A Comprehensive Study of Optical Character Recognition

Ravi Raj, Andrzej Kos
Faculty of Computer Science, Electronics, and Telecommunications
AGH University of Science and Technology
Al. Adama Mickiewicza 30, 30-059 Kraków, Poland
raj@agh.edu.pl, kos@agh.edu.pl

***Abstract*—In recent decades recognition of characters has become a most important research topic for computer vision researchers or scientists. One of the major techniques for character recognition is optical character recognition (OCR) which plays a vital role in recent years in the development of various methodologies for the recognition of characters of several languages alphabets. Currently, OCR technology is utilized by most of the applications of scanning documents and makes them readable for the users such as google translate, which has been developed for translating the language from one language to the other language. But the rate of accuracy and timing to perform the specified task is still a problem. This paper presents a brief description of the OCR technology, the timeline of its development, some major applications of this technology, and its future perspective in our daily life. Moreover, this article provides an overview of this fascinating research topic for the early-stage researchers of computer vision.**

***Keywords*—OCR; Artificial Intelligence; Computer Vision; Pattern Recognition; Character Recognition.**

I. Introduction

Optical character recognition is the mechanical or electronics transformation of images of handwritten, printed, and typed text into the text of machine-encoded type whether from a photocopy of the document, a scene-photo, a scanned document, or from a subtitle text which has been written on any image. It is mostly used as a data entry operator from typed, printed, and handwritten paper data records such as bank statements, business cards, mail, computerized receipts, static-data printouts, invoices, passport documents, or any other suitable documentation which is a common technique of digitizing printed, handwritten, or typed texts so that they might be electronically stored more compactly, searched, displayed online, edited, and utilized in the processes of machine such as text to speech extraction, machine translation, key data, cognitive computing, and text mining [1]. OCR technology is an area of research in artificial intelligence, machine learning, pattern recognition, and computer vision.

OCR technology is a well-known technique that is utilized to convert the words or letters written by hand or typed into a digital format [2]. This is an automated task performance algorithm that is utilized by various institutions for the recognition of text characters from the images containing text. Fig. 1 illustrates the working functions of the OCR technique to perform their tasks.

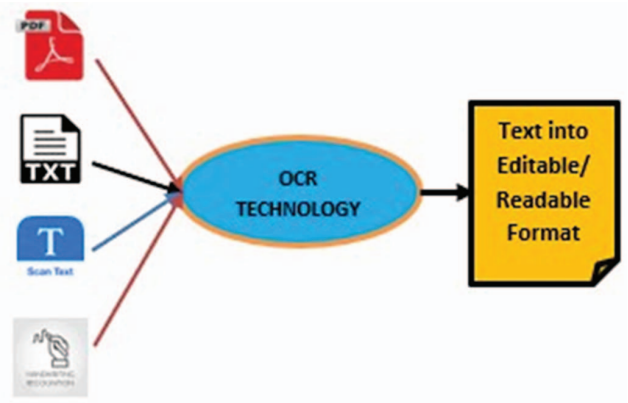

Fig.1. The working function of the OCR- technique

II. Historical Background

Working on OCR technology is started long back in the 18th century, but the major research work is started in the 19th century. OCR finds its origins back in telegraphy. During the first world war, scientist E. Goldberg discovered a machine that can read text characters and transform them into the code of telegraph. In the 1920s, physicist E. Goldberg went a further step and generated the first system of electronic document retrieval. IBM later acquired the United States patent authority for its "Statical Machine" [3]. Thus, this technology is playing an important role in character recognition for the last 100years. Table. 1 demonstrates the timeline of historical developments of OCR technology. This table is illustrating some major and important time-period of historical developments of OCR technology from 1870 to 2016.

III. Literature Survey

In the last century from 1920 to 2020, from the time when firstly OCR tools were discovered; there are various research has been carried out. J. Leimer (1962) [5] discusses and describes the development of an experimental OCR system, the recognition of the character part of the IBM 1418 OCR. The state of the art of OCR technology has been discussed in this article, which mainly focuses to identify the quality of the printed text, rate of performance of the system, discrimination ability to identify different symbols, and sensing the shape of symbol to get higher accuracy of recognition. P. M. Hall (1968) [6] discusses the methodology of one manufacturer towards the

Copyright © 2022 by Department of Microelectronics & Computer Science, Lodz University of Technology

Authorized licensed use limited to: University Of Minnesota Duluth. Downloaded on September 09,2025 at 20:52:55 UTC from IEEE Xplore. Restrictions apply.

TABLE I.
SOME MAJOR HISTORICAL DEVELOPMENTS [4]

Time Period	Major Developments
1870-1931	The first ideas of OCR are conceived. Tauschek's reading machine and Fournier d'Albe's Optophone are discovered as devices to assist blind people.
1931-1954	First tools of OCR are discovered and implied in industry, capable to calculate Morse code and understand text out rowdy. The first company was created to sell OCR tools; the company is known as Intelligent Machines Research Corporation.
1954-1974	The first OCR inbuilt portable device Optacon is developed. These devices are utilized to digitize postal addresses and Reader's digest coupons. For scanning, some special typefaces were designed.
1974-2000	Scanners are widely used for reading passports and price tags. Some companies such as ABBYY, Kurzweil computer products Inc., and Caere Corporation are created. The first OCR software of Omni-font is developed, able to understand any text documents.
2000-2016	OCR software has become online available for free, along with products like Google Drive, Web-OCR, and Adobe Acrobat.

issue of developing the machines of character recognition for using them for general purpose. This article mostly focuses on the discussion to enhance the accuracy and speed of the OCR tool. M. D. Freedman (1974) [7] illustrates the OCR systems and their corresponding rate of performance. The procedure of operation, scanning methodology, quantization, recognition logic, context correction, and feature extraction has been discussed widely in this article. P. Comelli et al. (1995) [8] present a system for car license plates recognition. This research has the ambition to develop a system to automatically read car number plates written in Italian crossing through a toll plaza. A-frame grabber card and a CCTV camera have been utilized to get vehicle rear-view image. The rate of recognition is approximate 91% from the analysis of three thousand images gathered by the camera in this experiment. A. Zramdini et al. (1998) [9] present an approach based on features of global typography is presented to the extensively neglected issue of recognition of font. This experiment is done with 280 fonts of the set. The rate of recognition is nearby 97 percent. These experiments have given robustness to text content and language of the document and its reactivity to the length of text. N. Arica et al. (2002) [10] present a novel analytical scheme, which utilizes a series of segmentation of image and algorithms of recognition, is demonstrated for the problem of recognition for the offline characters of cursive handwriting. This technique mostly corrects all errors generated by the segmentation and Hidden Markov Model ranking phases by optimizing a measure of information in an efficient search algorithm of the graph. The recognition rate in this experiment is almost very high. I. Goirizelaia et al. (2008) [11] present a multi-agent prototype, "Demotek" for an electronic system of voting based on OCR technology. The Demotek has been improved by using the technique of N-version programming, including those in novel capabilities and security. This article illustrates how the vote data transmission and voter's system of authentication could further improve and simplify the process of electoral by integrating these novel abilities to the electronic system of voting utilizing programming of N-version. I. Ahmad et al. (2017) [12] propose a methodology of automating feature extraction directly by utilizing a stacked denoising autoencoder from raw values of the pixel of ligature images. This network of deep learning has not been utilized so far before for the Urdu characters recognition. The rate of accuracy is approximate 96 percent. J. Memon et al. (2020) [13] review the state of the art of OCR technique and by providing the possible directions of research by illustrating research gaps. This article provides a systematic literature survey and covers almost every area of OCR techniques. The main objective of this article is to provide an overview of the research which has been conducted on the recognition of handwritten characters and gives an outlook for future research scope. M. Mohd et al. (2021) [14] presents a Quranic OCR system that depends on a convolutional neural network accompanied by a Recurrent neural network. Here six models of deep learning are created to illustrate the effect of multiple representations of the performance and the input and output and the accuracy of the models and analogize gated recurrent unit and long short-term memory. This experiment's outcome is a Quranic OCR model which can recognize the diacritic text of the Quranic images. The proposed system has a rate of accuracy is 98 percent on the data of validation.

IV. MAJOR STAGES OF OCR TECHNOLOGY

OCR technique performs the recognition by following the steps such as preprocessing, segmentation, feature extraction, classification, and post-processing respectively [15]. This technique is very helpful to digitized into readable-text for machine across multiple language from huge number of datasets for paper-based documents, which will make machine eligible to store huge dataset. Fig. 2 demonstrates every step taken by the OCR to perform their recognition tasks.

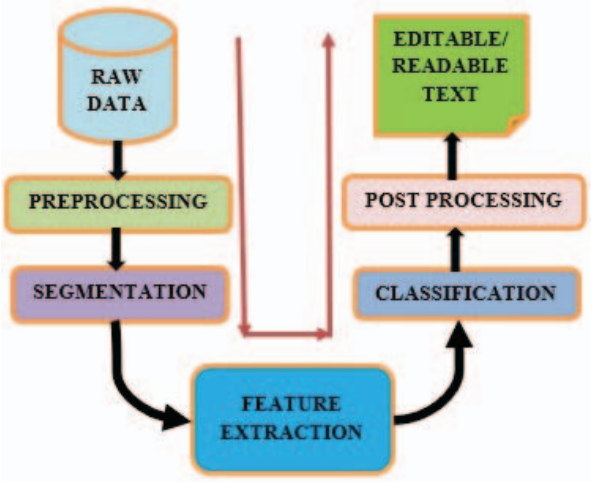

Fig. 2. Process of recognition by OCR technology

Authorized licensed use limited to: University Of Minnesota Duluth. Downloaded on September 09,2025 at 20:52:55 UTC from IEEE Xplore. Restrictions apply.

OCR technology has become an essential part of businesses, privacy, security, social services, and many more other things. Datasets is the main part to get higher accuracy rate from the OCR technology because high quality and quantity of datasets is required to process tasks to get better accuracy. Datasets of high quality is also responsible for the performance efficiency of the proposed OCR technology in this paper. It follows some major steps towards accomplishing their recognition tasks; Some of them are explained below.

1. *Preprocessing:* It is the first and primary step to start the recognition process through OCR technology. Quality of the final output is mostly related on this step because in this step we can remove most of the drawbacks associated with dataset. The main purpose of this step is to remove noise or undesired characteristics in an image without losing desired information. Process of preprocessing is required on color, binary document, or grey-level images having text or graphics. Color images processing is expensive with respect to black and white images thus we utilize grey or binary images. This step can reduce the noise and inconsistent data.
2. *Segmentation:* It is the most important step carried out by the OCR technique. This step is for the text line segmentation from the images. Segmentation of text from an image file combines segmentation of line, segmentation of word, and then segmentation of character. It is the process of text component isolation from the backgrounds of images. Segmentation of documents is a crucial pre-processing step in the working of the OCR system.
3. *Feature Extraction:* It is the process of extracting the features pertinent from alphabets or objects to make feature vectors. These feature vectors can be used by classifiers to recognize the unit of input together objective unit of output. The features can be of two types of structural features and statistical features. Topological or Structural features are devoted to the character's geometry set to be examined. Global or statical features are generally extracted and averaged in sub-images like meshes. The main aim of feature extraction is to identify a feature set, which optimizes the rate of recognition with the minimum number of elements [16].
4. *Classification:* It is the process of allocating inputs data with respect to information detected to their analogizing class to formulate groups with qualities of homogeneous type while isolating various inputs data into several classes. The classification stage can be done with the help of neural networks, statical techniques, support vector machines algorithms, template matching, and integration of classifiers. This stage can be done on the area of features put away in the space of feature. It is the final step towards the recognition of characters through OCR technology because characters are classified in this step in readable format. After this step another step is required to clean the output data.
5. *Post-Processing:* It is the final stage of the OCR system which is performed to remove the spelling and grammatical errors from the input data due to OCR system flaws before delivering the output data from the system. Post-processing has steps of data cleaning for digitized documents. It is a simple way of correcting errors related to human writing mistakes.

V. Some Major Applications

OCR systems can be applied in different cases such as for legal documentation, banking, word processing, and businesses. For a long period of storing legal documents such as loan documents, police FIR, etc. can be performed by the OCR technology with machine intelligence cooperation's. The bank check might be reviewed automatically by OCR technology integrated with Artificial intelligence to prove it is verifying and legitimate the cash you want to deposit [17]. The major purpose of the OCR system is to block damage and misuse of data of information contained in this technology by embodying fully the problem of how the modules of the internal process are preferred and how the information data is saved and converted to the editable or readable format [18]. OCR systems have been developed into various types of OCR applications domain-specific, such as invoice OCR, receipt OCR, check OCR, documents of legal billing OCR. These OCR technologies can be utilized in the following ways:

1. Automatic recognition of number plate.
2. Recognition of Traffic signs.
3. The technology of assistive for assisting blind and visually disabled peoples.
4. Extracting information of business cards into a list of contact.
5. Enabling scanned documents searchable by changing scanned documents to pdf format.
6. For the recognition of passports and extraction of information at airports.
7. Automatic extraction of main information from the insurance documents.
8. For the entry of data in business documents such as invoices, cheques, bank statements, and many more.
9. Enabling electronics images searchable coming from printed documents, such as Google books.
10. For defeating systems of CAPTCHA anti-bot.
11. Converting handwritten documents in real-time for controlling the computers.

OCR technology applications can provide various benefits to us in various ways. It enhances the editability, searchability, accessibility, storability, translatability, and backups of every document. This technology is also beneficial in the communications of two different language-speaking people. Recently, Google translated system also utilizes OCR technology to perform their tasks. Thus, OCR is very useful in our daily life and by its application, we can ease our life.

Authorized licensed use limited to: University Of Minnesota Duluth. Downloaded on September 09,2025 at 20:52:55 UTC from IEEE Xplore. Restrictions apply.

VI. Future Research Perspectives

OCR will be an important and valuable tool for full-fill in gaps whenever an electronic document generated by the application cannot be created. Ultimately, the paperless official works don't exist yet and extraction of data is still important that might augment the processing of e-document. OCR is an active and important area of research that focuses on the development of a computer system that can process and extract characters from images automatically [19]. This technology has various opportunities for researchers and technologists to work in OCR, such as in enhancing accuracy and reducing time to recognize text. The major challenges for researchers to enhance accuracy are due to different writing styles, different alphabets of various languages, and several languages in the world. Some other challenges such as different fonts of text, quality of scanned images, and quality of paper used for writing text also need to be overcome to enhance the accuracy. Thus, researchers can work extensively to remove these challenges towards getting very high accuracy.

VII. Conclusion

This paper is providing important and deeper insights about OCR technology to the researchers or technologists who are going to start their careers in the fields of computer vision, machine learning, artificial intelligence, and deep learning. A wide range of literature survey on the development of this technique and some major applications of this technique has been discussed in this paper. The major issues related to OCR are the rate of accuracy and time of operation, which means accuracy should be higher and time of operation should be lesser. This article provides the principle of working of OCR technology and future perspectives of this technique have been also provided. Further research is required which should be focused on the rate of accuracy and time by mostly focused on the quality of data.

Acknowledgment

The research was supported financially by the AGH University of Science and Technology, Krakow, Poland, subvention no. 16.16.230.434.

References

[1] "Optical character recognition", From Wikipedia, the free encyclopedia, 5th March 2022, available at- https://en.wikipedia.org/wiki/Optical_character_recognition (accessed on 7th March 2022).

[2] N. Islam, Z. Islam, and N. Noor, "A Survey on Optical Recognition System", Journal of Information & Communication Technology-JICT, vol. 10, Issue 2, pp. 1-4, Dec 2016, available at- https://arxiv.org/ftp/arxiv/papers/1710/1710.05703.pdf

[3] M. Edmonds, "A brief history of Optical Character Recognition (OCR)", Pitney Bowes, 2020, available at- https://www.pitneybowes.com/content/dam/pitneybowes/uk/en/shipping-and-mailing/e-invoicing/Blog_E-invoicing-The_Brief_History_of_OCR.pdf (accessed on March 11, 2022).

[4] "Timeline of optical character recognition", From Wikipedia, the free encyclopedia, 14th Feb. 2022, available at- https://en.wikipedia.org/wiki/Timeline_of_optical_character_recognition (accessed on March 11, 2022).

[5] J. Leimer, "Design factors in the development of an optical character recognition machine," in IRE Transactions on Information Theory, vol. 8, no. 2, pp. 167-171, February 1962, doi: 10.1109/TIT.1962.1057696.

[6] P. M. Hall, "A practical optical character-recognition system," in Electronics and Power, vol. 14, no. 4, pp. 149-153, April 1968, doi: 10.1049/ep.1968.0125.

[7] M. D. Freedman, "Advanced technology: Optical character recognition: Machines that read printed matter rely on innovative designs. Systems and their performance are compared," in IEEE Spectrum, vol. 11, no. 3, pp. 44-52, March 1974, doi: 10.1109/MSPEC.1974.6366398.

[8] P. Comelli, P. Ferragina, M. N. Granieri and F. Stabile, "Optical recognition of motor vehicle license plates," in IEEE Transactions on Vehicular Technology, vol. 44, no. 4, pp. 790-799, Nov. 1995, doi: 10.1109/25.467963.

[9] A. Zramdini and R. Ingold, "Optical font recognition using typographical features," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 8, pp. 877-882, Aug. 1998, doi: 10.1109/34.709616.

[10] N. Arica and F. T. Yarman-Vural, "Optical character recognition for cursive handwriting," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 6, pp. 801-813, June 2002, doi: 10.1109/TPAMI.2002.1008386.

[11] I. Goirizelaia, T. Selker, M. Huarte, and J. Unzilla, "An Optical Scan E-Voting System based on N-version Programming," in IEEE Security & Privacy, vol. 6, no. 3, pp. 47-53, May-June 2008, doi: 10.1109/MSP.2008.57.

[12] I. Ahmad, X. Wang, R. Li, and S. Rasheed, "Offline Urdu Nastaleeq optical character recognition based on stacked denoising autoencoder," in China Communications, vol. 14, no. 1, pp. 146-157, Jan. 2017, doi: 10.1109/CC.2017.7839765.

[13] J. Memon, M. Sami, R. A. Khan, and M. Uddin, "Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR)," in IEEE Access, vol. 8, pp. 142642-142668, 2020, doi: 10.1109/ACCESS.2020.3012542.

[14] M. Mohd, F. Qamar, I. Al-Sheikh, and R. Salah, "Quranic Optical Text Recognition Using Deep Learning Models," in IEEE Access, vol. 9, pp. 38318-38330, 2021, doi: 10.1109/ACCESS.2021.3064019.

[15] A. M. Sabu and A. S. Das, "A Survey on various Optical Character Recognition Techniques," 2018 Conference on Emerging Devices and Smart Systems (ICEDSS), 2018, pp. 152-155, doi: 10.1109/ICEDSS.2018.8544323.

[16] P. Purkait, "Optical Handwritten Character/Numerical Recognition", 9th North-East Workshop on Computational Information Processing, available at- https://www.isical.ac.in/~vlrg/sites/default/files/Pulak/Off-Line%20Handwritten%20OCR.pdf (accessed on March 13, 2022).

[17] "What is OCR (Optical Character Recognition): Overview, How it works, Application", by Simplilearn, 18th November 2021, available at- https://www.simplilearn.com/what-is-ocr-optical-character-recognition-article (accessed on Nov 13, 2022).

[18] S. Kim, J. Park, and Y. . -B. Kwon, "An Embedded OCR Software Architecture for Enhancing Portability," Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), 2007, pp. 1004-1008, doi: 10.1109/ICDAR.2007.4377066.

[19] K. A. Hamad, and M. Kaya, "A Detailed Analysis of Optical Character Recognition Technology", International Journal of Applied Mathematics, Electronics, and Computers, vol. 4, pp. 244-249, 2016.

Authorized licensed use limited to: University Of Minnesota Duluth. Downloaded on September 09,2025 at 20:52:55 UTC from IEEE Xplore. Restrictions apply.