

# Machine Learning, Deep Learning, and OCR: Revitalizing Technology

Mada Center

Optical character recognition tools are experiencing a quiet revolution as aspiring software providers merge OCR with Machine Learning and Deep Learning. Therefore, data capturing software is instantaneously capturing information and understanding the content. In practice, this means that Machine Learning and Deep Learning tools can check for mistakes independent of a human user providing efficient fault management.

Until now, OCR has contributed to helping business owners to automate the processing of handling physical documents. When it comes to people with functional limitations, virtually all people with visual impairments to learning disabilities use OCR technology provided by various entities. Today, OCR programs are still used to transform handwritten or printed text into machine-encoded text so that it can be retrieved on a computer. OCR programs make copies of documents like receipts, bank statements, passports, and other forms of documentation that needs managing.

Technology is being revitalized with the introduction of Artificial Intelligence, Machine Learning, and Deep Learning. Software developers are working on robust solutions and upgrades to existing OCR devices. Until now, OCR users only option to increase the reliability of scans is to manually measure and evaluate the process results. With the introduction of Machine Learning and Deep Learning, solutions will automatically conduct the evaluation while pulling insights from the text and understanding the meaning of the converted text. In other words, they can process document content more accurately.

As visionary technology providers blend OCR with Machine Learning and Deep Learning, these tools are experiencing a quiet revolution. As a result, software for data capture is simultaneously collecting data and understanding the information, which implies, in practice, that Machine Learning and Deep Learning tools can search for errors independent of a human-user, which will result in simplified and effective fault management.

The text in a grainy photograph can be read by today's Deep Learning and machine learning-driven OCR systems, such as the Google Vision API, even if it is thin, in a weird font, upside down, or partially obscured. This is made possible through probabilistic analyses of which letters are likely to occur where considering the context of the scene. While machine learning offers pioneering results in the extraction of information, extraction of receipt data, and freedom from templates, deep understanding helps to gain insights into the transformed data and algorithms to learn from continuous feedback generated by corrections to the extracted data to create better results over time.

Deep learning and machine learning innovation in OCR has helped overcome reading challenges for individuals with dyslexia, ADHD, and Irlen Syndrome while enabling visually impaired people by using high accuracy image-based PDFs with text-to-speech technology and deriving meaning from the converted phrases.

When it comes to the Arabic language, the accuracy rate for OCR is very low, making the technology effectively unusable on a wide scale. For People with Disabilities, namely people with visual disabilities, this means a low availability of accessible digital content in the Arabic language. Furthermore, it translates into that the means to create such material through OCR is not available as well.

The Mada Innovation Program has worked on a use case to develop an OCR with improved Arabic language support with key benefits like improved access to digital documents for Arabic speaking People with Disabilities and superior accuracy for Arabic OCR that can be used across multiple disciplines. This will also allow the conversion, management, and privacy of big data in Arabic.

Deep Learning and Machine Learning embedded OCR tools are sleeping giants on the broader topic of digital transformation. As Deep Learning and Machine Learning are widely embraced as disruptive new technology that has automated manual processes, its growth has led modern businesses to raise their expectations of what can be achieved by automation. Those using deep learning and machine learning integrated OCR engines to search for errors and meanings are beginning to outpace OCR engines that need to be controlled by human users.

This marks a revolution for users with functional limitations easing the usage at schools, the workplace, and home with various settings which will improve education and empower the user to take up challenging tasks, courses, and positions. While AI-based OCR tools may not be as desirable as other transformative technologies, they will predictably have a significant influence.