Research on Application Value Analysis of Real Estate Registration Based on Big Data Mining

Yipin Yan¹*, Yuan Ren²
¹Chongqing College of Architecture and Technology, Chongqing, China, 401331
*Corresponding author e-mail: yanyipin@cqrec.edu.cn

Abstract. With the rapid development and improvement of big data, data mining, cloud computing and other technologies, real estate registration has introduced many emerging technologies. As an advanced computer multimedia information processing technology, image recognition can automatically recognize pictures in the real estate registration system using features learning, pattern recognition, and machine learning, extract text, patterns and other information on the pictures, and save these information content in the database, it realizes the functions of retrieval, query and data pre-examination of real estate registration, and improves the intelligence, information and sharing of the real estate registration system, which has important functions and significance.

Keywords: Image Recognition, Real Estate Registration System, Feature Extraction, Data Mining

1. Introduction
In recent years, with the continuous development of the big data industry, the country has also attached great importance to it and has successively issued a series of policy documents [1-3]. As an emerging means of production, digital information has become a precious and rare resource, and data mining has also become a new format, such as text retrieval, image recognition, video tracking, etc., which increase the speed of data processing. It has realized the informatization, intelligence and automation of human society [4].

As an important property related to the people's livelihood, real estate registration has attracted great attention from all walks of life. In order to improve work efficiency and quality, the competent authority relies on the four levels of the nation to conduct networked real estate registration information management. Under the premise of ensuring the safe sharing and transmission of data information, it conducts real estate registration big data research and practice, and conducts statistics, analysis and mining of massive data. And artificial intelligence processing can give full play to the resource and asset attributes of real estate registration big data [5, 6]. In the process of real estate registration, information carriers are mostly stored by pictures. This article attempts to introduce big data mining technology. In the registration process, artificial intelligence image recognition is introduced to improve work efficiency and image recognition and processing efficiency.
2. Real estate registration data integration

In order to ensure the smooth development of the municipal real estate registration work, it is necessary to obtain the original data of real estate such as state-owned land use rights, collective land ownership, collective construction land use rights, rural homestead use rights, housing property rights and other spatial information, registration books and historical registration files. Integrate to form a unified real estate registration database. The integrated technical plan is mainly positioned to guide the real estate registration agency to sort out, sort out and build a database of the original real estate registration data.

2.1. Database content

According to the types and characteristics of real estate registration data, and in accordance with unified database standards, the real estate registration information database can be divided into real estate registration basic database, registration business database, query database, comprehensive analysis database, metadata database and other databases.

1) Basic database for real estate registration. The basic database is mainly responsible for the management, maintenance and update of the real estate registration information database. Design the database structure in accordance with unified standards, and conduct unified management and maintenance of registration results, processes, essentials and other information. On this basis, support the registration business, information query business and comprehensive analysis database. The basic real estate registration database contains spatial data, attribute data and archive data of real estate units. It not only contains current data, but also historical data, supporting the retrospective of historical data.

2) Real estate registration business database. Through the registration information collection tool, the daily registration business operation real-time record registration information is stored in the registration business database, and imported into the basic database to realize dynamic updates; the registration business database is the direct support of the registration business, and the registration business review is through direct calling of the registration business. The database information is compared and checked. Based on different types of real estate registration processes and data requirements, the registration business database is integrated, and the database structure is rationally designed to meet the timeliness and security requirements of massive real estate data registration.

3) Query the business database. The query business database is a subset of the basic database, which organizes information according to service objects such as government departments at all levels and the public to ensure that information queries are conducted in strict accordance with user rights. The query business database is based on the basic database and the registered business database data, and uses automated data extraction, conversion, and correlation technology to convert and extract the basic and business data in real time or regularly, and enter the query business database. The architecture design of the query business database should consider factors such as effectiveness, security, and query speed of massive data.

4) Comprehensive analysis library. The comprehensive analysis database is the real estate registration information data warehouse, mainly for the needs of real estate registration departments to carry out comprehensive analysis of registration information, establish an index system and dimensional model, and extract relevant information from the basic database through extraction, conversion, and leading to support comprehensive analysis.

The main content of the real estate registration database is divided into attribute data, graphic data, and scanned documents. Attribute data includes parcel information and real estate registration business information; graphic attributes include parcel graphic information, land graphics, contract management rights graphics, forest rights graphics, real estate and structure graphics; scanned documents are electronic scanned copies of various attachment materials.

2.2. Data organization

Data should be organized into a logically seamless whole, and continuous entities can be separated into different storage spaces and storage units for storage in physical storage.
(1) Division of data organization

When designing a database, according to different business applications, data organization is divided into operational data and analytical data. Their organizational form actually originates from and acts on two systems: operational systems and analytical systems.

In the design of the real estate registration information database, multi-dimensional information query requirements for data are generated for different systems, forming a separation of the data environment for transaction processing and the data environment for data analysis.

Operational systems center on traditional databases for daily business processing. For example, in the real estate registration information system, the real estate registration department handles each registration business.

The analytical system analyzes the associations and laws behind the data with the data warehouse as the center, and provides a reliable and effective basis for the decision-making of real estate registration. For example, in the real estate registration comprehensive analysis system, the real estate registration authority analyzes the registration data, realizes the near real-time summary of massive data, and generates various basic reports by day, month, quarter, and year.

The users of the operational system are usually the specific real estate business registration personnel, and the processed data is the detailed information of the real estate registration business, and its goal is to achieve business registration; while the users of the analytical system are usually personnel engaged in management or data analysis, including The information is the macro information of real estate registration rather than specific details, and its purpose is to provide supporting information for decision makers in various departments.

The distinction between operational data and analytical data has clarified the boundary between the analytical environment and operational environment for data processing, and has developed from a single database-centric data environment to a database-centric business processing system and Database-based analysis system.

(2) The difference between operational data and analytical data

The data is divided into operational data and analytical data, mainly showing the following 4 aspects. One is that the operational data needs to use complex relationships to ensure the speed, consistency and real-time performance of business data processing; the second is to conduct a large amount of data analysis in the operational data, which will affect the processing speed and performance of the transaction; When complex queries are executed in operational data, it will affect the efficiency of analysis and the execution of decisions due to the slow speed; finally, operational data will affect the consistency of data analysis due to frequent changes.

Operational system is the processing of query, addition, deletion and modification of basic data. It is based on database, while analytical system is more suitable for data analysis and processing based on data warehouse. The historical, derived and comprehensively refined data in the analytical system all come from the underlying database on which the operational system depends. Analytical system data requires one step more data multidimensionalization or pre-comprehensive processing than operational system data to establish different levels of statistical data to meet the requirements of rapid statistical analysis and query. In addition to the differences in data and processing, the interface style and data access methods of analytical system front-end products are also different from those of operational systems. Analytical systems mostly adopt methods that are easy for non-data processing professionals to understand (such as multi-dimensional reports and statistical graphics). ) query and data output are intuitive and flexible; and the operating system is mostly a fixed form frequently used by operators, and the query and data display are also relatively fixed and standardized.

3. Analysis of image recognition function of real estate registration system

There are many real estate registration operation procedures, and the number of input objects is also very large. The traditional manual and semi-manual processing mode. Window workers use scanners and high-speed cameras to collect images. The efficiency is low and the quality is not high, which is not conducive to improving government agencies. Service Level. Therefore, in order to improve the
efficiency and level of real estate registration, this article conducts a prototype analysis of the real estate registration system and uses advanced artificial intelligence technology to analyze the image recognition function of the real estate registration system. The main functions are as follows:

3.1. Image acquisition function
During the real estate registration process, the applicant can pass the content of the document through multiple channels, including the Internet, WeChat, Email, etc. After joining the artificial intelligence, the applicant can directly upload the application materials to the platform through the Internet and other channels, and import these images to the system.

3.2. Image preprocessing function
During the image recognition of the real estate registration system, due to the uneven light of the photos taken by the applicant, a lot of noise data is generated during image recognition. Binary method, threshold method, etc. need to be used to preprocess the image to reduce the noise data during image recognition, obtain the image content more accurately.

3.3. Image recognition function
Image recognition can use machine learning techniques such as support vector machines, BP neural networks, genetic algorithms to extract features, and then classify the feature content, divide the image into background content and target objects, and then process the image and combine these content with the database. Comparing the saved patterns in, it can better identify the content submitted by the applicant.

3.4. Image data saving
After the image recognition process is completed, the recognized content can be classified and saved in the database, and then the recognition results can be submitted to the system. The system will complete the pre-review through artificial intelligence. If the pre-review is passed, the paper can be submitted at the processing window to achieve one-window processing. Knot.

4. Real estate registration system image big data mining
The image recognition function of the real estate registration system uses automated processing technology to receive real estate registration images from the system, and then use the recognition function embedded in the system to analyze the image content. The image recognition of the real estate registration system can use cloud computing and service-oriented architecture technology to provide users with a concurrent processing platform and provide a set of interface components for system user operations. These components have strong portability, concurrency, and accessibility. It is very easy to operate and deploy, and has important functions. Specifically, the system processing architecture is shown in Figure 1.
5. Research on key technologies of image recognition processing in real estate registration system

The real estate registration system hides image recognition technology. These technologies are also an important research direction of current artificial intelligence. They can effectively complete image information collection, feature detection, feature extraction, pattern matching, semantic analysis and other technologies to obtain the image content in the real estate registration system. Has an important role and significance. Specifically, there are many key technologies in the real estate registration system, including four aspects: feature detection, feature extraction, pattern matching, and semantic analysis.

The image recognition of the real estate registration system can use cameras and sensors to collect external video images, text and voice information, or directly accept network upload data, and then use feature detection technology to preprocess this information to obtain video images or voice text. Target object, calibrated the outline size of these target information. For example, in an image or a frame of video, it contains very rich content, and there are many features of objective objects, such as histogram features, various templates, color features, structural features, etc., feature detection can be valuable. The information is picked out, and these features are used to realize object detection. At present, there are many image detection methods that can be used for image recognition, such as neural networks, support vector machines, Adaboost, etc. These algorithms are a classification method that can integrate weak classification methods together to build a strong classification method. Image detection can use the detection algorithm to select some rectangular features of the target object, build a strong classifier according to the weighted voting method, and then through training and learning, the strong classifiers can be connected in series to form a cascade structure of stacked classifiers, improve the detection speed of the classifier. Feature extraction is another key technology introduced in image recognition. Since there are many features contained in video images and voice texts, these features can be extracted after detection. Feature extraction is the process of image recognition learning and modeling. It is also called image representation description. The feature extraction methods can be divided into two categories after classification. They are based on knowledge representation, and the other is based on algebraic features or statistical learning. Characterization method. The knowledge-based representation method can obtain characteristic data according to the contour shape and distance measurement characteristics of the target object. The types of this characteristic distance measurement include curvature angle, Euclidean distance, etc. There are
many contour characteristics of the target object, and these local features can be decomposed one by one. The geometric structure between local features extracts common knowledge features. The characterization method based on algebraic features and statistical learning is to use the density of the target object pixel for statistics, and then can make full use of the strict rules of statistics to achieve feature separation and extraction. The current feature extraction methods based on algebra include K-means, density clustering, Spectral clustering, support vector machine.

Given a set of images \( I = \{i_1, i_2, \ldots, i_n\} \), a weight \( w(i_j) \) is assigned to each page \( i_j \), which is used to indicate that the importance of page access is different. The weight calculation formula is as follows:

\[
w(i_j) = \frac{A Time(i_j)}{\sum A Time(i_j)}
\]

Among them: \( A Time(i_j) \) is the average access time of image \( i_j \); \( \sum A Time(i_j) \) is the sum of the average access time of all images; \( 0 \leq w(i_j) \leq 1, j = \{1,2,\ldots,n\} \).

After the features are extracted, image recognition can match these features with the information stored in the pattern library. Matching is also a process of pattern search and comparison. Image recognition can set a pattern matching operation threshold. For example, similarity can be used as a metric, and then a threshold can be set. After the threshold is set, comparisons can be made. If the similarity exceeds the specified threshold, the pattern matching is determined to be successful. At present, the most commonly used algorithm for pattern matching is the BM algorithm. This algorithm is an improvement of the exact string matching algorithm. It introduces the key function of fuzzy matching. It can compare data from right to left. At the same time, it uses heuristic comparison rules. It is the bad character rule and the good suffix rule to determine the distance to jump to the right. The BM algorithm achieves fast pattern matching and good concurrency. It is more suitable for the current real estate registration system and plays an important role.

After successful pattern matching, image recognition needs to convert machine language into natural language that people can understand, so that semantic analysis technology can be used to provide people with the recognition results and help people make decisions, such as issuing next operation instructions. In the process of semantic analysis, the system needs to review each character identified, and compile the characters. The compilation can implement logical processing according to the calculations specified by the system, translate into natural language, and realize human-computer interaction.

6. Conclusion

With the increase in real estate registration, the number of photos taken by scanners and cameras is increasing. This article relies on big data mining technology, builds image recognition functions, mines hidden object features in images, and uses pattern matching for content analysis to further improve real estate registration Efficiency, aimed at providing theoretical support for real estate registration, has an important role and significance.

References
[1] Nora D, Sonia P. Challenges for the registration of vaccines in emerging countries: Differences in dossier requirements, application and evaluation processes[J]. Vaccine, 2018, 36(24):3389-3396.
[2] Ashton G P, Harding L P, Parkes G M B, et al. Application of hot - stage microscopy direct analysis in real time mass spectrometry (HDM) to the analysis of polymers[J]. Rapid Communications in Mass Spectrometry, 2019,5(8):1-10.
[3] Bilbao-Terol A, Arenas-Parra M, Canal-Fernandez V, et al. Multi-criteria analysis of the GRI sustainability reports: an application to Socially Responsible Investment[J]. Journal of the
Operational Research Society, 2018, 69(10):1576-1598.

[4] Al Ghaithi R, Mori J, Nagy Z, et al. Evaluation of the Total Thrombus-Formation System (T-TAS): application to human and mouse blood analysis[J]. Platelets, 2018, 4(8):1-8.

[5] Fanny, Weber-Goericke, Markus, et al. A quantitative meta-analysis of fMRI studies investigating emotional processing in excessive worriers: Application of activation likelihood estimation analysis.[J]. Journal of Affective Disorders, 2018, 3(2):105-113.

[6] Dittrich R, Butler A, Ball T, et al. Making real options analysis accessible for climate adaptation. An application to afforestation as a flood management measure in the Scottish Borders[J]. Journal of Environmental Management, 2019, 245(9):338-347.