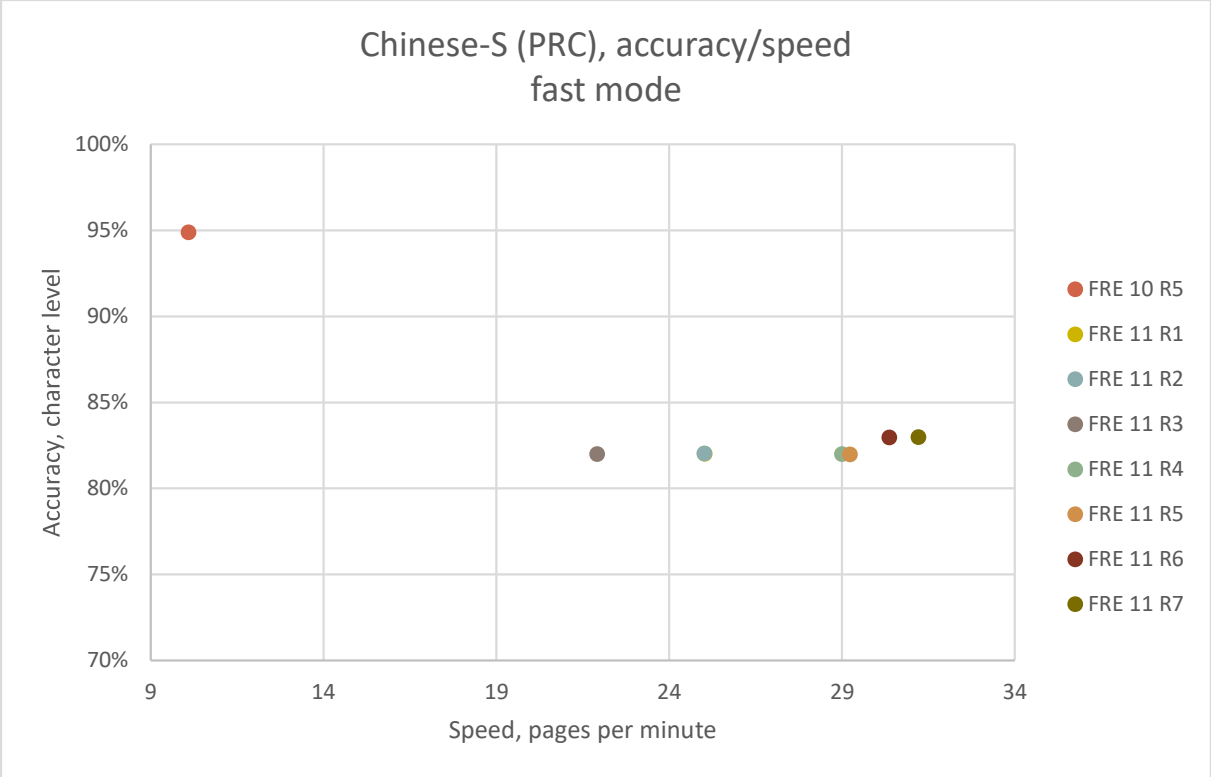
KOREAN



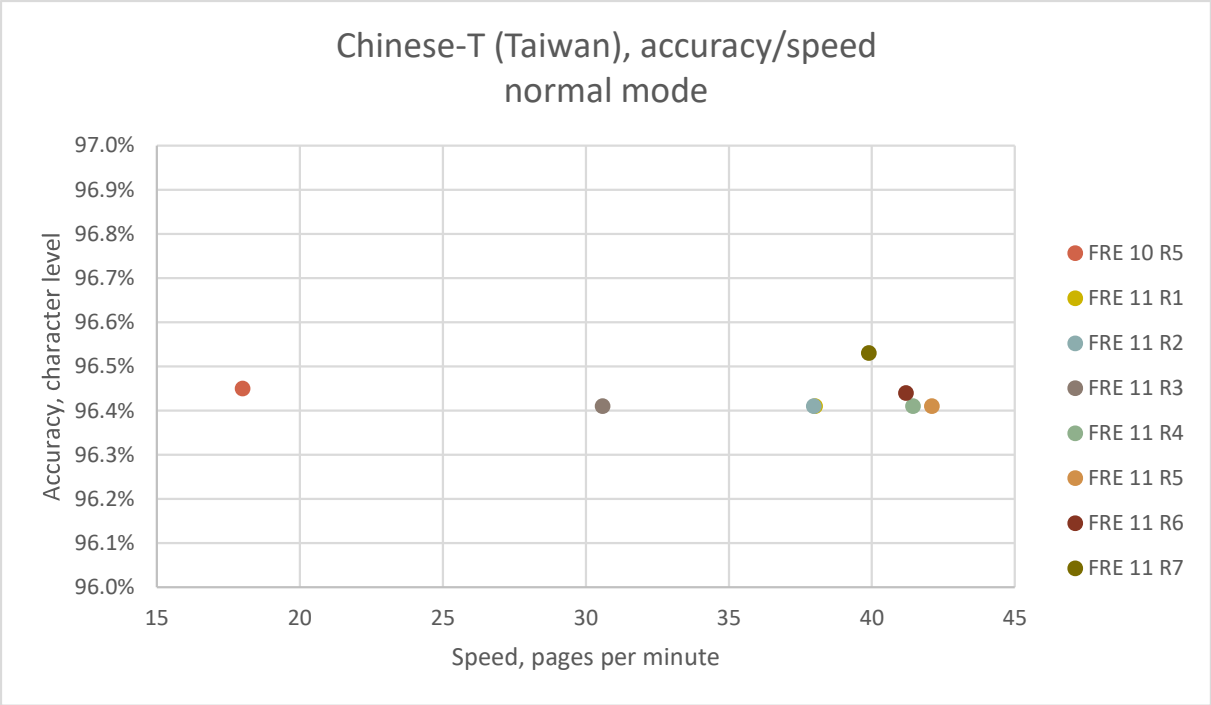


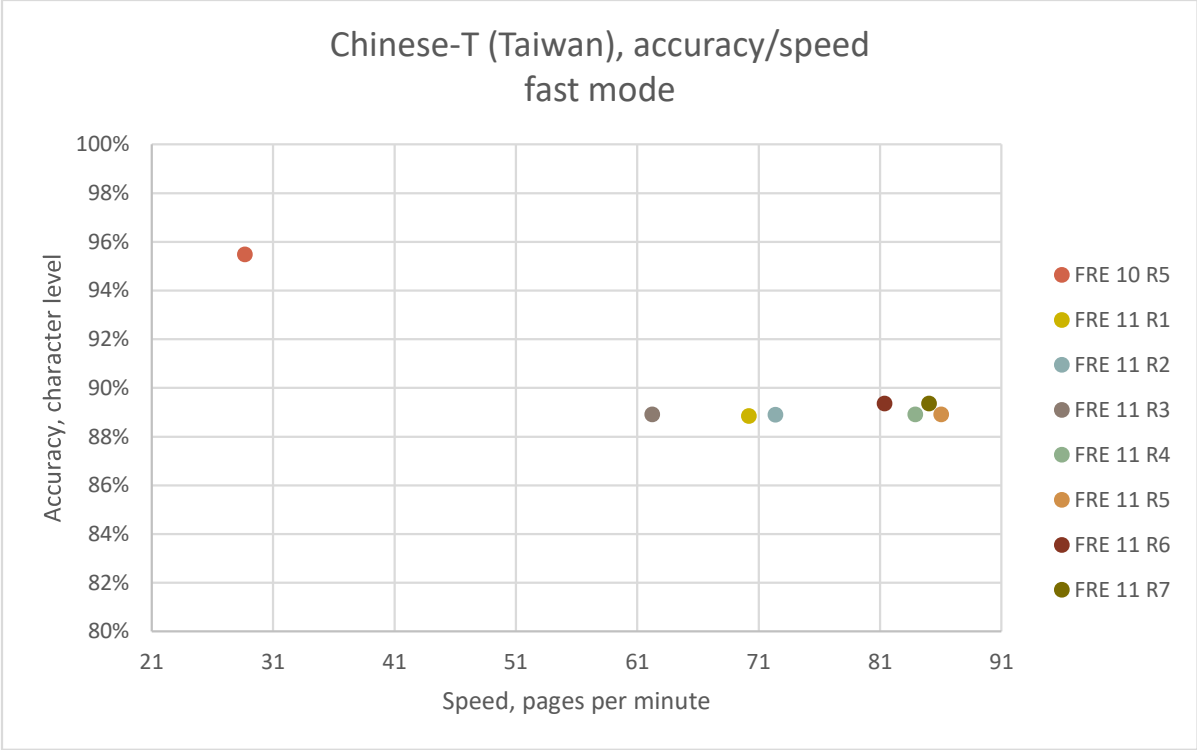
CHINESE PRC



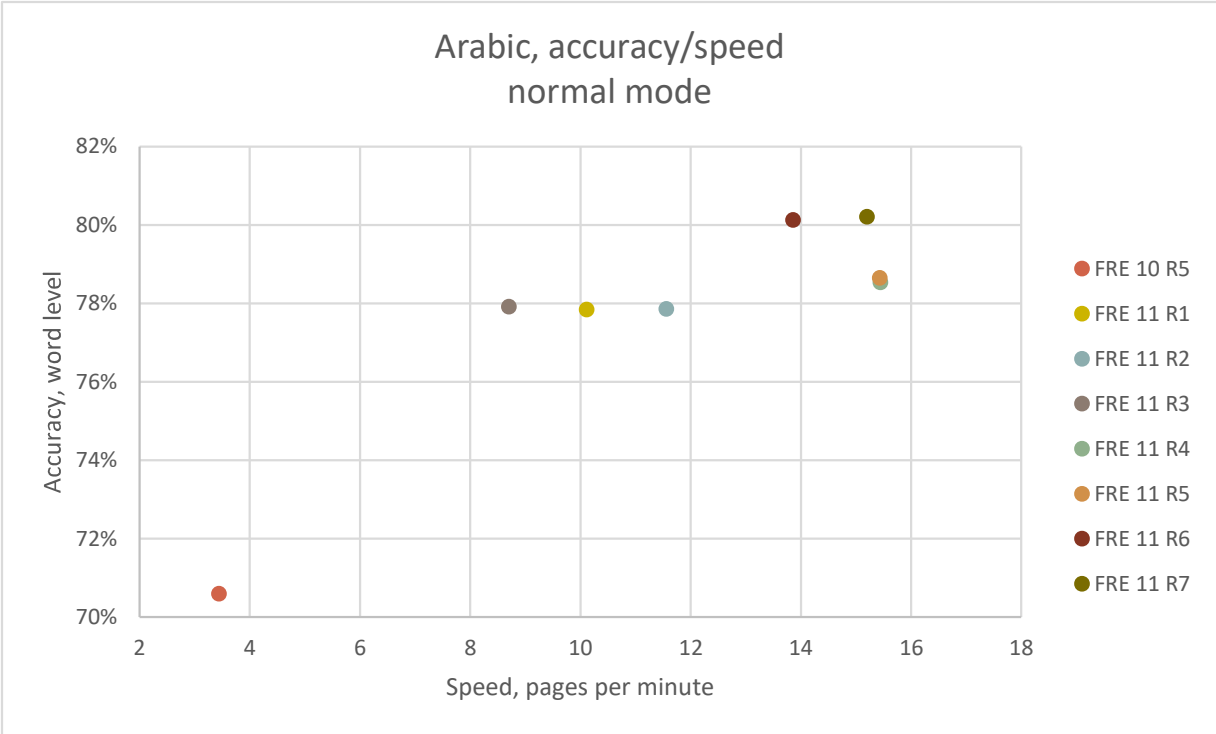


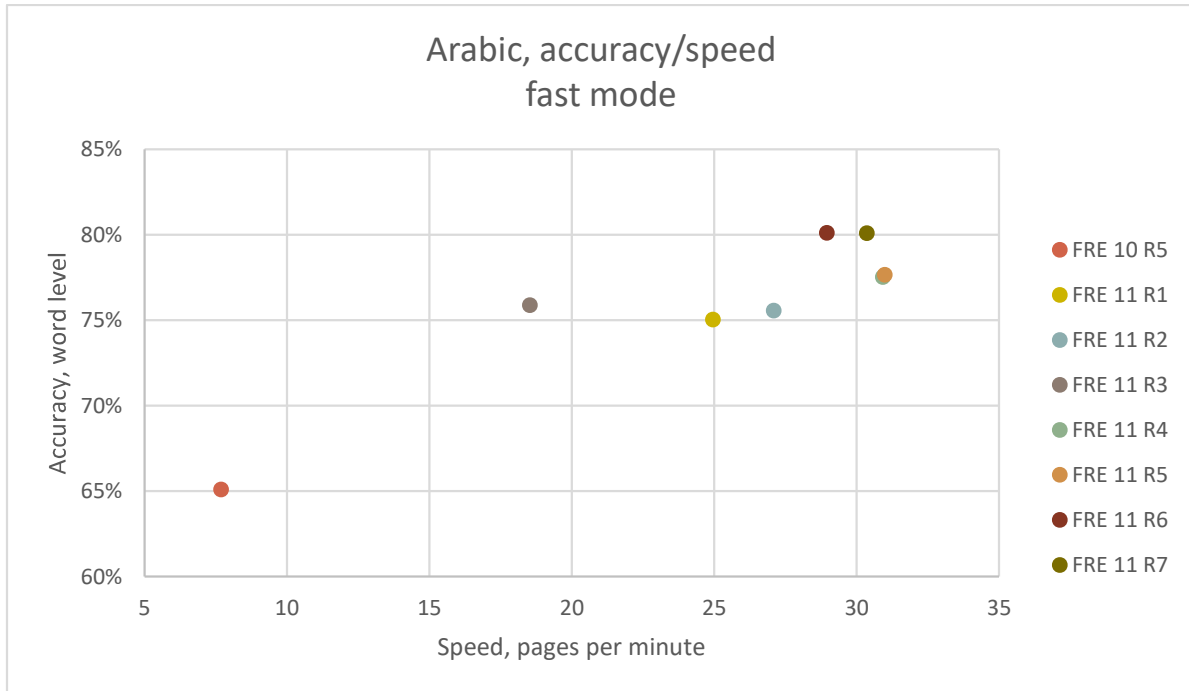
CHINESE TAIWAN





ARABIC





FIXED BUGS

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

Four-point scale will help you to evaluate the severity of each issue, enabling you to make informed decision 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 of the feature functionality to the internal specifications, memory leaks or data corruption.
Minor	A bug that leads to feature malfunctioning or affects minor functionality of the software. Minor bugs can include recognition errors, missing or lost objects, wrong color detection, incorrect document analysis, license counter errors, etc.
Trivial	A cosmetic issue that does not affect the functionality of the product but 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	Internal program error occurs when document structure is being read and IWord::get_Region is being used.	API	526665	RU
Critical	In multi-processing mode on Java the program crashes with 'Colnitalize has not been called' error after a long time of work when a lot of threads process the documents in infinite cycle. Root cause: when big amount of Java-threads is created, Java reuses internal threads of OS to its Java-threads. In this case two Java-threads can process using one thread of OS. And when one of the Java-threads finishes its work and deinitalize COM, the second one is still trying to use this thread. So, an error occurs. Workaround: not to use unlimited number of threads	API	542063	RU
Critical	An error 'Access violation' occurs during the processing our customer's input file without skew correction during image preparation and without pictures detection during layout analysis.	Recognizer	535364	EU
Critical	Use method FRPage::CorrectGeometricalDistortions, then on trying to look for vertically oriented text (PageAnalysisParams::DetectVerticalEuropeanText = true) internal program error occurs on our customers' PDF file.	Document Analysis	529645	EU
Critical	If region is laid on the customer's image and AnalyzeRegion method is used, internal program error occurs.	Document Analysis	511192	EU
Critical	In case of Network license usage, an error occurs when connection with server is lost after Engine has been loaded even if the interface for managing the connection breakdowns is used. Root cause: some problem situations with server were not processed in the right way and this is why the OnDisconnect() method was not called in this cases.	API	528727	US
Critical	Internal program error occurs on opening customer's image file with incorrect color masks.	Image	519469	RU
Critical	Internal program error occurs on the analysis stage of image processing on specific customer's image. Workaround: process image without choice of variants of page layout (ProhibitModelAnalysis = true)	Document Analysis	512057	EU
Critical	Internal program error occurs during the document analysis of specific customer's image with active orientation correction but without skew correction.	Document Analysis	537863	EU
Critical	Internal program error occurs when a document analysis is processing on specific customer's document.	Barcodes	537010	EU
Critical	An internal program error occurs when one tries to get a ParagraphAlignment value for Chinese text.	API	548896	EU
Major	'Cjk.*' resource files are demanded by FRE on processing with automatic language detection even though recognition languages set does not include any of CJK languages.	Document Analysis	509942	EU
Major	Throwing an exception when IParagraph::NextGroup() method	API	543033	US

	gets SF_BackgroundColor or SF_BaseLineRise flags in StyleFlagMask parameter. Root cause: an internal enumerator for recognizing the mask for style parameters didn't contain necessary elements.			
Major	Orientation is fixed only for the first 30 pages of the document when the page flushing policy is set to automatic mode (IFRDocument::PageFlushingPolicy = PFP_Auto), FREngine is set to distribute analysis and recognition of multi-page documents to CPU cores in automatic mode (IMultiProcessingParams::MultiProcessingMode = MPM_Auto) and the page orientation is detected during page preprocessing (IPagePreprocessingParams::CorrectOrientation = true). Workaround: set ABBYY FineReader Engine not to distribute analysis and recognition of multi-page documents to CPU cores. (IMultiProcessingParams::MultiProcessingMode = MPM_Sequential)	API	535634	US
Major	In case of processing PDF file using only text layer, when the pages of the document are recognized sequentially (MPM_Sequential) with image skew correction (CorrectSkew = TSPV_Yes), the output file doesn't contain a text layer. Workaround: use CorrectSkew = TSPV_Auto and/or MPM_Parallel mode Root cause: in technologies when the flag CorrectSkew is active, the copy of document is created with loss of information about text layer.	API	530146	US
Major	In case of multi-page documents, when the images are being recognized with BatchProcessor in parallel processes, pages of different documents can be mixed in output of the method. Root cause: while the BatchProcessor is used together with ExportFileWriter, the following situation was possible: the page has already been deleted from indices of unissued pages set, but has not been deleted from the document yet. The next pages are put to the indices set at and to the temp document with the same indices. The page order is mixed.	API	529263	US
Major	Memory leaks in the scenario: FRE initialization, image processing in cycle with default settings, FRE deinitialization.	Recognizer	535895	EU
Minor	In case of loading image with saving the coordinates on the image after modifications (PrepareImageMode::KeepOriginalCoordinatesInfo = false), on saving the recognized text from the PlainText object into an XML file, Export saves coordinates on original image instead of transformed image. Root cause: SaveToAsciiXMLFile() method couldn't recognize what coordinates should be used during the export: coordinates on a non-deskewed image or coordinates on deskewed image; and always saved coordinates on a non-deskewed image.	API	529807	US
Minor	In case of export to XML with words variants saving, several characters for one wordRecVariant have same coordinates. Root cause: export has no information about character variants' rectangles that are not equal to the basic variant's rectangle. So, output file contains same rectangles for every characters' variants.	Export	523884	US
Minor	In case of export to ALTO format in mode with saving characters'	API	520823	US

	coordinates on image with skew correction (AltoExportParams::WriteNonDeskewedCoordinates = FALSE), <Shape> elements are missed.			
Minor	When the document is processed using Farsi language, most of characters have CharConfidence = -1 Root cause: for Farsi language the normalized weight was not calculated correctly	Recognizer	512578	EU
Minor	Mobile phone number subscribed as 'wireless' recognized as BCFT_Phone, not BCFT_Mobile. Root cause: 'wireless' was not related to mobile phone number.	Document Analysis	538150	US
Minor	Images in the customer's document are detected even with settings when pictures are not detected during the layout analysis.	Document Analysis	528220	EU
Minor	QR-code is not detected on customer's document processing. Root cause: the edges of the dots are too noisy.	Barcodes Detector	518784	RU
Minor	A part of the text is not recognized on customer's document. Root cause: this area is recognized as an image.	Document Analysis	513141	3A
Minor	The output DOCX file contains not all of horizontal lines.	Synthesis	508370	US
Minor	Numbers like '60.01' at the beginning of the raw are synthesized as elements of the list during the export to rtf format. Root cause: Synthesis recognized numerators like '01' or '7.05' as list elements.	Synthesis	540540	3A
Minor	Numbers like '60.01' at the beginning of the raw are synthesized as elements of the list during the export to rtf format.	Synthesis	540540	3A
Minor	In case of export to PDF with embedded fonts and no predefined filters, the text layer is narrower than characters border.	Export	546487	EU
Trivial	The default value of IBackgroundLayer::IsPicture property is FALSE, but it is specified to be TRUE in documentation.	Help	546581	US
Trivial	The tab character is always replaced by one character during the processing. Root cause: an internal flag was active with no opportunity to switch it off.	API	537658	3A
Trivial	Incorrect sample in the 'Document Classification' article. Classify method must return ClassificationClasses object instead of StringsCollection.	Help	520210	EU
Trivial	The EnginePool sample returns 'There are no free licenses' when the Trial license with CPU limit is used.	Samples	451719	US
Trivial	Generated linearized PDF/A-1a file doesn't pass the compliance test in Preflight.	Export	546220	US
Trivial	In case of Runtime installation of FREngine, the SN parameter is not obligated and should be deleted from several articles related to installation in documentation.	Samples	545923	EU

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 FALSE or DEFAULT whereas default value is FALSE. (HelpDesk request 416986)

'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 is planned to be implemented in the next release with low priority because the scenario is not correct and there is a workaround for it.

THE PROCESS HANGS WHEN THE MULTIPROCESSING METHODS WORKS ON THE MACHINE WITH ONE CPU.

The machine has one core, `IMultiProcessingParams::MultiProcessingMode = MPM_Parallel` and `IMultiProcessingParams::RecognitionProcessesCount = 2` settings are set, operation system has radio System Properties->Advanced->Performance->Settings...->Advanced->Processor scheduling set to Background services value. The process hangs.

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:

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".

R7 GM PATCH –

PART NUMBER, BUILD NUMBER

Part#	1041/90
Build#	11.1.16.94 with Export.Pdf.dll from 11.1.16.96

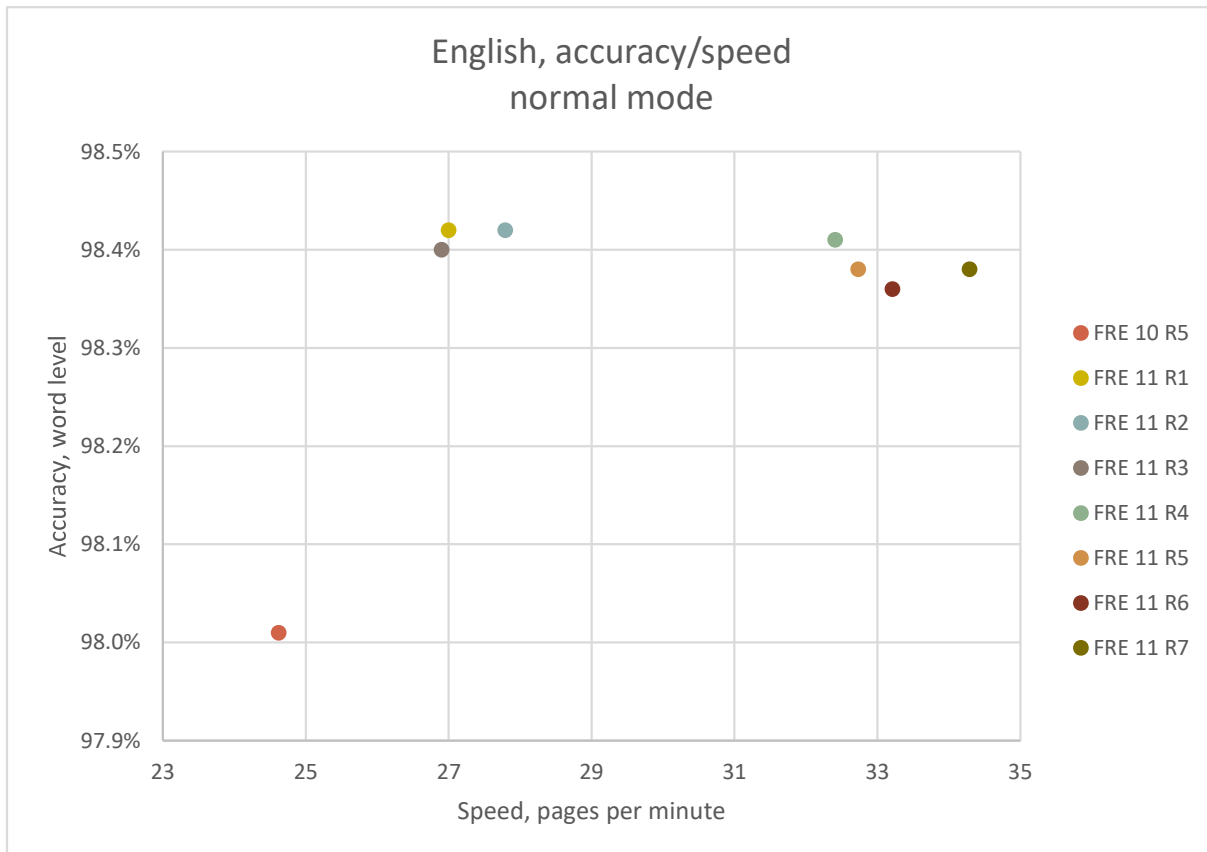
NEW FEATURES AND IMPROVEMENTS

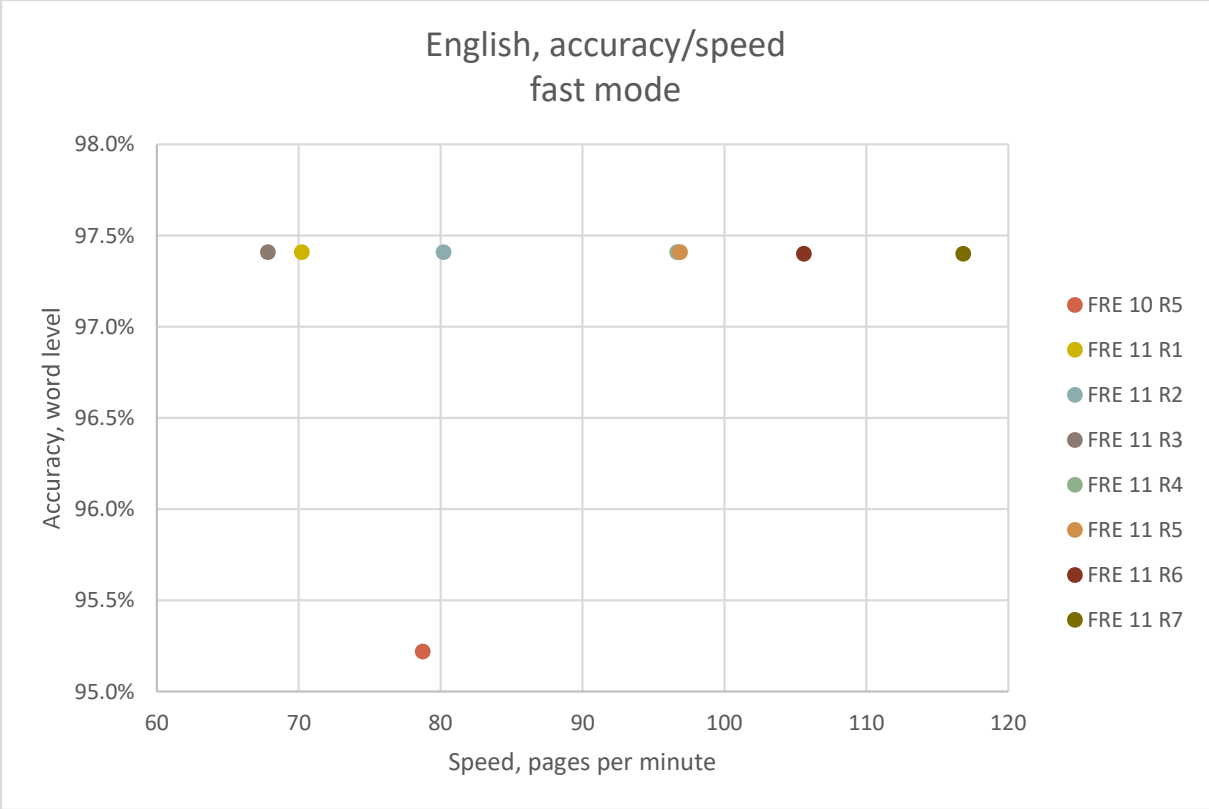
None

PERFORMANCE RESULTS

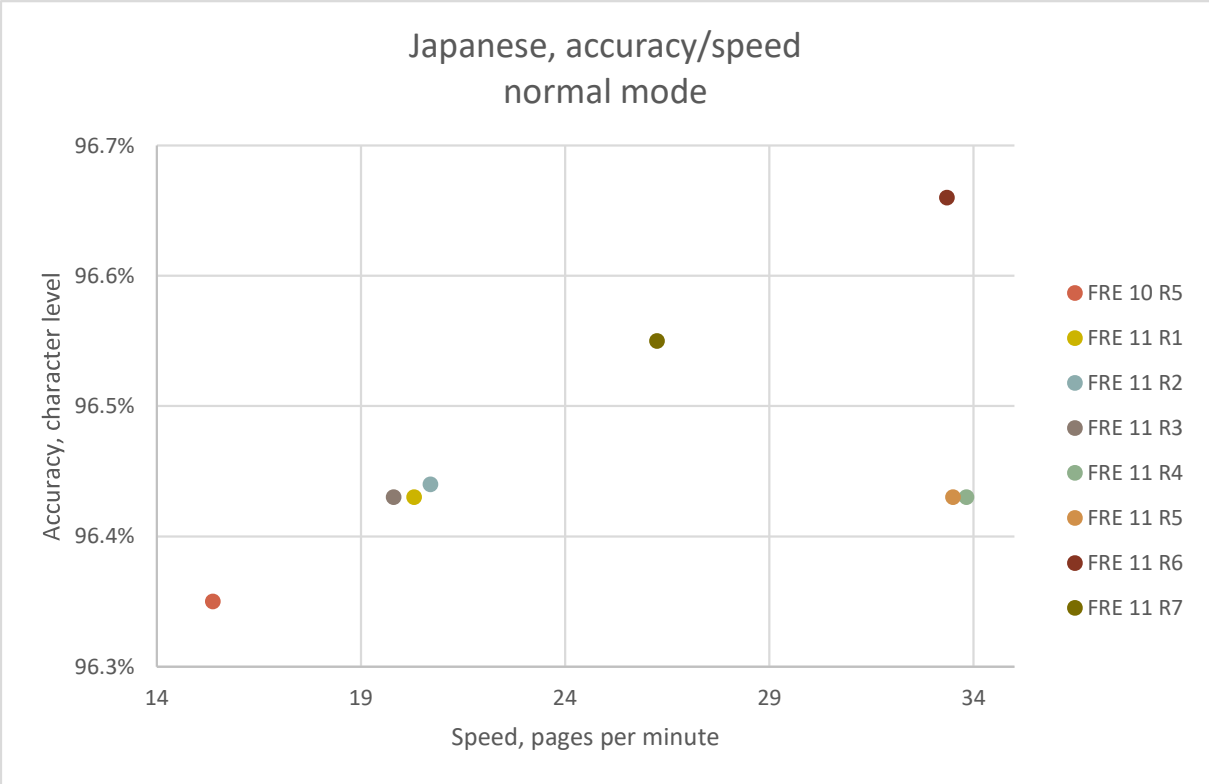
This section contains performance results of the release comparing to the previous releases. The processor of the testing machine is Intel® Core™ i5-4690 CPU (3.50GHz, 4 physical cores) with 8GB of RAM.

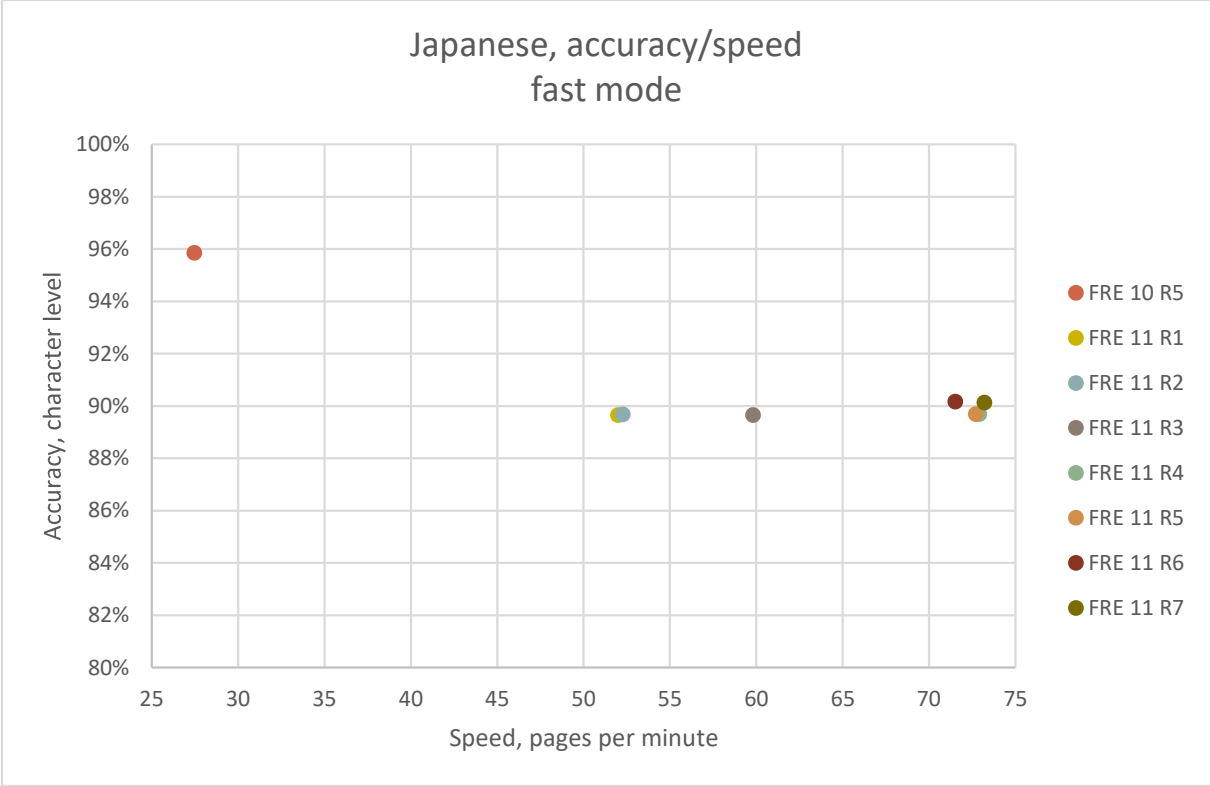
ENGLISH



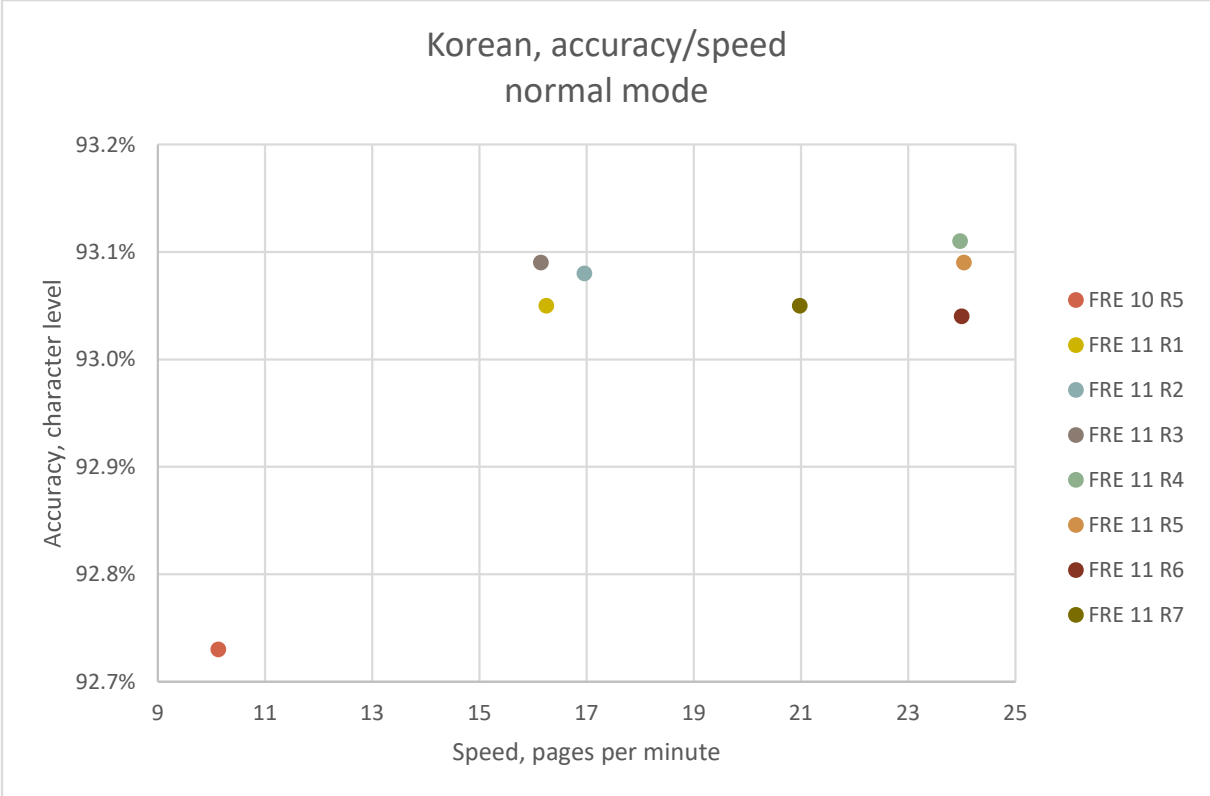


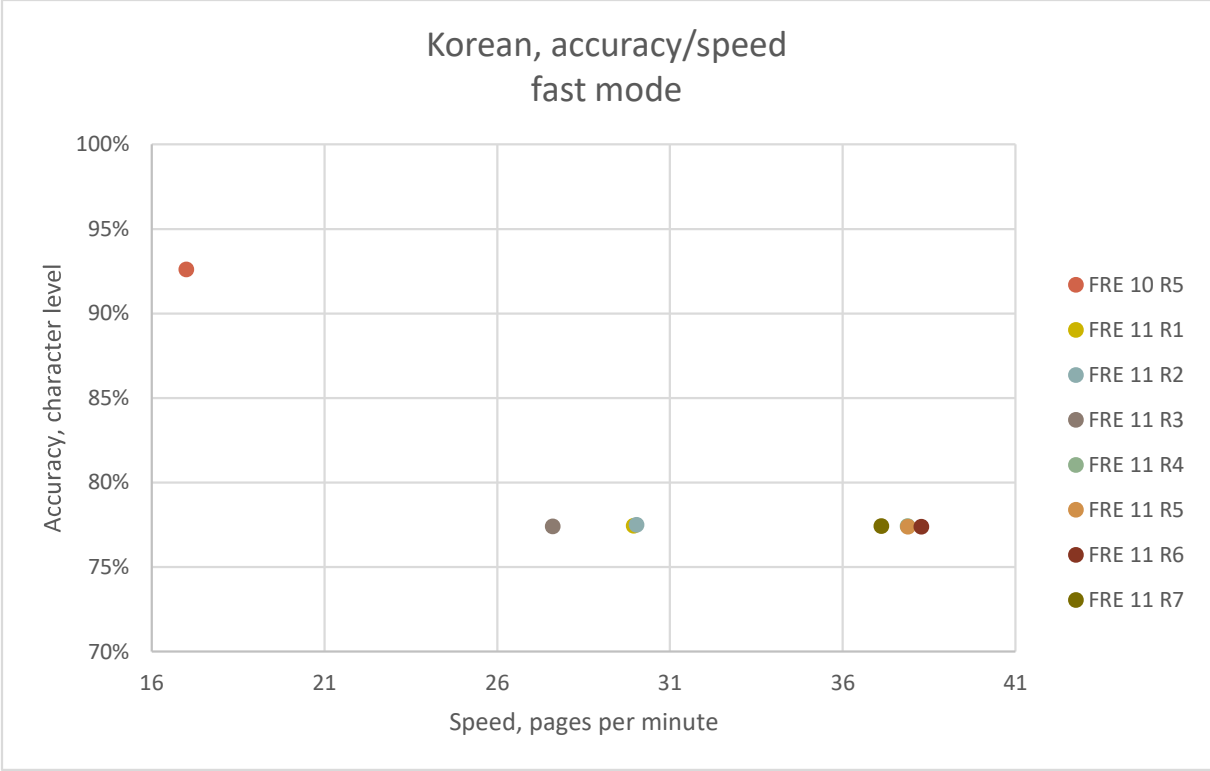
JAPANESE



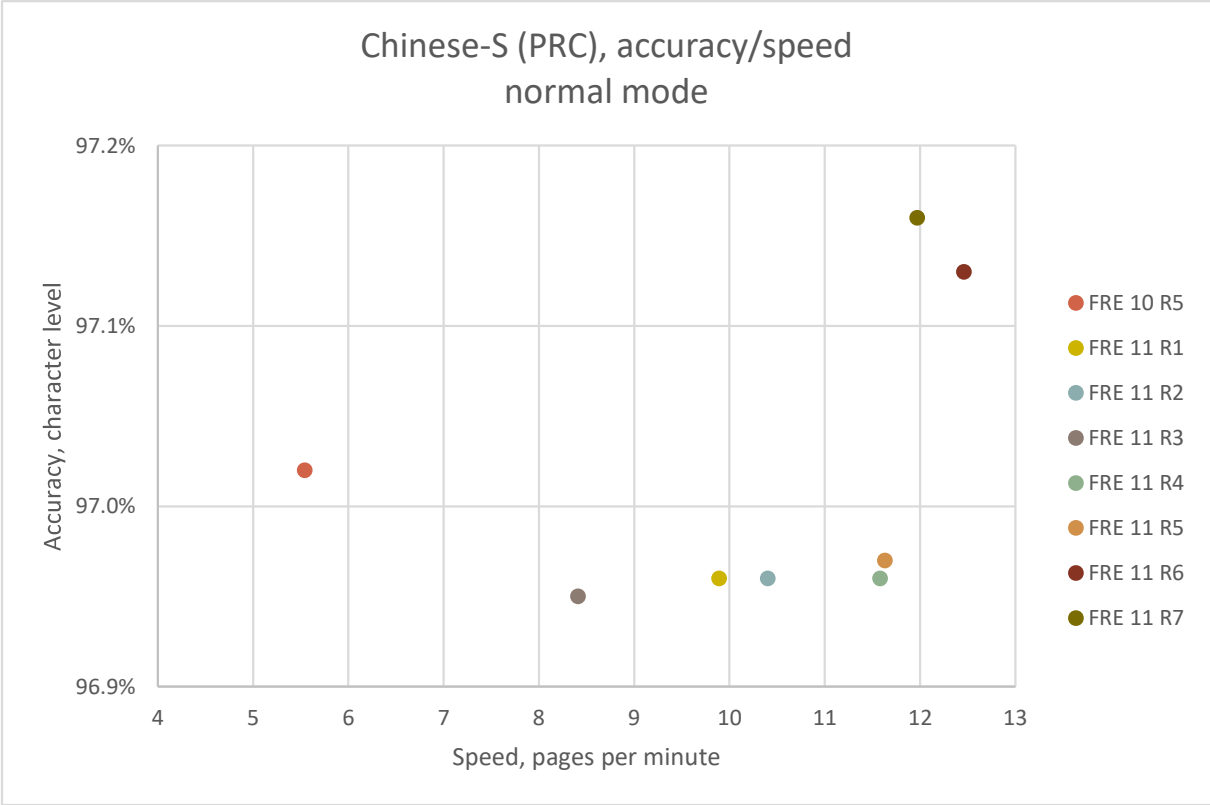


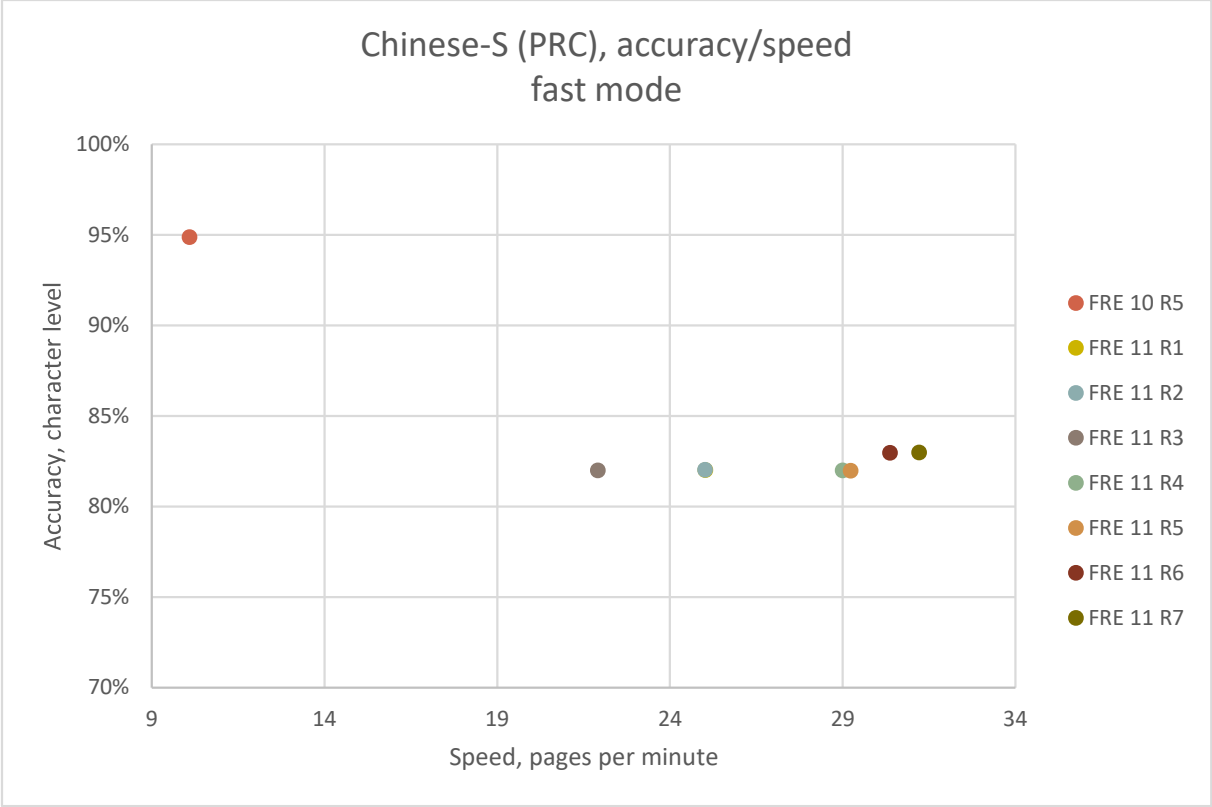
KOREAN



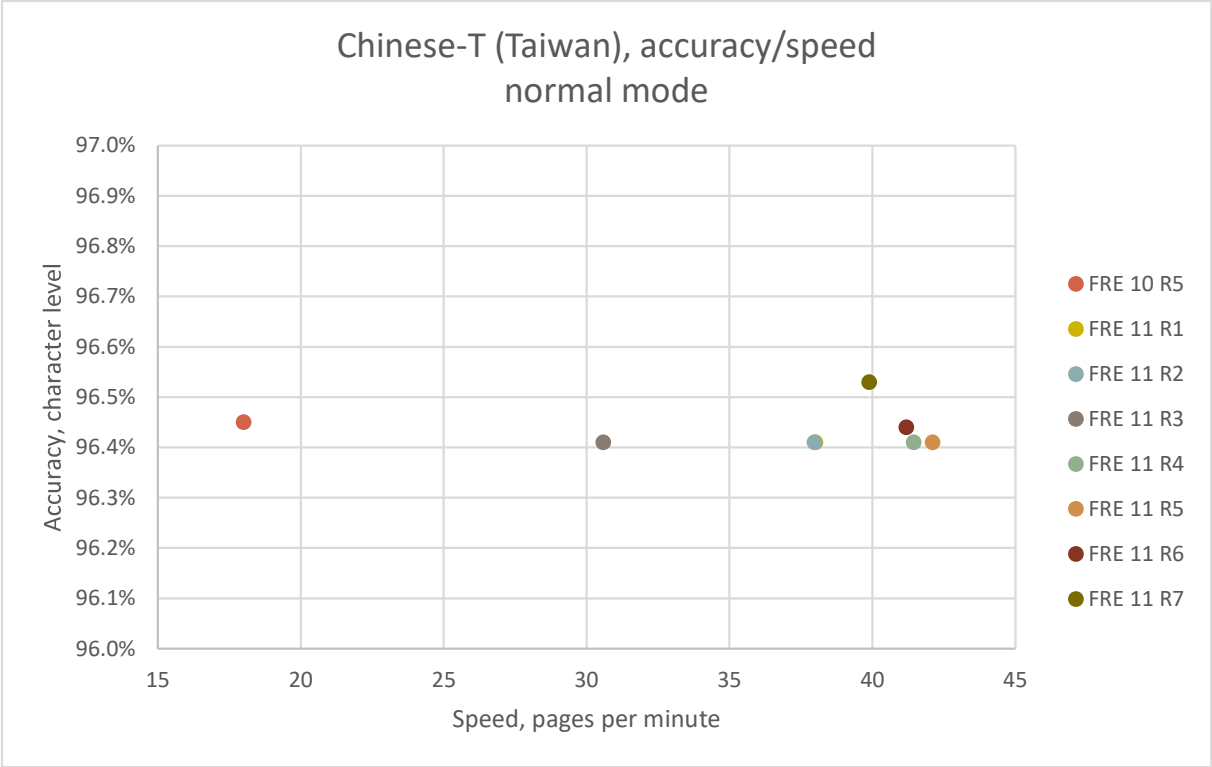


CHINESE PRC

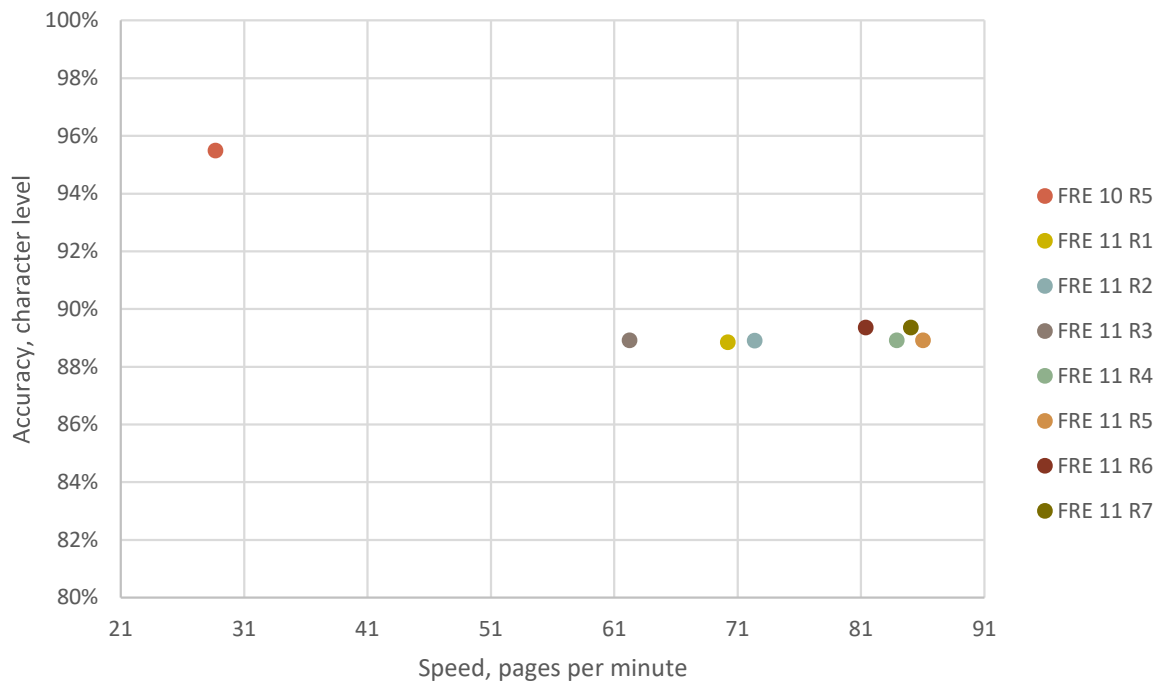




CHINESE TAIWAN

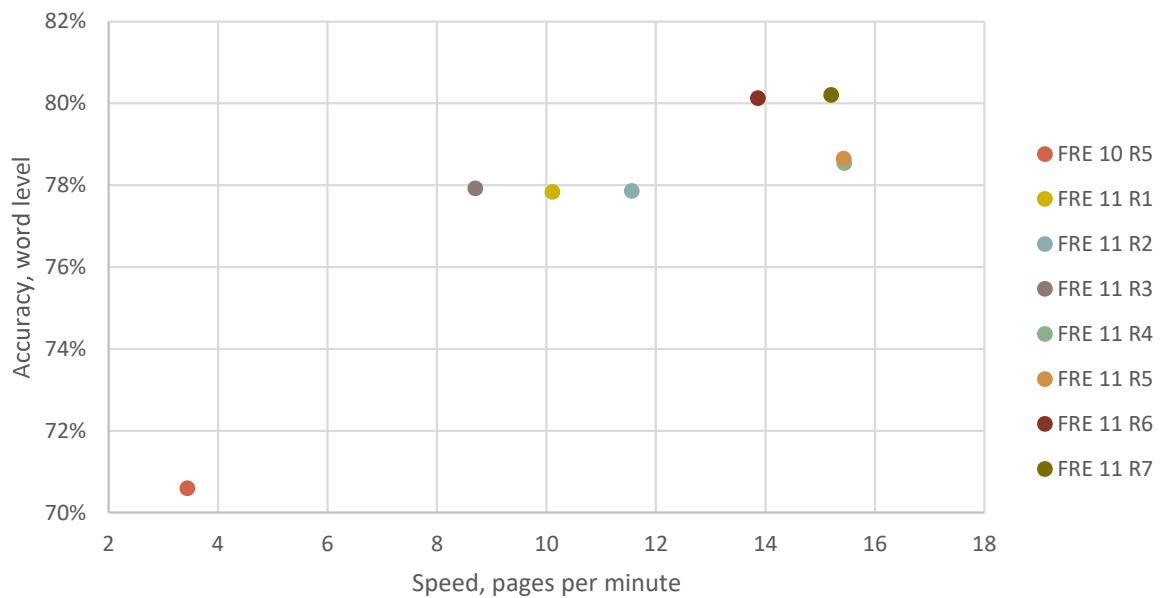


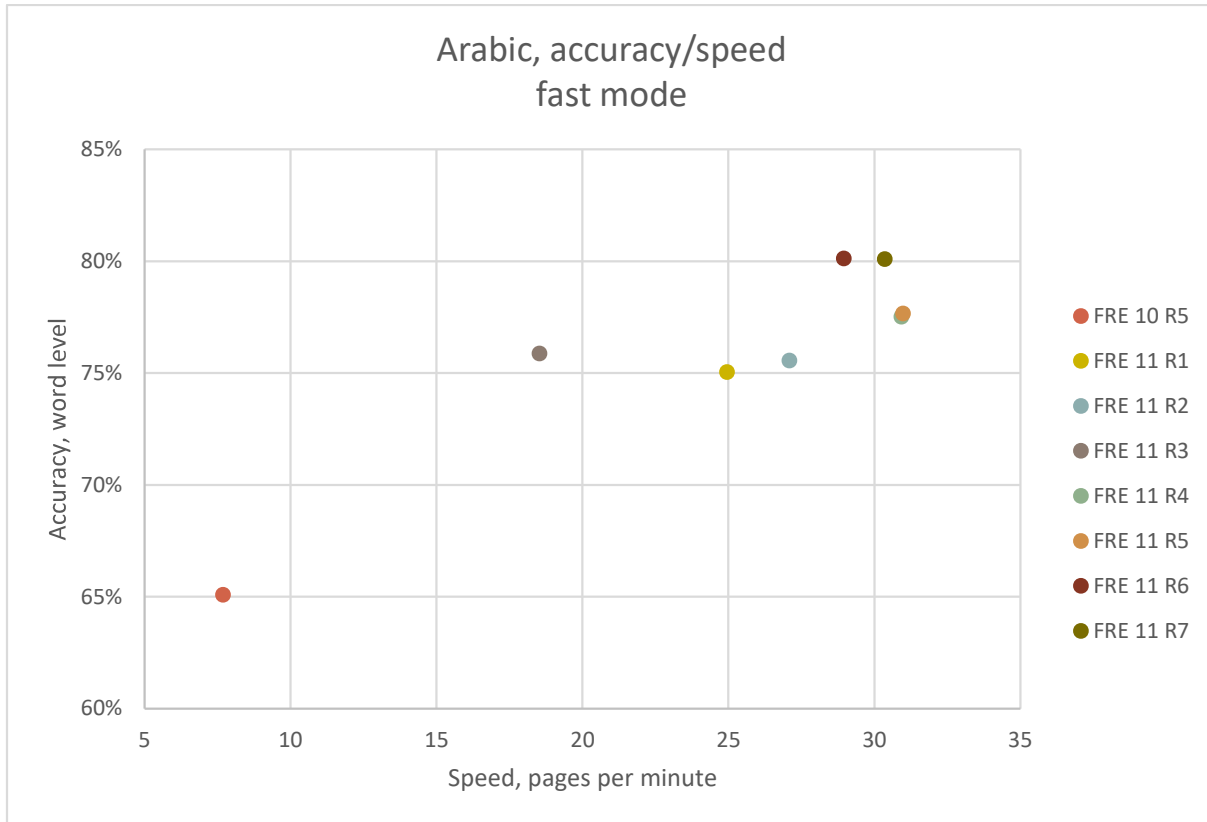
Chinese-T (Taiwan), accuracy/speed
fast mode



ARABIC

Arabic, accuracy/speed
normal mode





FIXED BUGS

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

Four-point scale will help you to evaluate the severity of each issue, enabling you to make informed decision 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 of the feature functionality to the internal specifications, memory leaks or data corruption.
Minor	A bug that leads to feature malfunctioning or affects minor functionality of the software. Minor bugs can include recognition errors, missing or lost objects, wrong color detection, incorrect document analysis, license counter errors, etc.
Trivial	A cosmetic issue that does not affect the functionality of the product but 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
Minor	The time of exported file's creation/modification displayed in Adobe Reader is not equal to this file's creation/modification time displayed in file properties.	Export	530461	US

KNOWN ISSUES AND WORKAROUNDS

Same as R7.

R8 GM –

PART NUMBER, BUILD NUMBER

Part#	1041/104
Build#	11.1.19.48

CHANGE IN BEHAVIOR

We have changed the pattern of working with redundancy time. From now in LicensingSettings.xml there is no opportunity to set redundancy live time (RedundancyLiveTime parameter) and redundancy synchronization time (RedundancySynchronizationTime parameter).

Redundancy live time was fully moved to serial number parameters. Redundancy synchronization time is an internal parameter that was available for editing by mistake.

NEW FEATURES AND IMPROVEMENTS

ORIENTATION AND SKEW CORRECTION USING INJECTTEXTLAYER METHOD

All settings in the PrepareImageMode and PageProcessingParams object such as CorrectSkew and CorrecrOrientation work when passed to InjectTextLayer method.