**What is OCR (Optical Character Recognition)?**

Optical Character Recognition (OCR) is the process that converts an image of text into a machine-readable text format. For example, if you scan a form or a receipt, your computer saves the scan as an image file. You cannot use a text editor to edit, search, or count the words in the image file. However, you can use OCR to convert the image into a text document with its contents stored as text data.

**Why is OCR important?**

Most business workflows involve receiving information from print media. Paper forms, invoices, scanned legal documents, and printed contracts are all part of business processes. These large volumes of paperwork take a lot of time and space to store and manage. Though paperless document management is the way to go, scanning the document into an image creates challenges. The process requires manual intervention and can be tedious and slow.

Moreover, digitizing this document content creates image files with the text hidden within it. Text in images cannot be processed by word processing software in the same way as text documents. OCR technology solves the problem by converting text images into text data that can be analyzed by other business software. You can then use the data to conduct analytics, streamline operations, automate processes, and improve productivity.

**How does OCR work?**

The OCR engine or OCR software works by using the following steps:

**Image acquisition**

A scanner reads documents and converts them to binary data. The OCR software analyzes the scanned image and classifies the light areas as background and the dark areas as text.

**Preprocessing**

The OCR software first cleans the image and removes errors to prepare it for reading. These are some of its cleaning techniques:

* Deskewing or tilting the scanned document slightly to fix alignment issues during the scan.
* Despeckling or removing any digital image spots or smoothing the edges of text images.
* Cleaning up boxes and lines in the image.
* Script recognition for multi-language OCR technology

**Text recognition**

The two main types of OCR algorithms or software processes that an OCR software uses for text recognition are called pattern matching and feature extraction.

**Pattern matching**

Pattern matching works by isolating a character image, called a glyph, and comparing it with a similarly stored glyph. Pattern recognition works only if the stored glyph has a similar font and scale to the input glyph. This method works well with scanned images of documents that have been typed in a known font.

**Feature extraction**

Feature extraction breaks down or decomposes the glyphs into features such as lines, closed loops, line direction, and line intersections. It then uses these features to find the best match or the nearest neighbor among its various stored glyphs.

**Postprocessing**

After analysis, the system converts the extracted text data into a computerized file. Some OCR systems can create annotated PDF files that include both the before and after versions of the scanned document.

**What are the types of OCR?**

Data scientists classify different types of OCR technologies based on their use and application. The following are a few examples:

**Simple optical character recognition software**

A simple OCR engine works by storing many different font and text image patterns as templates. The OCR software uses pattern-matching algorithms to compare text images, character by character, to its internal database. If the system matches the text word by word, it is called optical word recognition. This solution has limitations because there are virtually unlimited font and handwriting styles, and every single type cannot be captured and stored in the database.

**Intelligent character recognition software**

Modern OCR systems use intelligent character recognition (ICR) technology to read the text in the same way humans do. They use advanced methods that train machines to behave like humans by using machine learning software. A machine learning system called a neural network analyzes the text over many levels, processing the image repeatedly. It looks for different image attributes, such as curves, lines, intersections, and loops, and combines the results of all these different levels of analysis to get the final result. Even though ICR typically processes the images one character at a time, the process is fast, with results obtained in seconds.

**Intelligent word recognition**

Intelligent word recognition systems work on the same principles as ICR, but process whole word images instead of preprocessing the images into characters.

**Optical mark recognition**

Optical mark recognition identifies logos, watermarks, and other text symbols in a document.

**What are the benefits of OCR?**

Data scientists classify different types of OCR technologies based on their use and application. The following are a few examples:

**Simple optical character recognition software**

A simple OCR engine works by storing many different font and text image patterns as templates. The OCR software uses pattern-matching algorithms to compare text images, character by character, to its internal database. If the system matches the text word by word, it is called optical word recognition. This solution has limitations because there are virtually unlimited font and handwriting styles, and every single type cannot be captured and stored in the database.

**Intelligent character recognition software**

Modern OCR systems use intelligent character recognition (ICR) technology to read the text in the same way humans do. They use advanced methods that train machines to behave like humans by using machine learning software. A machine learning system called a neural network analyzes the text over many levels, processing the image repeatedly. It looks for different image attributes, such as curves, lines, intersections, and loops, and combines the results of all these different levels of analysis to get the final result. Even though ICR typically processes the images one character at a time, the process is fast, with results obtained in seconds.

**Intelligent word recognition**

Intelligent word recognition systems work on the same principles as ICR, but process whole word images instead of preprocessing the images into characters.

**Optical mark recognition**

Optical mark recognition identifies logos, watermarks, and other text symbols in a document.

**What are the benefits of OCR?**

The following are major benefits of OCR technology:

**Searchable text**

Businesses can convert their existing and new documents into a fully searchable knowledge archive. They can also process the text database automatically by using data analytics software for further knowledge processing.

**Operational efficiency**

You can improve efficiency by using OCR software to automatically integrate document workflows and digital workflows within your business. Here are some examples of what OCR software can do:

* Scan hand-filled forms for automated verification, reviews, editing, and analysis. This saves the time required for manual document processing and data entry.
* Find the required documents by quickly searching for a term in the database so that you don't have to manually sort through files in a box.
* Convert handwritten notes to editable texts and documents.

**Artificial intelligence solutions**

OCR is often part of other artificial intelligence solutions that businesses might implement. For example, it scans and reads number plates and road signs in self-driving cars, detects brand logos in social media posts, or identifies product packaging in advertising images. Such artificial intelligence technology helps businesses make better marketing and operational decisions that reduce expenses and improve the customer experience.

**What is OCR used for?**

The following are some common OCR use cases in various industries:

**Banking**

The banking industry uses OCR to process and verify paperwork for loan documents, deposit checks, and other financial transactions. This verification has improved fraud prevention and enhanced transaction security. For example, BlueVine is a financial technology company that provides financing to small and medium-sized businesses. It used Amazon Textract, a cloud-based OCR service, to develop a product for small businesses in the US to quickly access Paycheck Protection Program (PPP) loans as part of the COVID-19 relief stimulus package. Amazon Textract automatically processed and analyzed tens of thousands of PPP forms per day so that BlueVine could help several thousand businesses get funds, saving over 400,000 jobs in the process.

**Healthcare**

The healthcare industry uses OCR to process patient records, including treatments, tests, hospital records, and insurance payments. OCR helps to streamline workflow and reduce manual work at hospitals while keeping records up to date. For example, the nib Group provides health and medical insurance to over 1 million Australians and receives thousands of medical claims per day. Its customers can take photos of their medical invoice and submit them through the nib mobile app. Amazon Textract processes these images automatically so that the company can approve claims much faster.

**Logistics**

Logistics companies use OCR to track package labels, invoices, receipts, and other documents more efficiently. For example, the Foresight Group uses Amazon Textract to automate invoice processing in SAP. Manual entry of these business documents was time-consuming and error-prone because Foresight employees had to enter the data in multiple accounting systems. With Amazon Textract, Foresight software can read characters more accurately across many different layouts, which increases business efficiency.

**How can AWS help with OCR?**

AWS offers two services that can help you implement OCR in your business:

[Amazon Textract](https://aws.amazon.com/textract/) is a machine learning (ML) service that uses OCR to automatically extract text, handwriting, and data from scanned documents such as PDFs. It can read thousands of different documents in multiple layouts and formats at high speed. When it extracts information from documents, Amazon Textract returns a confidence score for everything it identifies so that you can make informed decisions about how you want to use the results.

[Amazon Rekognition](https://aws.amazon.com/rekognition/) can analyze millions of images and videos within minutes and augment human visual review tasks with artificial intelligence. You can use Amazon Rekognition APIs to extract text from both images and videos. You can extract skewed and distorted text from images and videos of street signs, social media posts, and product packaging.

Get started with OCR on AWS by [creating an AWS account](https://portal.aws.amazon.com/gp/aws/developer/registration/index.html?nc2=h_ct&src=header_signup) today.

https://aws.amazon.com/what-is/ocr/