



Article

Development of a Mobile Application to Buy Books through Visual Recognition

Antonio Sarasa-Cabezuelo

Departamento Sistemas Informáticos y Computación, Universidad Complutense de Madrid, 28001 Madrid, Spain; asarasa@ucm.es

Abstract: Mobile devices have become the most used tool for a large number of tasks that we regularly perform such as relating them, searching for information, and in particular for making purchases. A situation that is frequently repeated in many areas is discovering an object that belongs to another person but we would be interested in being able to acquire it. However, the problem arises of knowing where to buy it. For example, this happens with the clothes that other people are wearing. Today, technology offers recognition mechanisms that can help solve this problem. This article presents an Android app that can recognize a book based on an image and offer places where it can be purchased. For this, Google technology was used to recognize objects from images and it has been combined with the information provided by Google Books to find stores that sell recognized books. In this way, a system has been created that makes it easier for any user to identify and purchase books that they discover at any given time.

Keywords: image recognition; Android app; Google books API; cloud vision API



Citation: Sarasa-Cabezuelo, A. Development of a Mobile Application to Buy Books through Visual Recognition. *Knowledge* **2022**, *2*, 41–54. <https://doi.org/10.3390/knowledge2010003>

Academic Editor: Vasile Palade

Received: 17 November 2021

Accepted: 22 December 2021

Published: 6 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent decades, the use of mobile devices has transcended the initial objective of facilitating wireless communication to become a tool used for other tasks such as the search for information, the development of social life, economic transactions, or consumption. For these tasks, different applications can be installed on the devices and that makes it easier for the user to carry out these types of tasks.

Different reasons explain the success and rapid expansion of mobiles as a tool to access different services than wireless communication itself [1]. One of these reasons is the ability to use any of these services anywhere and at any time. This flexibility is more complicated with other devices such as if you use a laptop or a conventional computer. Likewise, it is necessary to highlight the immediacy to access the information in real-time, or the simplicity of knowledge to be able to use this type of application [2].

There are a wide variety of services [3] that can be accessed through mobile applications. A type of service that is repeated in many areas is the search for a consumer item to acquire. For example, some apps allow the purchase of music, food, clothing, and other utensils. In all of them, user interaction is similar [4]. The application offers a means to search for the item, and if it is found in the catalog, then there is the possibility of purchasing it online. A common characteristic in this type of interaction is that the user knows what he is looking for [5]. For example, for the purchase of books, there are applications such as Quiero Libros [6] or My Library [7] that allow the purchase and sale of books, the registration of new books by scanning the barcode or ISBN code [8], the search for books by title, author or ISBN as well as its ordering or classification, or the creation of personal libraries, or more powerful platforms such as Google Books that, in addition to services similar to the previous ones, allows more sophisticated searches with a greater number of filters, and information and details of it are displayed for each work (and in many of them a preview of its content or even, in some cases, the possibility of reading it in its entirety and

downloading it, as well as links to platforms where they can be purchased), and provides an API that can be used to send requests with search criteria and receive the search result in JSON format.

However, a different situation is one in which the user does not know what he is looking for. For example, when he hears a song that he would like to purchase but does not know the musician or the name of the song. This situation is repeated with other elements, such as clothing, food, and others. For these cases, it would be interesting to have a mechanism that allows a reverse search, in which the content is known but the representative of the content is not. A search mechanism of these characteristics requires the ability to recognize the content to find the representative, and it will depend on the type of content. It does not have the same difficulty and the same technological requirements to recognize a song as it does to recognize a type of clothing or food. For example, book recognition could be implemented in various ways. One way would be through the textual content of the book [9], for which it would be necessary to take some text from it and carry out a search on book texts until the match was found. The main problem with this form of search is that it is computationally very complex compared to other types of searches [10]. Another more appropriate search method for this case is to perform a search using an image of the book cover [11]. The image would be used to perform recognition of the text that is in the image and extract it to implement the search by title, author, and any other data that may facilitate its location. Computationally, this form of search is more efficient than the previous one [12].

In this article, a mobile application is presented that implements a value-added service that allows the identification of books from the cover image, and offers as a result the book that has been recognized and stores in which they can be purchased.

To implement an application, an essential component is the use of an application that is capable of recognizing text in an image. Applications that allow you to recognize text in images are called OCR (Optical Character Recognition). Some of its main characteristics are the ability to obtain data such as chromatic characteristics, the relationship between an image and geographic data, to know what objects are in it, to know the context that is perceived in the image, or to discover if it appears as text in the image and knows what it is, where it is, or what it means. Its implementation is based on the use of machine learning algorithms that allow a program to be able to recognize the existence or not of certain elements in an image. To do this, a learning process is followed that consists of training the algorithm with a large number of images in which the elements to be recognized are present, whether they are objects, colors, or letters, indicating what they are and where they are. The algorithm learns to recognize what these elements are, and later, it can detect them in new images. Currently, there are online applications that offer this type of recognition service through APIs. In this sense, the service consists of sending an image to the application through the API, receiving the object that has been recognized as a response. Some examples of these types of applications are the Google Cloud Vision API [13] that detects objects and faces, reads printed and hand-written text, classifies them into predefined categories, and assigns a set of metadata to images; Microsoft's Computer Vision is a set of 3 APIs that provide services to detect and extract handwritten or printed text that appears in images (Read API detects the text content of an image and converts the identified text into a sequence of characters readable by a machine being optimized for images with large amounts of text; OCR API that works similar to the previous one but runs synchronously and is not optimized for large documents; and Recognize Text API that works similar to OCR but that runs asynchronously and uses updated recognition models); Amazon's Amazon Recognition [14] facilitates image and video analysis with deep learning technology being able to identify objects, people, text, scenes and activities in images and videos, in addition to detecting any inappropriate content and providing high precision facial analysis and capabilities of face search to detect, analyze and compare faces; or the open source software Tesseract [15] that supports various image formats and is capable of reading multipage documents. The possibilities offered by text recognition

are very varied and applicable to different areas. For example, in the medical field, it facilitates the digital transition from handwritten medical reports to electronic records, such as the Savana system [16], in the field of circular economy to retrieve knowledge from patents on how to recycle and reuse a waste [17], in the field of forensic analysis to associate the authorship of documents with people [18], or in the field of digital letters to digitize analog books or documents [19]. In some cases, it becomes a key piece to preserve analog knowledge and in others to obtain value by automating tasks.

A type of application similar in objectives to the one proposed in this article are price comparators such as Google Shopping [20] that allow the price comparison of products similar to one provided by the user. As a result, a set of products ordered by price is obtained, making it possible to consult each product's description, photos, or reviews and information on the prices of different sellers. Other similar apps are Idealo or Kelkoo. Idealo [21] allows the comparison of prices of products from Amazon, eBay, and other famous marketplaces using the name of the desired product as an input element and obtaining, as a result, a set of products together with their price, the description of the product, the number of offers they have, or reviews made to the products. Kelkoo [22] is another price comparator that allows the user to see the evolution of prices over time, and to set alerts to purchase the product when the price is lower. Finally, an application similar to the one described in this article is ASOS [23]. It is an online search engine specialized in clothing. It allows you to search for a product based on its image and display all the items that match the search, being able to compare between the different sellers and find the product that best suits the customer's wishes.

The article is structured as follows. In Section 2 the materials and methods of the application are presented. Next in Section 3, the results are described. Next, Section 4 presents the conclusions and a set of lines of future work.

2. Materials and Methods

The main objective of this application is to offer a simple, fast and intuitive search service for books that provides the most relevant information about them and offers a set of value-added services compared to traditional search engines. This objective is specified in the following more specific objectives:

- Provide a simple search engine for books based on title and author.
- Offer a search engine that takes the image of the book cover as input.
- Show a list with the possible results to choose the exact book that the user refers to see relevant information about the book at a glance, such as a title, author, description, publication date, publisher, number of pages, or its ISBN.
- Offer the possibility of reading a preview or the entire book online.
- Download the book in digital format.
- See book purchase links from the most common sellers.
- Have a favorites section.

For this, the development of a mobile application has been proposed that allows photographing the cover of a book and automatically shows all the information related to that work, as well as a preview of it, links to the main digital stores, and even the possibility to read it in its entirety online or download it in PDF or EPUB format to read in an e-book, as well as other functions such as adding titles to a favorites list.

With respect to the implementation of the application. It has been carried out using a client-server architecture (see Figure 1) where the client is the multiplatform mobile application and the server provides the client with the data corresponding to the books and their searches, as well as the rest of the services.

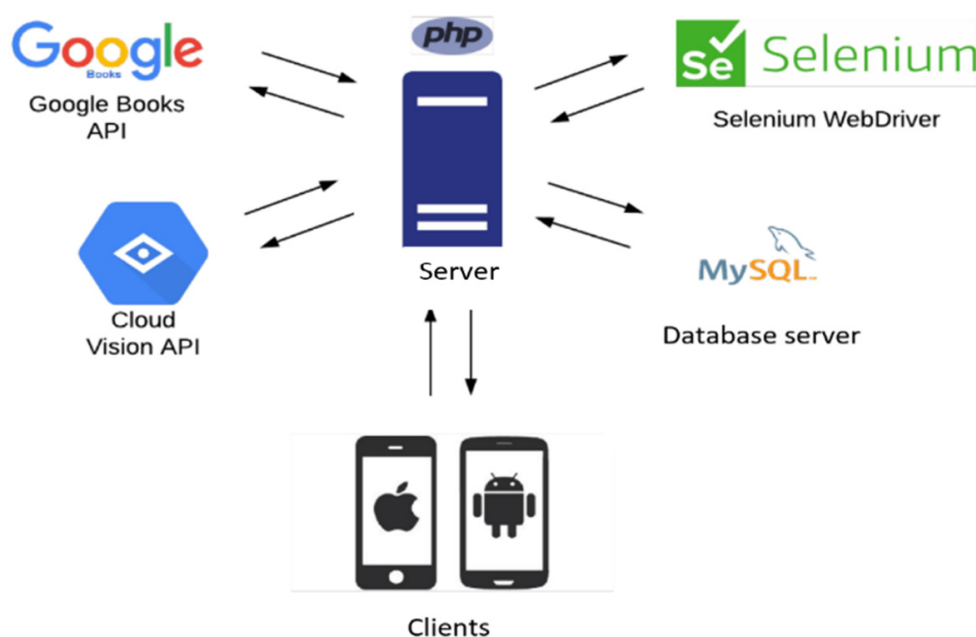


Figure 1. Architecture of the application.

The mobile application has been implemented in C # classes using the Xamarin platform [24], and the server, which corresponds to all the processing of requests, searches, and communication with the database, has been implemented in PHP files on an Apache server.

On the other hand, Google Cloud Platform [25] is a platform that facilitates access to numerous paid or free APIs to perform tasks related to computing, storage, and databases, data analysis, machine learning, administration tools, or managed infrastructure. In this project, the Vision API is used that offers machine learning models in predefined categories as well as object and face detection, identification of printed and handwritten text, or metadata mapping. To use the API, an image of the searched book must be provided, which is processed and relevant information is obtained from the image such as the text that appears in it, which identifies the author and the title of the work.

Selenium [26] is an API that is capable of automating tasks in a browser such as processes, queries, and other actions. In this sense, using this API it is possible to search for books in different stores. For each store, a request is made to the Google search engine whose request incorporates as keywords the name of the store, the title of the book, and its author. For each of the searches, the results obtained are consulted by examining its URL to check if it is really the website on which the requested copy is advertised or, on the contrary, the result is not the desired one. When the verified results corresponding to the web pages where the book is advertised are obtained, they are transformed into a JSON object that will be returned to the application so that the user can access those links.

Finally, the Google Books API [27] is used, which allows automating most of the operations that can be performed interactively on the Google Books website, such as performing full-text searches and retrieving information about books, visibility, and availability of e-books, as well as managing personal bookshelves. In the project, the author and title extracted from the user-supplied image are sent to the API to find books that match the supplied data. The API returns the books that match the search by providing data such as a description, the publisher, the number of pages, or the link to an online reader where a fragment or, sometimes, the entire book can be read. This data will be displayed in the application as a search result.

This architecture offers the following advantages:

- Server-centric administration: the client has minimal administration needs.
- Centralization of the resources: the resources of all the users are in a single server, thus avoiding the inconsistency and redundancy of the databases.

- Improved security: the chances of improper access are reduced, as there is a centralized authentication mechanism.
- Scalability of the installation: the network and its operation are not affected by the incorporation or elimination of users.

3. Results

Next, the main functionalities of the developed application will be explained.

3.1. Login and Registration

Once the application is started, the user will see the login screen, where they can log in or register (see Figure 2) if it is the first time using the application. To do this, you must click on the link shown at the bottom and that will take you to the registration screen. On the Registration screen, you must enter your email and a password, which you will be asked to repeat, and which must be at least 8 characters long. If any of the two conditions is not met, you will be shown an error message, and if the registration has been successful, you are notified by a message that an email has been sent to you from which you can activate your user to be able to log in. Once the user has registered, they will be able to return to the Login screen by clicking on the arrow that appears in the upper left corner or the Back button of the system, and they must enter their credentials to start the session.

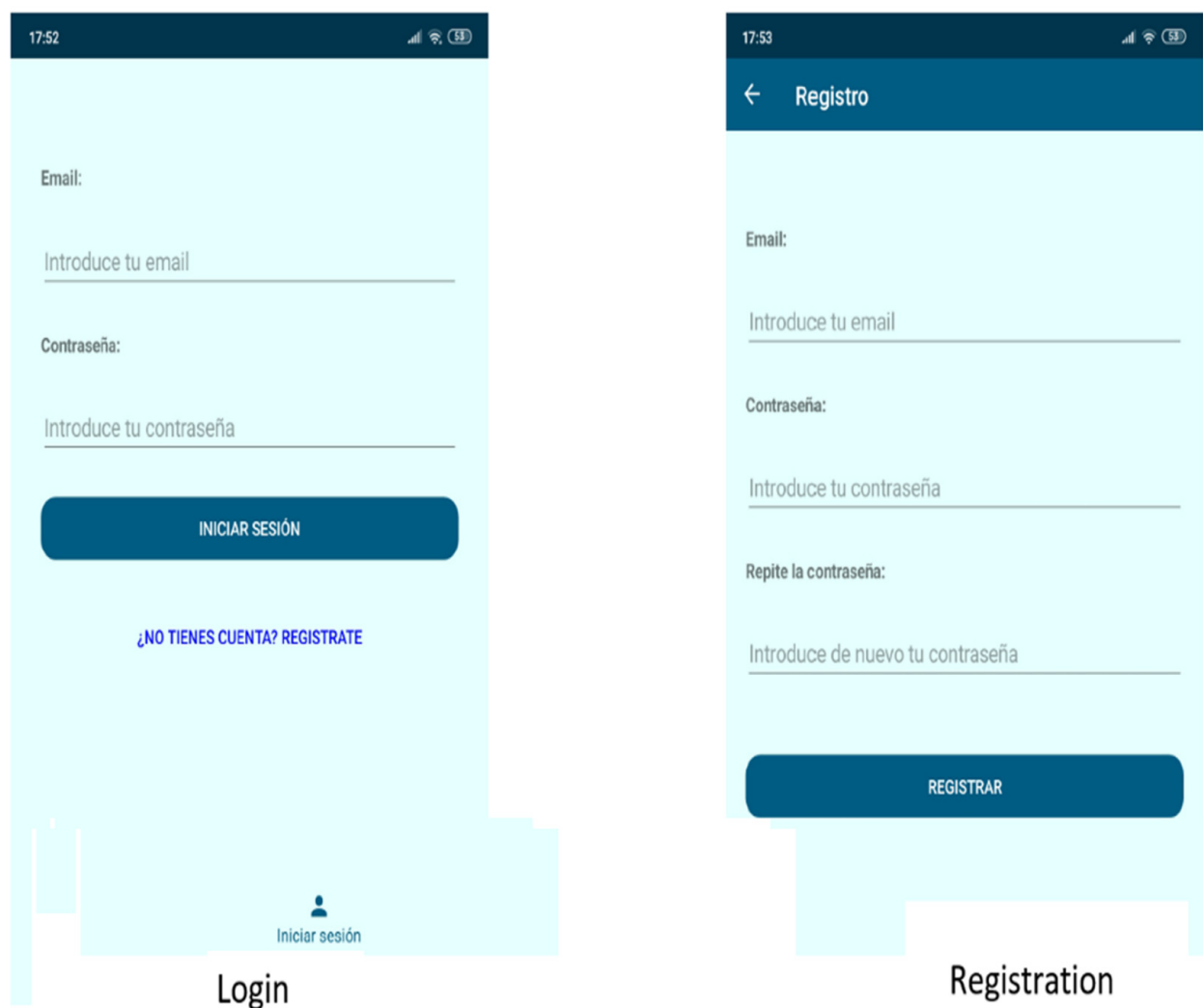


Figure 2. Architecture of the application.

3.2. Search by Image

When you log in, the application will show the Search tab, although you can see two more tabs that correspond to the favorites and profile sections (see Figure 3). In this sense, you can navigate between them by clicking on the name of the one you want to visit or by sliding your finger across the screen.

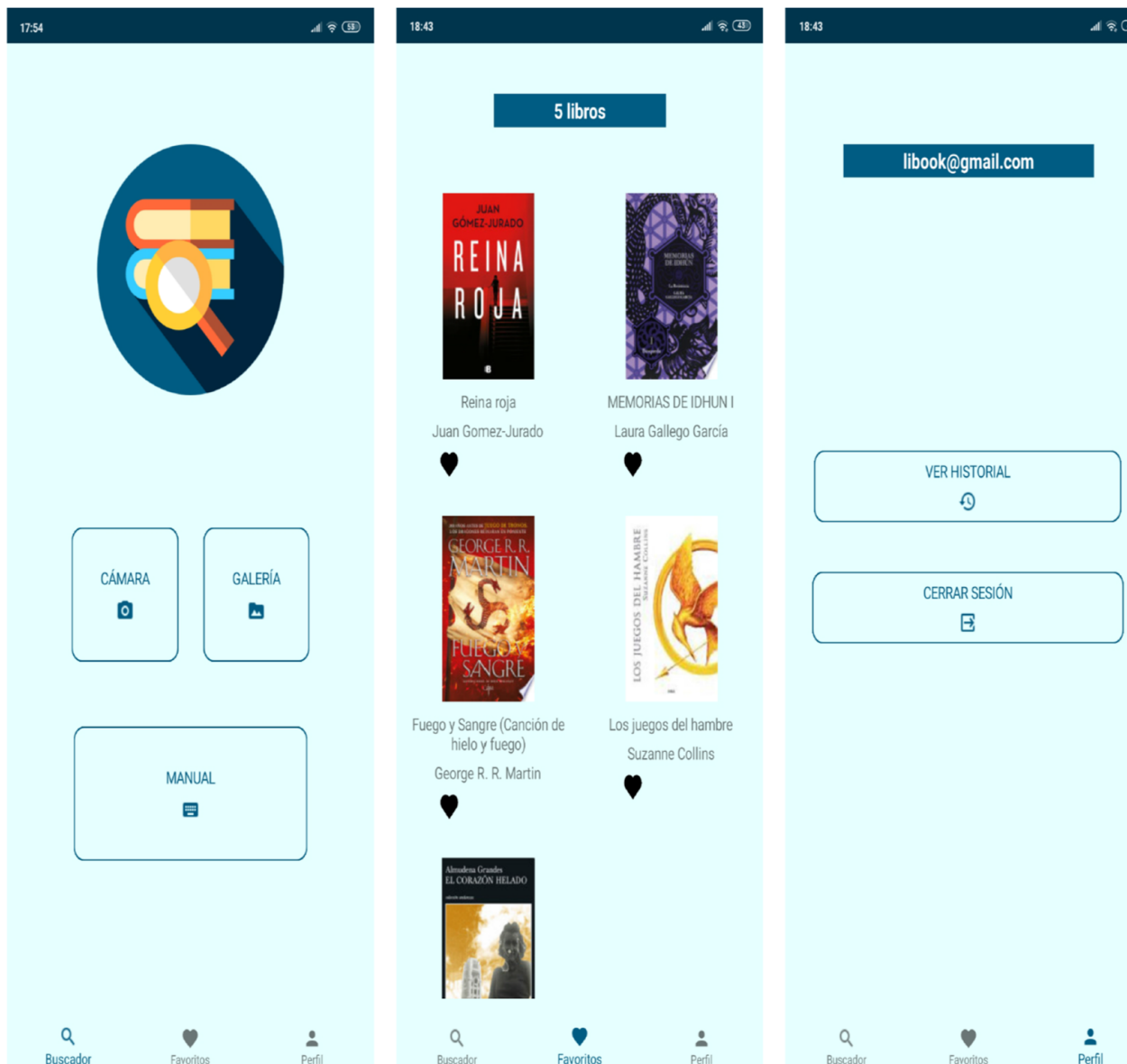


Figure 3. Application tabs.

In the first of the tabs are the three book search options, which can be through an image, taken with the camera or from the device's gallery, or by writing the title or author of the book manually. To search by image (see Figure 4), the user must click on one of the two available options and take or choose the photo. Next, a spinner will be displayed indicating that the request is being processed and the different search results will appear quickly. When the list of results has appeared, the user will be able to select the book they want to see it in detail.

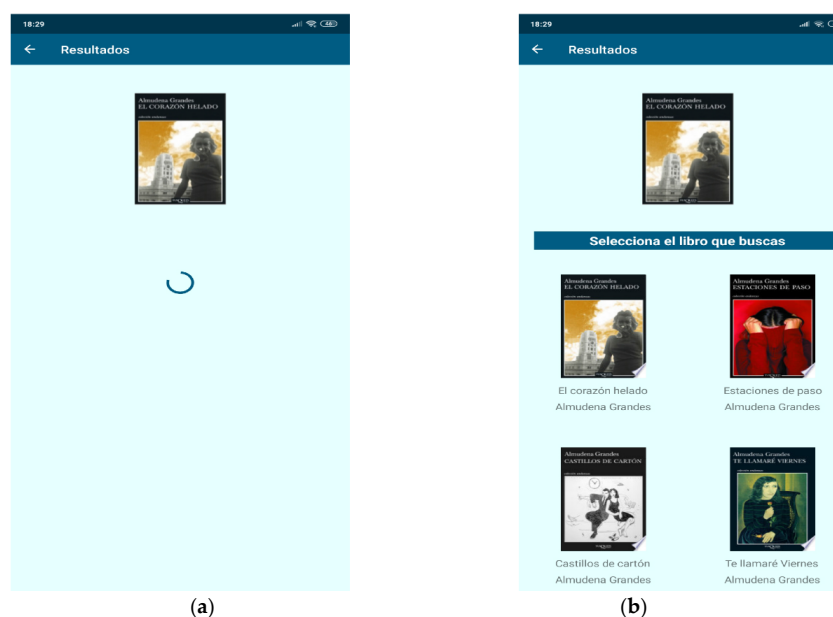


Figure 4. Search for books by image: (a) Loading the results, (b) Results list.

3.3. Manual Search

The other option to search for books is manual (see Figure 5), in which the user can enter the data related to the book they are looking for in a form. You can choose to indicate only the book, the author, or both. Once entered, you must click on the button that indicates Search, and again, a spinner will appear indicating that the search is being carried out and, later, the list with the results. On the results screen, the search parameters are indicated.

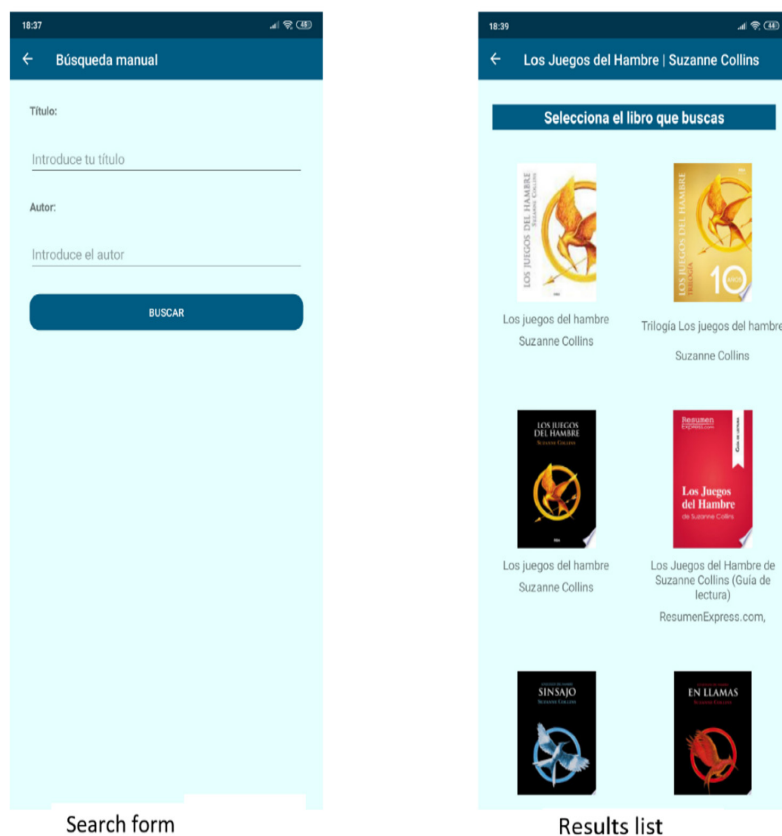


Figure 5. Manual search.

3.4. Querying a Book

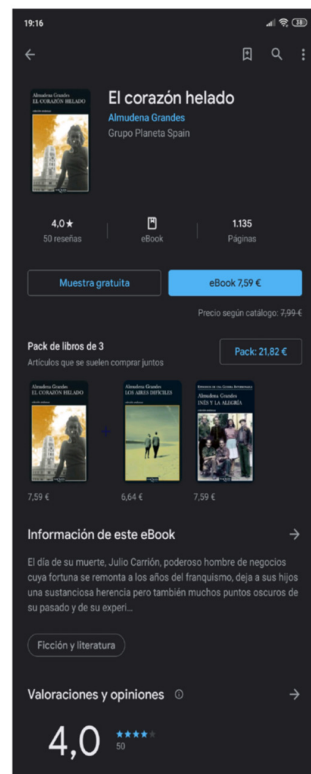
When the user is in front of a list of books, be it the results of a search, the list of his favorite books, or the history, he can select any of them to see it in detail. When you click on a book, a new screen will open with its characteristics (see Figure 6). In it, you can see the title of the book in the upper bar, an image of its cover, the title, and the author, which in turn is a link that performs a new search on that author, a button in the form of a heart to add or remove the copy from favorites, the description of the work, a table with technical details such as the publisher, the date of publication, its ISBN or the number of pages, links with purchase options for that copy and the possibility of download it in PDF or EPUB format. Although these are all the data that can be consulted about a book, not all of them will always be visible, since only those data that the specific book that is being visited will appear.



Figure 6. Querying a book.

On this screen (see Figure 7), the user can consult the different stores where the option to purchase that copy is given. To do this, you must click on the link corresponding to the

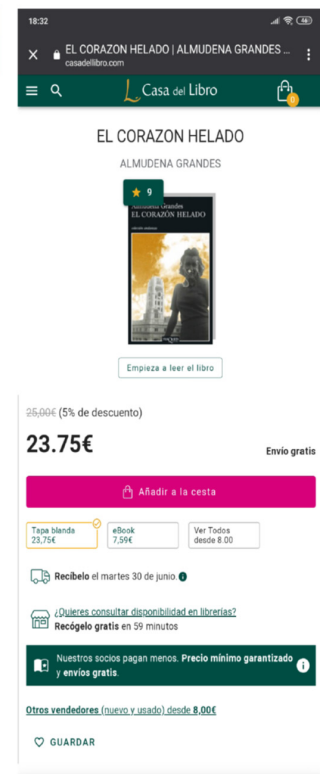
platform you want to consult. Doing so will open a browser within the application that will display the indicated website.



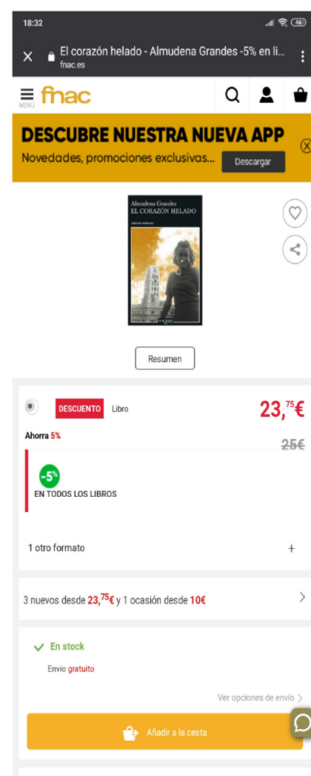
(a) Google Play



(b) Amazon



(c) Casa del Libro



(d) Fnac



(e) El Corte Inglés

Figure 7. Buying options for books.

This screen offers other functionalities (see Figure 8) such as searching for other books by this author, for which you will have to press the name of the author of this book to bring up a new screen with the results, or reading a preview of the book, or the full book if this option is available. If the latter occurs, it is indicated in the results view, as well as whether a book is available for download. When it is, you will find the section below the links to purchase the book, with an icon in PDF or EPUB that will start downloading the copy when you click on it.

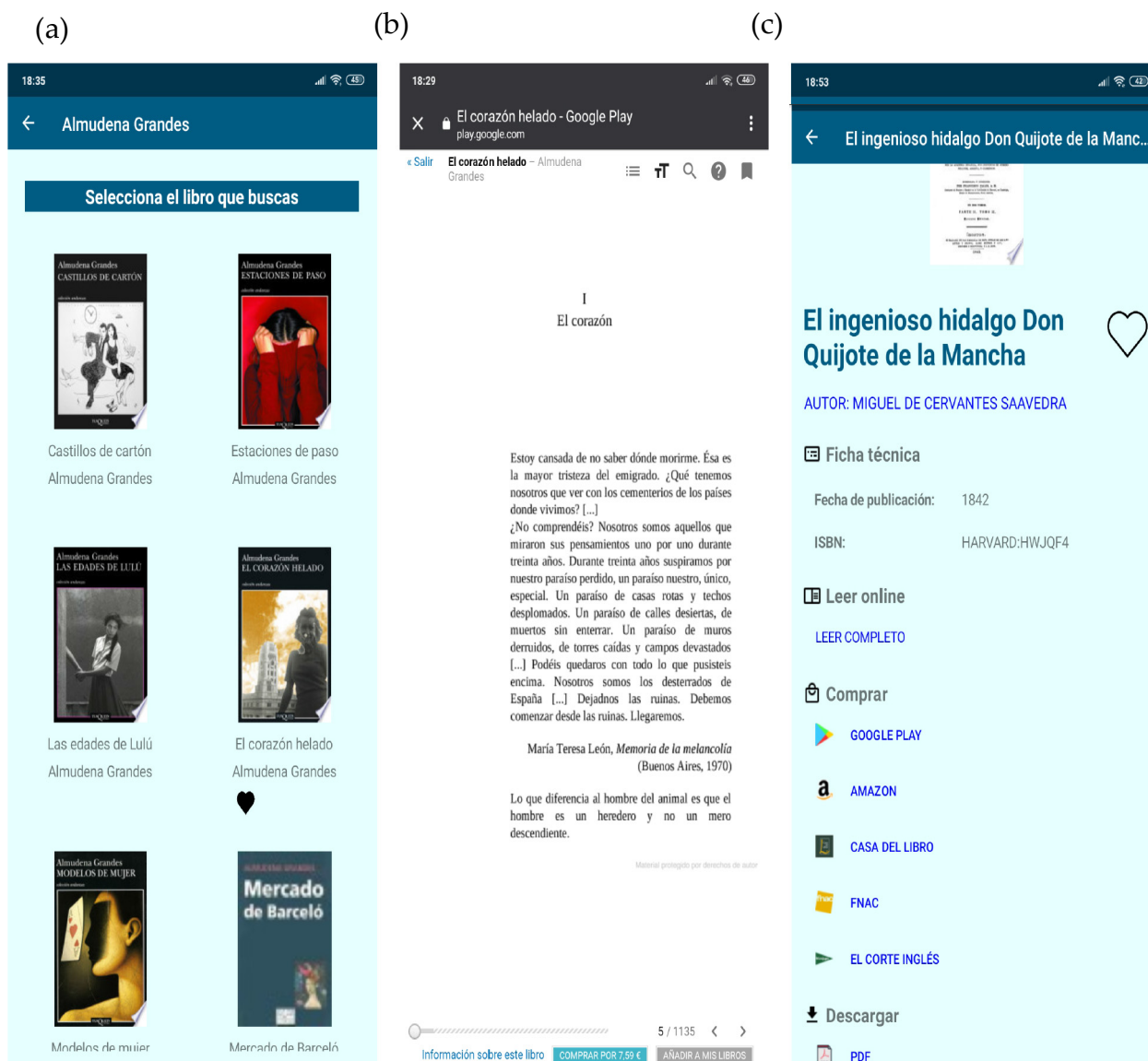


Figure 8. Functionalities in the querying of a book: (a) Search for a book by its author, (b) Preview a book, (c) Download a book.

3.5. Favorites

The user has a Favorites section (see Figure 9), in which he can consult those copies in which he has marked the heart-shaped button to include them in this category of favorites. This section shows a list with the books that the user has as favorites or a message that indicates that they have not yet marked any book as a favorite otherwise. In this list, you will be able to select any book to open the information screen of that book, and from which you will be able to perform all the usual functions, among which is to unmark the book of favorites by clicking on the heart icon again.

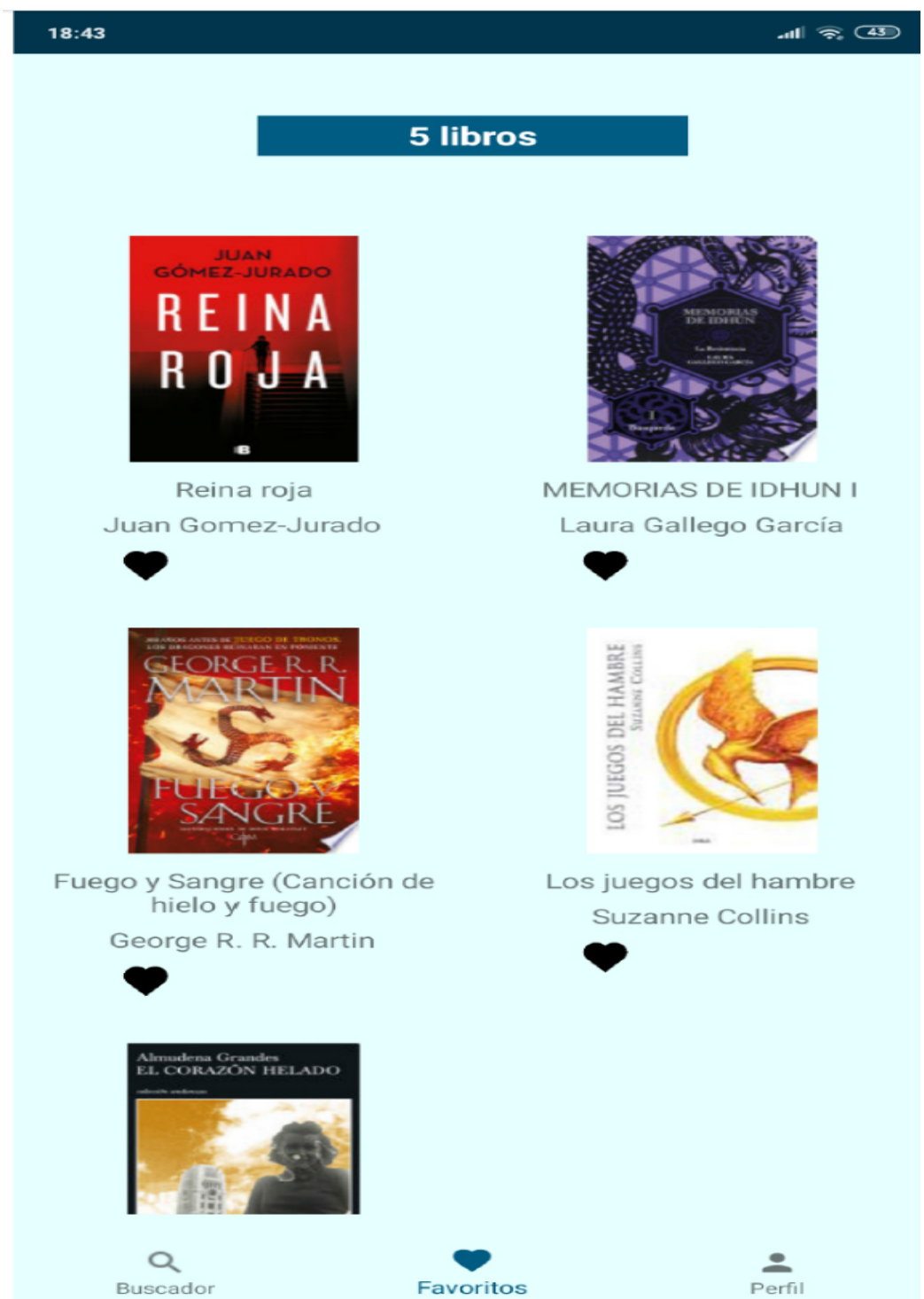


Figure 9. Favorites.

3.6. Historical Information

It is one of the two options available in the Profile section. To access it (see Figure 10), the user must click on the History button, showing a list of the latest books he has viewed, arranged chronologically starting with the most recent. In this list, the user can click on any of these copies to open the book, consult its details and choose any of its options.



Figure 10. Historical information.

4. Conclusions

In this project, a multiplatform mobile application has been described that implements a value-added service combining different existing services for the identification and search of information about books using an image of the book cover. It also offers information about stores where you can buy the books you are looking for. The application has been developed using Microsoft's Xamarin platform as well as the Vision API of Google Cloud Platform that allows the recognition of text on images of books and the Google Books API that allows the search of books. Finally, it has been possible to offer users different purchase links for books using the Selenium WebDriver API, which automates Google searches by consulting the sale link of a specific book in the main online stores. The originality of the application is not in the recognition of the images, since it is a problem that is beyond the scope of the project. However, what is original about this application is the value-added service that it implements by combining different functional components, to obtain a service that goes beyond the originals. On the other hand, the architecture of the application allows its extensibility to incorporate more services in a simple way.

Regarding future lines of work, the application can be improved by implementing new functions such as a password reset function in case the user forgets it, a more attractive and modern user interface, extending the usage statistics offered by the application, expand the history where the user can see the last books he has visited with other interesting points such as the most consulted author, a record of the books that have been downloaded, or the number of books that have been consulted in a period of time specified. Likewise, the favorites section could be improved by adding the possibility of organizing books in collections, in such a way that the user could have collections of books organized by genre, author, extension, or those he has already read. Finally, it could add functionality to recommend books similar to those you are consulting, copies that are also visited by other users with similar tastes, or which are the most visited works. To do this, machine learning algorithms would be used on the information collected about which books each user visits completely anonymously. Likewise, another possible application would be to use bibliometric analyzes [28] about the books that have been recovered or that have been most searched so that trend lists can be created on books and offered to recommend books to users. Finally, comment that no measures have been taken about the performance of the application in the sense that is discussed in [29,30] since the recognition functionality is a component external to the application that is simply used to combine it with others and therefore it is not possible to modify their behavior or improve it. However, as future work, it is proposed to use machine learning algorithms to create recommendations for users. In this context, a study will be carried out on the performance of the application. Likewise, it is proposed as future work, to carry out an analysis of the performance of the application with respect to the number of users that it is capable of supporting or the speed of response.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: I would like to thank Iván Monterrubio Cerezo for implementing the system described in the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Uğur, N.G.; Turan, A.H. Mobile Applications Acceptance: A Theoretical Model Proposal and Empirical Test. *Int. J. E-Adopt.* **2019**, *11*, 13–30.
2. Henze, N.; Pielot, M.; Poppinga, B.; Schinke, T.; Boll, S. My app is an experiment: Experience from user studies in mobile app stores. *Int. J. Mob. Hum. Comput. Interact.* **2011**, *3*, 71–91.
3. Adrakatti, A.F.; Mulla, K.R. A realistic approach to information services on mobile apps. *J. Access Serv.* **2017**, *14*, 7–15.

4. Shim, K.; Yim, J. Design and implementation of a mobile shopping app specializing in regional product. *Int. J. Multimed. Ubiquitous Eng.* **2016**, *11*, 43–54.
5. Megalingam, R.K.; Vishnu, S.; Sekhar, S.; Sasikumar, V.; Sreekumar, S.; Nair, T.R. Design and implementation of an android application for smart shopping. In Proceedings of the 2019 International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, India, 4–6 April 2019; pp. 0470–0474.
6. Silva, M.; Morais, D.; Mazedo, M.; Teixeira, L. Mobile Applications in Cultural Heritage Context: A Survey. In *Multidisciplinary Perspectives on New Media Art*; IGI Global: Hershey, PA, USA, 2020; pp. 189–216.
7. Mansouri, A.; Asl, N.S. Assessing mobile application components in providing library services. *Electron. Libr.* **2019**, *37*, 49–66.
8. McCarthy, G.; Wilson, S. ISBN and QR barcode scanning mobile app for libraries. *Code4Lib J.* **2011**, *13*.
9. Mu, Z.; Jiang, L. Online Bookstore Management System Based on Android. In Proceedings of the 2018 International Conference on Virtual Reality and Intelligent Systems (ICVRIS), Changsha, China, 10–11 August 2018; pp. 498–500.
10. Dinesh, R.; Pravin, S.A.; Aravindhan, M.; Rajeswari, D. Library access system smartphone application using android. *Int. J. Comput. Sci. Mob. Comput.* **2015**, *4*, 142–149.
11. Roy, M.B.; Kumar, N. Application of mobile technology in Library services. *Int. J. Inf. Mov.* **2017**, *2*, 168–172.
12. Chen, Y.; Sun, Y.; Zhang, F. A Crowd-Sourcing Mobile Platform for Textbook Selling and Exchange Using Information Retrieval. In *Proceedings of the 2019 Conference CS & IT Conference*; AIRCC Publishing Corporation: Tamil Nadu, India, 2019.
13. Reis, A.; Paulino, D.; Filipe, V.; Barroso, J. Using online artificial vision services to assist the blind—an assessment of microsoft cognitive services and google cloud vision. In *Trends and Advances in Information Systems and Technologies*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 174–184.
14. Mishra, A. *Machine Learning in the AWS Cloud: Add Intelligence to Applications with Amazon SageMaker and Amazon Rekognition*; John Wiley & Sons: Hoboken, NJ, USA, 2019.
15. Agbemeny, A.S.; Yankey, J.; Addo, E.O. An automatic number plate recognition system using opencv and tesseract ocr engine. *Int. J. Comput. Appl.* **2018**, *180*, 1–5.
16. Graziani, D.; Soriano, J.B.; Rio-Bermudez, D.; Morena, D.; Díaz, T.; Castillo, M.; Izquierdo, J.L. Characteristics and Prognosis of COVID-19 in Patients with COPD. *J. Clin. Med.* **2020**, *9*, 3259.
17. Spreafico, C.; Spreafico, M. Using text mining to retrieve information about circular economy. *Comput. Ind.* **2021**, *132*, 103525.
18. Marfianto, A.; Riadi, I. WhatsApp messenger forensic analysis based on Android using text mining method. *Int. J. Cyber Secur. Digit. Forensics* **2018**, *7*, 319–327.
19. Malaperdas, G. Digitization in Archival Material Conservation Processes. *Eur. J. Eng. Technol. Res.* **2021**, *6*, 30–32.
20. Broos, S.; Marcos Ramos, J. Google, Google Shopping and Amazon: The Importance of Competing Business Models and Two-Sided Intermediaries in Defining Relevant Markets. *Antitrust Bull.* **2017**, *62*, 1–18.
21. Chopdar, P.K.; Korfiatis, N.; Sivakumar, V.J.; Lytras, M.D. Mobile shopping apps adoption and perceived risks: A cross-country perspective utilizing the Unified Theory of Acceptance and Use of Technology. *Comput. Hum. Behav.* **2018**, *86*, 109–128.
22. Zaim, H.; Ramdani, M.; Haddi, A. Fuzzy-based mining framework of browsing behavior to enhance E-commerce website performance: Case study from Kelkoo. com. In Proceedings of the 12th International Conference on Intelligent Systems: Theories and Applications, Rabat, Morocco, 9–10 May 2018; pp. 1–6.
23. McLean, G.; Wilson, A. Shopping in the digital world: Examining customer engagement through augmented reality mobile applications. *Comput. Hum. Behav.* **2019**, *101*, 210–224.
24. Hermes, D. *Xamarin Mobile Application Development: Cross-Platform c# and Xamarin. Forms Fundamentals*; Apress: New York, NY, USA, 2015.
25. Wankhede, P.; Talati, M.; Chinchamalathpure, R. Comparative study of cloud platforms-Microsoft Azure, Google Cloud Platform and Amazon EC2. *J. Res. Eng. Appl. Sci.* **2020**, *5*, 60–64.
26. Gojare, S.; Joshi, R.; Gaigaware, D. Analysis and design of selenium webdriver automation testing framework. *Procedia Comput. Sci.* **2015**, *50*, 341–346.
27. Costello, E.; Bolger, R.; Soverino, T.; Brown, M. Determining textbook cost, formats, and licensing with Google books API: A case study from an open textbook project. *Inf. Technol. Libr.* **2019**, *38*, 91–99.
28. Spreafico, C.; Russo, D.; Spreafico, M. Investigating the evolution of pyrolysis technologies through bibliometric analysis of patents and papers. *J. Anal. Appl. Pyrolysis* **2021**, *159*, 105021.
29. Botchkarev, A. Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Typology. *arXiv* **2019**, arXiv:1809.03006.
30. Naser, M.Z.; Alavi, A.H. Error Metrics and Performance Fitness Indicators for Artificial Intelligence and Machine Learning in Engineering and Sciences. *Archit. Struct. Constr.* **2021**, *2020*, 1–19. [[CrossRef](#)]