Nafath

Issue 15
November 2020

Machine Learning, Deep Learning and OCR Revitalizing Technology

Arabic Optical Character Recognition (OCR) Technology at Qatar National Library

Overview of Arabic OCR and Related Applications

Nafath 15

Content

Arabic Optical Character Recognition and Assistive Technology

Mada Center

Abstract -

As Optical Character Recognition (OCR) technology has gone through substantial improvements over the past decades, the Assistive Technology (AT) industry has utilized it as a crucial tool to serve as a foundation for developing breakthroughs in innovative AT solutions. AT for individuals with various kinds of disabilities has been developed through the use of OCR. These solutions have enabled Persons with Disabilities (PWDs) to be active members of society in several domains such as education, employment, and community.

Introduction

The advent of OCR based AT has had a transforming effect on PWDs in terms of areas like increased educational productivity and enhanced independence in performing daily tasks, leading to an improved quality of life. The ability of OCR technology to offer efficient and accurate conversion of paper and image documents to editable digitized formats has greatly influenced the potential of providing information in accessible formats suitable for use by PWDs. Several types of innovative AT solutions based on OCR technologies are available in the market today, some of which are as follows:

- **Technologies that help to overcome Learning Difficulties:** Individuals with learning difficulties like dyslexia, and Attention Deficit Hyperactivity Disorder (ADHD) often find it challenging to read printed materials as it becomes problematic to distinguish between characters and keep track of the flow of reading. Various AT solutions support the creation of digital documents from scanned documents and images through the use of OCR technology. These digitized documents can be processed automatically and converted into accessible materials tailored fit for the needs of the student with learning disabilities. Once converted to digital format, specialized AT (e.g., TextHelp, Clicker, Kurzweil, etc.) can support features like text-to-speech and text-tracking tools that help users visually track the words being read out to them by the software. The software features can also include additional useful options like creating a user-configurable library, dictionary, highlighting, adjustable word and sentence tracking colors, and customizable backgrounds.
- **Technologies that enable Individuals with Visual Impairment to read independently:** Individuals with visual impairments like low vision or blindness can often encounter challenges to read printed materials. AT solutions can automate the process of converting printed materials to accessible format using specialized software or hardware solutions. The first step in the conversion process relies on using OCR technologies to identify the characters in the reading

materials and convert the content into a digital format. Once converted, the digitized document can be used by AT solutions (e.g., Duxbury, EZ converter, etc.) to create reading materials in an accessible format like large print, text-to-speech, and braille. Without accurate OCR technology, the translation of scanned documents into digital format would be unreliable and thus, hinder the ability for those with visual impairments to read independently.

- **Wearable technologies that enable Individuals with Visual Impairment to identify key objects and improve their ability for independent living:** In recent years the concept of integrating accessibility features into smart glasses to improve the lives of individuals with visual impairment has been explored. The inclusion of a camera and computing chip in smart glasses allows them to be an ideal platform for learning OCR based AT solutions. Smart glass technologies like Envision, NuEyes, and OrCam have integrated OCR based features that enable the identification of print-based information like product expiry dates, restaurant menus, and printed bills. These glasses are also equipped with audio output which allows text-to-speech feedback of OCR identified information.
- **Technologies that enable Individuals with Physical Disabilities to access print materials:** Accessing conventional reading materials like books and newspapers can be challenging for Individuals with Physical Disabilities as it requires sufficient dexterity to perform tasks like flipping the document pages. In such cases, the preferred access method is to have the reading materials available in an electronic format (e.g., HTML, PDF, etc.). OCR based AT solutions like OpenBook allow the creation of documents in an electronic format from scanned printed documents and graphics-image based text.

The development of OCR technologies in other languages like Arabic has progressed considerably over the past decade. This has allowed the AT industry to create innovative OCR based solutions localized in the Arabic language as the improved OCR accuracy meant a more reliable outcome from the AT solution. Currently, there are a handful of OCR based solutions commercially available in the Arabic language, some of which have around 96% or more accuracy. Examples of OCR solutions that support the Arabic language include:

- **Sakhr OCR**

Sakhr OCR solution is capable of identifying complex fonts (including cursive writing), diacritics, position-dependent character shapes, overlapping, and non-standard fonts in the Arabic language. Sakhr OCR converts scans of both Arabic and Arabic-script based languages and can recognize Arabic text with an output accuracy of up to 99%.

- **ABBYY FineReader**

ABBYY FineReader is an OCR solution that includes features like digital conversion of Arabic scanned documents through applying intelligent document layouts, image enhancement, barcode recognition, and command line integration. ABBYY FineReader can directly convert the contents of printed documents into editable Microsoft Word, Excel, or PDF format.

- **ReadIRIS**

ReadIRIS offers an accurate Arabic OCR recognition rate through its OCR software. The OCR engine supports the recovery of text from printed materials to various file format (e.g., Word, Excel, PowerPoint, or PDF). ReadIRIS supports the digitization of Arabic paper documents and thus, paves the way for AT solutions to create Arabic content in an accessible format.

Mada Center has continuously been invested in supporting the development of AT based on Arabic OCR technologies. This continues to be achieved primarily by supporting innovators and entrepreneurs through the Mada Innovation Program. A major contribution of Mada Center towards its commitment of supporting the evolution of Arabic OCR is the development of the **"Arabic Money Reader App"**, which recognizes Qatari Riyal currency notes using the mobile phone camera. Mada Center has also been committed to supporting the digitization of Arabic language reading materials in accessible format available worldwide. This objective is achieved by collaborating with international partners like **Bookshare**, which hosts one of the largest platforms of accessible reading materials for individuals with print disabilities.

State of the Art in Arabic OCR: Qatari Research Efforts

Mada Center

Introduction

Since the mid-1940s, there has been extensive research and publications on character recognition. With most of the published work being on Latin characters, and Japanese and Chinese characters emerging in the mid-1960s. Despite almost a billion people worldwide using Arabic characters for writing (Arabic, Persian, and Urdu), Arabic character recognition research, starting in the 1970s, is sparse.

This may be attributed to:

- Inadequate journals, books, conferences, funding, and interaction between researchers.
- The lack of utilities like Arabic text databases, dictionaries, programming tools, and supporting staff.
- Delayed onset of Arabic text recognition.
- The techniques developed for other writings cannot be successfully applied to Arabic writing due to the unique attributes of Arabic script.

Arabic OCR challenges

Arabic is written from right to left, which presents many challenges to the OCR developer, which include (Al-Badr 1995; Attia 2004):

The connectivity challenges

Arabic text can only be scripted cursively, i.e., graphemes are connected and only interrupted at limited characters or at the end of the word. This necessitates that any Arabic OCR system undergoes a traditional grapheme recognition task and a more rigorous grapheme segmentation (see Figure 1). To complicate things, both tasks are mutually dependent; therefore, must be done simultaneously.

Figure (1): Grapheme segmentation process illustrated by manually inserting vertical lines at the appropriate grapheme connection points

The dotting challenge

Dotting is extensively used to differentiate characters sharing similar graphemes. Figure (2) shows small differences between members of the same set. Whether the dots are eliminated before the recognition process, or recognition features are extracted from the dotted script, dotting is an area of confusion – therefore, recognition errors – in Arabic font-written OCR systems, especially when using devices such as photocopiers.

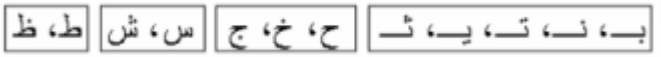

Figure (2): Examples sets of dotting-differentiated graphemes

- The multiple grapheme cases challenge

Due to the connectivity in Arabic orthography, the same grapheme representing the same character can have multiple variants according to its position within the Arabic word segment (Starting, Middle, Ending, Separate) as exemplified by the four variants of the Arabic character "ع" shown in bold in Figure (3).

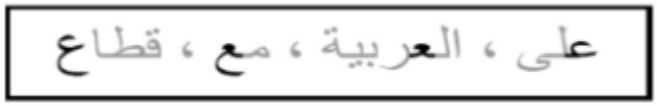

Figure (3): Grapheme "ع" in its 4 positions; Starting, Middle, Ending & Separate

The ligatures challenge

Certain compounds of characters at particular positions of word segments are represented by single atomic graphemes called ligatures; found to some extent in most Arabic fonts. Traditional Arabic font contains around 220 graphemes and Simplified Arabic contains around 151 graphemes. Compared to English with 40 or 50 graphemes. A broader grapheme set means higher ambiguity for the same recognition methodology; hence, more confusion. Figure (4) illustrates some ligatures in Traditional Arabic.

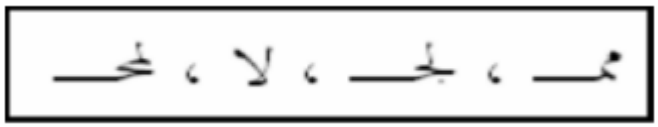

Figure (4): Some ligatures in the Traditional Arabic font

The overlapping challenge

Characters in a word may overlap vertically even without touching, as shown in Figure (5).

Figure (5): Some overlapped Characters in Demashq Arabic font

Size variation challenge

Different Arabic graphemes do not have a fixed height or a fixed width. Moreover, neither the different nominal sizes of the same font scale linearly with their actual line height nor the different fonts with the same nominal size have a fixed line height.

- The diacritics challenge

Arabic diacritics are used only when they help resolve linguistic ambiguity of the text. The problem of diacritics with font written Arabic OCR is that their direction of flow is vertical while the main writing direction of the body Arabic text is horizontal from right to left. (See Figure (6)) Similar dots and diacritics are a source of confusion of font written OCR systems; due to their relatively larger size they are usually pre-processed.

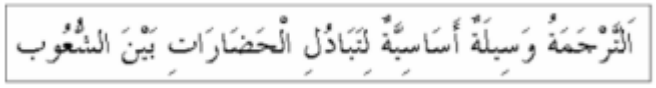

Figure (6): Arabic text with diacritics

Qatari Research Efforts

Arabic Language Technologies Group at Qatar Computing Research Institute (QCRI) is leading research on OCR in Qatar. They are dedicated to promoting the Arabic language by conducting world-class research in Arabic language technologies. Ensuring that the Arabic language flourishes in the digital world is a primary focus area. Some current research projects address the challenges related to the lack of content and extracting that content.
QCRI strives to become the regional and global leader in Arabic language technologies – in the areas of search, information retrieval and analysis, multilingual language processing, advanced machine translation, and leading efforts to increase and enrich Arabic language content online.

Moreover, QCRI's initiatives also examine challenges in retrieving content, making it accessible, and enabling information flow across language barriers. In this regard, development is underway to process Arabic in the search domain such as the use of morphological word analysis, named entity recognition, and data learning technology to detect relevant content for more elaborate analysis. In addition, developing proofing

tools such as typographical checks and language identification for local Arabic dialects and Arabic written using Latin characters.

A major effort at QCRI goes into improving machine translation. Combining an Arabic "Speech-to-Text" engine that permits instantaneous transcription of videos with a machine translation system allows access to broadcast news and news distributed over the web. Future research will concentrate on applications such as lecture translation.

QCRI has established projects related to e-education and making non-native language material accessible. The development of an Arabic language supported e-reader and assistive language tutor are examples that will directly impact society and learning.

Some of the projects run by the Arabic Language Technologies Group at QCRI include:

QATIP – An Optical Character Recognition System for Arabic Heritage Collections in Libraries

The Qatar Computing Research Institute team worked on an end-user oriented QATIP system for OCR in such documents. The recognition is based on the Kaldi toolkit and sophisticated text image normalization. The QATIP interface for libraries consists of a graphical user interface for adding and monitoring jobs and a web API for automated access. It also uses a novel approach for language modeling and ligature modeling for continuous Arabic OCR. The QATIP system was tested on an early print and a historical manuscript and report substantial improvements – e.g., 12.6% character error rate with QATIP compared to 51.8% with the best OCR product (Stahlberg 2015, 2016).

PrepOCRessor

The QCRI Preprocessing Tool for Arabic OCR was developed for preprocessing document images for optical character recognition. A set of image processing operations is chained such that the output of each operation serves as an input to the next one. The tool supports batch processing for high parallelism and scalability. PrepOCRessor is intended to be used in combination with the recognition toolkit Kaldi and supports file formats for feature sets (.ark,t) and forced-alignments (.al) for seamless integration. Though the focus is on Arabic script, the tool has been successfully used for other writing systems, e.g., Latin in the ICDAR2015 Competition HTRtS on historic documents.

References

Stahlberg, F., & Vogel, S. (2015, September). The qcri recognition system for handwritten arabic. In *International Conference on Image Analysis and Processing* (pp. 276-286). Springer, Cham.

Stahlberg, F., & Vogel, S. (2016, April). QATIP--An Optical Character Recognition System for Arabic Heritage Collections in Libraries. In 2016 12th IAPR Workshop on Document Analysis Systems (DAS) (pp. 168-173). IEEE.

Al-Badr, B., & Mahmoud, S. A. (1995). Survey and bibliography of Arabic optical text recognition. Signal processing, 41(1), 49-77.

Overview of Arabic OCR and Related Applications

Mada Center

Optical Character Recognition (OCR) is a generic term used to characterize technologies that recognize text within scanned documents, and photos, to help convert them into a digital format. OCR technology is used to convert virtually any kind of images containing written text (typed, handwritten, or printed) into machine-readable text data. OCR has become a key area of interest over the past two decades with respect to implementing projects related to digitize historic documents (e.g., newspapers, manuscripts, constitutional bills, letters, etc.). The importance of OCR technologies has become even more widespread with the advent of the internet which serves as a resource of multilingual information based on digital textual data.

While OCR technology has undergone several improvements over time and achieved close to a hundred percent accuracy in languages based on Latin scripts (e.g., English), there have been major challenges in enhancing OCR accuracy for languages based on right-to-left reading stylized scripts (e.g., Arabic, Persian, Urdu, etc.). Arabic is the first language for more than 400 million people worldwide and Arabic speaking readers represents a major proportion of internet users who are potentially interested in accessing Arabic digital resource. Hence, the importance of optimizing Arabic OCR technology is significantly vital for improved information and knowledge sharing within society.

The fundamental challenges involving Arabic OCR is the fact that recognition accuracy is harder to achieve primarily due to the following characteristics of the Arabic script set:

- **Character Position:** An Arabic character may have one to four unique shapes depending on its position within a word (i.e., isolated, initial, middle, and end). The OCR solution must be able to effectively identify the concerned Arabic character irrespective of its position within a word.
- **Dot and No-Dot Character:** Certain Arabic characters may have the presence of dots above or under them which can impact the outcome of the final character or word. There may be one to three dots used with the character to determine the final word.
- **Dot Character Baseline:** The presence of a dot within a character is in relation to a baseline as the dot used with the Arabic character may be located above or below the baseline (where applicable). The baseline is significant in developing Arabic OCR systems as it helps to classify Arabic characters into two classes: dot character above the baseline and dot character below the baseline.
- **Zigzag-Shaped Character:** Another distinguishing characteristic of Arabic script is the presence of Hamza, a zigzag-shaped mark (ء) with some Arabic characters which can pose challenges for the OCR systems to recognize the character or word accurately.
- **Loop-Shaped Character:** Several Arabic characters have a loop shape, such as Saad (ص), Dhad (ض), Fa ((ف, Meem (م), and Qaf (ق). An obstacle

for Arabic OCR is to be able to accurately recognize Arabic characters that contain a loop shape.

- **Diacritics:** Some Arabic text may be written with diacritical marks accompanying each character which makes it challenging for the OCR engine to identify the character effectively because this affects the graphical analysis of the image.

Over the past decades, researchers and scientists have worked on developing various databases of Arabic handwritten words to serve as a reference for OCR developers to create solutions for identifying textual shapes, characters, and reconciliating them into a digital text format. In 2002, a database on Arabic handwritten words (IFN/ENIT-database) was made available to the community. In September 2006, a summit on Arabic and Chinese Handwriting Recognition was held at College Park, MD in the US where experts from both research fields presented their actual work. From that time, intensive research on Arabic script recognition started and has resulted in a big step forward today.

The most common application of OCR technology is converting printed paper documents to machine-readable digital text format. Some other areas of application for OCR technology are as follows (but not limited to):

- Data Entry Automation
- Indexing Documents for Search Engines
- Automatic Vehicle Plate Number Recognition
- Voucher Code Scanning
- Office filing system
- Self-service stores / e-Kiosk
- Digitizing handwritten documents, books, and manuscripts
- Assistive Technology

OCR technology has a key role to play in the development of Assistive Technologies to help improve the lives of People with Disabilities (PWDs). As such PWDs, primarily Visually Impaired individuals, cannot use their Assistive Technology to read digital content without the accurate utilization of OCR technology. With improved Arabic OCR, PWDs can enjoy greater access to digital documents and improve their quality of life across education, employment, and other aspects of daily living. Additionally, the availability of digital text is vital to make printed information accessible to PWDs because this enables the creation of information in other accessible formats such as audio, large print, and Braille. Digital text is especially helpful for struggling readers, including those who have learning difficulties such as dyslexia.

Mada Center has a significant role to propel the improvement of Arabic OCR and the development of innovative accessible OCR based solutions. This is done by supporting relevant innovators and entrepreneurs through the Mada Innovation Program to successfully develop their Assistive Technology solutions and prepare them to be market-ready for Qatar and the Arabic speaking region. Mada Center strives to increase the number of ICT Digital Accessibility solutions to adequately serve the growing needs of PWDs in Qatar and the region.

Examples of Optical Character Recognition Tools

Oussama El Ghoul
Mada Center

Optical Character Recognition (OCR) refers to computer processes for converting images of printed, typed, or handwritten texts into text files. A computer requires OCR software to perform this task. This allows retrieving the text in the image and to save it in a file that can be used in a word processor for enrichment and stored in a database or on another medium that can be used by a computer system. Today, there many OCR engines that are used including Google Drive OCR, Tesseract, Transym, OmniPage, etc. Many of them are paid, however, some are accessible for free.

Recognizing Arabic text is a popular research topic, a significant amount of research efforts is being invested to increase the accuracy rate of Arabic OCR by using different approaches and technologies. In 2002, a system to recognize Arabic text using a neural network was developed using a set of moment invariants descriptors. An artificial neural network (ANN) is used for classification [1]. The study has shown a high accuracy rate of 90% [2]. Another research project made in 2017 used a database of 34,000 characters, 70% are used for training the machine learning, 15% for the testing phase, and 15% for validation. The project has achieved a 98.27% recognition rate [3]. In 2018, a project aiming to recognize Arabic handwriting used a dataset of greater than 43,000 handwritten Arabic phrases, 30,000 used for training and 13,000 used for the testing stage. The recognition result showed a 99% rate of accuracy [4].

A number of tools and services have emerged in the market as a result of advances in such research. The quality, accuracy, and precision of OCR tools have become more effective and improved over the years. Today, from simple to complex, there is a wide range of OCR solutions available for use. Some of these tools may need programming skills to make them work while others are ready to use off-the-shelf solutions. Depending on their features and accuracy, the solution costs may vary, while some OCR tools are available to use for free too. Details of the most known OCR resources in the market are provided in the table below:

Name	Founded Year	License	Online	Programming Language	SDK	Arabic Language
QATIP	2016	Free	Yes	Unknown		Arabic
Google Cloud Vision	2016	Proprietary	Yes	Unknown	Yes	Arabic; Modern Standard + More than 200
Tesseract	1985	Apache	No	C++, C	Yes	Arabic + More than 100
ABBYY FineReader	1989	Proprietary	Yes	C/C++	Yes	Arabic + 192
Asprise OCR SDK	1998	Proprietary	Yes	Java, C#,VB.NET, C/C++/Delphi	Yes	Arabic not supported + 20
AnyDoc Software	1989	Proprietary	No	VBScript		Arabic not supported
CuneiForm	1996	BSD variant	No	C/C++	Yes	Arabic not supported
Dynamsoft OCR SDK	2003	Proprietary	Yes	C/C++	Yes	Arabic + 40
OmniPage	1970s	Proprietary	Yes	C/C++, C#[15]	Yes	Arabic + 125
Ocrad	2003	GPL	Yes	C++	Yes	Latin alphabet
SmartScore	1991	Proprietary	No	–		Music
Microsoft Office Document Imaging	–	Proprietary	No	–		Arabic
Puma.NET	2006	BSD	No	C#	Yes	Arabic not supported + 28
ReadSoft	–	Proprietary	No	–		Arabic not supported
OCRFeeder	2009	GPL	No	Python		Arabic not supported
OCRopus	2007	Apache	No	Python		All languages using Latin script (other languages can be trained)

References

[1] Muna Ahmed Awel, Ali Imam Abidi, Review on optical character recognition, International Research Journal of Engineering and Technology (IRJET), p-ISSN: 2395-0072, Volume: 06 Issue: 06 | June 2019

[2] M. M. Altuwaijri and M. A. Bayoumi, “Arabic text recognition using neural networks,” pp. 415–418, 2002.

[3] N. Lamghari, M. E. H. Charaf, and S. Raghay, “Hybrid Feature Vector for the Recognition of Arabic Handwritten Characters Using Feed-Forward Neural Network,” Arab. J. Sci. Eng., vol. 43, no. 12, pp. 7031– 7039, 2018.

[4] N. A. Jebril, H. R. Al-Zoubi, and Q. Abu Al-Haija, “Recognition of Handwritten Arabic Characters using Histograms of Oriented Gradient (HOG),” Pattern Recognit. Image Anal., vol. 28, no. 2, pp. 321–345, 2018.

The optical recognition technology in assistive technology devices

Mada Center

Optical Character Recognition (OCR) plays a vital role in converting printed materials into digital text files. The use of OCR technology (OCR) has transformed productivity, increased engagement and motivation, and, most importantly, accelerated learning. The technology has also played a crucial role in improving the lifestyle of individuals requiring support and people with disabilities. OCR technology provides effective, efficient, and accurate document processing that converts paper or image original texts to editable formats and then, translates it into multiple languages.

Digital text is one of the numerous formats that make printed information accessible to a broader audience; other forms include audio, large print, and Braille. Digital text is incredibly accommodating for struggling readers, including those who experience learning differences such as dyslexia. The digital format makes it feasible for readers to see words on a screen and hear them read aloud at the same time, which promotes more ways to interlock with the information. It can also help children amplify independent reading skills.

Standalone OCR Devices:

The latest OCR technology comes integrated with various hardware like a smart reading pen, handheld magnifier devices, standalone CCTV devices, and braille devices. The basic functionality is the same but the accuracy varies with the device. The device camera scans the document and the OCR software then converts the images into identified characters and words and creates temporary files, including the text's characters and page design. The identification process considers the logical form of the language. Solutions like Ebot Pro or Compact 10 HD portray all the key components of a flawless OCR device while also having other features. Such a system can recognize that a word spelled incorrectly at the beginning of a sentence is an error and can fix the error.

OCR systems use a dictionary and implement spell-checking techniques comparable to those found in many other word processors. The synthesizer in the OCR system then speaks the recognized text and the knowledge is stored in an electronic format. In certain OCRs, these temporary files can be translated to forms that can be retrieved by commonly used computer applications such as word processors, spreadsheets, and databases. A visually impaired or blind individual can access the scanned text by using adaptive technology devices that magnify the computer screen or provide text to speech or braille output.

Future – Machine Learning and Deep Learning:

The future of OCR technology is rebooting using artificial intelligence-based machine learning and deep learning technologies; these new technologies are not limited by the rules-based character matching of existing OCR software.

The latest updates in OCR technology will have a neural network that mirrors human brain function to confirm that the algorithms don't have to depend on historical patterns to determine accuracy and the benefit will be that it can derive meaning from the recognized text and will help to automate manually intensive tasks such as document classification, data extraction, and storage. With this adaption, a disabled user will also benefit from enriched data, deeper analysis, and thoughtful recommendations.

Arabic Optical Character Recognition (OCR) Technology at Qatar National Library

Hany A. Elsawy Abdellatif

Optical Character Recognition (OCR) is the practice of extracting text from images. The process itself is becoming popular in terms of usage and research, as it spans multiple areas of science, including image processing, machine learning, information retrieval, and artificial intelligence.

In layman's terms, it is the only way to copy, use, and index the text from a scanned image. The benefits vary from a simple copy/paste process, citation, search within, text annotation, and tagging. Moreover, it perfectly meshes with the modern algorithms of text mining, morphological search, text automatic translation, text summarization, linked data, and indexing tools.

Using OCR images is the ultimate value-add to scanned documents—it brings the documents to life and allows users to discover every bit of information stored within.

At Qatar National Library, multiple techniques and algorithms to OCR text have been developed; these methods include both human operators and automatic SDKs (Software Development Kits) and APIs (Application Programming Interfaces). Adittionally, QNL built an accurate yet scalable system that will efficiently streamline the operation, as it harmonizes the roles and responsibilities between humans and machines to reach the maximum quality of the extracted text.

While we use OCR for a wide array of languages, we are most proud of our achievements with regards to the Arabic text. Since the start of machine learning algorithms, and even with the most modern OCR executables, Arabic text remains a formidable challenge. The multiple shapes and sizes of the Arabic font, in addition to its diacritics, dot usage, cursive characters, and the changing of character shapes based on their location inside a word were all factors that reduced the quality of the OCRed text.

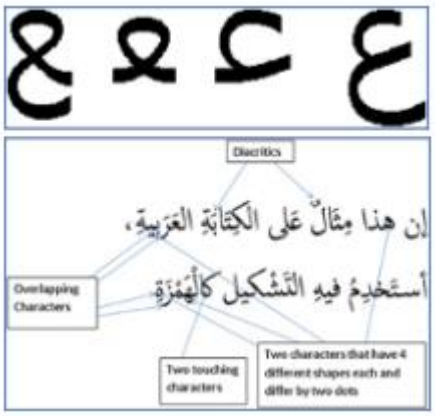

FIGURE 01 Arabic OCR challenges

With the proper tools and algorithms, QNL built universal libraries that cover 99% of the printed Arabic text based on shape, quality, and size and we have engineered a clear workflow to enhance the quality of the images at the start by raising the DPI, smoothing the edges, refining smudged printing and removing the noise. With this image-enhancement process and the trained machine learning libraries, our OCR accuracy achieves 99 percent character level accuracy in Arabic.

FIGURE 02 shape classification

This allowed QNL to index the output text using robust yet sophisticated Arabic text lexical analyzers and offered that to our patrons with just a single click. Our patrons need only to access the Library's Digital Repository and enjoy the "search within" feature, which is expected to immensely help improve the quality of research in Arabic studies, such as art, history, science, and philosophy, to name just a few.

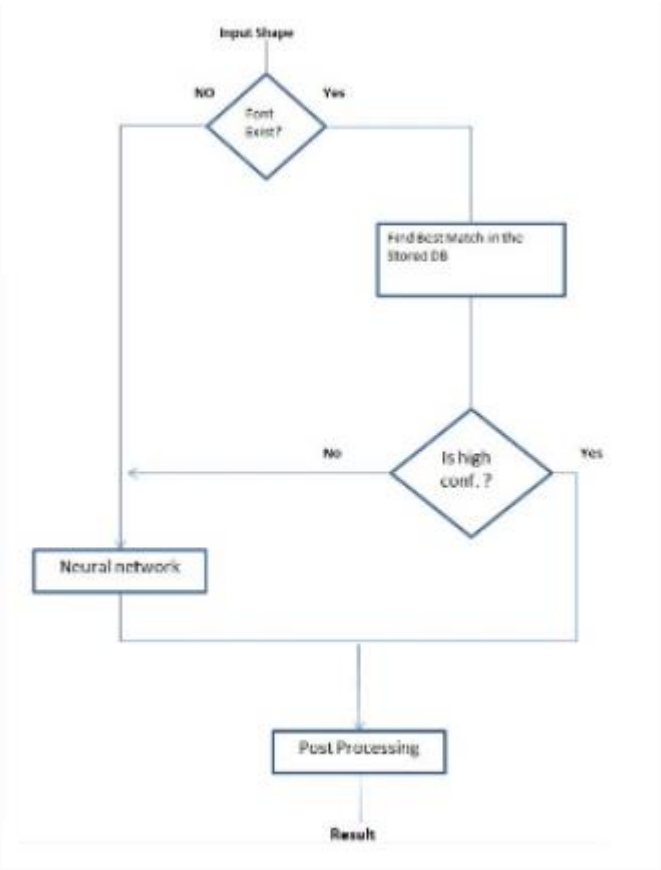

FIGURE 03 Arabic Text Recognition flow

The Library's state-of-the-art digitization facility makes Arabic content from its Heritage Library and other institutions available online, increasing the availability of Arabic content worldwide.

The Library harnesses the expertise of a fully trained international team, and laboratories equipped with cutting-edge technology, to undertake various processes of digital preservation. The Library offers the services of bulk digitization, large-format scanning and image stitching, on-site digitization, E-Pub creation, Optical Character Recognition (OCR), 3D Photography, audiovisual digitization, and long-term preservation.

In addition to ongoing efforts to digitize the Library's collections of rare books, manuscripts, maps, and photographs, the Library's Digitization Center is working on digitization and Arabic OCR projects with other heritage collections in Qatar and international institutions, including:

- New York University (NYU) project: This joint project applies optical character recognition to more than 8,000 Arabic books in NYU library collections, which will also be available on Qatar National Library's online platforms.

- Doha Historical Dictionary of Arabic Language: The Library contributes to the field of optical recognition of Arabic characters, which will assist research on the etymology and meaning of Arabic words.
- Museum of Islamic Art: A Memorandum of Understanding outlined possible collaborations, including a project to digitize **164** of the rarest books and manuscripts in the museum as well as its and library collections, including the Latin OCR.
- Al Shaqab Horse Collection: the Library digitized over 50,000 images from Al Shaqab's Horse Collection.
- Ottoman Archive: 1,600 digital images of heritage documents related to the Gulf region from the Ottoman Archive have been processed, to be made available on the Library's online platforms.
- Qatar Traditional Architecture Photographic Collection: The Library digitized a collection of 1793 photographs from a 1985 French archeological expedition to Qatar that produced a comprehensive record of traditional 19th-century architecture.

The Digitization Center follows international best practices and guidelines, including the Federal Agencies Digitization Guidelines Initiative (FADGI), Metamorfoze Preservation Imaging Guidelines, ISO- 19264, and the IFLA guidelines for digitization projects. This enabled the Library to digitize **10,277,367** pages from various collections including **4,957,546** Arabic pages from Qatar National Library's Heritage Collection, and **2,782,016** pages from the Online Arabic Collection of New York University.

Libraries play an important role in preserving heritage for future generations, and digitization and its sophisticated processes and operations go a long way in ensuring this is done. Furthermore, Qatar National Library is committed to the preservation of heritage not only of the region but of the Islamic world as a whole. We have come a long way in building a reliable process for digitization and OCR Arabic content for the benefit of spreading rich Arabic knowledge and heritage; we are committed to working harder to fulfill that goal.

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 resultsy. 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.

Smart Apps for PWDs using OCR

Shahbaz Ahmed
Mada Center

Optical Character Recognition (OCR) is a technology that allows you to transform different types of documents into editable and searchable content, recognize and transform text or print documents, such as scanned paper documents, PDF files, or digital camera images, and convert them into digital text documents. The OCR program extracts and converts the characters into machine-readable data.

With the fundamental technology being Optical Character Recognition (OCR) technology, there are many creative assisted learning technologies available.

Using OCR technology has transformed efficiency in the classroom, increased student participation and motivation, and, most significantly, accelerated learning. Also, this innovation has played a crucial role in improving the lifestyle of persons with disabilities.

To convert a physical copy of a document into an interactive (or soft) format, OCR technology can be used. For instance, if you scan a multi-page document and want to convert it into a digital image format such as a TIFF file, you can load the document into an OCR program that identifies the text and converts it into an editable text file.

When creating, editing, and reusing different documents, an advanced, powerful OCR program helps you to save a lot of time and effort.

Groundbreaking examples of OCR technology-based solutions were developed to support people with disabilities. These solutions helped reduce reading difficulties for people with dyslexia, allowed visually impaired people to read mail and fill out forms independently, made learning tools and lessons accessible to learners with learning difficulties and reading difficulties through text-to-speech and gave visually impaired people access to text-to-speech image-based PDFs.

examples of these apps include:

1. ClaroPDF

As PDF files are basically images of documents, they pose a problem with the basic technologies of text-to-speech. ClaroPDF is an app that can recognize and interpret text in pictures. It retains the formatting of the original text, unlike most OCR apps. It includes synchronized highlighting, text-to-speech, annotation tools, audio and video notes addability, and Dropbox integration.

2. SnapType Pro

For students with dyslexia, workbooks and photocopied worksheets may be troublesome. For fill-in-the-blank and matching exercises, the formatting is always lost during regular OCR, an issue that makes it hard to use AT to insert responses. By giving users the ability to overlay text boxes on worksheet images, SnapType solves the problem. A keyboard can then be used by students to position their answers in the correct spaces.

3. Snapverter

For Google ChromeTM, Snapverter is an easy-to-use add-on for Read&Write that converts classroom papers and files into readable PDF documents for easy sharing through Google Drive and aloud reading.

4. Prizmo

With advanced editing, OCR, and text-to-speech, scan any image with text into PDF. In order to enhance their learning environment or reading skills, Prizmo may be used by low vision students or students struggling with dyslexia in classrooms.

5. Voice Dream Reader

Voice Dream Reader is a flexible iOS and Android application for reading. This app is beneficial for anyone, including people with various disabilities, with Dyslexia friendly font, text and audio synchronization, adjustable font size and color variations, as well as complete voiceover support.

6. Aipoly Vision

Aipoly Vision utilizes artificial intelligence and OCR to help low-vision people better understand their surroundings. Users point the app at an object and simply press a recognition button. Once users encounter a sign or document, they can switch to the “read text” button to read it out loud. It can also read text in multiple languages.

7. Digit Eyes

Digit Eyes was created for the visually impaired shopper. This software audibly reads the manufacturer barcode and the product’s name. For household products, users may also record their own labels.

Making Social Media Accessible for All – Twitter

Oumer Seid
Mada Center

Making social media Accessible

This article will showcase some of the ways Twitter has been designed to be accessible to people with disabilities, thereby enabling everyone to share and access content in a manner that makes the most sense to them.

The article is part of a Nafath series that focuses on the different ways major social media platforms implement the fundamentals of accessibility and universal design to their websites and apps. At a time when the use of these platforms has increased, replacing traditional media outlets and, at times, even workplace collaboration tools, it is important to ensure that there are ample resources to enable people with disabilities to access them and use them at par with the rest of the world.

About Twitter

Twitter is an American microblogging and social networking platform that users post messages known as tweets on and interact with. Registered users can upload, like, and retweet tweets, but unregistered users can only read them. Via a browser, Short Message Service (SMS), or its mobile device application program (‘app’), users access Twitter. Twitter has more than 25 offices worldwide and is headquartered in San Francisco, California. Tweets were initially limited to 140 characters but were doubled to 280 characters in November 2017. For most accounts, audio and video tweets remain restricted to 140 seconds.

In March 2006, Twitter was founded and was launched in July of that year. By 2012, 340 million tweets a day were posted by over 100 million people, and there was an average of 1.6 billion daily search requests. It was one of the ten most-visited websites in 2013 and was described as “The Internet SMS.” Twitter had more than 321 million active monthly users as of 2018.

Twitter Accessibility

With over 285 million users with visual impairments around the globe, an accessible Twitter will undoubtedly have a profound impact on this community.

Twitter mostly has text-based tweets, which are accessible by default and readable by screen readers, but often people use photographs and videos to tweet, and before publishing, they must be checked for accessibility. When you tweet images using the iOS or Android Twitter app or on twitter.com, you have the option to write image descriptions so that more individuals, including those who are blind or low-

vision, can access the content. Good image representations are concise and informative, helping people understand what is happening in a photo.

Twitter is built in a way that allows users with screen readers to quickly access the app. This means that a blind person can access the entirety of the app by audible signals by triggering the voiceover function on iOS or Apple computers. For Windows (JAWS or NVDA) and Android (Talkback) platforms, this function is also available.

Here are some tips for making Twitter as accessible as possible:

Image Descriptions

There is an option to create a description while tweeting using pictures.

This is done by going to **Settings** > **Display and Sound** > **Accessibility** and turning on **Compose Image Descriptions**.

The next time a picture is attached to one of the tweets, the Add Description button at the bottom of the tweet will appear. Clicking it will take the user to an Image Description screen, where they will be able to add a picture description of up to 420 characters, often known as alternative text or alt text for visually impaired users.

However, the description will not appear as part of the main update. People who are visually impaired will have access to the description via their assistive technology (e.g., screen readers and braille displays).

Color Contrast

Twitter enables users to improve the color contrast between the text and the background colors of the platform, making it easier to read text.

The web app’s accessibility setting now has a new toggle button for “increase color contrast”. When switched on, the button activates the high contrast colors for the User Interface UI components. The high contrast mode makes using the Twitter Web app easier for people with visual impairments.

Users can activate the feature from **Settings** > **Accessibility** > **Vision** > **Increase color contrast**.

Using indicators in tweets

Indicators must be used before hyperlinks to ensure that visually impaired users know what to expect. “[PIC] for images, [VIDEO] for videos, and [AUDIO] for audio.”

Practicing social tagging etiquette.

In their messages, if a user uses @’s or #’s, he must always place these at the end of a tweet. This is usually good practice on social media in general, but it is also useful for those who use screen readers. Additionally, it is better to capitalize the first letter of each word (this is called camel case) while using hashtags that are compound words, like “#AssistiveTechnology”.

Using text-speak.

Twitter restricts the number of characters, but that does not mean that common abbreviations such as U, tho, K, etc. are clever to use. These would sound weird to read by a screen reader, so it is advisable to use full words. Similarly, it is best to either avoid acronyms where possible or type them out following the abbreviation.

Conclusion

Major social media platforms implement the fundamentals of accessibility and universal design to their websites and apps in different ways. At a time when the use of these platforms is increasingly taking the place of traditional media outlets and workplace collaboration tools, it is important to ensure that there are plenty resources to enable the persons with disabilities to access them and use them equally with the other users in the world. Twitter is built in a way that allows users with screen readers to quickly access the app, it mostly has text-based tweets, which are accessible by default and readable by screen readers, but often people use photographs and videos to tweet, and before publishing, they must be checked for accessibility and add alt text and closed captioning when needed.