

Mobile application to digitize handwritten patient records in Peruvian public hospitals

Wilfredo Sebastian Romero Villa, Bachelor of Science in Information Systems Engineering¹, Franco Alonso Gregorini Machuca, Graduate in Information Systems Engineering², and Richard Nivaldo Copaja Cornejo, Systems Engineer³
^{1,2,3} Department of Information System, Peruvian University of Applied Sciences, Peru, u20171a226@upc.edu.pe, u201718179@upc.edu.pe, pcsircop@upc.edu.pe

Abstract- This research has shown that public hospitals need a digitization model that allows the availability of medical records for patient medical consultation within a repository that unifies their medical information.

This article presents MEDREC APP, a mobile application aimed at public hospitals of the Ministry of Health (MINSA), in order to expedite the immediate obtaining of the medical history in digital format, and to visualize their medical care. The main function of this solution is based on Optical Character Recognition (OCR) as a text extraction process by capturing images of handwritten medical history formats and registering them within the mobile application in digital format.

As part of the validation process, indicators were defined, surveys and interviews were conducted with our users of the application: the medical staff. The results obtained show that the average time to obtain a medical history is between 30 and 15 minutes. By including the mobile solution to be presented, the time was reduced to 6 minutes, which means a reduction of 9 minutes, equivalent to approximately 70%.

Keywords: Public Hospitals, Medical Records, Digitization, Optical Character Recognition, Optical Character Recognition, Digitalization

Digital Object Identifier: (only for full papers, inserted by LACCEI).
ISSN, ISBN: (to be inserted by LACCEI).
DO NOT REMOVE

Mobile application to digitize handwritten patient records in peruvian public hospitals

Wilfredo Sebastian Romero Villa, Bachelor of Science in Information Systems Engineering¹, Franco Alonso Gregorini Machuca, Graduate in Information Systems Engineering², and Richard Nivaldo Copaja Cornejo, Systems Engineer³
^{1,2,3} Department of Information System, Peruvian University of Applied Sciences, Peru, u20171a226@upc.edu.pe, u201718179@upc.edu.pe, pcsircop@upc.edu.pe

Abstract—This research has shown that public hospitals need a digitization model that allows the availability of medical records for patient medical consultation within a repository that unifies their medical information. This article presents MEDREC APP, a mobile application aimed at public hospitals of the Ministry of Health (MINSA), in order to expedite the immediate obtaining of the medical history in digital format, and to visualize their medical care. The main function of this solution is based on Optical Character Recognition (OCR) as a text extraction process by capturing images of handwritten medical history formats and registering them within the mobile application in digital format. As part of the validation process, indicators were defined, surveys and interviews were conducted with our users of the application: the medical staff. The results obtained show that the average time to obtain a medical history is between 30 and 15 minutes. By including the mobile solution to be presented, the time was reduced to 6 minutes, which means a reduction of 9 minutes, equivalent to approximately 70%.

Keywords— Public Hospitals, Medical Records, Digitization, Optical Character Recognition, Optical Character Recognition, Digitalization

I. INTRODUCTION

Today, healthcare institutions have a potential to adopt different technologies that influence clinical research. Paper review can not only become cumbersome but is also subject to human error [1]. According to [2], within the national setting, the problem presented by the Ministry of Health (MINSA) is precisely the management of patient information, in view of the fact that most clinical centers still use a handwritten clinical history. Thus, in [3] it is stated that a recurrent problem in health centers around the world is the lack of initiative to carry out digitalization projects.

On the other hand, Law No. 30024 created the National Registry of Electronic Health Records (RENHICE), which led to the implementation of information systems that will safeguard the information in the patient's medical records [4]. Therefore, alternatives such as the Electronic Health Record Information System SIHCE (e-qhali) have emerged for certain first level health facilities [5]. However, according to [6] the Multiannual Investment Programming Office (OPMI) of MINSA, in its report on the analysis of gaps in infrastructure

and equipment in the health sector in recent years, mentions that at least 76% of the information systems do not respond adequately to the needs of health personnel.

II. RELATED WORK

In the present scientific article [7] bases its proposed solution on the use of optical character recognition for the preservation of historical documents. So, in that sense, the main problem is that historical documents can be difficult to read due to paper and ink degradation. That is why the authors propose a recognition system consisting of three stages: image preprocessing, character segmentation, and character recognition. In the preprocessing stage, operations are performed to improve image quality, such as noise removal and contrast enhancement. In the segmentation stage, individual characters in the image are identified and separated. In the recognition stage, machine learning techniques are used to identify the characters and convert them into text. Finally, the results achieved an average recognition rate of 98.67%, making it effective for ancient historical documents in English. This system can be useful for the preservation and analysis of valuable historical documents.

In [8], the feasibility of digitization was proposed in the aftermath of the COVID-19 pandemic, which meant an acceleration in the digital transformation, efficiently storing all physical medical records. In summary, the article discusses how the digitalization of hospitals has become increasingly important due to the need to minimize physical contact. The study focused on three hospitals in the Czech Republic and analyzed the implementation of digital tools for healthcare, such as online consultations, telemedicine, and the use of mobile devices for patient monitoring. The authors highlight a crucial point regarding the challenges faced by hospitals in implementing these technologies, such as resistance to change and concerns about data privacy. The study's results indicate that hospital digitalization can be effective in improving the quality of healthcare for patients. However, the authors emphasize the need for clear policies and strategies to ensure the successful implementation of digitalization in hospitals.

Likewise, in [9] it is focused that character recognition solutions are fully viable for mobile devices being mainly results the scanning of images. In fact, the authors use a visualization system called MyOcrTool that is used to generate associative images of Chinese characters on smart devices. The main aim of the system is to provide an intuitive and visually attractive way for students to learn and memorize Chinese characters. In the evaluation of the system, a study was conducted with 30 participants, and it was found that the system significantly improved the students' ability. Therefore, digitalization is not only applied to the healthcare sector but also to the field of education.

Similarly, in [10] analyzed the feasibility of OCR-based image scanning for COVID-19 patient records to efficiently manage the storage and analysis of medical records. This article focuses on the development of a system for extracting medical information from paper forms used in COVID-19 assessments. The aim is to automate the data collection process to reduce the workload of healthcare professionals and improve the efficiency of the healthcare system. The authors used image processing and machine learning to extract relevant information from the forms, such as the patient's symptoms, medical history, and COVID-19 test results. The effectiveness of the system was evaluated using a dataset of real forms, and it was found that the system achieved an accuracy rate of 93.6% in extracting medical information. Like the previously mentioned article, it has a positive impact on the digitization of documents in the healthcare sector by allowing for faster and more accurate data collection.

In relation to [11] the issue of digitization of handwritten characters by applying classification algorithms in image processing is addressed, so it seeks to generate identification patterns of the numbers and the alphabet with thousands of images that can help to train the neural network and thus predict and classify the images. Likewise, it is important to mention that the architecture it presents is cloud storage using mobile devices. The results show that the proposed system achieves an average accuracy of 97.8%, making it a viable option for handwritten character recognition applications on mobile devices. This system has great potential to improve the accuracy and speed of text input on mobile devices, which could significantly enhance the user experience in a variety of applications. Finally, it was analyzed that in [12] performed image processing on structured and unstructured documents, managing to implement an alignment algorithm to be used by any optical character recognition engine. Therefore, the study proposes a comparative approach for automatic alignment of documents in environments where multiple OCR engines are used. This study is of great utility since, based on the results, we can have references to apply a specific OCR engine for applications that require processing of large volumes of documents.

The implementation of digitalization in the healthcare sector has been analyzed, and the article [13] aims to improve the barriers that exist in the development of digitalization of medical information, as well as to enhance resource allocation and strategic pathways that benefit both patients and healthcare professionals. This research identifies five main dimensions of barriers: digitized processes, data sharing, infrastructure, regulations, and operational issues. To provide solutions, the DEMATEL method and the IRA-NRM model were used, revealing that the lack of qualified personnel, non-standardized data sets, and inadequate understanding of laws and regulations were the most significant resistance factors. The results of the analysis indicate a need to focus on simplifying technological products for practical use by healthcare professionals.

One of the main considerations for the use of digitalization in hospitals is discussed in article [14], which focuses on electronic health records (EHRs) in China and the United States. Data from 2007 to 2018 from the Chinese Health Information Management Association (CHIMA) and data from 2008 to 2017 from the American Hospital Association's Health Information The results showed that the average adoption rate of sampled hospitals in China increased from 18.6% to 85.3%, compared to the increase from 9.4% to 96% in American hospitals from 2008 to 2017. The main conclusion is that the implementation of digitalization must be gradually incorporated to ensure a feasible transition for those working with electronic health records. This article [15] presents a set of methods for converting historical scans into their textual representation for efficient information retrieval with a minimal number of manually annotated documents. The study focuses on optical character recognition (OCR) and various methods for evaluating the localization of elements or words on the page. Therefore, the authors assert that it is possible to create an efficient OCR system for historical documents with a small amount of annotated training data.

III. METHOD

With the present research we proposed a mobile application that is easily accessible and allows our user, the medical staff, mainly to digitize and visualize the physical medical records of the hospitals of the Ministry of Health (MINSAs). The name of the solution is MEDREC APP. As a new user you will register and log in to the application giving you access to the main menu, which you can choose between patient registration or medical record registration.

The patient registration is done only once, then you can scan the physical document of the medical history, which will encrypt the patient's data, then the medical staff can search for the patient either by his full name or by his national identity card (DNI).

Finally, it will be possible to filter by a specific date in the clinical history or the last consultations made to the patient. As part of the solution, an integrated backend and frontend architecture was proposed to support the final deliverable, which is the mobile application. The architecture developed and certified will be presented below.

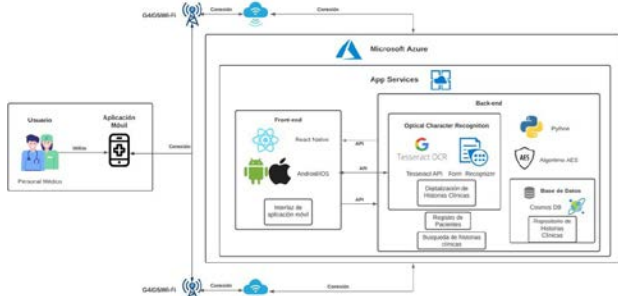


Fig. 1. Integrated architecture

A. Test scenario preparation

For the development of the solution, it was proposed to use the design structure of mobile applications attached to the Ministry of Health, in that sense, after an analysis, it was possible to implement an easy and understandable user experience. Likewise, regarding the cloud repositories, a benchmarking was performed and after this process Cosmos DB was chosen, thus achieving an implementation according to our database model, which is NoSQL.



Fig. 2. Login interface

For the image processing of character recognition, it was done using Azure Form Recognizer, which was coded in the backend of the application using the Python programming language to extract handwritten text from medical records, training it in each execution and thus digitizing the physical document.



Fig. 3. Image processing monitoring interface

Additionally, the framework used was React Native, in order to make it cross-platform for Android and IOS operating systems, and through the use of JavaScript it allows the connection with the backend through secure hypertext transfer protocol (HTTPS) connections.

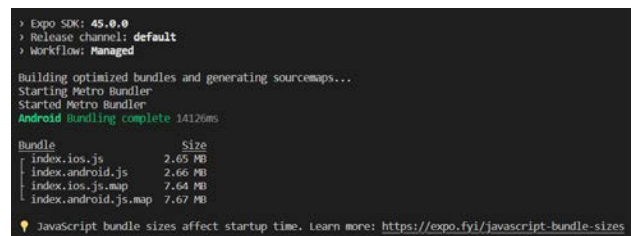


Fig. 4. Image processing monitoring interface

B. Process to follow

The following methodology was proposed for the validation of our final deliverable.



Fig. 5. Validation process

C. Definition of indicators

The following methodology was proposed for the validation of our final deliverable.

TABLE I
INDICATORS TO TEST

INDICATORS	
IND01	Average turnaround time in the availability of medical records
IND02	Average time spent searching for patient information
IND03	Overall percentage of user satisfaction with the mobile application

The table above shows the three proposed indicators for a quantitative measurement of the solution.

The purpose of the validation was to get the perspective of the users of the application, which is why a rating scale from 1 to 5 under the Likert scale was used to measure satisfaction with the application, as well as response and search times.

IV. EXPERIMENTAL SETTINGS

For the validation process, the following requirements will need to be taken into account, for which the minimum technical details that the smartphone mobile device must have are specified.

- Version: Android 5.0 (API level 21), it is recommended to have updated to the latest android version.
- RAM: 4 GB
- Camera: 8MP
- Network: 4G/5G, it is recommended to have a stable connection to the Wi-Fi network.
- APK: Yes, you will need to install it and grant it all the required permissions. For the current version of the application (1.0.0) you must have a storage space of 65 MB.

A. Process description

The development of the mobile application is aimed at health professionals working in public hospitals of the Ministry of Health (MINSA), which is why the corresponding tests were carried out with professionals from the “Maria Auxiliadora” Hospital. The activities involved were the explanation of the purpose of the registration project, the usability of the mobile application and finally the completion of the satisfaction survey. The test environment is not limited to a specific location, only the smartphone was requested to have a stable connection to its internet service provider. Accordingly, a mobile device was provided with the Android Application Package (APK).

B. Preparation of the surveys

Because the end users are health professionals, questions were developed based on the healthcare environment in which the patient’s medical records are most commonly used, so the questions allowed us to find out the perception of our users through Google Forms. The questions, both quantitative and qualitative, will be related to the indicators presented above. The results were analyzed and were important feedback that was considered for the evaluation of the results.

TABLE II
Questions to Survey

QUESTIONS	
Q1	Was it easy to digitize the medical records?
Q2	How satisfied are you with the MEDREC APP mobile application?
Q3	Compared to the manual search of the physical documents, how efficient was the search of the digitized medical records?
Q4	Do you find the design and flow of the application easy to use?
Q5	Do you consider that this application can help in the digitization of the Ministerio de Salud (MINSA) medical records?

The table above shows five questions asked to medical professionals in order to

C. Mathematical model

Given the quantitative data obtained, the use of the arithmetic mean will be taken into consideration, which is why the respective equation will be used on the results of the survey of health professionals on the use of the mobile application called MEDREC APP.

It is important to mention that the sample data obtained for the calculation of the time averages are in minutes and the optimization data is specified as a percentage to two decimal places. It is worth mentioning that this mathematical model is used to obtain the average value, which is obtained by adding up all the data and dividing the result by the total amount of data (n), which can be represented as follows:

$$\bar{X} = \frac{\sum_{i=1}^n \cdot X_i}{n} \quad (1)$$

The data will be obtained by means of this formula, which, as mentioned above, works directly on the total value. That is why for the validation process, this mathematical model will be applied on the basis of the survey conducted, taking into consideration the application called MEDREC APP.

V. RESULTS

The following is a summary of the results obtained in the surveys conducted for each indicator.

TABLE III
Satisfaction Survey Results - Health Professional

Indicators	Percentage expected	Percentage obtained
IND01	≤ 80%	90.8%
IND02	≤ 80%	89.3%
IND03	≤ 80%	88.4%

The table above shows the three proposed indicators, the second column shows the expected percentage of compliance, the third column shows the percentage of compliance of the indicator by the proposed solution.

With the results of the survey, in which 26 health professionals participated, whose main requirement was that in their daily professional practice they had to complete the clinical history by hand, it was verified that the required goal was achieved, given that the percentage of compliance with each of the defined indicators shows satisfactory compliance. It is important to mention that as part of the validation process the main condition was that the focus should be on the area of medical care for direct consultation with the patient.

For a better understanding, some graphs were elaborated to represent the results obtained, thus, in figure 6, it was sought as an indicator that the average response time for the availability of medical records, at the beginning a maximum value of up to 30 minutes was registered, but when the validation process was carried out, an average time of 6 minutes was achieved, thus establishing a reduction of more than 70 percent, also as a

percentage obtained from the indicator was 90.8 percent, achieving to exceed the initially proposed objective.

IND 1: Average turnaround time in the availability of medical records

TABLE 4
Data collected from surveys

N°	AVAILABILITY TIME OF THE MEDICAL RECORD - AS IS (MIN)	AVAILABILITY TIME OF THE MEDICAL RECORD - TO BE (MIN)	OPTIMIZATION OF MEDICAL RECORD AVAILABILITY TIME
1	19	4	79%
2	28	8	71%
3	23	4	83%
4	27	6	78%
5	26	5	81%
6	17	8	53%
7	21	4	81%
8	22	6	73%
9	20	7	65%
10	25	4	84%
11	24	8	67%
12	29	5	83%
13	18	7	61%
14	16	6	63%
15	15	5	67%
16	30	4	87%
17	17	8	53%
18	26	6	77%
19	23	5	78%
20	19	7	63%
21	20	4	80%
22	29	6	79%
23	21	5	76%
24	18	8	56%
25	27	4	85%
26	22	7	68%

The table above shows the 26 surveys conducted. The second column indicates the time in minutes that the medical staff takes to physically obtain the medical history. The third column indicates the time in minutes that the mobile solution can digitally obtain the medical record. And finally, the fourth column shows the time that was optimized by using the proposed solution.

INDICATOR CHART N° 01

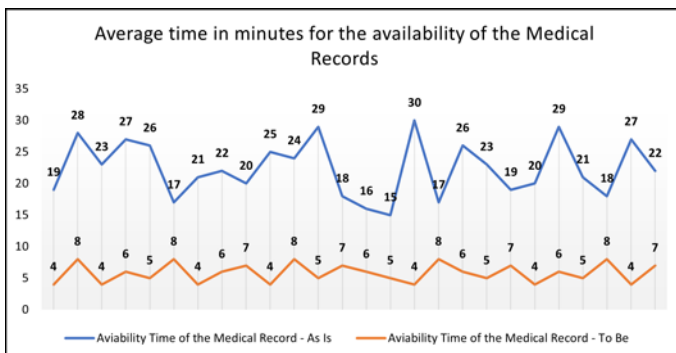


Fig. 6. Graph of the average time in minutes for the availability of the Medical Record

Along the same lines in figure 7, we looked for an indicator for the average patient information search time, which we had as a record that the maximum value reached was 15 minutes and with the proposed solution it reached an average of 8 minutes, thus establishing a reduction of more than 50%, also the percentage obtained for the indicator was 89.3%, so it was also an objective that exceeded expectations.

IND 2: Average time spent searching for patient information.

TABLE 5
Data collected from surveys

N°	MEDICAL RECORD SEARCH TIME - AS IS (MIN)	MEDICAL RECORD SEARCH TIME - TO BE (MIN)	OPTIMIZATION OF MEDICAL RECORD SEARCH TIME
1	12	8	33%
2	10	7	30%
3	15	9	40%
4	13	6	54%
5	10	9	10%
6	15	7	53%
7	14	8	43%
8	11	10	9%
9	13	8	38%
10	10	7	30%
11	11	9	18%
12	14	6	57%
13	12	8	33%
14	11	7	36%
15	12	10	17%
16	14	6	57%
17	13	9	31%
18	15	7	53%
19	10	8	20%
20	13	9	31%
21	11	6	45%
22	14	10	29%
23	15	7	53%
24	10	8	20%
25	12	9	25%
26	13	7	46%

The table above shows the 26 surveys conducted. The second column shows the time in minutes that the medical staff takes to look up the patient's information in the physical medical record. The third column indicates the time in minutes that the medical personnel search for the patient's information in the mobile solution. And finally, the fourth column shows the time that was optimized by using the proposed solution.

INDICATOR CHART N° 02

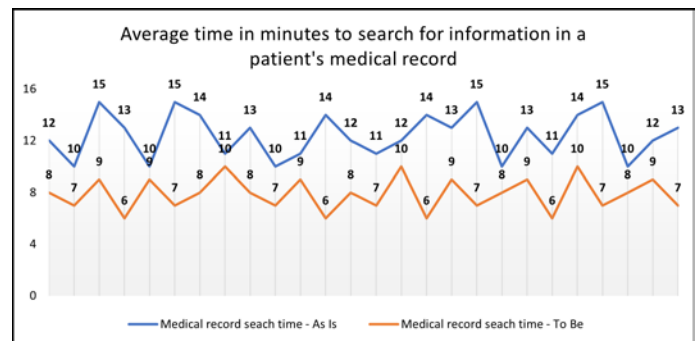


Fig. 7. Graph of the average time in minutes to search for information in a patient's medical record.

Finally, indicator number three, by means of figure 8, the general percentage of user satisfaction with the mobile application was established, for this we took into consideration the likert scale, so from the defined objectives and based on ISO 25000 to have software quality as a priority, the results obtained is on average that the user is satisfied with the application shown, highlighting the usability of the same.

IND03: Overall percentage of user satisfaction with the mobile application

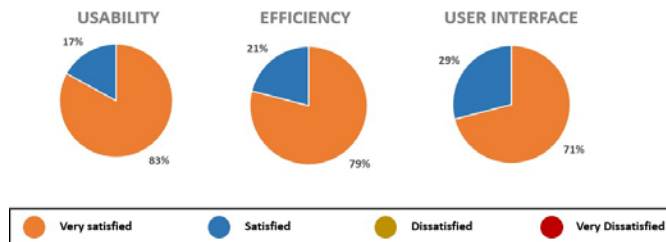


Fig. 8. Overall percentage of user satisfaction with the mobile application

VI. CONCLUSION

The data obtained, after the validation process, lead to the result that the digitisation of medical records in Ministry of Health hospitals is feasible. This was achieved through the solution developed MEDREC, whose main objective was the immediate availability of medical records in the public health sector, for which the data obtained is that the average time recorded was 30 to 15 minutes and with the solution an average of 6 minutes was obtained, thus achieving a significant reduction of 73%, with this we are solving not only the time delay in medical care, but the patient can have their medical records, in any health centre corresponding to MINSA. Likewise, given that priority was given to compliance with ISO/IEC 25000, under the pillars of usability and efficiency, an average user satisfaction rate of 78% was achieved. Finally, it should be noted that Optical Character Recognition (OCR) technology was the main means by which medical records were digitised, thus being a factual application for the automation and reduction of processes, such as medical care.

REFERENCES

- [1] Das, A., Kammari, P., Vadapalli, R., & Basu, S. (2020). Big data and the eyeSmart electronic medical record system - An 8-year experience from a three-tier eye care network in India. *Indian Journal of Ophthalmology*, 68(3), 427–432.
- [2] Bayona, L. (2019). Radiografía de la Historia Clínica en Perú. [Tesis de maestría, Universidad Politécnica de Valencia]. Repositorio de la Universidad Politécnica de Valencia. <https://riunet.upv.es/bitstream/handle/10251/128913>.
- [3] Lu, W. C., Tsai, I. C., Wang, K. C., Tang, T. A., Li, K. C., Ke, Y. C., & Chen, P. T. (2021). Innovation resistance and resource allocation strategy of medical information digitalization. *Sustainability (Switzerland)*, 13(14). <https://doi.org/10.3390/su13147888>.
- [4] Plataforma Digital Unica del Estado Peruano, “Ley N° 30024 — Gobierno del Peru,” 2013. <https://www.gob.pe/institucion/minsa/normas-legales/240527-30024>.
- [5] Gobierno del Perú. (2020). Lineamiento que establecen las metas de implementación del SIHCE (e-Qhali), así como la remisión por parte del ministerio de salud de las bases de datos que sustenten dicha implementación. https://www.mef.gob.pe/contenidos/archivos-descarga/Lineamientos_RD007_2020EF5001.pdf
- [6] Ministerio de Salud. (2022). Diagnóstico de brechas de infraestructura y equipamiento del sector salud. <https://www.minsa.gob.pe/Recursos/OTRANS/08Proyectos/2022/diagnosticobrechas-infraestructura-sector-salud-2022.pdf>
- [7] Sathya Narayanan, V., & Kasthuri, N. (2021). An efficient recognition system for preserving ancient historical documents of English characters. *Journal of Ambient Intelligence and Humanized Computing*, 12(6), 6275–6283. <https://doi.org/10.1007/s12652-020-02201-w0>.
- [8] Zimmermannova, J., Pavlik, L., & Chytilova, E. (2022). Digitalisation in Hospitals in COVID-19 Times—A Case Study of the Czech Republic. *Economies*, 10(3). <https://doi.org/10.3390/economies10030068>.
- [9] Rai, L., & Li, H. (2021). MyOcrTool: Visualization System for Generating Associative Images of Chinese Characters in Smart Devices. *Complexity*, 2021. <https://doi.org/10.1155/2021/5583287>.
- [10] White-Dzuro, C. G., Schultz, J. D., Ye, C., Coco, J. R., Myers, J. M., Shackelford, C., Rosenbloom, S. T., & Fabbri, D. (2021). Extracting Medical Information from Paper COVID-19 Assessment Forms. *Applied Clinical Informatics*, 12(1), 170–178. <https://doi.org/10.1055/s-0041-1723024>.
- [11] Weng, Y., & Xia, C. (2020). A New Deep Learning-Based Handwritten Character Recognition System on Mobile Computing Devices. *Mobile Networks and Applications*, 25(2), 402–411. <https://doi.org/10.1007/s11036-019-01243-5>.
- [12] Tomovic, S., Pavlovic, K., & Bajceta, M. (2021). Aligning document layouts extracted with different OCR engines with clustering approach. *Egyptian Informatics Journal*, 22(3), 329–338. <https://doi.org/10.1016/j.eij.2020.12.004>.
- [13] Martínek, J., Lenc, L., & Kral, P. (2020). Building an efficient OCR system for historical documents with little training data. *Neural Computing and Applications*, 32(23), 17209–17227. <https://doi.org/10.1007/s00521-020-04910-x>.
- [14] Liang, J., Li, Y., Zhang, Z., Shen, D., Xu, J., Zheng, X., Wang, T., Tang, B., Lei, J., & Zhang, J. (2021). Adoption of electronic health records (EHRs) in China during the past 10 years: Consecutive survey data analysis and comparison of Sino-American challenges and experiences. *Journal of Medical Internet Research*, 23(2). <https://doi.org/10.2196/24813>.
- [15] Dalal, A. K., Piniella, N., Fuller, T. E., Pong, D., Pardo, M., Bessa, N., Yoon, C., Lipsitz, S., & Schnipper, J. L. (2021). Evaluation of electronic health record-integrated digital health tools to engage hospitalized patients in discharge preparation. *Journal of the American Medical Informatics Association*, 28(4), 704–712. <https://doi.org/10.1093/jamia/ocaa321>.