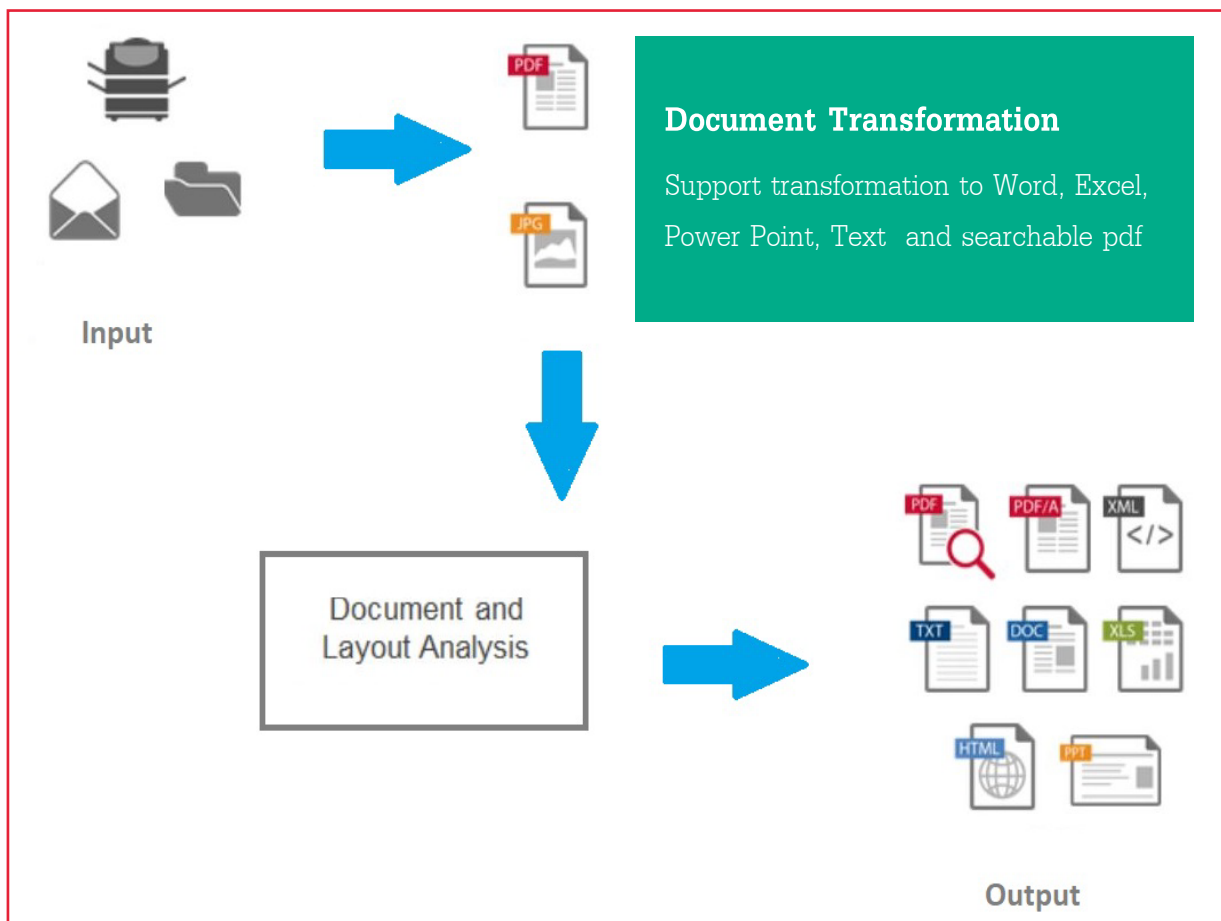# E.S.
(OCR enterprise scanner)

OCR enterprise scanner mark-3 datasheet

What is OCR?

Digital character recognition is in many use cases such as Optical Character Recognition (OCR), Intelligent Character Recognition (ICR), Intelligent Image Recognition (IIR). When you scan a document that has text on it, the scanner simply takes a picture of the page and saves it as an image. Text saved in an image cannot be edited or searched, so the data it contains is useless to other software applications.
In order to transform this information into an editable format that you can search, copy, and modify without retyping it manually, you need Optical Character Recognition (OCR) software.

Document Transformation

Support transformation to Word, Excel, Power Point, Text and searchable pdf

Input

Document and Layout Analysis

Output

Recognition Languages

- Software be able to recognizes the following languages.
- 93%-98% accuracy range after training for standard fonts.
- Default be able to recognition language by bookmarking the corresponding link. - Support image size 300 DPI , fix format and size 14 or higher, grayscale

Afrikaans	Amharic	Arabic
Assamese	Azerbaijani	Azerbaijani Cyrilic
Belarusian	Bengali	Tibetan
Bosnian	Breton	Bulgarian
Catalan	Cebuano	Czech
Chinese Simplified	Chinese Traditional	Cherokee
Welsh	Danish	German
Dzongkha	Greek Modern	English
English Ancient	Esperanto	Math Equation
Estonian	Basque	Persian
Finnish	French	Frankish
French Middle	Irish	Galician
Greek Ancient	Gujarati	Haitian
Hebrew	Hindi	Croatian
Hungarian	Inuktitut	Indonesian
Icelandic	Italian	Italian Ancient
Javanese	Japanese	Kannada
Georgian	Georgian Ancient	Kazakh
Khmer	Kirghiz	Korean
Korean Vertical	Kurdish Sorani	Kurdish Kurmanji
Lao	Latin	Latvian
Lithuanian	Luxembourgish	Malayalam
Marathi	Macedonian	Maltese
Mongolian	Maori	Malay
Burmese	Nepali	Dutch
Moldavian	Norwegian	Occitan
Oriya	Panjabi	Polish
Portuguese	Pushto	Quechua
Romanian	Russian	Sanskrit
Sinhala	Slovakian	Slovenian
Sindhi	Spanish	Spanish Ancient
Albanian	Serbian	Serbian Latin
Sundanese	Swahili	Swedish
Syriac	Tamil	Tatar
Telugu	Tajik	Tagalog
Thai	Tigrinya	Tonga
Turkish	Uighur	Ukrainian
Urdu	Uzbek	Uzbek Cyrilic
Vietnamese	Yiddish	Yoruba

Recognition languages

Standard fonts formats

English: - Calibri

- Cambria

- Garamond

- Didot

- Georgia

- Helvetica

- Arial

- Book Antiqua

- Trebuchet MS

*Including Roman numeral, any of the symbols used in a system of numerical notation based on the ancient Roman system.

Recognition languages

Standard fonts formats

Thai :
- TH Sarabun PSK
- TH Krub
- TH Niramit AS
- TH Kodchasal
- TH KoHo
- TH Chakra Petch
- TH Baijam
- TH Fah Kwang
- TH Sarabun

*Including Thai numeral, any of the symbols used in a system of numerical notation based on Thai system.

Use our OCR software

Without installation, recognize text and characters from PDF scanned documents (including multipage files), photographs and digital camera captured images be able to process online.

OCR provides Representational State Transfer (REST) APIs, allowing you to call APIs using HTTPS.

Optical Character Recognition and Template Creation

We provide a template creation service that allows the digital interpretation of any PDF document.

The key components of automated document processing is the ability to take a paper or scanned (PDF) document and extract from it the data that the enterprise systems require for their business processes. Different languages, especially non-European ones that use complex and interdependent character sets were interpreted with difficulty, if at all. Recently there has been a quantum leap in capability, in part because of the introduction of Artificial Intelligence and machine learning in the interpretation of scripts.

Classifying Logos in Images with Convolutional Neural Networks Machine learning algorithm call CNNs are used for analyzing images. Train and deploy deep learning models for offline OCR (optional character recognition) on mobile and embedded devices in record time.