Security of Infocommunication Systems and Internet of Things

2024 Vol 2, No 1

https://doi.org/10.31861/sisiot2024.1.01008

Received 29 July 2024; revised 22 August 2024; accepted 28 August 2024; published 30 August 2024

Analysis of Methods for Classification and Aggregation of Textual Data From Images

Bohdan Popovych and Ganna Zavolodko*

Department of Multimedia and Internet Technologies and Systems, National Technical University 'Kharkiv Polytechnic Institute', Kharkiv, Ukraine

*Corresponding author (E-mail: anna.zavolodko@khpi.edu.ua)

ABSTRACT This study investigates modern methods of text recognition from images, specifically comparing optical character recognition and intelligent character recognition. The technologies of machine learning, including convolutional and recurrent neural networks, are compared based on criteria such as accuracy and efficiency in processing handwritten and printed texts. The advantages and limitations of existing solutions for forming digital documents from images containing various handwriting styles and low-quality text images are analyzed. Key challenges associated with processing multilingual texts are identified, and future prospects for the development of text recognition technologies are discussed.

KEYWORDS text recognition, machine learning, data processing automation, multilingual texts, comparative analysis.

I. INTRODUCTION

This is the most valuable resource in today's world where digital technologies play a significant role in many aspects of modern life. Despite that it is easier to type on a keyboard, many people still prefer writing their notes by hand even if it is done on traditional paper or digital tablets. In the context of this there is an increasing demand for technology to convert image-based text into structured data sets more efficiently. The development of intelligent character recognition (ICR) systems creates opportunities for merging the analog and digital worlds.

The recent developments in this area allow for very accurate text recognition which is critical for the preservation and processing of huge amounts of information. This is particularly true about documents containing important scientific, medical and historical data that should be accessible digitally. Technologies including optical character recognition and intelligent character recognition help to effectively transform handwritten texts into machine-readable formats. This makes it much easier to link manual entries with digital databases, enabling researchers to access them later and use them accordingly.

Incidentally, this is an issue that is greatly worth tackling since we would like the important information to be preserved and more easily accessible. In consequence, classifying and aggregating of visual text data from images in a world where massive data is generated on daily basis has consequently been required to be incorporated into creative solutions.

This implies that the use of technology for handwriting recognition in modern information systems is not only helping to store information but also to make it retrievable. That way digital technologies will now be used widely across various areas of operations leading to data processing efficiency increase.

II. MAIN APPROACHES

The process of text recognition has a long history, which began with old traditional optical character recognition (OCR) methods and has lately extended to smart ICR systems. These technologies have been used in many aspects of human life simplifying the process of transforming images into texts. In document processing, OCR is utilized to convert printed information into machine-readable form. However, it encounters difficulties when it comes to handwritten texts or poor quality documents. Although OCR is a basis for many modern recognition systems, its performance heavily depends on the quality of the source material [1].

Contemporary text recognition methods involve application of neural networks such as convolutional neuronal networks (CNN - convolutional neural networks) and recurrent neuronal networks (RNN - recurrent neural network). Such approaches significantly improve recognition accuracy owing to their capacity to interpret the context and structure of a written document. For example, for recognizing hand-written words, ICR involves use of machine learning techniques that can be trained to recognize different writing styles hence improving recognition accuracy [2].

Developments in text recognition, however, involve the creation of smart systems that can learn to adapt to changing situations and improve their performance. An ICR is a good example of such tools it not only recognizes printed but also written texts making it very useful when dealing with manual documents [3, 4].

Nevertheless, despite the considerable progress made in the field of text recognition there are few unresolved issues. One of them is how difficult it is to handle handwriting as it may be different in style and quality. Additionally, processing multilingual texts and symbols represented by different alphabets poses new problems for researchers. The future of ICR lies in increased accuracy and speed of recognition [5, 6].

Adopting the concept of OCR we can define that it is the conversion of printed or handwritten text into digital form. OCR is applied for fast input of texts into the computer systems and for decreasing the amount of spaces

for the storage of the documents. However, there are certain challenges associated with conventional OCR techniques and those include poor recognition with the different font types and the page layout [7]. An upgrade to OCR is ICR that employs the use of machine learning and artificial intelligence to make this handwriting recognition even better. ICR is carried out by sub-dividing each document into characters and by using computer vision before initiating the final recognition process by applying trained models on the huge array of handwriting recognition data models. ICR also adapts NLP methods to evaluate context of the text, enabling precise identification of character even in unfavourable circumstances [8, 9]. The CNNs are employed in recognizing the handwritten characters because of the effect exhibited by the networks in extracting the image features. CNNs consist of three main layers: Specifically, convolutional layer, pooling layer and fully connected layer are involved in the neural networks. The convolutional layer determines primary elements in the image to be processed such as edges and corners, and the fully connected layer differentiates between the type of the image based on these features. The employment of CNNs enhances the performance of handwritten text recognition especially for images with blurry backgrounds or complicate background [7].

RNN and its modified forms like LSTM (Long short-term memory) and BLSTM (Bidirectional Long Short-Term Memory) are used for recognizing the data sequences like handwritten text. The ability of RNNs to take into account preceding characters when recognizing subsequent ones is advantageous when processing digital documents and especially handwritten ones. The high recognition accuracy is explained by the nature of RNN and its modifications as they can take into account the context of the sequence of characters in the text [8].

CNN and RNN combination has be used to improve the performance of handwritten text recognition because each of these provides unique benefits. For instance, CNN is employed to obtain image features, and RNNs to process these features sequentially and classify them. This approach has good accuracy, thus making it possible to utilize when involves handling numerous documents with different formats and font types [8].

III. LATEST STUDIES

Article «Handwritten Character Recognition Using CNN» [7] written by Madhu M Nayak and Vaidehi D deals with the recognition of characters using CNN. The authors pay the most attention to designing and assessing the recognition system that should recognize the characters extracted from the optical images of handwritten text or directly input characters. Preprocessing techniques are distinguished, which improve the work of the system and increase the relevance of the proposed approach to such fields as digital processing of handwritten copies, automation of post office work, and increasing the functionality of the accessibility tools. It is unique in the area of OCR because of the methods of processing and the application of deep learning to enhance the identification of characters. The authors employ the EMNIST dataset that covers the set of handwritten characters, and thus, the model can be trained under different scenarios. In the course of this work, they managed to achieve high identification rates, which at some level of training exceeded 86% and this point to a high possibility of using CNN in the sphere of OCR. The findings of this work open a chance to apply this method in commercial and educational products where quickly recognising the handwritten text is significant. Therefore, this research is a significant contribution towards the development of effective and complex handwritten text recognition that belongs to the sphere of modern technological advancements.

According to the article «How Intelligent Character Recognition (ICR) is Overcoming OCR Limitations in Document Processing» [9], the history of document automation is described with the focus on the ICR benefits over OCR. The authors explain how the ICR technology entails the use of AI and machine learning in the recognition of handwritten texts. Namely, the ICR can employ various NLP methods for context aware recognition of the text and, therefore, accurately recognize rather complex character forms. The main weaknesses of the OCR are closely related to the strict framework for recognizing characters donated and used, which makes it difficult to work with semi-structured and handwritten texts. While the latter one uses machine learning algorithms to detect a number of specific features of the character, in comparison to ICR is much more flexible and accurate in front of various types of documents. This supports complex document inputting such as invoices, receipts, legal documents, and research notes among others. The scenarios using ICR are related to data entry, invoices, and receipts, and also for the recognition of the handwritten notes. ICR is much more accurate and faster compared to OCR, as it provides a spectrum of usable solutions for the document recognition in different areas of business. Other concerns raised in the article regards important aspects regarding the ICR's execution and these are; models must be trained well on large data sets and adequate measures should be taken to secure sensitive information. Therefore, it can be noted that the article engages with the assertion that ICR is the major enhance to the document processes technologies that assist business entities to enhance their data processing credibility, expediency, and velocity to create better data processing decisions and drive success.

In the article «Self Intelligence with Text Recognization» [8], Subodh L. Vasankar, Hershad Mahajan, Deovrat Deshmukh, and Hemant Munot are devoted to creating infrastructure that would allow machines to analyze certain video streams containing text images and then analyze the meaning of this text and program themselves in accordance with the text contents. The work's principal outcome is an algorithm and the software developed, which is capable of identifying and analyzing text characters in real-time and perform programmed actions based on the analyzed text. One of the features of this development is usage of a certain protocol involving an initial and a final protocol with inserted instructions written in C programming language and printed in a certain font. The authors explain stages of the

image processing as text detection, text cropping and creating an image file into text. The biometric features especially image samples are passed through a character recognition algorithm that helps differentiate between them and identify the relevant character, this makes it most suitable for specifying usage conditions. Furthermore, there is an analysis of programming: programming a microcontroller considering the text received, enabling the machine to perform the instructions assigned in the identified text. Future opportunities in the research provide for the broad use of this technology in the future, by constructing machines that can learn on their own, and perform various functions according to text prompts. This could tremendously enhance human-robot interaction; the robots would be able to perform numerous activities that call for identification of textual content.

Some of the conventional techniques to optical character recognition which can easily be implemented by an automated OCR system are:

According to the classification to the methods of automatic text recognition, OCR can be considered as one of the first methods developed. It can turn printed text into machine readable format making the process of document scanning far easier. The main stages of OCR operation include:The main stages of OCR operation include:

1. Preprocessing the Image: Filtering out the noise,чё making the text aligned and increasing contrast.

2. Segmentation: Blending of letters into characters, or words in the case of a word image.

3. Character Recognition: Selecting each characters based on the templates set.

4. Postprocessing: Punctuation of the text, deletion and addition of some words.

TABLE 1. Comparison of OCR and ICR.

Aspect	OCR	ICR
Basic Principle	Template or rule-based, without AI	Adaptive using AI, minimal human intervention
Human Supervision	Requires frequent human supervision	Reports only in case of anomalies
Adaptability to Document Layout	Suitable for fixed-layout documents	Trained for frequent layout changes
Template/Rule Requirements	Manual creation of templates/rules/layouts	Does not require templates or rules
Integration with ERP Systems	Integration can be more complex	Outputs easily integrate with ERP systems
Accuracy Improvement	Depends on the supporting database	Improves accuracy over time with AI

Using the above table one is in a position to see how ICR is better than OCR in every way. ICR, due to the efficient integration of artificial intelligence, requires little supervisory inputs and handles layout changes in a number of documents within an efficient manner that it does not require the creation of templates or rules. In contrast to OCR, which has to be supervised daily and operated based on templates, ICR is interfaced with ERP more effectively and develops with every iteration based on the data received. Summing up, the above-mentioned benefits explain why ICR is more effective and versatile compared to OCR in text data processing.

Text recognition in today's world employs high-performance machine learning computation techniques that boost both the speed of recognition as well as accuracy. These include:

1. Convolutional Neural Networks (CNN). Alarm and signal processing is the primary application of this kind of layers.

2. Recurrent Neural Networks (RNN). Applicable for examining sequences of characters for instance, used in recognizing handwritten texts.

3. Hybrid Models. CNN and RNN integrated into the design to increase the recognition results.

TABLE 2. Comparison of text recognition methods.

Methods	Accuracy	Speed	Limitations
CNN	~90%	~500 cps	Large data requirement for training
RNN	~85%	~300 cps	Sensitivity to long sequences
Hybrid Models	~95%	~400 cps	Complexity in implementation and tuning

The table shows the comparative analysis of CNN, RNN and the hybrid model where we can notice its pros and cons at the same time. CNNs give high recognition accuracy, but in exchange, they need more data for training and have moderate speed. RNNs also involves high accuracy and the relative speed is moderate but is NOT suitable for long data sequences. The combination of CNNs and RNNs is very accurate, and very fast; nevertheless, they are difficult to implement and set precisely. The conclusions from the table stress that, although the ensembling of two approaches is harder to achieve, its outcome could be significantly better than the outcome of each of the models separately which is why the hybrid models should be seen as the best way to go when it comes to text recognition [10].

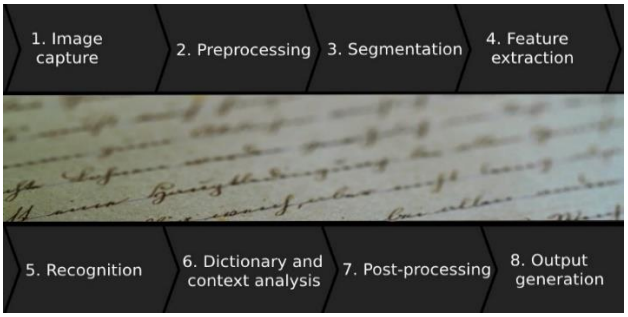

FIG. 1. How IRC works.

The paper reveals that Intelligent Character Recognition is among the most superior methodologies in the field of text recognition. It even allows the identification of not only printed text but also text in handwritten form, which enlarges the possibilities of its usage. The main advantages of ICR include:

1. Adaptability. About the ability to learn and get used to different handwritten texts.

2. Accuracy. Recognition accuracy is very high even for complicated handwritten texts.

3. Integration. Compatibility with the existing other systems to process as well as analyse the input data.

Table 3 shows that ICR has a lot of key parameters on which it outperforms the rest. It cannot be utilized to recognize printed texts like OCR, but it is accurate in recognizing printed and hand-written texts with much better recognition accuracy. It also has much higher flexibility in handling all document formats than OCR, whose adaptability is much low. While advanced functionality of ICR is traded off for longer processing times, OCR has retained high sensitivity to image quality. Results indicate advantages of using ICR as a solution that is more accurate and adaptable in text recognition, considering operation in high document variability environments [11, 12].

TABLE 3. Comparison of ICR and traditional OCR methods.

Parameter	OCR	ICR
Type of Text	Printed	Printed and Handwritten
Accuracy	80-95%	90-99%
Adaptability	Low	High
Limitations	Sensitive to image quality	High computational resource requirement

V. CONCLUSION

In this regard, it is necessary to comprehend that the existing and more recent methods of text recognition, including CNN, RNN and ICR, improve the recognition, as well as the efficiency of current handwriting and printed text recognition. Based on the presented advantages of ICR technology, the ability to recognize texts with different handwriting while using low-quality for text input, the continuation of research on the given technology can be considered a priority. Notable among them is the capacity to deal with multidimensional texts, which opens up the application of the tool to several disciplines. Nonetheless, some issues have not been solved, like difficulties in the straight text's recognition due to different handwriting styles, or a relatively slow speed of recognition.

AUTHOR CONTRIBUTIONS

B.P., G.Z. – formal analysis; B.P., G.Z. – conceptualization, methodology; B.P. – investigation; B.P. – writing-original draft preparation, writing-review and editing; G.Z. – supervision, validation.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose. The authors have no conflicts of interest to declare that are relevant to the content of this article.

REFERENCES

[1] Wikipedia, "Intelligent Word Recognition." [Online]. Available: https://en.wikipedia.org/wiki/Intelligent_word_recognition. [Accessed: Jul. 29, 2024].

[2] ABTO Software, "Intelligent Character Recognition (ICR) of Handwritten Text." [Online]. Available: https://www.abtosoftware.com/intelligent-character-recognition-icr. [Accessed: Jul. 29, 2024].

[3] Shufti Pro, "Demand for OCR Technology Increasing in ID Verification." [Online]. Available: https://shuftipro.com/demand-for-ocr-technology-increasing-in-id-verification. [Accessed: Jul. 29, 2024].

[4] Label Your Data, "What is ICR Technology." [Online]. Available: https://www.labelyourdata.com/what-is-icr-technology. [Accessed: Jul. 29, 2024].

[5] AWS, "What is OCR." [Online]. Available: https://aws.amazon.com/what-is/ocr/. [Accessed: Jul. 29, 2024].

[6] Shufti Pro, "Intelligent Character Recognition (ICR) Software: One Step Ahead of OCR." [Online]. Available: https://shuftipro.com/intelligent-character-recognition-software. [Accessed: Jul. 29, 2024].

[7] M. M. Nayak and D. Vaidehi, "Handwritten Character Recognition Using CNN," *International Journal of Computer Engineering and Technology (IJCET)*, vol. 15, no. 3, pp. 219-229, 2024.

[8] S. L. Wasankar, H. Mahajan, D. Deshmukh, and H. Munot, "Self Intelligence with Text Recognization," Government College of Engineering, Amravati, India, 2024.

[9] Techopedia, "How Intelligent Character Recognition (ICR) is Overcoming OCR Limitations in Document Processing." [Online]. Available: https://www.techopedia.com/how-intelligent-character-recognition-icr-is-overcoming-ocr-limitations-in-document-processing. [Accessed: Jul. 29, 2024].

[10] F. M. Shiri, T. Perumal, N. Mustapha, and R. Mohamed, "A Comprehensive Overview and Comparative Analysis on Deep Learning Models: CNN, RNN, LSTM, GRU," pp. 9-11, 2023.

[11] R. Smith, "An Overview of the Tesseract OCR Engine," in *Ninth International Conference on Document Analysis and Recognition (ICDAR)*, pp. 629-633, 2007.

[12] S. Malakar and P. Roy, "A Study on the Impact of Intelligent Character Recognition (ICR) on Digitizing Handwritten Documents," *International Journal of Advanced Research in Computer Science and Software Engineering*, pp. 15-20, 2018.

Bohdan Popovych

Bohdan, 23 years old, studied and obtained his bachelor's and master's degrees at NTU "KhPI." In 2023, he researched the use of AI in web applications while working on his master's thesis. In 2024, he joined the team at ReMnemo, a high-potential startup that successfully completed an acceleration program. At ReMnemo, he is engaged in the application of AI in various projects.

ORCID ID: 0009-0009-6802-1549

Ganna Zavolodko

Ganna, 46 years old, Ph.D., Associate Professor at NTU "KhPI," IEEE Senior Member; CEO and co-founder of ReMnemo.

ORCID ID: 0000-0003-0000-8910

Аналіз методів класифікації та агрегації текстових даних з зображення

Богдан Попович, Ганна Заволодько*

Кафедра мультимедійних та інтернет-технологій і систем, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна

*Автор-кореспондент (Електронна адреса: anna.zavolodko@khpi.edu.ua)

АНОТАЦІЯ У статті досліджуються сучасні методи розпізнавання тексту з зображень, зокрема порівнюються оптичне розпізнавання символів (ОРС) та інтелектуальне розпізнавання символів (ІРС). Розглянуто технології машинного навчання, включаючи згорткові нейронні мережі (ЗНМ) та рекурентні нейронні мережі (РНМ), які використовуються для підвищення точності та ефективності обробки рукописних і друкованих текстів. Вивчено переваги та обмеження існуючих рішень для формування цифрових документів з зображень, що містять різні стилі почерку і текст на зображеннях низької якості. Зокрема, ІРС демонструє високу адаптивність до змінних умов, що робить його ефективнішим у порівнянні з традиційними методами ОРС. Значна увага приділена здатності ICR обробляти багатомовні тексти, що розширює можливості її використання у різних сферах діяльності. У статті також аналізуються основні виклики, пов'язані з обробкою рукописного тексту різних стилів, а також необхідність покращення швидкості розпізнавання. Представлені результати свідчать про високий потенціал використання ЗНМ та РНМ у задачах ОРС, а також про перспективність гібридних моделей, що поєднують переваги обох підходів для досягнення дуже високої точності розпізнавання. Особливо важливим є застосування цих технологій у таких галузях, як цифрова обробка рукописних документів, автоматизація поштових послуг та розширення інструментів доступності. Ці висновки підкреслюють важливість подальших досліджень та розвитку технологій розпізнавання тексту для покращення ефективності обробки даних та інтеграції рукописних текстів у сучасні інформаційні системи. Дослідження показує, що технології машинного навчання та глибокого навчання можуть значно покращити точність розпізнавання тексту, зменшуючи необхідність людського втручання та прискорюючи процес обробки даних. У статті також розглядаються практичні аспекти впровадження цих технологій, зокрема, необхідність великих обчислювальних ресурсів для навчання моделей та забезпечення надійної роботи систем у різних умовах. Це дослідження є важливим внеском у розвиток технологій цифрової обробки тексту, що мають широкі перспективи застосування у різних галузях, включаючи науку, медицину, освіту та бізнес. Використання сучасних методів розпізнавання тексту дозволить значно підвищити ефективність обробки інформації та сприяти розвитку інноваційних рішень для роботи з великими обсягами даних.

КЛЮЧОВІ СЛОВА розпізнавання тексту, машинне навчання, автоматизація обробки даних, багатомовні тексти, порівняльний аналіз.